

COMPUTER

<http://www.phei.com.cn>

高等学校计算机基础及应用教材

数据库 技术基础及应用



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

刘国荣 编著

高等学校计算机基础及应用教材

数据库技术基础及应用

刘国燊 编著

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书是“高等学校计算机基础及应用教材”之一, 主要介绍数据库技术, 并力求体现三个特点: 基础性、实用性和先进性。全书共分 4 章。

第 1 章介绍数据库系统基本概念, 由浅入深地讲述数据库各种术语和数据库系统的体系结构。第 2 章是关系数据库基础理论, 以循序渐进的方式完整地讲述关系运算理论和关系规范理论。第 3 章以 SQL 语言为主线介绍 Access 2000 和 SQL Server 2000 两个软件的使用, 详尽讲解 SQL 语句和各种数据库对象操作, 并列举大量的实例。第 4 章主要介绍数据库应用系统的实用开发技术, 内容包括数据库应用系统的开发步骤, 使用各种流行工具和编程接口访问数据库 (ODBC, ADO, VB), 数据转换服务 DTS, 数据集成 (Access 和 SQL Server), 选择数据生成 Web 页面, 用 ASP 访问数据库等。

本书作为教材使用, 对知识点和实际操作的解释比较详细, 既适合高等学校计算机科学与技术专业本科生和大专相关专业的学生使用, 也可作为企、事业单位程序员和数据库管理员参考之用。

本课程在华南理工大学的计算机基础教学网公开教学, 学生可在网上选修。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有, 侵权必究。

图书在版编目(CIP)数据

数据库技术基础及应用/刘国荣编著. —北京: 电子工业出版社, 2003.9

高等学校计算机基础及应用教材

ISBN 7-5053-8238-1

I. 数… II. 刘… III. 数据库系统—高等学校—教材 IV. TP311.13

中国版本图书馆 CIP 数据核字 (2003) 第 082393 号

责任编辑: 刘宪兰 特约编辑: 明足群

印 刷: 北京牛山世兴印刷厂

出版发行: 电子工业出版社 <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销: 各地新华书店

开 本: 787×980 1/16 印张: 23.75 字数: 504 千字

版 次: 2003 年 9 月第 1 版 2003 年 9 月第 1 次印刷

印 数: 6 000 册 定价: 30.00 元

凡购买电子工业出版社的图书, 如有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系。联系电话: (010) 68279077。质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

高等学校计算机基础及应用教材

编委会

编委会顾问(按姓氏笔画排序)

史济民 艾德才 刘瑞挺 杨英清

编委会主任(按姓氏笔画排序)

李仲麟 韩国强

编委会主编(按姓氏笔画排序)

周雷如 黄钦胜

编委会委员(按姓氏笔画排序)

许兴存 刘国荣 官士鸿
陈立行 张 齐 张益良
贾德良

总 序

当今，很难找出一个不需要应用计算机技术的领域，这意味着几乎所有技术人员都需要学会使用计算机。对计算机了解的深度，不同领域、不同岗位上的工作人员有不同的要求。有些领域要求每个技术人员都对计算机技术有较深入的了解，例如高等学校的电子类专业都开设了大量计算机技术课程。有些领域的大多数工作人员只要对计算机应用有一般性了解就可以了。无论哪一类专业都要开发本领域的专用计算机应用系统，因而需要一大批既懂计算机技术又懂本领域业务的技术人才。造就这类人才的途径无非三种：一是设立专门的培养复合型人才的专业，由于行业种类、人员层次太多，需求不易量化，难以规划；二是由计算机专业人员深入某个应用领域，学习必要的专业知识并与该领域技术人员合作；三是各非计算机专业领域的技术人员（在学的或在职的）根据工作需要和个人志趣进一步学习有关的计算机技术。由途径二和途径三成长起来的计算机应用人员各有长处，可以互相补充，数量自然会由需求调节，无需行政规划。

编写本系列教材的主要目的是为选择上述第三条途径的人员（首先是高校非计算机专业的学生）提供一套比较系统又比较灵活、实用的学习材料。“系统”是指系列教材的编排从入门开始，循序渐进，涉及的预备知识均包含在教材中。“灵活”是指读者在读完系列教材中的“计算机实用技术”（内容大致相当于教育部提出的计算机基础教育第一层次的要求）、一门“程序设计”语言和“计算机软件与硬件基础”（基本属于第二层次的内容）后，可以根据需要任意选择学习其他课程。“实用”是指根据本系列教材的主要目标，取材注重应用，不追求完备。为了方便教学，每门教材都配有CAI课件，有些教材还有配套的习题集或实验指导，供任课教师和读者选用。

