

21世纪大学计算机规划教材



计算机科学 基础

◆ 陆汉权 主编 ◆ 冯博琴 主审



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY <http://www.phei.com.cn>



内 容 简 介

本书是浙江大学计算机基础课程建设成果。本书共10章,第1~3章重点介绍计算机的基本知识,数制、编码和数字逻辑基础,以及计算机的硬件组成。第4~6章介绍计算机的操作系统、软件、算法等,其中,算法是计算机科学的核心内容。第7章介绍数据库相关知识。第8章介绍网络,包括因特网的基本知识。第9章介绍有关计算机的一些高级主题,包括高性能计算、人工智能和计算理论。第10章讨论有关信息社会的问题,包括有关法律、版权、环境、健康、计算机安全、病毒、黑客等。本书免费提供电子课件。

本书适合作为大学计算机基础课程教材,也可以作为计算机专业的导论课程教材。

未经许可,不得以任何方式复制或抄袭本书的部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

计算机科学基础/陆汉权主编. —北京:电子工业出版社, 2011.8

21世纪大学计算机规划教材

ISBN 978-7-121-13907-9

I. ①计… II. ①陆… III. ①计算机科学—高等学校—教材 IV. ①TP3

中国版本图书馆CIP数据核字(2011)第124826号

策划编辑:章海涛

责任编辑:冉 哲

印 刷:

装 订:北京中新伟业印刷有限公司

出版发行:电子工业出版社

北京市海淀区万寿路173信箱 邮编 100036

开 本:787×1092 1/16 印张:17.25 字数:441.6千字

印 次:2011年8月第1次印刷

印 数:6000册 定价:36.00元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888。

质量投诉请发邮件至zltz@phei.com.cn,盗版侵权举报请发邮件至dbqq@phei.com.cn。

服务热线:(010) 88258888。

教师的集体贡献。浙江大学从2002年就开始将计算机科学作为大学计算机基础课程的核心，其教材也经过几次修改，本书就是近十年来教学实践的成果。作者一直坚持认为，大学不应该把如何操作计算机作为教学内容，而应该引导学生去认识计算机的原理，这样才能更好地在学习和专业中使用好计算机。例如，学习程序设计的目的，不是成为程序员，而是认识程序原理。一个非计算机专业的学生，肯定是计算机的用户，知道计算机能够做什么、不能够做什么，这应该是最基本的专业素质。因此任何专业的同学都需要学习计算机科学基础，这也是其专业需要。

本书作者近似固执地认为，课程不能通过“工程”的方法去建设，而是要依靠教师引导学生去思考。大学提供给大学生发现问题、讨论问题、争论问题、解决问题的平台，所以作者要说的是：教课及教师并不是教学的核心，特别在计算机课程中，需要“发现、讨论、争论”的问题很多，收获也更多。

本书为任课教师免费提供电子课件，有需要者可登录华信教育资源网<http://www.hxedu.com.cn>注册下载。

毫无疑问的是，由于计算机科学的特点，其概念和某些名词的定义可能随着时间流逝而变化，并不像数学、物理科学那样有极为严格的、不变的定义，甚至，从不同的角度表达问题可能有完全不同的说法，这些在书中是常见的。当然，由于作者的水平所限，错误肯定难免，希望使用本书的教师和学生能够不吝赐教。

教育部计算机基础教学指导委员会副主任委员、国家名师、西安交通大学冯博琴教授为本书的编写提供了极为细致的指导，并亲自担任本书的主审，在此作者深表感谢。也借此感谢作者在浙江大学的同事，他们为本书提供了许多宝贵的意见。电子工业出版社的章海涛策划编辑，他为本书的写作和出版付出了很多，在此仅以此文字表示真诚的谢意。同时也十分感谢本书的责任编辑冉哲为本书的顺利出版所付出的辛劳。

陆汉权

浙江大学计算机学院，luhq@zju.edu.cn

2011年于浙大玉泉

3.3.7	PC机的存储器	53
3.4	输入/输出系统	53
3.4.1	端口	53
3.4.2	输入设备	54
3.4.3	输出设备	55
3.4.4	其他输入/输出设备	58
3.5	系统连接和USB	58
3.5.1	系统总线标准	58
3.5.2	USB	60
3.5.3	接口和输入/输出方式	61
	本章小结	62
	习题3	63

第4章 操作系统..... 67

4.1	操作系统的定义	67
4.2	操作系统的类型	68
4.3	常见的操作系统	70
4.3.1	MS-DOS	71
4.3.2	Windows	71
4.3.3	UNIX/Linux	71
4.3.4	Mac OS	72
4.3.5	移动设备操作系统	72
4.4	操作系统的结构和组成	73
4.4.1	操作系统的层次结构	73
4.4.2	操作系统的组成	74
4.4.3	进程管理	74
4.4.4	存储器管理	77
4.4.5	设备管理	79
4.5	Windows操作系统	81
4.5.1	概述	82
4.5.2	系统结构	82
4.5.3	系统管理	83
4.5.4	启动和停机	84
4.6	文件管理	84
4.6.1	文件和文件系统	85
4.6.2	给文件命名	85
4.6.3	文件扩展名和通配符	86
4.6.4	常用的文件类型	87
4.6.5	目录结构	89
4.6.6	文件存取	89

即使很复杂的问题，也应该通过分解这些问题借助于计算机来帮助完成任务。如果用户能够有这样的能力，也就是说具有了“计算思维”。虽然这种说法未必准确，认识计算机，不但要知道计算机“如何工作”，更需要知道计算机“能够”做的工作。

