

高校非计算机专业计算机教材丛书

计算机软件基础

冯博琴 陈 俭 刘路放
刘跃虎 邓良松



COMPUTER

西安交通大学出版社

内 容 简 介

本书比较全面地介绍了计算机软件基础的几个主要部分。全书分四个部分,共 11 章。第一部分(1-2 章)介绍了操作系统的原理和常用的 DOS、UNIX 系统的使用,第二部分(3-5 章)介绍数据结构及部分典型算法,第三部分(6-10 章)介绍数据库基本概念,详细讨论了如何应用 FOXBASE 进行数据库设计,第四部分(11 章)介绍了软件工程的基本原理、方法。全书内容紧凑、深入浅出、注重实用。

本书按计算机考试大纲编写。可作为大专院校非计算机类专业计算机软件基础的教材,也可供工程技术人员作提高软件水平的参考书,成人教育和职业培训亦可用它作教材。

(陕)新登字 007 号

计算机软件基础

冯博琴 陈 俭 刘路放 刘跃虎 邓良松

*

西安交通大学出版社出版

(西安市咸宁西路 28 号 邮政编码:710049)

向阳印刷厂印装

陕西省新华书店经销

开本 787×1092 1/16 印张 19.375 字数 465 千字

1993 年 12 月第 1 版 1995 年 3 月第 3 次印刷

印数:8001—13000

ISBN7-5605-0628-3/TP·73 定价:15.90 元

序 言

在普通高等院校中对学生的计算机基础知识与应用能力的培养已经成为各学科各专业教学计划的重要组成部分。高等院校本、专科毕业生的计算机基础知识与应用能力的水平也已成为绝大多数用人单位选择录用人员的重要依据之一。我省各高等院校多年来在计算机基础课程的教学方面进行了精心的组织,工作在计算机基础课程教学第一线的广大教师、工程技术人员和实验室工作人员呕心沥血做了大量艰辛的工作。不过,由于人力、教材、计算机设备等等各方面条件的制约,各院校之间在计算机基础课程教学方面的进展还是不很平衡的。

为了把我省普通高等院校的计算机基础课程教学提高到一个更高的水平,陕西省教委非常重视普通高校计算机基础课程教学的课程建设和各学科各专业学生的计算机知识结构的研究和组织工作。在当前,教材建设又是一项当务之急的基础建设工作。为了使各院校在有关课程的教学内容方面能有一个比较合理、一致的基本要求,省教委邀请了省内近十所高校中工作在有关课程教学第一线的具有丰富教学经验的专家、教授经过认真切磋并参照国家教委工科计算机基础课程教学指导委员会制定的有关课程的教学基本要求,编写了一套“陕西省普通高校非计算机专业计算机教材丛书”。这套丛书包括:《计算机应用基础》(文科类);《计算机应用基础》(理工科类);《BASIC 语言程序设计》;《FORTRAN 语言程序设计》;《PASCAL 语言程序设计》;《COBOL 语言程序设计》;《C 语言程序设计》;《微机原理及接口技术》;《计算机软件基础》共九种教材。这不仅是一套适合普通高等院校作为有关课程教学使用的教材,也可作成人教育及各种专门培训班组织有关课程教学之用,当然也可作为社会上各行各业有关人员学习计算机基础课程的自学教材。这套教材普遍的特点是内容规范、取材精炼、便于组织教学和学生自学。我们相信,这套教材在大面积的使用过程中经过不断的听取意见和锤炼修改,定会成为一套受广大读者欢迎的好教材。

胡正家

1994.02.10

前 言

通俗地讲,应用计算机解决实际问题,第一步需要分析原来的问题各种现象,或处理过程,然后采用归纳、抽象等手段,建立起反映该问题本质的、比较精确的数学模型;第二步是对该数学模型在计算机上进行求解;最后对所得的解作理论和实际的验证。若不正确,则还须回到上面某一步重作,如此循环,直至满意。由此可见,数学建模是解题的第一步,它首先要搞清楚待解决的问题类型,然后根据具体情况采用有关专业知识和数学方法建立数模。虽然确有一些实际问题的数学背景十分简单,但目前还有许多问题无法解决或解答不精确。其原因主要是人们无法建立它的数学模型或者模型过于简化。目前已有部分高校开设了建立数学模型的专门课程,这是十分必要的。第二步是在计算机上求解,它也遇到了与上面类似的情况:虽有数学模型,但不能有效地使用计算机软硬件。这正是目前困扰许多非计算机专业人员的一个难题,这里提出了一个问题:怎样提高非计算机专业人员的素质?怎样加强大学生的计算机应用能力?