从上述关于内容的说明可见，系列教材中的“计算机实用技术”和一门语言的“程序设计”事实上是理工科非计算机专业学生的必修课，其余课程是选修课。

某些非计算机专业对计算机技术训练的要求与计算机应用专业的某些方向的要求交叉，各领域计算机应用系统开发所涉及的技术也有许多相同之处。因此，本系列教材中有一部分也可供计算机专业的相关课程选用。例如，本系列教材中的各种高级语言程序设计的教材，对于选择该语言作为第一门程序设计语言的任何专业都是适用的。又如“计算机组成原理”、“微型计算机接口技术”、“网络技术应用基础”等也可供计算机应用专业选用。

系列教材的作者都是有丰富教学经验的在职一线教师，以计算机系的教师为主，广泛征求相关专业教师的意见并且吸收部分相关专业教师参与编写。尽管经过反复讨论修改，但限于作者水平和其他条件的限制，在总体布局、内容取舍或其他方面一定还存在不足和值得商榷之处，敬请批评指正。

高等学校计算机基础及应用教材编委会

2002 年 11 月

前 言

21 世纪,人类社会已进入信息和网络时代。计算机应用已普及各行各业、深入到人类各种社会活动和生产活动中。在这种背景下,学校里不同的学科、不同的专业都把某些计算机课程列为必修课。编写“高等学校计算机基础及应用教材”的目的就是为了满足教学的需求,该系列教材的内容涵盖了本科生和大专生必须掌握的计算机软件、硬件的主干课程内容,而“数据库技术基础及应用”是这套系列教材中的一门软件课程,主要介绍数据库技术。

计算机学科正以前所未有的惊人速度向前发展,新的研究成果、新的实用软件不断涌现,流行软件的版本也不断更新。反观目前计算机课程的一些教材,未能做到与时俱进、及时更新内容,教材建设不尽人意。鉴于此,在本书的内容编排上,除了介绍数据库的基础理论外,还要摒弃一些过时的内容和产品,力求反映数据库技术方面的一些较新的成果。例如,对已经过时的层次模型、网状模型不予详述;对早期的产品如 FoxPro,因嵌套三层以上的 Select 语句已不能正常运行,其安全性、灵活性及连网能力等远逊于 SQL Server,本书也不予介绍。根据实际的需求,本书增加了一些数据库应用系统实用开发技术的内容,如 ODBC 的使用。目前 ODBC 已被广泛采纳,几乎每个流行的数据库都带有 ODBC 驱动程序,应用程序通过 ODBC 可以访问多操作系统平台上不同类型的数据库,包括异构数据库。学生若掌握 ODBC 的知识,可使解决实际问题的能力进一步得到提升。此外,本书还介绍了数据转换服务 DTS, ADO 控件, ADO 对象模型, VB 应用程序开发,用 Web 页面和 ASP 页面访问数据库等内容,努力追求理论与新技术(新软件)相结合。

希望学生学完本课程后,能够站在开发阵地的前沿上,充满自信地迎接新的挑战。

在本书的编写过程中,周霭如、刘艳霞、黄小兵老师仔细校对了全书;张庆、冼进老师对教材配置的习题提出了许多宝贵的意见;本书的责任编辑刘宪兰同志在全书的标准化方面做了大量严谨而细致的工作;在此,向他们致以衷心的感谢。

编著者

2003 年 8 月

目 录

第 1 章 数据库系统概论	(1)
1.1 信息、数据和数据处理	(2)
1.1.1 信息与数据	(2)
1.1.2 数据处理	(3)
1.2 数据管理技术的发展	(4)
1.2.1 人工管理阶段	(4)
1.2.2 文件系统阶段	(5)
1.2.3 数据库系统阶段	(8)
1.3 数据模型	(11)
1.3.1 数据与数据联系的描述	(11)
1.3.2 实体联系模型	(14)
1.3.3 结构数据模型	(17)
1.4 数据库 (DB)	(23)
1.4.1 三级模式结构	(23)
1.4.2 两层映射和两级数据独立性	(27)
1.4.3 DBA 及数据语言	(28)
1.5 数据库管理系统 (DBMS)	(32)
1.5.1 DBMS 的主要功能及其组成	(33)
1.5.2 用户通过 DBMS 查询数据的过程	(35)
1.6 数据库系统 (DBS)	(36)
1.6.1 DBS 的组成	(37)
1.6.2 DBS 的主要特征	(38)
本章小结	(39)
习题 1	(40)
第 2 章 关系数据库理论基础	(43)
2.1 关系模型的基本概念	(44)
2.1.1 关系的定义和基本术语	(44)
2.1.2 关系模型及其形式定义	(48)