按照专家的观点，计算思维（Computational Thinking）是指运用计算机科学的基础概念进行问题求解、系统设计以及人类行为理解等涵盖计算机科学之广度的一系列思维活动。计算思维建立在计算过程的能力和限制之上，由人或机器执行。计算思维的本质是抽象（Abstraction）和自动化（Automation）。如果说，数学思维是“对象与关系”，那么计算思维则表现在“状态和过程”上。所谓抽象，就是要求能够对问题进行抽象表示、形式化表达（这些是计算机的本质），设计问题求解过程达到精确、可行，并通过程序（软件）作为方法和手段对求解过程予以“精准”地实现，也就是说，抽象的最终结果是能够机械式地一步步自动执行。

表1-2中列出了计算思维倡导者之一的周以真教授（卡耐基·梅隆大学计算机系）所总结的有关计算思维是什么和不是什么的描述。可以看到，这里面提出的观点已经超越了计算机科学本身，也就是说，它成为一种自觉的、自然的思维方式。换句话说，具有计算思维能力，也就是人们理解计算机处理问题的能力应该成为一种常识，如同人们对数学的理解一样。

表1-2 计算思维的特征

	计算思维是什么	计算思维不是什么
1	是概念化	不是程序化
2	是根本的	不是刻板的技术
3	是人的思维	不是计算机的思维
4	是思想	不是人造物
5	是数学与工程思维的互补与融合	不是空穴来风
6	面向所有的人，所有的地方	不局限于计算学科

事实上，无论用什么观点进行分析和解释，一个儿童从对计算机的喜爱到离不开计算机，最终他对计算机的认识肯定会随着年龄的增长、阅历的提升和知识的积累，自然而然地知道如何运用计算机去帮助他解决遇到的各种问题。当然，理想的情况是，他使用计算机处理问题的选择每次都是正确的。

本章小结

本章介绍了有关计算机及其相关的一些基本知识。

计算机是一种现代化的信息处理工具，它对信息进行处理并提供结果，其结果取决于所接收的信息及其相应的处理算法。

计算机科学技术研究的内容包括计算机的设计、制造，以及运用计算机进行信息的获取、表示、存储、处理、控制等的相关理论、原则、方法和技术。计算机科学研究现象并揭示其规律，而计算机技术则侧重于研制计算机和使用计算机进行信息处理的方法和手段。



$(75.612)_8, (64A.C3F)_{16}$

11. 什么是原码？什么是补码？什么是反码？为什么要定义原码、反码和补码？能够对十进制数定义反码和补码吗？

12. 写出下列各二进制数的原码、补码和反码：

$0.11001, -0.11001, 0.11111, -0.11111$

13. 在计算机中如何表示小数点？什么是定点表示法和浮点表示法？

14. 若将一个无符号的二进制数向左或向右移动 n 位，则所得到的数和原数之间是什么关系？

15. 设有一台浮点计算机，使用16位表示浮点数，最高位为符号位，尾数为8位定点小数，阶码位为7位，其中阶码的最高位为阶码符号位，问这个浮点数能够表示的数范围是多少？

16. 什么是ASCII码和Unicode编码？

17. 什么是逻辑运算，基本逻辑运算有哪一种？

18. 列出下列函数的真值表：

$$(1) F = \overline{A}B + A\overline{B}$$

$$(2) F = ABC + \overline{ABC}$$

$$(3) F = A + B + C$$

$$(4) F = \overline{A}BC + A\overline{B}C + AB\overline{C}$$

19. 试证明下列逻辑表达式是成立或者不成立的。

$$(1) \overline{AB} + \overline{AB} = \overline{AB} + \overline{AB}$$

$$(2) A + BC = (A+B)(A+C)$$

$$(3) A + \overline{AB} = A + B$$

四、思考题

1. 数据表示不但是一个技术问题，也是一个社会问题。使用编码对各种对象进行标记，但必定会产生某些有意无意的歪曲。例如，对某一社会问题进行问卷调查，而问卷中的答案使用等级，给每个等级赋值进行统计分析，你认为这种“量化”是合适的吗？为什么？

2. 计算机中为了表示数，有多种编码方法。你可以通过书籍、杂志或者网络查找有关资料，了解以下编码：BCD8421，BCD2421，余3码，格雷（Gray）码，EBCD。以BCD8421为例，说明它是如何进行加法计算的。

3. 多媒体数据格式尚无统一的标准。请收集有关资料，汇总一下音频数据格式有多少种，各有什么特点。

4. 数据压缩是计算机中最常用的方法，较为著名的压缩软件有RAR、ZIP、ISO等。请找一找相关软件和资料，归纳其数据压缩使用的技术。图像压缩采用的是另一种技术路线，也请你对其进行总结，看看有多少种数据压缩格式。为了便于网络传输，又使用了哪些图像数据格式？

5. 请查找并收集有关存储器的资料，看看各类存储器采用了哪些逻辑设计技术。



4. 协处理器

20个世纪80年代, 计算机科学家就开始考虑将处理器承担的特定任务分离出来, 交由专门设计的处理器去完成。例如, 用于完成数值运算的数学处理器, 用于图形处理的处理器等, 这一类处理器不单独工作, 它们在CPU的协调下完成任务, 因此称为协处理器。现在高性能的处理器中都带有能处理浮点运算的协处理器。

