

电子计算机应用系列教材

中小型计算机 数据库简明教程

冯玉才 卢正鼎 曾昭苏 编著



科学出版社

3
1

电子计算机应用系列教材

中小型计算机数据库简明教程

冯玉才 卢正鼎 曾昭苏 编著

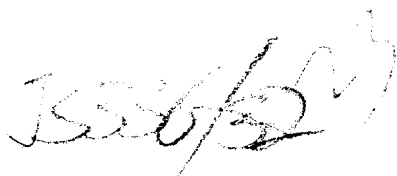
科学出版社

(京)新登字092号

内 容 简 介

本书全面而简明地论述数据系统的基本概念、基本方法和基本原理,介绍数据库技术的最新研究动态和研究成果.在叙述和内容上注重理论联系实际,以通俗易懂的语言和例子说明各种概念和技术.具有简明、通俗、实用的特点.全书共分六章,主要内容包括:数据库的实现与使用,数据库管理系统 Rdb,数据库的最新研究动态等.

本书可作为高等学校非计算机专业的本科生教材,亦可作各种专科班、培训班和从事数据库应用设计的工程技术人员的教科书或参考书.



电子计算机应用系列教材
中小型计算机数据库简明教程

冯玉才 卢正鼎 曾昭苏 编著

责任编辑 唐正必

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

华中电子信息产业开发集团公司激光照排

国防科工委印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1994年1月第一版 开本:787×1092 1/16

1994年1月第一次印刷 印张:13 1/4

印数:1—4000 字数:293000

ISBN 7-03-003549-6/TP·269

定价:11.30元

“电子计算机应用系列教材”主持、组织编著单位

主持编著单位:

国务院电子信息系统推广应用办公室

组织编著单位:

广东、广西、上海、山东、山西、天津、云南、内蒙古、
四川、辽宁、北京、江苏、甘肃、宁夏、江西、安徽、
河北、河南、贵州、浙江、湖北、湖南、黑龙江、福建、
新疆、广州、大连、宁波、西安、沈阳、武汉、青岛、
重庆、哈尔滨、南京等35省、市、自治区、计划单列市

电子振兴

计算机领导小组办公室

科技工作

“电子计算机应用系列教材”联合编审委员会名单

(以姓氏笔划为序)

主 编 审 委 员:

王长胤* 苏世生 何守才 陈有祺 陈莘萌* 邹海明* 郑天健
殷志鹤 童 颀 赖翔飞 (有“*”者为常务主编)

常务编审委员:

于占涛 王一良 冯锡祺 刘大昕 朱维华 陈火旺 陈洪陶 余 俊
李 祥 苏锦祥 佟震亚 张广华 张少润 张吉生 张志浩 张建荣
钟伯刚 胡秉光 高树森 徐洁盘 曹大铸 谢玉光 谢育先 韩兆轩
韩培尧 董继润 程慧霞

编审委员:

王升亮 王伦津 王树人 王振宇 王继青 王翰虎 毛培法 叶以丰
冯鉴生 刘开瑛 刘尚威 刘国靖 刘晓融 刘德镇 孙令举 孙其梅
孙耕田 朱泳岭 许震宇 何文兴 陈凤枝 陈兴业 陈启泉 陈时锦
邱玉辉 吴宇尧 吴意生 李克洪 李迪义 李忠民 迟忠先 沈林兴
肖金声 苏松基 杨润生 芮福德 张志弘 张银明 张 勤 张福源
张翼鹏 郑玉林 郑 重 郑桂林 孟昭光 林俊伯 林钧海 周俊林
赵振玉 赵惠溥 姚卿达 段银田 钟维明 袁玉馨 唐肖光 唐楷全
徐国平 徐拾义 康继昌 高登芳 黄友谦 黄 侃 程锦松 楼朝城
潘正运 潘庆荣

秘 书 组:

秘 书 长:胡茂生
副秘书长:何兴能 林茂荃 易 勤 黄雄才

序

当代新技术革命的蓬勃发展,带来社会生产力新的飞跃,引起整个社会的巨大变革.电子计算机技术是新技术革命中最活跃的核心技术,在工农业生产、流通领域、国防建设和科学研究方面得到越来越广泛的应用.