2.2	关系运算	(59)
2.2.1	关系代数	(60)
2.2.2	关系演算*	(70)
2.2.3	关系运算安全约束及查询优化简介	(76)
2.3	关系模式设计	(78)
2.3.1	关系模式设计问题	(78)
2.3.2	函数依赖	(80)
2.3.3	关系模式的分解*	(91)
2.3.4	关系模式的范式	(100)
	本章小结	(113)
	习题 2	(114)
第 3 章	SQL 及其应用	(121)
3.1	SQL 概述	(122)
3.1.1	SQL 发展概况	(122)
3.1.2	SQL 数据库体系结构	(122)
3.1.3	SQL 的组成及语法描述约定	(123)
3.1.4	Access 2000——SQL (J-SQL) 的使用	(124)
3.1.5	SQL Server 2000——SQL (T-SQL) 的使用	(129)
3.2	SQL 的数据定义	(137)
3.2.1	数据库的创建、修改和删除	(137)
3.2.2	表结构的创建、修改和删除	(141)
3.2.3	视图的创建、修改和删除	(151)
3.2.4	索引的创建和删除	(154)
3.2.5	SQL Server 其他数据库对象的定义	(157)
3.3	SQL 的数据查询	(165)
3.3.1	SELECT 语句的基本句法	(166)
3.3.2	SELECT 语句的完整句法	(177)
3.4	SQL 的数据更新	(189)
3.4.1	插入数据——INSERT 语句	(189)
3.4.2	修改数据——UPDATE 语句	(192)
3.4.3	删除数据——DELETE 语句	(194)
3.4.4	可更新视图	(196)
3.5	T-SQL 游标的使用	(201)
3.5.1	声明游标	(202)

3.5.2	打开游标	(206)
3.5.3	使用游标操作数据	(206)
3.5.4	关闭游标及删除游标	(212)
3.6	SQL 的数据保护	(212)
3.6.1	事务机制与数据的可恢复性	(212)
3.6.2	锁定机制与并发控制	(216)
3.6.3	数据库备份和还原	(221)
3.6.4	数据复制*	(231)
3.6.5	安全性控制	(243)
本章小结		(257)
习题 3		(259)
第 4 章	数据库应用系统的开发	(265)
4.1	数据库应用系统的开发步骤	(266)
4.1.1	规划	(266)
4.1.2	需求分析	(267)
4.1.3	概念模型设计	(268)
4.1.4	逻辑设计	(270)
4.1.5	物理设计	(272)
4.1.6	程序编制及调试	(274)
4.1.7	运行和维护	(275)
4.2	通过 ODBC 访问数据库	(275)
4.2.1	ODBC 的体系结构	(276)
4.2.2	创建 ODBC DSN	(277)
4.2.3	用 ADO 控件访问数据库	(282)
4.3	应用程序设计	(286)
4.3.1	直接编写 ODBC 应用程序*	(286)
4.3.2	VB 访问数据库编程接口概述	(289)
4.3.3	使用 VB 的 ADO 对象模型访问数据库	(290)
4.4	数据转换和集成	(308)
4.4.1	SQL Server DTS 的导入/导出向导	(309)
4.4.2	Access 与外部数据的连接	(316)
4.5	Web 页面与数据库*	(325)
4.5.1	Web 页面概述	(325)
4.5.2	用 Web 助手向导生成 Web 页面	(331)

4.5.3 用 ASP 访问数据库*	(341)
本章小结	(358)
习题 4	(359)
附录 A 有关算法方面的习题参考答案	(361)
参考文献	(367)

第1章

数据库系统概论

从20世纪60年代开始,数据库技术不断地持续发展,已经成为计算机科学的一个重要分支。数据库具有数据结构化、冗余度低、数据独立性高、容易扩充和容易操纵等特点,较大的信息系统都是建立在数据库设计之上的。目前,数据库管理系统作为通用的系统软件,不仅在大、中、小型计算机上,甚至在微型计算机上都成为必备的基本配置。数据库应用也从最初的一般管理渗透到各行各业,在国民经济发展中发挥着越来越大的作用。