5. 内部高速缓存器

计算机在运行过程中, CPU与存储器之间的数据交换最为频繁。CPU是计算机中速度最快的部件, 而存储器的工作速度要低于CPU, 这就意味着, CPU读取存储器中的数据需要有等待过程。计算机科学家研究发现, 如果把要执行的程序代码和数据先行调入CPU暂存, 就能提高CPU的运行效率。为此设计者在CPU内部开辟出一个存储区域, 用于存放将要执行的程序代码, 它和CPU以相同的速度运行。这个CPU内部存储器叫做高速缓存器 (Cache)。

在CPU内的那一部分存储器叫做L1缓存 (Level 1 Cache)。后来在CPU外部和存储器之间也使用了一种速度快于存储器但稍慢于CPU的L2缓存 (Level 2 Cache)。最新的CPU技术将L1、L2缓存都设计在CPU芯片内, 外置的是L3缓存 (Level 3 Cache)。缓存对CPU的运行效率有直接影响, 对各级缓存的调度也是由CPU负责完成的。

与CPU性能相关的还有其他指标, 例如, 工作温度范围、电源电压范围、芯片封装材料及结构等。由于CPU的高集成度使得芯片内部升温很快, 因此需要在芯片的外壳上加装散热器或者小型散热风扇。现在的计算机能够自动测试CPU内部温度并提供了过温保护。

3.2.4 CISC和RISC

作为商品化的产品, 目前市场上有两种类型的处理器系统: CISC (Complex Instruction Set Computer, 复杂指令集计算机) 和RISC (Reduced Instruction Set Computer, 精简指令集计算机)。RISC和CISC是两种完全相反的设计方法, 其设计目的都是为了提高计算机的性能。从名字上可以知道, 它们的主要区别是处理器所拥有的指令的数量不同。

1. CISC

CISC体系的设计思路就是使用数量和种类较多的指令, 包括复杂指令。它的优点是, 程序设计比较容易, 因为每一个简单的或者复杂的操作都有相应的指令可以实现。PC机上使用的Intel公司的系列处理器就是CISC体系结构。

2. RISC

处理器设计的另一种观点是, 如果处理器只包含那些常用指令且指令的长度和执行时间都相同, 那么程序控制逻辑将得到简化, 处理器的速度也会得到很大的提升。按照这种观点设计出来的处理器被称为RISC。

典型的CISC有300条以上的指令, 而RISC一般使用的指令数目在100条以内。典型的RISC处理器是Apple公司生产的PowerPC, 还有Sun公司生产的Ultra SPARC [Sun公司于1982年成立, 公司的名字是Stanford University Network 的缩写。1995年, Sun公司在C语言的基础上开发了Java语言。2010年被Oracle (甲骨文) 公司收购]。

使用RISC能否达到提高计算机性能的目的, 视计算机承担的任务而定。如果以计算为主, 则可能使用CISC更好。但对于计算简单的任务, 如网络服务, RISC的性能表现更好。

3.3 存储器

计算机硬件的第二个子系统是存储器。存储器的一个重要特性是复制性，即其存放的数据被取出后，原来存放的数据依然存在，因此可以被反复取用。组成存储器系统的是多种类型的存储元件。本节主要介绍存储器的存储模式，介绍各种存储元件（器）的原理和存储器系统的主辅结构，以及PC机的存储器系统。

3.3.1 存储模式和存储容量

如图3-3所示，存储器有内存和外存两个部分。无论内存、外存采用何种存储原理，其本质都是存储数据，因此计算机采用统一的存储模式。

计算机存储模式规定，存储单元以字节为单位，1字节由8位二进制位组成。存储器由若干字节组成，每字节都有一个唯一的标识叫做存储器地址，且存储器地址也按二进制位模式进行标识。例如，10位二进制位地址能标识存储单元 2^{10} 个，即1024个，如图3-8所示，图中分别列出了存储单元和对应的二进制位单元地址。

存储容量即存储器中存储单元的总数，也叫做字节数，或者称为地址空间。由于存储容量较大，因此用量词表示，常用的存储单位量词如表3-1所示。这里，存储器的千、兆等量词并非是存储器的实际大小，只是一种较为直观的近似表示。

十进制地址	二进制位单元地址	单元内容
0	0000000000	0 1 0 1 0 1 0 1
1	0000000001	1 1 0 0 1 1 0 0
2	0000000010	1 0 1 1 0 1 0 0
	⋮	
1021	1111111101	0 0 1 1 0 0 1 1
1022	1111111110	1 0 0 1 0 0 1 1
1023	1111111111	0 1 1 0 0 0 1 0

图3-8 存储器地址和单元

表3-1 存储单位量词

量 词 缩 写	量 词 描 述	实际字节数	近 似 表 示
KB (KiloByte)	千字节	$2^{10}=1\,024$	10^3
MB (MebiByte)	兆 (百万) 字节	$2^{20}=1\,048\,576$	10^6
GB (MegaByte)	吉 (十亿) 字节 (千兆)	$2^{30}=1\,073\,741\,824$	10^9
TB (TeraByte)	太拉 (万亿) 字节 (兆兆)	$2^{40}=1\,099\,511\,627\,776$	10^{12}

程序为了获取数据，需要对存储单元地址进行定位操作，因此地址数据也是程序处理的对象。如同在第2章中介绍的用数制、编码表示的位模式一样，存储器地址也采用位模式。计算机将地址数据当做无符号数，因为不可能有负的地址码。

3.3.2 内存

如图3-3所示，与CPU经内总线连接的存储器是“内存”。程序被执行时，其指令代码和运行数据主要存放在内存中，因此也将其称为主存储器，简称主存。内存由半导体存储器组成，它具有运行速度快，体积小特点。半导体存储器有RAM和ROM两种类型。