党的十一届三中全会以来,我国计算机应用事业的发展是相当迅速的.到目前为止,全国装机量已突破三十万台,十六位以下微型计算机开始形成产业和市场规模,全国从事计算机科研、开发、生产、应用、经营、服务和教学的科技人员已达十多万人,与1980年相比,增长了近八倍.他们在工业、农业、商业、城建、金融、科技、文教、卫生、公安等广阔的领域中积极开发应用计算机技术,取得了优异的成绩,创造了显著的经济效益和社会效益,为开拓计算机应用的新局面作出了重要贡献.实践证明,人才是计算机开发应用的中心环节.我们必须把计算机应用人才的开发与培养放在计算机应用事业的首位,要坚持不懈地抓住人才培养这个关键.

从目前来看,我国计算机应用人才队伍虽然有了很大的发展,但是这支队伍的数量和质量还远不适应计算机应用事业发展的客观需要,复合型人才的培养与教育还没有走上规范化、制度化轨道,教材建设仍显薄弱,培训质量不高.因此,在国务院电子信息系统推广应用办公室领导、支持下,全国35个省、市、自治区、计划单列市计算机应用主管部门共同组织118所大学和科研单位的400多位专家、教授编写了全国第一部《电子计算机应用人才培训大纲》以及与之配套使用的电子计算机应用系列教材,在人才培训和开发方面做了一件很有意义的工作,对实现培训工作规范化、制度化将起到很好的推动作用.

《电子计算机应用人才培训大纲》和“电子计算机应用系列教材”贯穿了从应用出发、为应用服务,大力培养高质量、多层次、复合型应用人才这样一条主线.培训大纲总结了近几年各地计算机技术培训正反两方面的经验,提出了计算机应用人才的层次结构、不同层次人才的素质要求和培训途径,制定了一套必须遵循的层次化培训办学规范,编制了适应办学规范的“课程教学大纲”.这部培训大纲为各地方、各部门、各单位制定人才培训规划和工作计划提供了原则依据,为科技人员、管理人员以及其他人员学习计算机技术指出了努力方向和步聚,为社会提供了考核计算机应用人才的客观尺度.“电子计算机应用系列教材”是培训大纲在教学内容上的展开与体现,是我国目前规模最大的一套计算机应用教材.教材的体系为树型结构,模块化与系统性、连贯性、完整性相兼容,教学内容注重实用性、工程性、科学性,并具有简明清晰、通俗易懂、方便教学、易于自学等特点,是一套很好的系列教材.

这部培训大纲和系列教材的诞生是各方面团结合作、群策群力的结果,它的公开出版和发行,对计算机应用人才的培训工作将起到积极的推动作用.希望全国各地区、各部门、各单位广泛运用这套系列教材,发挥它应有的作用,并在实践中检验、修改、补

充和完善它。

通过培训教材的建设，把培训工作与贯彻国家既定的成人教育、函授教育、电视教育和科技人员继续工程教育等制度相结合，逐步把计算机应用人才的培训工作引向规范化、制度化轨道，为培训和造就大批高素质、多层次、复合型计算机应用人才而努力奋斗，更好地推动计算机应用事业向深度和广度发展。

李祥林

1988年10月17日

前 言

数据库技术作为70年代数据处理的研究成果已非常成熟,在信息领域中发挥了巨大作用.

我国高等学校计算机专业开设这门课已有十几年的历史.最近几年,许多非计算机专业也都纷纷开设这门课,而各种计算机专业的短训班、培训班、夜大、专科班也都开设这门课.但是其目标和要求是不一样的.计算机专业开设这门课的目标是使学生掌握数据库技术的基本概念、基本理论、基本技术,以及系统的实现策略,而后者只需使学生了解使用数据库系统的知识和技术.但由于缺乏非计算机专业的数据库教材,因而他们只有借用计算机专业的教材,致使教学效果不甚理想.本书试图针对这一类读者,写出一本适合于他们教学所用的教材.

鉴于这一目标,本书在写法和内容上都与传统的数据库教材有很大差别.首先避免了使用枯燥的技术介绍概念、术语;其次在内容上做了较大变动,省略了较为原始的中国极少使用的IMS系统,只介绍层次模型的基本概念和原理,省略了关系理论的定理、推论的冗长证明,只介绍有实用价值的算法和结论.为增强实践教学,特介绍VAX上的Rdb系统及数据库设计,以及实现技术.而将传统教科书中的数据库保护一章的内容简化后融合到第一章中,使读者读完第一章后对数据库系统有个完整的概念.最后一章让用户了解当今数据库的最新研究动态.