一个老练的计算机软件人员必须具备哪些最基本知识和能力呢?第一,他应能为自己设计和创造一个良好的软硬件开发环境,能充分调动计算机系统资源,因此他对操作系统的原理、使用运用自如,掌握各种软件工具。第二,善于分析研究计算机加工的数据对象的特性,能选择合适的数据结构和算法,编写高质量的程序。第三,掌握以低投入获得高质量软件产品的方法和技术。第四,熟练掌握数据库、网络等技术,等等。

本书旨在提高非计算机类专业学生的软件素质,以缩短与软件人员的距离。本书内容仅选取上面提到诸项的最基础部分。考虑到学时和读者的特点,不追求计算机学科体系完整,尽量做到深入浅出,讲求实用。全书分四部分,第一部分包括 1-2 章,介绍操作系统原理及使用;第二部分包括 3-5 章,内容是数据结构和算法;第三部分包括 6-10 章,介绍数据库概念和 FOXBASE 的使用;第四部分仅有一章,介绍软件工程。

本书主要按照计算机等级考试大纲要求编写。它也是西安交通大学从事计算机教育的部分教师长期教学积累的总结。参加编写的有陈俭(1-2 章)、刘跃虎(3-5 章)、刘路放(6-10 章)、邓良松(11 章),全书由冯博琴主编,他还参加了 3-5 章的编写。西安电子科技大学计算机系金益民教授审阅了本书,西安交通大学出版社为本书出版做了大量工作,使本书能尽早与读者见面,在此一并表示感谢。

由于作者水平有限,编写时间匆促,加之对非计算机类专业学生的软件基础所包括的内容和讲授方法有待进一步研讨,因此本书的遗漏和谬误之处一定不少,恳望同行和读者不吝赐教。

作者

1994 年 1 月

目 录

第一章 操作系统原理基础

§ 1.1 绪论	(1)
§ 1.2 文件系统管理	(3)
1.2.1 文件的概念及分类	(3)
1.2.2 文件系统及其功能	(5)
1.2.3 文件系统的组织	(5)
1.2.4 文件系统的保护与使用	(13)
§ 1.3 进程管理	(14)
1.3.1 进程的概念	(14)
1.3.2 进程的性质	(14)
1.3.3 进程的调度与控制	(15)
1.3.4 UNIX 中的进程管理	(16)
§ 1.4 存储器管理	(18)
1.4.1 存储器管理的任务	(18)
1.4.2 单一连续存贮管理	(19)
1.4.3 多连续区存贮管理	(20)
1.4.4 页式存贮管理	(21)
1.4.5 段式存贮管理	(22)
1.4.6 DOS 的主存管理	(22)
1.4.7 UNIX 的主存管理	(23)
§ 1.5 设备管理	(24)
1.5.1 设备的分类与管理任务	(24)
1.5.2 主处理机与设备的通讯方式	(25)
1.5.3 设备的分配与驱动	(26)
1.5.4 DOS 的设备管理	(27)
1.5.5 UNIX 的设备管理	(27)
§ 1.6 作业管理	(28)
1.6.1 作业及相关概念	(28)
1.6.2 作业的调度	(29)
1.6.3 作业的控制	(30)
1.6.4 DOS 的作业管理	(32)
1.6.5 UNIX 的作业管理	(32)
第一章习题	(32)

第二章 操作系统的使用

§ 2.1 DOS 系统构成	(34)
2.1.1 DOS 的概念与发展	(34)
2.1.2 DOS 的系统构成	(35)
§ 2.2 DOS 启动的工作流程	(37)
2.2.1 DOS 的两个特殊文件	(37)
2.2.2 DOS 的系统引导记录	(37)
2.2.3 启动 DOS 的工作流程	(38)
§ 2.3 DOS 的命令格式与功能键	(38)
§ 2.4 DOS 命令的使用	(41)
2.4.1 文件管理命令	(41)
2.4.2 目录管理命令	(43)
2.4.3 系统维护命令	(44)
2.4.4 数据备份命令	(46)
2.4.5 系统状态检测与设置命令	(48)
2.4.6 其他命令	(50)
§ 2.5 改变 DOS 的配置文件	(51)
§ 2.6 DOS 的深入应用与技巧	(56)
2.6.1 预置工作环境	(56)
2.6.2 输入输出重定向的应用	(58)
2.6.3 批处理命令及使用技巧	(59)
2.6.4 几个提高工作效率的命令	(60)
2.6.5 DOS 中断处理与系统调用	(62)
§ 2.7 汉字 DOS 及使用基础	(65)
2.7.1 汉字 DOS 的概念	(65)
2.7.2 汉字信息的输入输出	(65)
§ 2.8 UNIX 操作系统及其使用	(66)
2.8.1 UNIX 简介	(66)
2.8.2 UNIX 的构成与特点	(67)
2.8.3 UNIX 命令的使用	(68)
第二章习题	(72)