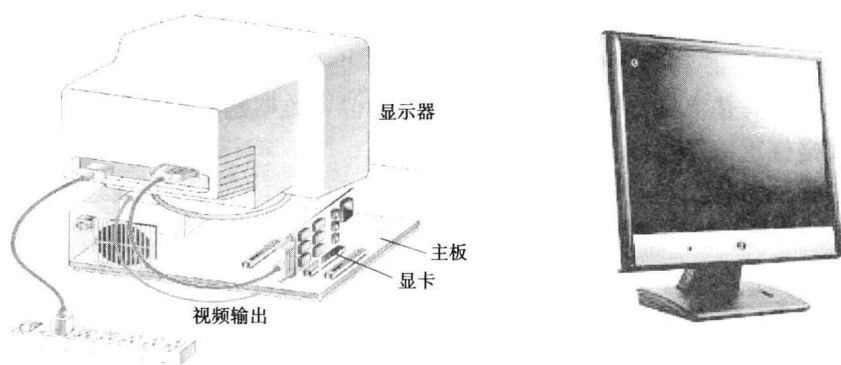


图3-22 CRT、LCD显示器

PC机使用的集成主板上带有显示器接口，这种结构的显示技术占用了主存的一部分空间作为显存。而独立显卡目前多为AGP（Accelerated Graphics Port 加速图形端口）显卡，用于高性能图形和视频，如图3-23所示。如图3-4所示的显卡插槽就是专门用于AGP显卡的。早前还有其他类型的显卡，如VGA卡等，现在很少使用了。

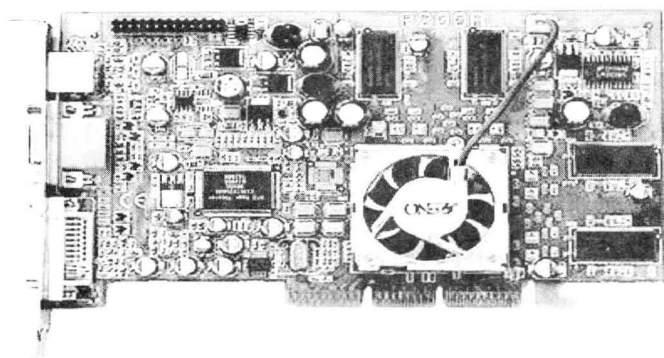


图3-23 AGP显卡

（1）CRT显示器

CRT（Cathode Ray Tube）显示器使用与普通电视机类似的玻壳显像管，在显像管内通过高速电子束轰击屏幕上的“亮点”，通过扫描屏幕上的各个亮点形成图像的输出。彩色显示器有RGB（红绿蓝）三色电子束，对应于屏幕上的三种颜色亮点，控制各电子束的能量使得被轰击的点发出不同亮度的光，以显示不同色彩。

CRT显示器采用的扫描模式为逐行扫描，直至扫完整个屏幕的所有行——叫做“帧”扫描。帧扫描频率（也叫垂直扫描频率）是显示器的重要指标。显示器的另一个主要指标是尺寸，它是指显示屏对角线的长度，通常使用英寸表示，如17英寸、21英寸等。

CRT显示器的体积较大且功耗也较大。

（2）LCD显示器

LCD（Liquid Crystal Display）显示器又叫做液晶显示器，目前已经是PC机标准的配置，也被称为平板显示器，这是相对于CRT显示器而言的。它的体积小且功耗低。早前LCD显示器专用于便携式笔记本电脑。LCD显示器通过激发液态晶体单元发光产生显示图像。

（3）其他类型的显示器

现在还有使用LED（Light Emitting Diode，发光二极管）显示器、等离子（Plasma）技

及使用何种端口等，就必须明确规定出来。这些规定就是外部设备的总线标准。

内部三总线是芯片级的信号通路，而外部设备与主机间的连接则是“系统”级的，因此也将外部总线称为系统总线（System Bus）。

任何厂商生产的任何设备，只要按总线标准设计，这些设备就可以被计算机所支持，这类设备叫做即插即用（Plug and Play，PnP或PNP）设备。如果设备不被计算机直接支持，设备厂家就需要提供设备驱动程序（Driver Program）并安装在计算机上。一旦使用这个设备，计算机就将使用这个设备驱动程序去控制该设备完成相应的输入或输出任务。对PnP设备，计算机使用的是标准的设备驱动程序。

我们已经知道，外设连接是通过如图3-4所示的主板的端口或接口到主机的。其中，内存条直接插入主板上的内存插槽中。外存接口是连接磁盘、光盘的，其标准就是ATA/IDE和SATA。外设端口连接的键盘、设备、显示器等也都采用了统一的标准。其他设备可以通过电路插卡插在“接口总线插槽”上，也称为扩展槽（和AGP卡类似），与主机连接。这个扩展的插槽支持的标准就是PCI总线。

1. PCI总线

PCI（Peripheral Component Interconnect）始于20世纪90年代初，目前还是微机的总线标准之一。它支持32位和64位数据传输，支持5V电源，也支持3.3V的低功耗应用。由于这种设计不依赖于处理器，因此它也适合PC机的多种机型。

早前的主板上有多达6 ~ 8个PCI插槽，而今天的主板只保留了一两个插槽。其原因是，现在使用PCI进行系统扩展的设备越来越少，如果有的话也大多数是早期生产的。

2. PCMCIA

PCI设计为适合台式机的物理尺寸，而在笔记本电脑中使用的是体积如信用卡大小的插卡总线标准PCMCIA（Personal Computer Memory Card International Association Industry Standard Architecture，个人计算机存储卡国际协会工业标准结构）。这种插卡的一头插笔记本电脑扩展槽，另一头为连接外设的端口。如图3-27所示是一个PCMCIA插卡的实物图。