本书由冯玉才主编,确定章节结构,对全稿进行统一整理,并编写第一、第二和第六章;卢正鼎编写第三、四章;曾昭苏编写第五章.由于作者水平有限,书中不妥、不当之处,请读者批评教正.

编者

1992年6月

目 录

第一章 数据库系统论概	(1)
1.1 一个管理信息系统的实例	(1)
1.1.1 实例的原型	(1)
1.1.2 手工处理	(2)
1.1.3 自由管理方式	(3)
1.1.4 文件管理方式	(3)
1.1.5 数据库管理方式	(6)
1.2 数据库系统的本质——变换	(9)
1.3 数据库系统的功能和特性	(13)
1.3.1 数据独立性	(13)
1.3.2 数据安全性	(14)
1.3.3 数据完整性	(14)
1.3.4 数据一致性	(15)
1.3.5 数据共享	(15)
1.3.6 控制冗余	(15)
1.3.7 集中管理	(16)
1.3.8 并发控制	(16)
1.3.9 故障恢复	(16)
1.4 数据库管理系统(DBMS)	(17)
1.4.1 数据描述语言	(18)
1.4.2 数据操纵语言	(18)
1.4.3 数据库实用程序	(19)
1.5 数据库系统的工作环境	(19)
1.5.1 硬设备	(20)
1.5.2 软件资源	(20)
1.5.3 数据库管理员	(20)
1.6 数据库系统的工作过程	(21)
小结	(22)
习题一	(23)
第二章 数据模型	(24)
2.1 实体及其联系	(24)
2.2 数据模型	(25)
2.3 实体-联系模型	(25)
2.4 层次模型	(26)
2.5 网状模型	(28)
2.5.1 网状模型的基本概念	(29)

2.5.2	DBTG 主要概念	(30)
2.5.3	DBTG 系统的描述语言	(31)
2.5.4	DBTG 系统的 DML	(31)
2.6	关系模型	(32)
2.6.1	关系模型的概念	(32)
2.6.2	关系模型的特点	(34)
2.6.3	关系模型的描述	(35)
2.6.4	关系模型的数据操纵语言	(36)
2.6.5	关系代数	(37)
2.7	关系数据库设计理论	(40)
2.7.1	问题的提出	(40)
2.7.2	函数依赖	(41)
2.7.3	函数依赖公理	(42)
2.7.4	闭包的计算	(42)
2.7.5	函数依赖集的最小集	(43)
2.7.6	关系模式的分解	(44)
2.7.7	关系模式的规范化	(46)
2.7.8	候选关键字的求解算法	(48)
2.7.9	图论在关系理论中的应用	(51)
习题二		(54)
第三章	数据库设计	(55)
3.1	数据库的筹划	(55)
3.1.1	可行性论证	(55)
3.1.2	确定设计目标	(56)
3.1.3	系统选择	(56)
3.1.4	制定开发规范	(58)
3.2	数据库设计实现运行综述	(59)
3.2.1	数据库系统的生命周期	(59)
3.2.2	数据库设计过程	(60)
3.3	需求说明与分析	(62)
3.3.1	数据信息与事务处理信息	(62)
3.3.2	调查	(62)
3.3.3	分析	(64)
3.4	概念设计	(66)
3.4.1	概念模式的表达工具——E-R 图	(66)
3.4.2	构造局部 E-R 图	(67)
3.4.3	构造整体 E-R 图	(72)
3.5	实现设计与物理设计	(73)
3.5.1	E-R 图到逻辑模式的转换	(73)
3.5.2	物理设计简介	(75)
3.5.3	逻辑模式的数据描述语言	(77)
3.6	基于函数依赖的关系模式规范化图解法	(81)