第三章 线性数据结构

§ 3.1 数据结构概述	(74)
3.1.1 数据和数据结构	(74)
3.1.2 算法的描述及评价	(76)
§ 3.2 线性表	(78)
3.2.1 线性表的逻辑结构	(78)
3.2.2 线性表的顺序存贮结构	(79)
3.2.3 线性表的链式存贮结构	(81)

3.2.4 几个问题的讨论	(87)
§ 3.3 栈和队列	(92)
3.3.1 栈	(92)
3.3.2 队列	(96)
§ 3.4 串和数组	(103)
3.4.1 串	(103)
3.4.2 数组	(107)
第三章习题	(110)
第四章 非线性数据结构	
§ 4.1 树结构及其基本概念	(112)
§ 4.2 二叉树结构	(113)
4.2.1 二叉树的定义	(113)
4.2.2 二叉树的链式存贮结构	(114)
4.2.3 几种特殊的二叉树	(114)
§ 4.3 二叉树的遍历	(117)
§ 4.4 树、森林与二叉树的转换	(119)
§ 4.5 图及其基本概念	(120)
§ 4.6 图的存贮结构	(122)
4.6.1 邻接矩阵	(122)
4.6.2 邻接表	(122)
§ 4.7 图的遍历	(125)
4.7.1 深度优先遍历连通图	(125)
4.7.2 广度优先遍历连通图	(127)
§ 4.8 有关二叉树、图的几个问题讨论	(128)
第四章习题	(134)
第五章 查找与排序	
§ 5.1 简单查找方法	(137)
5.1.1 顺序查找	(137)
5.1.2 折半查找	(138)
5.1.3 分块查找	(139)
§ 5.2 树表查找	(140)
§ 5.3 哈希查找	(141)
5.3.1 哈希表的建立	(142)
5.3.2 处理冲突的方法	(142)
5.3.3 哈希查找	(144)
§ 5.4 平均查找长度的计算举例	(145)
§ 5.5 简单排序方法	(147)
5.5.1 简单插入排序	(148)
5.5.2 简单选择排序	(148)

5.5.3 冒泡排序	(149)
§ 5.6 快速排序	(150)
§ 5.7 归并排序	(152)
第五章习题	(154)
第六章 数据库系统概述	
§ 6.1 引言	(156)
§ 6.2 什么是数据库	(156)
§ 6.3 数据模型	(158)
§ 6.4 数据库的构成	(162)
第六章习题	(164)
第七章 关系数据库及其数学基础	
§ 7.1 关系及其基本术语	(165)
§ 7.2 关系运算	(168)
§ 7.3 关系模型的数学定义	(172)
§ 7.4 关系代数	(175)
§ 7.5 关系的规范化理论	(177)
第七章习题	(181)
第八章 FOXBASE 系统总论	
§ 8.1 FOXBASE	(183)
§ 8.2 FOXBASE 的基本元素	(183)
§ 8.3 数据库的基本操作	(188)
§ 8.4 选择、投影和范围	(194)
§ 8.5 排序和索引	(196)
§ 8.6 多个工作区	(203)
§ 8.7 数据库结构操作命令	(210)
§ 8.8 设计一个数据库系统(一)	(212)
第八章习题	(218)
第九章 向数据库索取更多信息	
§ 9.1 列出漂亮的表格	(219)
§ 9.2 计数、求和、平均与汇总	(221)
§ 9.3 宏代换函数与内存变量文件	(223)
§ 9.4 格式文件	(227)
§ 9.5 定义和输出报表及标签	(231)
§ 9.6 怎样与其他程序设计语言交流数据	(233)
§ 9.7 设计一个数据库系统(二)	(236)
第九章习题	(238)
第十章 应用程序设计	
§ 10.1 建立应用程序	(239)
§ 10.2 程序控制结构	(240)

§ 10.3 过程.....	(243)
§ 10.4 设计一个数据库系统(三).....	(244)
第十章习题.....	(249)
第十一章 软件工程	
§ 11.1 软件工程概述.....	(250)
§ 11.2 软件的需求定义.....	(253)
§ 11.3 软件的设计.....	(263)
§ 11.4 软件的编程.....	(279)
§ 11.5 软件的测试.....	(281)
§ 11.6 小结.....	(296)
第十一章习题.....	(296)
参考书目	

第一章 操作系统原理基础

§ 1.1 绪论

一、操作系统的基本作用

一个计算机系统可以划分为硬件资源和软件资源两大部分。硬件资源包括中央处理器(CPU)、存储器(内存和外存)和各种外部设备(输入输出装置)。软件资源包括各类程序和数据。操作系统是计算机软件资源中的重要组成之一,是一种系统程序。它的基本职能有两个,第一,负责监督、控制和协调计算机系统中的各类软硬件资源;第二,建立操作人员与计算机间的联系。概括起来讲,就是操作系统的资源管理作用和信息转换作用。现代计算机系统只有在操作系统的有效管理下,才能充分发挥系统的整体效益。也只有通过操作系统提供的功能,计算机终端用户方能得到系统的有效服务。