除了上述的扩展总线外，实际上曾经有过多种输入/输出的技术标准，例如，串行端口的RS-232标准曾广泛用于电话线连网的数据交换，还有并行端口曾长期是打印机的技术标准。表3-4列出的就是常用的接口总线及其主要技术指标。

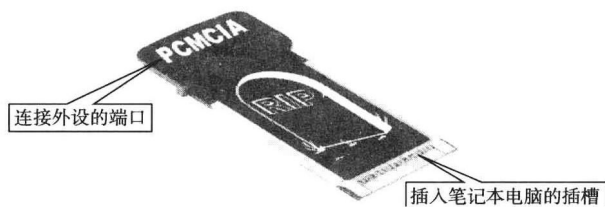


图3-27 PCMCIA插卡

表3-4 常用的接口总线

名 称	格 式	连接设备数	线 长 度	速 率	用 途
USB 2.0	串行	127	4m	1.5 ~ 480Mbps	各种外设
RS-232	串行	2	12 ~ 25m	20 ~ 115kbps	Modem, 鼠标等
RS-485	串行	32	1000m	10Mbps	数据获取和控制系统
IrDA	红外	2	1.5m	115kbps	打印机等

如智能手机，也叫做袖珍PC，即PPC（Packet PC），通常专指使用Windows Mobile操作系统的移动设备。

③ Symbian OS，这是Nokia和Sony Ericsson等手机生产商联合开发的智能手机操作系统，常用于Nokia和Sony Ericsson的手机上。Symbian（塞班）OS支持使用流行的计算机程序设计语言编程，曾在智能手机中占据很大的市场。

④ Android，这是Google公司收购了原开发商Android后，联合多家制造商推出的面向平板电脑、移动设备、智能手机的操作系统。Android是基于Linux开放的源代码开发的，且仍然是免费系统。

⑤ iOS，这是Apple公司为其生产的移动电话iPhone 开发的操作系统。它主要用于Apple的i系列数码产品，包括iPhone、iPod touch、iPad以及Apple TV。原本这个系统名为iPhone OS。

我国的移动电话服务商也推出了各自的手机操作系统，使用在其推广的手机产品上。

4.4 操作系统的结构和组成

在介绍了操作系统的基本作用和几种常见的操作系统之后，本节介绍操作系统的结构和组成。本质上，各种类型操作系统的功能基本相同，其结构也差不多，只是实现方法不同。操作系统的结构基于软件的层次结构，把操作系统分为内核（Kernel）和用户接口（Shell）两层。操作系统的组成包括：进程管理、存储管理、文件管理和设备管理。

4.4.1 操作系统的层次结构

操作系统中定义了它的内核层以及它与用户之间的接口，这种结构如图4-4所示，这是根据系统设计计划的，由UNIX定义并实现，为各种操作系统所采用。

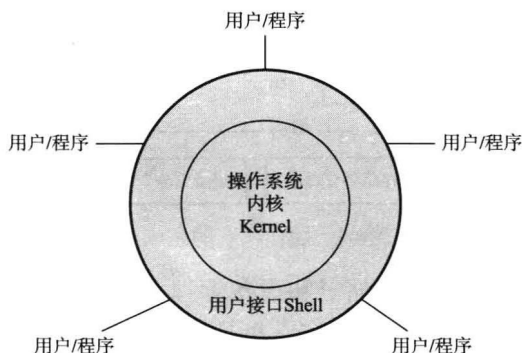


图4-4 操作系统的内核和用户接口结构

1. 操作系统内核

操作系统的内核程序叫做Kernel，它有一个操控计算机各种资源的基本模块，如文件管理。Kernel的另一个部分是设备驱动（Device Driver），与设备进行通信并执行设备操作。例如，接收来自键盘的输入并存放到内存的指定位置，供用户或其他程序使用。

Kernel的第三个部分是内存管理。为实现多任务（程序）处理，内存管理把当前程序调入内存运行，然后将另外一个程序调入内存替代前一个程序，或者将内存分成几个部分，分

3. 进程同步和死锁

进程管理的另一个主要问题是同步，要保证不同的进程使用不同的资源。如果某个进程占有另一个进程需要的资源而同时请求对方的资源，并且在得到所需资源前不释放其占有的资源，如图4-8所示，那么就会导致发生死锁，也就是进程不能够实现同步。

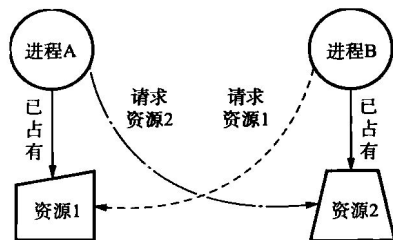


图4-8 进程死锁

解释死锁的一个例子是，在一条只能容纳一辆车通过的窄桥上，同时有相向方向的车开上了桥。解决死锁问题的方法之一是，当某个资源不空闲时，让需要这个资源的进程处于等待状态。另一种方法是，限制进程占有资源的时间。

现代操作系统尽管在设计上已经考虑防止死锁的发生，但死锁并不能完全根除。发生死锁会导致系统处于无效等待状态，因此必须撤销其中的一个进程。例如在Windows中，可用“任务管理器”（见图4-9）终止没有响应也就是无效的进程。