数据库技术有一套颇具规模的理论体系,它是一门应用广泛的实用技术。不论计算机专业或非计算机专业的学生和技术人员,掌握好这门技术,一定可在计算机应用开发的实践中找到自己的位置。

本章以传统的集中式数据库系统为背景,介绍数据库中的基本概念及有关名词术语,使读者对数据库技术的概貌有所了解。

1.1 信息、数据和数据处理

在系统地介绍数据库系统之前,首先要准确地把握好信息、数据和数据处理这几个基本概念。

1.1.1 信息与数据

信息与数据是既紧密联系又相互区别的一对术语。

信息是现实世界各种事物的特征、形态以及不同事物间的联系等在人脑里的抽象反映。对这些经抽象而形成的概念,人们可以认识理解,可以加工传播,可以进行推理,从而达到认识世界、改造世界的目的。时至今天,人们已经越来越认识到研究和利用信息的重要性。能源、物质和信息已并列为人类社会活动的三大要素。

可以毫不夸张地说,人类的社会活动是受信息支配的,特别是在信息极度膨胀的今天尤为如此。例如,天气预报说:“今天阴天,有雨,气温下降”,人们知道这条信息后,就会多穿点衣服御寒,出门时带上雨具。再以一个从事商品生产的企业为例,他们要花时间组织人力从事市场信息的收集和整理工作;企业的决策者分析这些经过综合整理的信息后,做出决策:是增产还是减产或是停产,并据此指挥生产部门的生产活动。可见,信息在人类社会活动中有着非常重要的地位。

数据是表达信息的载体。换言之,为了表达信息(抽象概念),必须使用某些符号,这些符号就叫数据。广义地讲,数据是一切文字、符号、声音、图像等有意义元素的组合。

例如,电视台要传播上述的那条天气预报信息,是通过在屏幕上显示文字、图标以及广播员用声音来传播的。这些文字、图标、声音就是载荷信息的数据。因此,数据是信息的载体,而信息是数据的内涵。信息依赖数据来表达,数据则生动具体地表现信息。这就是为什么说“信息与数据是既紧密联系又相互区别的一对术语”的原因。同一信息可以有不同的数据表示形式,同一数据也可能有不同的解释。

对数据这一术语,在这里再做一些解释。第一,不能把数据简单地与数字等同起来。数值数据仅是数据之中的一种类型,在数据库系统里处理得更多的是非数值数据,包括文字、图形、声音等。第二,数据在计算机中存储和处理时,都转换成计算机能够识别的符号,即只用0和1两个符号编码的二进制串来表示。例如,数值数据12,在计算机内用一个字节(即8个二进制位)表示为00001100。又如,英文字母A,在计算机以其内部码表示为01000001,直到要在屏幕显示或在打印机上印出时,才把内部码转换成A的点阵字模输出。对于声音和图形等类型的数据,也是按同一原理转换成二进制串进行处理的。了解数据在计算机内部形式和外部形式的区别,将有助于后续内容的学习。最后,尽管信息

与数据两个术语严格地讲是有区别的,但在很多场合下,往往无需严谨地加以区分也不致引发误解。因此,很多时候对这两个术语的使用较为随意。例如,可以说“数据处理”,也可以说“信息处理”,反正,数据处理本质上就是信息处理。

1.1.2 数据处理

现实世界事物的存在特征和变化状态有其自身规律,各种事物之间也存在着种种联系。基于这一事实,人们可从获知的信息中,经过加工推理,得到新的信息,并以此作为社会活动的决策依据。这一过程就叫做**信息处理**。由于信息是用数据表示的,所以信息处理具体地体现在数据处理上。

数据处理是指对数据的收集记载、整理、组织、存储、检索、计算/加工、维护、传送等一系列活动的总和。其基本目的是从已知数据出发,参照各种相关数据,经过计算/加工,得到新的数据;新的精炼数据(表示新的信息)反映事物的本质和特征以及事物间的内在联系。这些有价值、有意义的数据成为人们决策的依据。

数据处理的特点是数据量大,数据间的逻辑联系错综复杂;但计算相对比较简单,较少涉及复杂的数学模型和计算方法。因此,数据处理过程中的主要矛盾不是计算,而是如何把数据组织管理好。数据库技术正是瞄准这一目标,为有效地管理数据、最大限度地减轻程序员的负担,提高数据处理的效率而逐渐完善起来的计算机软件技术。