3.6.1	函数依赖图	(81)
3.6.2	函数依赖左部均为单属性时的图解法	(82)
3.6.3	函数依赖左部均为多属性时的图解法	(83)
3.7	关系模式规范化的 Smith 图解法	(85)
3.7.1	属性的依赖关系与依赖表	(86)
3.7.2	Smith 依赖图及化简法则	(87)
3.7.3	规范化模式的构成	(91)
习题三		(93)
第四章	数据库实现与使用	(94)
4.1	数据采集、模拟试验、数据加载	(94)
4.1.1	数据的采集	(94)
4.1.2	模拟试验与加载	(94)
4.2	应用程序的编制	(99)
4.2.1	用户如何使用数据库	(99)
4.2.2	数据库应用程序的开发	(100)
4.2.3	数据库应用程序语言基础	(101)
4.3	数据库维护	(104)
4.4	数据库开发维护工具	(105)
4.4.1	一般开发工具	(105)
4.4.2	计算机辅助数据库设计工具	(106)
习题四		(107)
第五章	关系数据库管理系统 Rdb	(108)
5.1	Rdb 简介	(108)
5.1.1	关系数据库管理系统 VAX Rdb/VMS	(108)
5.1.2	Rdb 的特色	(109)
5.2	Rdb 语言元素	(110)
5.2.1	Rdb 值表达式和条件表达式	(110)
5.2.2	Rdb 记录选择表达式	(113)
5.3	Rdb 数据库的创建	(115)
5.3.1	定义数据库	(116)
5.3.2	定义域	(116)
5.3.3	定义关系	(118)
5.3.4	定义约束	(120)
5.3.5	定义索引	(120)
5.3.6	定义视图	(121)
5.3.7	Rdb 数据库的创建方法	(122)
5.4	Rdb 数据库的保护	(123)
5.4.1	访问控制表 ACL	(123)
5.4.2	ACL 表的创建	(124)
5.4.3	Rdb 检查 ACL 表的方法	(125)
5.4.4	改造缺省的 ACL 表	(125)
5.4.5	改变保护	(128)

5.4.6 删除保护	(128)
5.5 Rdb 数据操纵	(129)
5.5.1 Rdb 数据库的装入	(129)
5.5.2 Rdb 数据库的打开	(130)
5.5.3 Rdb 记录流的形成	(131)
5.5.4 Rdb 单关系检索	(132)
5.5.5 Rdb 多关系检索	(134)
5.5.6 用视图使询问固定化	(135)
5.5.7 Rdb 数据库的更新	(136)
5.5.8 Rdb 事务	(137)
5.6 在应用程序中使用 Rdb	(144)
5.6.1 注意事项	(145)
5.6.2 用终端接口 RDO 开发 Rdb 询问	(145)
5.6.3 主变量和语句标志	(146)
5.6.4 数据类型的一致性	(147)
5.6.5 处理 Rdb 运行时产生的错误	(149)
5.7 Rdb 数据库的重构	(155)
5.7.1 改变关系、域、数据库	(155)
5.7.2 删除关系、域、数据库	(156)
5.8 Rdb 数据库的维护	(156)
5.8.1 Rdb 数据库的恢复方法	(157)
5.8.2 Rdb 数据库的性能改进	(159)
习题五	(163)
第六章 数据库的最新研究动态	(164)
6.1 分布式数据库	(165)
6.1.1 分布式数据库的特色	(165)
6.1.2 为什么要有分布式数据库	(166)
6.1.3 分布式数据库的体系结构	(167)
6.1.4 分布式数据库的透明层次	(167)
6.1.5 分布式数据库设计	(169)
6.1.6 分布式数据库查询	(170)
6.1.7 分布式数据库的事务管理	(170)
6.1.8 分布式数据库的并发控制	(170)
6.1.9 分布式数据库的目录管理	(171)
6.2 专家数据库	(171)
6.2.1 专家数据库的研究动机和目标	(172)
6.2.2 EDS 的用户集团	(173)
6.2.3 人工智能与数据库技术的发展	(173)
6.2.4 EDS 的系统结构	(175)
6.2.5 知识——数据模型和方法论	(177)
6.2.6 逻辑和数据库	(182)
6.2.7 Prolog 和关系数据库的接口	(184)

6.2.8 语义查询优化	(186)
6.2.9 智能数据库接口	(186)
6.2.10 结论	(186)
6.3 演绎数据库	(187)
6.3.1 演绎数据库的基本概念	(187)
6.3.2 演绎数据库的基本结构	(187)
6.3.3 演绎数据库的研究现状	(188)
6.3.4 已经实现的演绎数据库系统	(190)
6.3.5 演绎数据库与知识库、智能数据库	(190)
6.4 工程数据库	(191)
6.4.1 工程数据库的应用环境	(191)
6.4.2 工程数据库的特点要求	(191)
6.5 多介质数据库	(194)
6.5.1 多介质数据库管理系统的功能要求	(194)
6.5.2 多介质 DBMS 的体系结构	(194)
参考文献	(197)