多任务处理就是多进程。进程管理是Kernel的工作，但是用户也能够对进程进行干预。例如，Windows通过“任务管理器”将程序运行情况呈现给用户，以使得用户在必要时对任务进行操控。

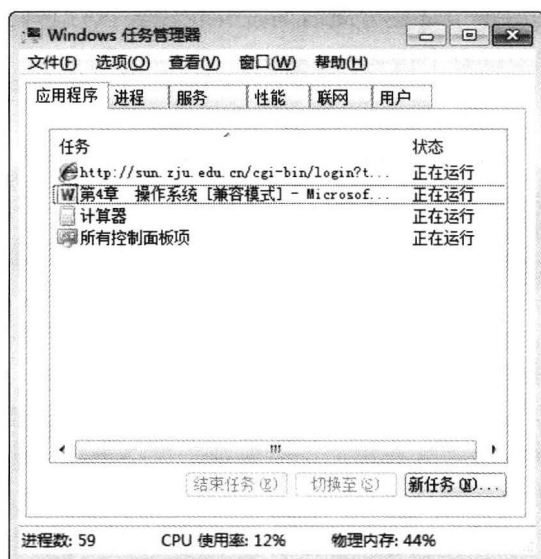


图4-9 Windows的任务管理器

Windows任务管理器将用户进程冠以“应用程序”，而显示的“进程数”则包括了系统为管理机器运行而执行的、属于操作系统本身的那些程序。

4. 线程

“线程”（Thread）是进程概念的延伸。如果程序只有一个进程就可以处理所有的任务，那么它就是单线程的。如果程序可以被分解为多个进程共同完成程序的任务，那么被分解的不同进程就叫做线程，也叫轻量级进程（Light Weight Processes）。线程有几种模式，如单线程模式、单元线程模式和自由线程模式。

我们可以将程序想象成一个搬家的过程：从一所房子搬到另外一所房子。单线程模式需

4.5.1 概述

在微机系统中，Windows占有绝对的市场份额。微软公司从1985年开始至今已经陆续发布了Windows 1.03、2.0、3.0、3.1、95（又名Chicago）、98、NT、2000系列、XP、2003、Vista、2008，以及2009年发布的Windows 7。2011年的统计数据表明，PC机目前主要使用的是Windows XP和Windows 7，而Vista只有10%左右的占有率。

对Windows的评价有很大争议，主要表现在普通用户和专家之间的评价是不同的。但这似乎并没有影响它成为占有最大市场份额和拥有最多用户的操作系统软件。Windows的重要性还在于它使得计算机的操作、应用变得非常容易，非专业人员也能够使用计算机。

Windows基于图形用户界面和支持即插即用等特性，运用了多种先进的技术，如内存交换技术、多线程技术等。Windows也能处理多媒体信息，内置了多种网络协议，用户能够很容易地使用局域网和因特网。

Windows提供了应用程序接口（API）、设备驱动程序开发工具，为开发基于Windows的应用程序提供了极大的便利，因此有极为丰富的各种应用系统，这也是Windows得以流行的主要原因。

Windows是一个系列产品，从服务器到智能手机都提供了支持。Windows在安全性上一直备受批评，用户需要不断地从微软网站上下载“补丁”程序进行更新（Update）。除了系统本身外，它的用户之多以及它的垄断地位也使一些有意者（如黑客）刻意寻找它的漏洞，实施攻击。

4.5.2 系统结构

和大多数操作系统一样，Windows也使用了内核和外壳模型结构。它的操作系统内核代码运行在处理器特权模式下，被称为内核模式（Kernel Mode），能够访问PC机的硬件和系统数据，而用户和应用程序被设置运行在非特权模式（也叫用户模式，User Mode）下。

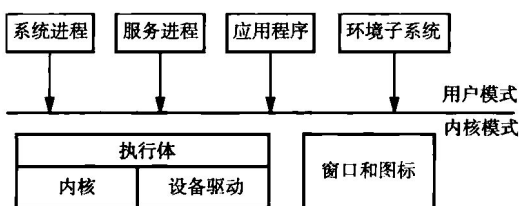


图 4-14 Windows结构示意图

操作系统设计的主要部件，如处理器管理、内存管理、I/O管理等，都运行在各自独立的进程中，且有自己独立的内存地址。但是Windows把大多数管理和控制代码都运行在内核模式下。它的结构如图4-14所示。用户模式下的几个部分说明如下。

系统进程是操作系统固定的或硬件需要的程序，如用户登录等。

服务进程是提供系统服务的，如任务管理器、假脱机打印等。

应用程序是指用户运行的程序，例如用户上网使用的IE浏览器，编辑文档使用的Word等。Windows支持多种类型的应用程序，如Windows 95/98/2000/XP等旧版本下开发的应用程序，在Windows 7下运行这些程序需要使用管理员登录并选择兼容方式。

环境子系统实现了操作系统环境的支持，早前微软还支持可移植操作系统接口（Portable Operating System Interface，简称POSIX，其中X表示其对UNIX API的传承），Windows XP之后就不再支持其他操作系统的子系统。

在内核模式下，其执行体（Executive）包括了操作系统的内存、进程、线程管理，以及

的图标，以帮助用户识别文件类型。双击文件图标，Windows将根据文件的类型决定采取何种操作：如果被选择的是程序文件，就执行它；如果是数据文件，就启动其关联程序打开它。例如，选择了一个Word文档文件，Windows将启动Word程序打开该文档。