我们可以从两个方面理解一个操作系统的作用。首先,操作系统的信息转换作用。本质上,一台计算机的硬件装置是些只能表示0或1两种状态的开关电路。它所能表达的信息或知识,形式上是非常低级的。人们期望所表达的知识形式或描述任务要求能尽量符合日常习惯,即尽可能是高级的。显然,在人易于理解的高级表达形式与机器的低级表达形式之间存在着相当大的距离。因此,有必要引入沟通二者相互理解的“翻译”。从这一意义上讲,计算机系统中所有的系统软件均具有这种作用。然而,操作系统是与计算机硬件联系最为密切的系统软件,它也是一切其他系统软件的工作基础。其次,操作系统的资源管理作用。如上所述,计算机系统中存在着各类资源,包括各种硬件功能部件以及各种软件程序,它们的有机结合形成了总体系统。功能部件各有特色,它们的性能有高有低,工作节奏有快有慢,与外部环境的联系有强有弱。因此,合理地分配、有效地协调这些系统资源的使用才能保证总体系统的正常工作,才能取得最佳的整体效益,充分提高系统的时间与空间的利用率。正是现代操作系统的这种管理职能,才使得计算机系统能有效地同时服务于多个用户和多个程序。

在一个计算机系统中,操作系统是与硬件资源联系最为紧密的软件层次,相对于直接针对硬件装置进行操作和编程而言,操作系统为一般用户使用计算机提供了一个更为高级的环境,因而降低了用户需要深入了解硬件结构方能使用计算机的需求。

从系统资源管理的角度讲,一个功能完备的操作系统应具备文件管理、进程管理、存储器管理、设备管理和作业管理五大管理职能。

二、操作系统的发展

在计算机系统的发展过程中,硬件资源和软件资源所占的比例并不是固定不变的。早期的计算机系统中,硬件资源所占比重远远大于软件资源。在此阶段,几乎不存在系统软件的概念,用户采用低级的代码形式编写应用程序。计算机的功能主要靠硬件装置来实现,系统中硬件资源比重很大,但软件资源的内容比较简单,外围设备很少。由人工直接控制系统硬件设备的使用,硬件资源为一个用户独占。在此阶段,计算机的程序员同时也是操作员,用户必须熟悉计算

机的硬件构成,仔细地为计算机工作安排每一操作步骤。不难理解,按这种方式使用计算机,不仅难以使用,容易出错,而且由于人工的大量干预,计算机系统的效率是极低的。此外,由于计算机使用的复杂性,一般人员对计算机只能“敬而远之”。

为了提高计算机的使用效率,首先想到的就是如何降低使用计算机过程中人工干预的程度,尽量让计算机保持不间断的工作。为此,一种自然的处理途径就是,变零散的单一任务处理为集中式批处理。初期的批处理仍然是一次处理一个程序,因此,称其为单道程序批处理方式。其基本特点是,将需由计算机处理且性质相同的程序先积累起来,按先后次序存贮到磁带上。形成批量后,再一次性递交计算机处理。批处理方式大大减少了人工操作控制过程中上机准备、出错调试等所花费的系统时间,使计算机系统有相对较长的连续运行时间。批处理方式的出现使个人独占系统资源向用户共享资源迈进了一步。为批处理的需要,相应地产生了批处理控制管理程序。这种管理程序的诞生,标志着计算机开始摆脱早期的手工操作方式,开始引入程序运行的自动管理机制。因此,一般把批处理控制管理程序看成是第一代操作系统。

第一代操作系统出现后,程序员与操作员的职责有了明显分工,程序员可不必再亲自上机操作,他们只需关心需解决的任务本身。而上机操作的工作则可交由受过训练的专门操作员完成。具有批处理服务特点的操作系统为计算机用户提供了一套控制命令,称为作业控制命令。操作人员通过合理组织和编排这些命令序列,写出作业说明书,来表达对程序(作业)进行加工的要求。计算机通过识别这些命令,对操作员递交的任务完成诸如装入、启动、停止等各种操作。当一个程序的任务处理结束后,操作系统会自动启动下一个待处理的程序而无需人工干预。

单道批处理管理程序出现不久,为进一步提高系统资源利用率,很快就产生了能处理多道程序的第二代操作系统。在单道程序操作系统中,下一个程序只能等待前一程序执行结束后方能投入运行,如果顺序处理的某一程序因故不能正常执行,则需重新装入另一程序。而采用多道程序处理技术,则允许一次将几道程序调入内存,由操作系统根据资源的占用情况,以及当前程序的执行状态,统一调度这些程序,使其共享系统资源,而不致于因一个程序的故障,降低系统的总体效率。第二代操作系统的出现,得益于计算机中断技术的发展,中断技术的引入使外部设备与中央处理机之间的独立性加强,大大改变了计算机系统中资源使用不充分和不均衡的情况。