第一章 数据库系统概论

诸位读者,当你拿到《中小型计算机数据库简明教程》这本书时,一定有这样的心理:什么是数据库?为什么要用数据库?如何使用数据库?也许你已经从其他地方了解到数据库方面的知识,你是否觉得它有点玄,令人费解?其实,数据库是一项实实在在的技术,它科学,巧妙,是用数据处理各种研究成果的结晶.为了说明这一点,让我们从一个实际例子说起,分别用文件技术和数据库技术来处理它,然后再来比较两种处理方法的优劣.你会很快悟出其中之奥妙.然后再上机用 dBASE(代表文件技术的特点)和其他的数据库管理系统(简称 DBMS,我们先用一下这个术语,以后再作介绍)来尝试一下数据处理的味道.在我们有了共同的语言之后,再结合上述提到的实例,逐一介绍数据库的各种概念、思想和处理技术,继而展开该书其他章节的内容,诸如数据模型,数据库设计、使用、维护等,再介绍一个新的 DBMS——Rdb 作为全书的结束.顺便说明一下,我们将用下面介绍的管理信息系统实例贯穿全书.

1.1 一个管理信息系统的实例

我们所举的例子是大家都熟悉的,尤其是学校的老师和学生更不会感到生疏.这就是学校的管理信息系统.

1.1.1 实例的原型

学校的管理主要是人事、科研、教学、财务、图书、医务、学生等管理.为了简明起见,我们假定每个处只有少量文件,文件中仅包括本部门最重要的、最典型的一些信息.图 1.1 是各个文件的结构.

图 1.1 是经过简化后的学校各主要部门信息管理的原始模型.用户的要求和信息内容是多种多样的,有时差别很大,甚至互相矛盾.有时我们深入到一个部门去了解其管理信息的要求、内容,会被那里的错综复杂的数据和报表弄得眼花缭乱,但是透过这些繁杂的表面现象,我们可以看出整个管理信息实际上包含两大方面的内容:

- (1) 信息流,处理数据的过程;
- (2) 物流,处理业务的过程.

物流伴随着信息流,信息流必须依附于物流.无论是何种处理方式,都必须弄清这两方面的问题.而手工处理的过程是计算机处理过程的基础,所以在用计算机管理信息前必须首先调查应用的实际背景和全貌,这就是第三章中要讲的需求调查和分析.以下我们分别介绍用手工处理、计算机自由管理、文件管理、数据库管理几种方式来解决这一问题的过程,然后再讨论它们的特点和进化,从而认识和理解什么是数据库技术的本质.

(1) 科研处

序号	项目名称	级别	完成情况	奖励	经费	项目负责人		
						姓名	职称	所在系

(2) 教务处

课程号	课程名	时数	教材	课表安排		任课教师			
				时间	教室	序号	姓名	职称	所在系

(3) 图书馆

登 录 号	分 类 号	书 号	作 者	出 版 社	价 格	套 数
-------	-------	-----	-----	-------	-----	-----

序 号	姓 名	职 称	所 在 系
-----	-----	-----	-------

序 号	登 录 号	借 期	还 期
-----	-------	-----	-----

(4) 财务处

序 号	姓 名	基 本 工 资	工 龄 工 资	薪 级 工 资	水 费	奖 励
-----	-----	---------	---------	---------	-----	-----

(5) 医院

序 号	姓 名	年 龄	性 别	职 称	单 位	身 高	体 重	血 压	心 率	肺 结 核
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------

(6) 人事处

自然状况					政治状况				业务状况		系名	基本工资	健康状况
姓名	年龄	性别	学历	职称	序号	政治面貌	简历	奖惩	科研	教学			

(7) 学生管理处

序 号	姓 名	性 别	年 龄	班 级	总 成 绩	所 在 系
-----	-----	-----	-----	-----	-------	-------

(8) 院办

系代号	系 名	系主任	专业个数	教师数	学生人数
-----	-----	-----	------	-----	------

图 1.1 学校管理信息系统原型

1.1.2 手工处理

目前,我国学校的各项管理基本上仍是手工方式.在这种方式下,要几百名管理人员进行大量的烦琐的管理工作.这种手工处理的效率很低,质量也很差.他们用手抄写大量的数据,绘制繁多的报表,用眼查看记录在纸上的成千上万的数据.有时为了统计一个