大多数程序在创建数据文件时，会给出数据文件的扩展名。例如，使用Word创建文档，在保存文件时，Word会自动提示加上.doc（.docx）扩展名。

文件名的用途之一是检索文件。在一个磁盘、CD或DVD盘中，如果有成千上万个文件，需要查找或检索其中的一个或者一部分特定的文件时，有两个非常有用的符号：“*”和“？”，称为通配符（或万能符），它们在查找文件时用来定位文件。

“*”代表它所在位置为任意个任意字符，“？”代表它所在位置为一个任意字符。例如，mytxt.*代表文件名为mytxt的所有文件；*.doc代表扩展名为.doc的所有文件；A*代表任何以A字母开头的文件名；A?.*代表开头字母为A，第二个字符为任意，文件名的长度为2个字符的所有文件；*.*代表所有文件；而???.*代表所有文件名长度为3个字符，扩展名为任意的所有文件。大多数操作系统都支持这两个通配符。

4.6.4 常用的文件类型

要熟练使用计算机就需要清楚文件的分类，熟悉常用的文件类型。目前，已有数以千计的文件类型，表4-3中给出了部分常用文件扩展名，这里再进一步介绍常用的文件类型。

1. 可执行文件

可执行文件就是程序，在Windows中称为应用程序。MS系统可执行文件的扩展名一般为“.exe”、“.com”、“.bat”。其中，习惯称为可执行文件的.exe文件取自于单词executive。而.com文件也叫做命令（Command）文件。批处理文件是文本文件，其内容为可执行文件或命令。

2. 数据文件

这里的数据文件特指文档（Document）、电子表格、数据库数据文件等。一般来说，程序执行所需要的或产生的其他文件都是数据文件。几乎每一个专门的应用程序都有自己所适用的数据文件。

数据文件本身不能被直接运行或操作，它们需要借助于相应的执行程序（应用程序）来运行或打开。操作系统建立数据文件关联机制，使得在打开某种类型的数据文件时，相应的应用程序就会自动启动。例如，Windows的文件“打开方式”（在文件名上单击鼠标右键，从弹出的快捷菜单中选择）。注意，只有数据文件才能够建立关联，而程序文件的右键快捷菜单中只有运行方式。

3. 图形图像文件

在2.6节中介绍了多媒体数据，包括图形/图像数据、音频/视频数据，它们也都是以文件形式存放的。常见的图形/图像数据文件如表4-4所示。有如此种类繁多的图形/图像格式，一方面是为了满足不同应用环境、不同图形/图像质量的需要，另一方面也是因为不同的开发商需要使用专门的格式保护其软件的知识产权。

Disks, 独立磁盘冗余阵列)。简单地说, RAID是由多个硬盘组成的硬盘组, 在用户看起来, 组成的硬盘组就像是一个硬盘。RAID技术的两大特点: 一是速度, 二是安全。由于这两项优点, RAID技术早期被应用于高级服务器中。RAID需要专用的RAID卡(适配器)。现在已有了从 RAID 0到RAID 6 共7种级别, 不同的级别代表着不同的存储性能、数据安全性和存储成本。有关RAID技术的介绍超出了本书的范围, 有兴趣的读者可以参考专门书籍或资料。

本章小结

本章主要介绍计算机中的核心软件——操作系统的有关概念、结构、组成方面的知识。操作系统是计算机和用户及其他程序之间的接口, 负责对计算机的所有资源进行管理。根据操作系统的发展, 有实时系统、单任务单用户、单用户多任务、多用户多任务、并行系统、分布式系统等类型。常见的操作系统有MS-DOS、Windows、UNIX和Linux、Mac OS等, 目前广泛使用的智能手机也有多种操作系统。

基于层次结构, 操作系统的核心为Kernel, 在Kernel和用户之间的接口为Shell。根据功能的观点, 操作系统由进程管理、存储器管理、设备管理和文件管理4个部分组成。

程序从被选中运行直到运行结束的整个过程称为作业, 进程是正在内存中运行中的作业。操作系统对多作业和多进程进行调度、控制以实现进程同步。线程被称为轻量级进程。进程管理器负责管理程序的运行。内存管理包括程序使用内存、在内外存之间进行数据交换和建立并管理虚拟内存。虚拟内存是在磁盘上开辟的一个空间, 并且按照内存结构进行组织和管理。

操作系统将设备分为块设备和字符设备。它将对设备处理的软件独立于设备, 对设备的操作由设备驱动程序完成。操作系统的设备管理提供设备调度、缓冲区、假脱机等多种服务。

Windows是一个在微机系统中最常用的操作系统, 它基于图形界面、面向对象和多任务系统, 各个应用程序共享Windows系统提供的所有资源。

文件将计算机的数据以一种抽象的、概念化的、易于理解的组织形式展示给用户, 文件系统是基于操作系统的。文件是一个存储在存储器上的数据的有序集合并标记以一个名字, 文件名及其扩展名用于文件的查找和关联文件的操作。操作系统对文件的管理采用目录结构。文件有顺序和随机两种存取方式。

文件的存储结构是基于操作系统的, 微软系统的文件存储结构包括FAT和NTFS系统。文件系统需要通过备份提高其安全性。

习题4

一、填空题

1. 操作系统的定义为: 操作系统是计算机硬件和用户(其他软件和人)之间的_____, 它使得用户能够方便地操作计算机, 能有效地对计算机软件 and _____ 进行管理和使用。



2. 按照使用环境和对程序的运行方式, 操作系统的类型有: _____、_____、单用户多任务、_____、并行系统、分布式系统等。

3. Windows操作系统支持单用户_____任务, 即可以有_____程序在Windows的支持下在机器中运行。

4. 移动设备(如手机)是嵌入了处理器的设备, 它们也需要操作系统的支持。例如, 常见的移动设备操作系统有Palm OS、Windows Mobile、Symbian OS、Android、iOS等。其中, 用于iPhone的是_____, 而用于PPC的是_____, 用于Nokia的是_____。而_____是免费的移动设备操作系统。