第二代操作系统的出现也引进了许多新的技术问题,如:多道程序的调度,内存空间的分配与保护,外部设备的管理等。70年代中期,随着硬件技术的发展尤其是大容量磁盘存贮装置的出现和这些技术问题的解决,以IBM公司的360系列机上操作系统的问世为标志,产生了具有计算机资源综合管理功能,设计更完善的第三代操作系统。

计算机操作系统与计算机硬件资源密切相关,因此,操作系统是针对不同的机器环境专门设计开发的。早期的操作系统主要是面向大中型计算机的,70年代后,微型计算机的面世,则出现了一批以微型机为硬件支撑环境的操作系统。随着微型机功能的不断增强,以往大型操作系统所具备的功能,不断地向微型机操作系统中渗透。目前,在微机领域中,最为流行的两大操作系统分别为DOS和UNIX。书中有关具体系统的例子主要依据这两个操作系统。

三、操作系统的分类

操作系统是一种系统程序,它只是计算机系统中软件资源之一。在研究和使用时,根据不同的目的和着眼点,可以把操作系统分为各种类型。

(1) 按适用面分

专用操作系统——指为特定应用目的或特定机器环境而配备的操作系统。包括一些具有操作系统特点的监控程序。

通用操作系统——指为通用计算机配备的、能为各种计算机用户服务的系统。通常提到的操作系统均是指通用操作系统。

(2) 按任务的处理方式分

交互式操作系统——指能为用户提供交互操作支持的操作系统。

批处理式操作系统——指以成批处理用户程序为特征的操作系统。它是相对交互式操作系统而言的,在批处理方式下,用户只能在一个批次处理完毕后,方能调试程序中可能存在的问题,或获得计算的结果。批处理方式着眼于提高计算机系统效率,而交互式则着眼于方便用户的使用。

(3) 按处理器使用特点分

分时操作系统——采用分配时间片的方法,使一个处理机为多个用户服务。

实时操作系统——指能够在期望的较短时间内即时响应用户要求或完成信息处理的操作系统。

(4) 按用户数量分

单用户操作系统——只能服务于单个用户的操作系统。如单机 DOS 系统。

多用户操作系统——能为多个用户服务的操作系统,如 UNIX 分时系统。

(5) 按硬件支撑环境和控制方式分

集中式操作系统——指驻留在一台计算机上或管理一台计算机的操作系统。

分布式操作系统——指用于管理分布式计算机的操作系统。

这些分类方法并无公认标准,也不是相互独立的。如分时操作系统本身是多用户操作系统,同时,它也属于交互式操作系统。上面给出的分类,只是希望读者了解实际工作中的有关术语,目的是便于全面理解操作系统的概念。

§ 1.2 文件系统管理

文件系统管理是操作系统五大资源管理功能之一。进程管理、内存管理、作业管理和设备管理更多地涉及到计算机的内部构成,这些管理功能的实现往往是计算机专业技术人员关心的内容。相比之下,和以应用为目的的用户关系最为密切的是文件系统管理。因此,本章重点介绍文件系统的概念以及典型的文件系统结构及管理方法。

1.2.1 文件的概念及分类

在计算机操作系统中,文件是指具有某种性质的信息集合。这一信息集合通常通过一个指定的名称,即文件名来区分。文件通常存放在计算机的外部存储设备中,如磁盘、磁带等。在一个存储设备上可以存放许多文件。通过文件名则可以对这些文件的内容进行读写操作。根据文件的不同用途或文件中信息的特征,可以对文件进行不同的分类。下面是一些经常提到的分类。

(1) 按文件的信息性质分

文本文件——文本文件中的信息是可用文字编辑程序直接处理的。这些信息按行存贮，输出的内容可供阅读。通常文秘人员使用计算机处理的文件均属此类文件。

代码文件——代码文件有时称二进制文件，文件中的信息往往是其他程序处理后的结果。这种文件的信息无法直接编辑处理，文件信息组织也无“行”的概念。

(2) 按文件的加工与被加工状态分

程序文件——程序文件进一步又分为源文件和可执行文件。源文件在信息的性质上是文本文件，在用途上则是各类计算机程序。源文件经编译后，产生可执行文件，可执行文件在信息性质上，则属于代码文件。

数据文件——数据文件是包含被加工信息的文件。在数据库系统中也称为数据库文件。

(3) 按文件的地位分

系统文件——系统文件通常是指随计算机系统提供的各类文件，或者是指与具体用户或应用领域无关的文件。如从文件的角度看，构成操作系统、编译程序、数据库管理系统等的文件均属系统文件。