在这里,对与数据处理概念有关的问题再做一些说明。

首先,数据处理的由来已久。早期人们使用的工具是算盘、手摇计算机等,这是手工数据处理阶段。之后,出现了穿孔机、卡片机、制表机等机器,通常把这种用机械设备进行数据处理的系统称为机械数据处理系统。直到使用电子计算机,才使数据处理工作发生革命性的变革,计算机存储容量大、计算速度快、处理能力强、自动化程度高,使用电子计算机的数据处理系统称为电子数据处理系统。这种崭新系统的效率和效益是以往任何一种系统都不能比拟的。本课程的内容均系基于电子数据处理系统而言的。

其次,用电子数据处理系统进行数据处理工作,涉及两个主要的问题:数据处理方式问题和数据管理问题。

数据管理是一切数据处理过程的基本环节,是所有数据处理任务必需的共性部分。如何组织管理数据,是数据库系统要研究和解决的问题。本章后述的内容将较为详细地进行讨论。

数据处理方式主要有集中处理方式、分散处理方式和分布处理方式三种。简单地讲,如果把要处理的数据作为一个完整的集合,所有的工作都由一台计算机来完成,称为集中处理方式;若把数据集合成多个部分,分别放在多台计算机上,对数据的管理和操纵都分开并且相互独立,彼此之间没有数据通信联系,这就称为分散处理方式。若把分散在多

台计算机上的数据通过网络连接起来,数据虽然分开存储,但逻辑上却仍是一个整体。每台计算机既可单独进行“局部处理”,多台计算机也可以合起来进行“全局处理”,如同集中处理方式一样,这就是分布处理方式。

数据处理方式与计算机的硬件系统结构和软件都有关系。假设在分散处理方式的基础上,仅简单地用网络电缆把计算机连接起来而没有相应软件支持的话,是不可能像分布处理方式那样把分散的数据作为一个整体来处理的。在设计一个应用系统时,到底采用哪种方式比较好,通常取决于数据处理规模的大小和数据处理的复杂程度。下面,从硬件系统结构的角度的分类,简单介绍不同数据处理方式的一些应用情况。

1) 微型机类

(1) 单台微型机独立处理方式——任何时候只能一人操作一台微型机。

(2) 多台微型机分散处理方式——通常,在一个单位中,不同的部门配置各自的微型机。各管各的数据,微型机之间的数据传递只能用磁盘交换。

(3) 多台微型机分布处理方式——多台微型机连网,配置相应软件,实现分布处理。

2) 大、中、小型计算机类

这些计算机与微型机相比较,显然功能强得多。特别地,一台主机可带多个终端。简单说,终端就是只有键盘和专用显示器的设备。这种一台主机带多个终端的情况又称连机。主机使用时间分片的方法轮流为多个终端用户服务。这些计算机相互也可以连网,包括和微型机连网。

(1) 单机**集中处理方式**——本书第1章就以这种单机集中处理方式为背景,介绍数据库系统的基本原理。

(2) 多机**分散处理方式**——数据分散,计算机的硬件资源和软件资源不能共享。

(3) 多机**分布处理方式**——由多台计算机组成分布式系统。

1.2 数据管理技术的发展

数据管理指的是如何在计算机内把数据管好,具体包括对数据的组织、编目、定位、存储、检索和维护等,它是数据处理任务的矛盾焦点。数据管理技术的优劣,直接影响数据处理的效率。数据管理技术是伴随着计算机硬件和软件的发展而不断发展的,大体上经历了三个阶段:人工管理阶段、文件系统阶段、数据库系统阶段。

回顾数据管理技术的发展过程,将有助于读者更全面地了解数据库的结构。

1.2.1 人工管理阶段

20世纪50年代中期以前,计算机的使用刚刚起步,计算机主要应用于科技计算。硬

件方面,计算机的外存只有磁带、卡片和纸带等,没有磁盘。软件方面,只有汇编语言,没有操作系统(OS, Operating System),没有数据管理的软件。

1) 数据不保存在计算机中

计算机主要用于科技计算,算题时输入数据,计算后输出结果数据。计算完毕数据与程序都“撤出”计算机。

2) 没有专用的软件对数据进行管理

数据管理要人工去做。程序员除了编写程序还要花费许多精力组织数据。举个简单例子,假设计算时要用到一批实验数值数据,程序员就要考虑:如何把这批数据输入计算机,数据怎样转换成二进制串的内部形式,数据在计算机内以什么形式存放(包括用多少个字节来表示一个数,整数和小数、正数和负数怎么表示等),放在计算机内的什么地方,计算时怎样取数,计算完成后结果数据放在哪里,数据输出时又如何把数据转换成人们看得懂的外部形式,等等。用计算机专业的术语来说,程序员不仅要决定数据的逻辑结构,而且还要设计数据的物理结构,包括存储方式、存取方法、输入输出方式等。此外,程序员还要在程序中表达如何对数据实施操作。