数据,他们要花费大量的时间,进行艰苦而琐碎的统计计算,翻阅大量的卷宗和资料,而得到的数据可能已经过时了.比如图书资料管理,大量的工作是办理图书的借还手续,以及整个图书馆图书、资料的分类,登记,存放.其工作十分繁重,整个图书馆需要上百个管理人员.他们成天忙于事务,无时间钻研业务,而且,工作的效果未必很好.这也是广大管理人员迫切要求使用计算机进行管理的原因.但是如何用计算机进行管理,却有多种不同的方式和策略,其效果也不大一样.归纳起来有三种管理方式:自由管理方式、文件管理方式和数据库管理方式.这三种管理方式将在以下三节分别予以介绍.

1.1.3 自由管理方式

50年代计算机主要用于科学计算,所用的语言是机器语言.这种应用的特点是计算多数据少.因此,数据未独立分出来,而是与程序交织在一起,形成一个不可分离的整体.程序员对于他所用的数据在内存、外存的位置要完全清楚,并由其应用程序完全控制.看起来用户对他的数据握有绝对的权利,可以任意处置,但这并非好事,相反会带来几个坏处:

(1) 程序员要把一大部分精力消耗在繁琐的数据存贮、存取细节上,无时间钻研业务;

(2) 数据与程序交织在一起,大大增加了程序的复杂程度.

这种方式对于拥有大量数据的数据处理系统更带来众多的麻烦和问题.这也正是当时数据处理在计算机应用中难以推广、所占比例极少的根本原因.要改变这种状态,只有将数据与程序进行分离.这就是下面要讨论的方式——文件管理方式.

1.1.4 文件管理方式

把数据与程序分离开来独立管理,这是计算机数据处理史上的第一个里程碑.数据处理后来能发展得这样迅速,以至达到今天的先进程度,应首先归功于这一技术.因为只有两者分离,才能使两者比较清晰地按各自的特点向前发展.这不仅有利于数据处理,也有利于程序设计技术的发展,比如结构化程序设计方法以及软件工程中用的一系列方法无不与此有密切的关系.很难设想,混杂着大量复杂数据的程序会有一个清晰的模块化结构.

既然数据已经脱离了程序,我们就可以集中精力研究数据处理的规律和形式.首先,像科学计算那样将数据随意地、杂乱地在内存中存放是再也行不通了.虽然书架上存放的书很多,但主人能信手取出所需的书;仓库中存放的货物很多,而管理员能毫不费劲地找到所要的货物.这都是因为书和货物都是分门别类、按照某种规律存放的.同样,数据在计算机中也应按照某种规律分类排序,并有统一的格式来组织存放.这就是文件管理的方式.通常,我们把结构相同、性质相近的同类数据的集合称为文件.文件的一个元素称为记录.记录是由一些项组成的,每个项表示了事物的一个特征.管理文件的软件叫文件管理系统.比如图 1.1 中每个处(或相应的部门)都有一个所需数据的框架,这个框架就是一个文件的结构形式或记录格式,其中的一栏就是一项,例如学生管理中的记录格式即图 1.1 中的(7),学号、姓名、年龄、性别等就是项,或称数据项,而对应记录格式的一组值就是一个记录值,全部记录值的集合就是一个文件,即学生文件.

数据组成文件以后就独立存在于程序之外了。每个文件有一个文件名,它存贮在外存器上。用户通过文件名可以在程序中逻辑地调用它。而具体的物理存贮用户程序就不用过问了。文件管理系统就是做这个工作的。它是应用程序和数据文件之间的一个接口。应用程序只有通过文件管理系统才能建立和存取文件。

现在我们来看看用文件管理方式如何来进行学校的管理信息系统。

首先我们要为每个处(或部)建立各自的文件结构,这里一共有 10 个文件,除了图书馆以外其他各处各有一个文件。这些文件是:

- 科研管理文件;
- 教务文件;
- 图书文件;
- 借书人文件;
- 借书登记文件;
- 工资文件;
- 体检表文件;
- 人事文件;
- 学生文件;
- 系文件。