5. 基于软件的层次结构, 操作系统可以分为Kernel, 即_____和外壳, 即_____。其中, 后者是负责接收用户或者应用程序的操作命令并将其解释后交由前者去执行。

6. 功能模块就是组成操作系统的组件, 因此按照功能的观点, 可以将操作系统分为4个功能模块, 即_____、_____、_____、_____。

7. 进程管理是操作系统的核心。现代操作系统把进程管理归纳为一个_____被选中后成为_____, 进而进入内存运行而成为_____, 运行结束后再次保存到磁盘上。

8. 支持多道程序处理的操作系统, 有时需要在内存和_____之间进行数据交换, 以便把程序的执行代码装载到_____中。

9. 输入/输出设备也叫做I/O设备。对设备的管理, 操作系统通常从中抽象出一些通用类型, 将其分为_____设备和字符设备。例如, 磁盘和键盘分别是这两类设备的例子。实际上, 操作系统并不直接使用设备, 而是通过使用_____程序来使用设备的。如果一个设备连接到计算机上而不能使用, 那么就需要安装这个程序。

10. 时钟是计算机中特殊且重要的设备, 它除了提供系统时间外, 也为操作系统复制某个进程垄断CPU提供作用。PC机中有两种时钟, 分别是_____和_____。

11. Windows有注册表、服务和管理规范3种管理进制。其中_____存放了计算机的系统和_____信息。通过执行_____命令可以打开注册表, 使用_____可以更新注册表。

12. 文件的扩展名给出了文件的基本属性, 例如, 扩展名为.com、.bat和_____的文件为应用程序, 而.txt扩展名表明是_____文件, 而.jpg扩展名为图片文件, .mpg扩展名为_____文件。

13. Windows的文件系统有两种存储结构, 一种是FAT, 即_____, 另一种是_____, 即新技术文件系统。

二、选择题

1. 操作系统是_____的接口。

- A. 用户和软件
- C. 主机和外设

- B. 系统软件和应用软件
- D. 用户和计算机

2. 通常, 任何软件都需要依赖其运行环境, 这个环境也叫做平台, 它是指_____。

- A. 硬件, 主要是指CPU
- C. 机器运行的操作系统

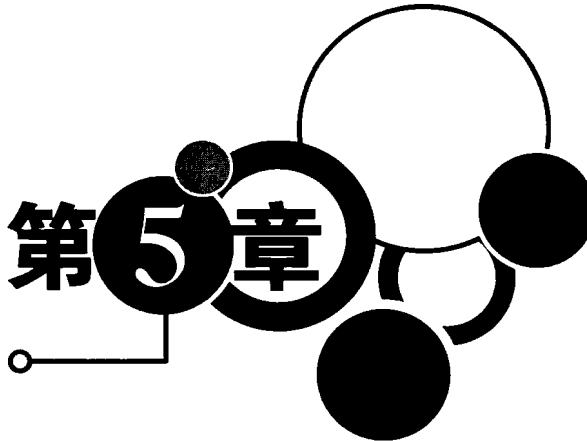
- B. 指机器的规模
- D. 机器使用的编程语言

3. 系统软件(System Software)的主要功能是_____。

- A. 管理、监控和维护计算机软、硬件资源
- B. 为用户提供友好的交互界面, 支持用户运行应用软件

第5章

算法基础



算法 (Algorithm) 是计算机科学的最基本的概念, 计算机科学研究的核心之一就是算法的研究。简单地说, 算法就是解决问题的一系列步骤。本章介绍有关算法的基本知识, 包括算法的概念、分类、特性以及算法的表示, 还介绍基本算法和有关数据抽象与数据结构方面的知识。

5.1 算法的概念

不是只有数学问题才有算法, 广义地说, 为解决问题而采用的方法和步骤就是算法。在计算机中, 算法是程序设计的基础, 算法的质量直接影响程序运行的效率。如果从机器完成任务的角度, 程序就是与机器兼容的算法的实现。在本书前述章节中, 已经有许多算法的例子, 例如在介绍操作系统时就提到的作业管理、文件管理、存储器管理等, 以及文件检索使用的索引方法、哈希方法都是算法。

相对于硬件, 算法及实现算法的计算机语言就是软件。在软件开发中, 核心工作是进行算法设计。在计算机被研制出来之前, 算法一直属于数学的范畴, 主要是指寻找特定的问题如何解决的方法。一个著名的例子是, 古希腊数学家欧几里得 (Euclid) 所发现的求两个正整数 A 和 B 的最大公约数问题:

第一步: 比较 A 和 B 这两个数, 将 A 设置为较大的数, B 为较小的数。

第二步: A 除以 B , 得到余数 C 。

第三步: 如果 C 等于 0, 则最大公约数就是 B , 否则将 B 赋值给 A , C 赋值给 B , 重复进行第二步、第三步。

欧几里得算法也被称为计算最大公约数的辗转相除法。根据图灵理论, 只要能够被分解为有限步骤的问题就可以被计算机执行。这里有两层的意思, 一是算法必须是有限步骤的, 二是能够将这些步骤设计为计算机所执行的程序。因此算法就是程序设计的基础。反过来, 如果一个问题不能被有效地分解为有限的步骤, 那就是说, 问题的解决方案是计算机所不能