应用文件——应用文件有时也称用户文件，即由某一领域的用户针对具体需要开发编写的程序等信息集合，如工资管理程序、报表打印程序等。

(4) 按文件的关系分

主数据文件——在数据处理中，数据文件由一个个记录形成，它们是数据加工的主要对象。相对于索引文件而言，称它们为主数据文件。如在一个职工档案管理系统中，包含职工档案的数据文件。

辅助文件——相对于主数据文件，为处理主数据文件而建立的有关文件，如为快速检索主文件中数据而建立的关键字索引文件。

(5) 按文件的保护权限分

含权限文件——文件作为信息资源，为了使其具有安全性，或为了管理上的方便，可以赋予各种保护或操作权限。最基本的权限有只读、只写、读或写等。

不含权限文件——可为任何人所读写的文件。

(6) 按文件的保存时间分

临时文件——在数据处理中临时生成的过渡性文件称为临时文件，当一次完整的数据处理任务结束后，它们由系统自动清除。

永久文件——在数据处理中，数据处理过程结束后仍保留的文件。

(7) 按对文件的存取方式分

顺序文件——指只能对其中内容按顺序依次访问的数据文件。

随机文件——指对其中指定位置的内容可随时访问的数据文件。

(8) 按文件的组织结构分

连续文件——指文件信息在存贮设备上占据了连续物理地址空间。

链表文件——文件信息分解存放在若干物理存贮块中，各物理块不必占有连续物理地址，通过指针建立起块间的联系。

(9) 按索引级别分

单重索引文件——索引文件是指由关键字与指针为记录构成的文件，由指针可迅速地找

到数据主文件中包含关键字信息的数据记录。若只有一级指针,则称为单重索引文件。

多重索引文件——当一个索引文件本身很大时,可以对它分类,并进一步建立索引文件,由此构成多重索引文件。

与操作系统分类一样,文件分类也无公认标准,上面的分类仅是为了帮助理解文件的概念,了解实际工作中常常提到的这些用语的含义。根据实际工作的需要,还可以列举出许多其他的文件类型,如根据源程序的语种称文件为 BASIC 文件、FORTRAN 文件、PASCAL 文件等。此外,不同的操作系统中也会引入各种专有的文件概念。如 UNIX 操作系统中,将文件分为普通文件、目录文件和特殊文件三大类。如此等等,不一一细举。

1.2.2 文件系统及其功能

在现代计算机系统中,文件的数量和种类越来越多。计算机需要根据使用者的要求快速地找出指定的文件,文件的内容需要在计算机的存贮与加工装置间频繁传递。如果将文件的这些管理任务全交由人工完成,显然是一个沉重的负担,不仅可靠程度差,而且会大大降低整个计算机系统的信息处理效率。实际上人工也难以完成。为此,有必要将所有的文件按系统方式加以组织,对其统一管理。由此形成了操作系统中的文件系统。

所谓文件系统是指按某种方式在计算机外部存贮器上组织的各种文件集合,以及实现对这一文件集合中指定信息存取、控制和管理和各种操作机制。因此,对文件系统的研究主要是对文件组织结构和在此结构上操作方法的研究。

引入文件系统后,许多基础性的操作均由文件系统的管理程序自动完成,从使用者的角度讲,它具有以下几个特点:

(1) 使用简单方便

由于文件物理结构和存贮设备与文件的使用相互独立,在用户眼里,文件只是具有名称等特征的连续信息集合,用户无需考虑这些信息在具体设备上如何存贮,可以用统一的概念和逻辑操作使用文件。

(2) 信息安全可靠

通过系统给文件施加各种特征信息,使得不同的用户有不同的使用权限。同时,管理程序能够提供文件备份或转贮的功能。

(3) 信息共享

对于重要的信息或系统文件,在文件系统的管理下可以避免重复占用存贮空间。尤其是在多用户环境下,文件系统提供的文件并发控制能力,使多个用户可同时访问同一文件。

1.2.3 文件系统的组织

一、文件的构成

文件是构成文件系统的基本单位。文件系统中,一个文件所包含的信息可划分为两部分。一部分称为特征信息,另一部分称为实际信息。实际信息也就是文件中的具体数据内容,属于文件的主体部分;特征信息包括文件的名称等,用于对文件实施检索与操作控制。

在文件系统中,文件的管理与日常工作中人工管理文件档案的方式是一致的,通过给文件指定名称、日期、密级等特征来识别文件和保护文件。从文件管理者的角度讲,其所关心的往往