各文件的格式即图 1.1 中所示的框架。

其次,我们用高级语言,如 COBOL,FORTRAN,BASIC 等来描述这些文件结构。由于绝大多数读者对流行的 dBASE 系统都有不同程度的了解,而 dBASE 本质上是一个文件系统,它的描述形式和使用形式与文件形式完全一样,其差别在于 dBASE 提供了终端接口,而文件系统无此功能。所以我们就借用 dBASE 来代替文件系统。

请你准备一张 dBASE 的盘,在操作系统状态下键入 dBASE,当屏幕上出现圆点以后,再键入建库命令 CREATE 及回车键:

CREATE

在输入文件名以后便开始定义文件结构。我们先定义科研文件,其他文件的定义由读者在计算机上完成。这里要说明的一点是,dBASE 不能处理组项(由几个项组成的项,如科研文件中的项目负责人),组项在应用程序中用报表生成系统 REPORT 命令或由用户自己来设计。在描述结构和装配数据时只定义单项(如项目负责人的姓名,职称,所在系等)。

以下是科研文件的描述格式:

FIELD NAME,TYPE,WIDTY,DECIMAL PLACES

001	序号,	N,	6
002	项目名称	C,	10
003	级别,	C,	6
004	完成情况,	C,	10
005	奖励,	C,	10
006	经费,	N,	10, 2
007	姓名,	C,	8
008	职称,	C,	6
009	所在系,	C,	10

在定义了文件结构以后,我们就可以装载数据了. dBASE 有多种命令可以完成此工作,你可以任用一种. 比方在定义完结构后,系统问:

INPUT DATA NOW ?

若你按下“Y”,则屏幕上出现了装载的环境:

RECORD # 0001 (你可开始键入数据)

序 号: 1

项目名称: 智能数据库

级别: 国家基金委员会

完成情况: 正在进行

奖励:

经费: 人民币 5 万

姓名: 史文志

职称: 教 授

所在系: 计算机

这样就完成了第一个记录的装入,你可继续装入其他的记录:

记录编号	序号	项目名称	级 别	完成情况	奖 励	经费(万元)	姓 名	职 称	所在系
00001	1	智能数据库	国家基金	正进行		5.00	史文志	教 授	计算机
00002	2	优化设计	七五项目	正进行		8.00	王大明	讲 师	机械系
00003	3	机器人	高科技	正进行		40.00	鲁 明	教 授	自控系
00004	4	建筑 CAD	省 级	已完成	一等奖	10.00	李 志	副教授	建筑系
00005	5	水电站自动管理	部 级	正进行		12.00	杨 成	副教授	电力系

在装入数据以后就可进行文件的各种操作,如查询,修改,删除等等. 你还可以用高级语言或 dBASE 的命令文件来编制你所需的事务处理程序.

同样你可通过进行其他部门的文件处理,来增加你对文件的感性知识.

通过上述文件的处理方式,我们可以看出文件管理系统有下列特点:

(1) 文件的专有性. 每个文件只属于特定的用户. 它的文件结构,文件内容,只有这个用户知道,它是针对该用户(例如上述科研文件是针对科研处)而设计的.

(2) 由于上述原因,所以文件的设置只对单个用户为最优.

(3) 文件的分散性. 由于各个用户并不考虑其他用户的使用要求,各自设计自己的文件,所以各文件是分散的,相互独立的. 文件与文件之间不存在任何联系.

(4) 文件系统只负责 I/O. 它对于文件内容不解释,不过问.

(5) 文件的逻辑结构等于物理结构,因而数据与程序间仍然有千丝万缕的联系. 这就是稍后要介绍的数据独立性问题.

(6) 数据存取以记录为单位.

(7) 数据间存在大量的重复,即冗余性. 我们看到在上述 10 个文件中,序号出现 7 次,姓名出现 6 次,职称出现 5 次,其他还有一些属性重复. 要知道这是一个人的情况,如果有 1000 个人,则重复就非常可观.

(8) 无集中管理. 文件系统只管这些文件的输入输出,而不保证这些文件的安全性、完整性.

文件系统的文件与应用的关系如图 1.2 所示.

文件系统比起自由管理方式来是大大进化了,最大的优点是数据与程序分离,这才使得尔后的数据处理技术得以独立地迅速地发展. 有了文件系统,用户省去了不少数据存