是文件特征信息。在操作系统中,文件的特征信息包括了若干项内容,由这些内容构成的特征信息集合通常称为文件的目录项。

不同的操作系统中,文件目录项包含的内容可能是不同的。例如,在 DOS 操作系统中,一个文件的目录项包含了如表 1-1 所示的基本项,共占用 32 个字节。其中主文件名占用 8 个字节,对应于 8 个字符,因此,它限定了在 DOS 操作系统下一个文件的主文件名长度不能超过 8。除了主文件名外,一个文件还可有 3 个字符长度的扩展名或后缀。通常可以用扩展的文件名来标识文件中的信息特征。如用 FOR、PAS、BAS 来表示对应的文件是 FORTRAN、PASCAL、或 BASIC 程序,用 EXE 来标识对应的文件是可执行文件等。在 DOS 中,主文件名与扩展文件名一起构成一个完整的文件名称,书写时,二者之间用“.”号分隔。目录项中的文件属性指出了文件的读写权限。DOS 下的一个文件可以指定五种属性:系统属性(S),只读属性(R),隐含属性(H),档案属性(A)和可读写属性(无特殊代表符号),这些属性通过不同的编码状态表示。一个文件可以有多种相互兼容的属性,如一个文件可以是只读的,同时它又是隐含的。通过给文件指定属性,就能对文件实施必要的操作控制和安全保护。变更时间项记录了对文件最近一次更改的时间,而变更日期记录了最近一次更改文件的日期。文件大小给出了文件中包含数据的字节数。关于起始簇号这一特征信息的用途将留待后面介绍。

表 1-1 DOS 文件目录项的构成

目录项中的起始字节号	目录项的含义
0	主文件名
8	扩展文件名
11	文件属性
12	(DOS 留用)
22	变更时间
24	变更日期
26	起始簇号
28	文件大小

在 UNIX 系统中,文件的特征信息集合包含的内容更多,除了文件名外,其他的特征信息如表 1-2 所示。

表 1-2 UNIX 中文件目录项的构成

目录项的代号	目录项的含义
i-mode	文件属性
i-uid	文件属主
i-gid	文件属组
i-size	文件大小
i-nlink	文件联接数
i-addr(40)	文件数据区索引
i-atime	文件访问时间
i-mtime	最近访问时间
i-ctime	文件创建时间

其中,文件属性的职能与 DOS 中相同,但权限设置得更加精细。通过文件属主目录项,可指出文件的创建者或主人,而通过文件属组则可将文件分类编组。文件大小指出文件的字节数。由于 UNIX 属于分时多用户操作系统,因此,一个文件可以为多个用户服务,通过文件联

接数可指出文件当前被占用的状态。目录项中最后三项给出了文件建立、更改和访问的时间。关于文件数据区索引项的作用也将在后面介绍。

二、文件系统的组织

文件的组织与外部存贮设备是密切相关的,目前大量使用的外部存贮设备是磁带和磁盘。在具体介绍文件系统的组织前,有必要先了解一些描述外部存贮设备的概念。在文件系统中,可以把多个文件组织成一个文件卷(volume)。卷是存贮介质的物理单位。一般一个文件卷就对应于一盘磁带或一个磁盘上的文件。块(block)是存贮介质上连续信息所组成的一个区域,它是计算机内外存交换信息的基本单位。每当从磁盘读取数据时,计算机将其一次一个块调入内存缓冲区,然后,进一步从中筛选出指定数据;反之,往磁盘存入数据时,先将它们保存在缓冲区中,当形成一块后,一次将其写入磁盘。

由此会提出这样的问题,一个数据块应为多大?不难理解,若一个块很小时,它调入内存后筛选的时间会缩短,但产生的问题是内外存交换信息的次数将增加,因此,增加了数据读取时间。反之,若一个块太大,则占用内存空间会过大,同时,使数据传送时间加长。因此,恰当确定块的大小是很重要的。块的大小一般通过实验确定,它与信息的获取与传送速率有关,不同的操作系统块的大小可能不同。

下面针对 DOS 和 UNIX,分别介绍文件系统的存贮组织。

1. DOS 文件系统的存贮组织

在 DOS 下,存贮文件的主要设备是磁盘。首先让我们了解一些与磁盘设备有关的知识。磁盘是一种直接存取的外部存贮设备,又叫随机存取存贮设备。因为信息是按块组织的,因此也叫块设备。这种设备上,每个物理记录有唯一确定的地址。存取任何一个物理记录所需的时间基本上是不变的。在实际应用中,磁盘的种类很多。从几何尺寸上分,有 8 英寸,5 英寸,3 英寸盘等;从存贮容量分,有数百 K 到数百兆的范围;从硬件特征分,有固定式和活动式;从磁记录方式分,有单密度,倍密度,高密度等;从技术特征分,有温式盘和非温式盘;从磁介质的基底分,有硬盘和软盘;等等。然而,无论如何划分,使用磁盘的共同之处是,使用前需要对磁盘进行格式化。经格式化后的磁盘划分了磁盘上存贮介质的块的大小,同时确定了块的物理位置。

在 DOS 下,磁盘对信息的组织和存取方式都是相同的。不过,对不同的磁盘设备,将有不同的技术参数。下面我们以 5 英寸软盘为例说明磁盘格式化的作用,剖析文件在盘上的存贮形式。

我们知道,厂家出售的磁盘只是一个在基片上覆盖着磁性存贮介质的圆盘。这种磁盘称为空(白)盘。空盘如同一张白纸,为了在上面记录和读取信息,自然的方式是在上面“打格子”,通过格子的位置可以定位具体信息内容。在矩形纸张上打格子是划出纵横线条,而在圆盘上,“格子”由围绕圆心的“线”和径向的“线”构成。在这里,围绕圆心的线称为磁道。对不同类型的磁盘,磁道的数目可能是不同的。磁道通常由外向内从 0 起顺序编号。由相邻两个磁道和相邻两条径线所围成的区域构成了磁盘上的一个“格子”。每个格子称为一个扇区,同理,对扇区也可顺序编号。一个磁盘上扇区数目的多少与磁盘类型和磁道的疏密有关。在空盘上“打格子”,即确定磁盘的扇区数,与具体的操作系统紧密相关。不同的操作系统会有不同的“格子”数目。虽然打格子的具体任务是根据使用者的要求完成的,但确定其大小的参数却不是用户随意可变更的。实际的磁盘产品种类很多,表 1-3 给出了几种软磁盘的磁道与扇区数的技术数据。

对于一个具体的扇区而言,存贮信息的多少也不尽相同。一个扇区的存贮量用字节数度

量,通常它是二的倍数。在 DOS 下,典型的值为 256,512 和 1024。为了确定一个磁盘的最大存贮量,可用下面的公式计算:

表 1-3 DOS 下常用 5 英寸软磁盘的磁道与扇区数

磁盘类型	每道扇区数	每面磁道数	总扇区数
单面	8	40	320
单面	9	40	360
双面	8	40	640
双面双密	9	40	640
双面高密	9	80	1400
双面高密	15	80	2400

存贮容量=(盘面数)×(磁道数)×(扇区数)×(扇区大小)

例如:对常用的双面双密,5 英寸软盘,计算可知它的存贮量大约为 360K。

存贮容量=(2)×(40)×(9)×(512)=368 640≈360K

DOS 中,文件系统的特征信息集中组织,而文件的主体部分则按链表方式组织。按照这种组织结构,一个文件在磁盘上占用的空间一般不是连续的物理记录。通过设置指针的方式,将不连续的物理记录“链”成一串,从而可访问到整个文件的内容。图 1-1 是链表结构的文件组织示意。采用链表结构组织文件的好处是灵活,文件易于扩充,存贮空间可按链表节点申请和释放,故空间利用率高。

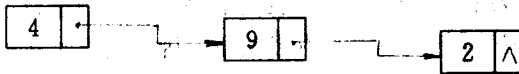


图 1-1 链表结构组织示意图

为了组织文件系统,DOS 对格式化后的磁盘上的存贮空间作了如图 1-2 所示的使用范围划分。其中除引导记录区外,其余部分均用于组织文件。

引导记录区
文件分配表区
根目录区
用户数据区

图 1-2 磁盘信息内容划分示意图

引导记录中存放了有关磁盘存贮介质的描述信息,以及引导操作系统启动的例行程序,详细的内容留待第二章 2.2.2 节介绍。(为了迅速地查找和调用它,系统固定地将引导记录存放在磁盘的 0 面 0 道 1 扇区。)

文件分配表(FAT)占用若干物理记录,它是 DOS 记录磁盘空间使用情况的信息区。实质上就是链表结构中的拉链指针区。文件分配表是 DOS 存取文件的重要一环,该表的项目个数与分配给文件主体的数据区拥有的物理记录数据相同,即每个扇区对应表中一项。表中每项通过编码的方式,反映对应扇区的状态,即是否已占用,以及该扇区是否有缺陷等。鉴于它的重要性,文件分配表实际采用备份技术,一式两份。当由于异常情况使其其中之一发生问题时,可由另一份加以恢复。

根目录区用于存贮文件卷的文件特征信息集合,即文件的目录项。DOS 中文件系统是按树状分类组织的,根目录区中只包含了处于根目录下的所有文件的目录项。因为,根目录区在磁盘格式化后确定了其占用的空间大小,而每个文件的目录项占用的字节数也是一定的,因此,在文件系统的根一级能存贮的文件数目是固定的。

在 DOS 下,为了组织数据的需要,引入了一个称之为簇(cluster)的数据存贮单位,一个簇也称为一个逻辑扇区,它对应于前述块的概念,是磁盘空间组织文件和存取信息的基本单位。视具体磁盘设备的类型,一个逻辑扇区可能包含一到多个物理扇区(即磁盘格式化所产生的扇