在这种人工管理数据的情况下,数据与程序不具有独立性。数据在存储上稍有改变,如增加一些数据,减少一些数据或者改变一下数据存储的位置,都必须修改程序。

3) 以应用为中心,数据面向应用

有一个应用,就要建立一个应用程序,并且要用一组数据与之对应。增加一个新的应用,也就增加一组数据。实际上,各应用程序所使用的数据不会完全没有联系。结果是,程序与程序之间出现大量的重复数据,如图 1.1 所示。

早期,由于没有磁盘,没有操作系统,也没有数据管理软件,谈不上普及计算机的使用。数据管理技术处于极其低下的水平。只有计算机的硬件和软件发展了,数据管理技术才能向前发展。特别是存储器的发展,为数据管理技术的发展提供了良好的物质基础。

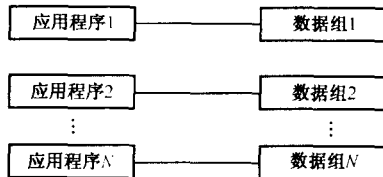


图 1.1 数据面向应用示意图

1.2.2 文件系统阶段

20 世纪 50 年代后期至 20 世纪 60 年代中期,计算机不仅用于科技计算,而且越来越多地用于信息处理。硬件方面,1956 年生产出第一台磁盘,外存有了磁盘、磁鼓等直接存取存储设备。软件方面,出现了高级语言和操作系统。操作系统中的文件系统就是专门管理外存的数据管理软件。

数据组织成文件形式存储在磁盘上。文件中最小的数据单位是**数据项**。若干相关数据

项的集合称为**记录**，一个记录能够表达一个具体的事物。若干记录的集合叫做**文件**。例如，一个文件可以是一个班学生的数据，其中每个记录表达一个学生的情况，而记录中的姓名、性别、年龄、住址等，就是一个个数据项。由于文件是根据数据所代表的意义组织起来的，所以能反映现实世界中的事物。

文件是操作系统管理数据的一个基本单位。文件可以命名，程序员在程序中通过文件名就能够以记录为单位来存取数据、操纵数据，不必关心数据的存储位置以及数据的存放方式。数据的物理结构细节完全由文件系统处理。当数据存储有了改变，维护工作通常由文件系统负责，程序员不一定要修改程序，从而大大节省维护程序的工作量。

磁带是**顺序存取存储设备**，存放数据时只能一个个顺序地存储，读取数据时也只能一个个顺序地读。犹如听卡式录音带一样，假设当前正在第一首歌曲的位置，你不能一下子跳到后面去马上播放第六首歌曲。磁盘是**直接存取存储设备**，它好比 CD 光盘，可以随意地直接播放它上面的每一首歌曲。显然，有了磁盘这种直接存取存储设备来存放文件，文件就可以组织成多种形式，对文件中的记录既可以顺序访问，也可以随机访问。

数据组织成文件，按名存取数据，可顺序访问，也可随机访问。既方便，又实用。显然，用文件方式管理数据是数据管理技术的一大进步，在后续的内容中将会看到，即使是数据库系统也是在文件系统的基础上发展起来的。

1) 数据可以长期保存在外存的磁盘上

对文件可进行检索、修改、插入、删除等操作。

2) 有专门的软件进行数据管理

数据的逻辑结构与数据的物理结构有了区别，举一个简单的例子予以说明，如图 1.2 所示。

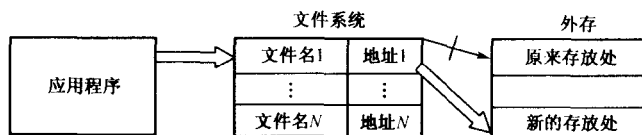


图 1.2 数据逻辑结构与物理结构区别示意图

这里假设应用程序所使用的文件改变了存储的物理位置。如果程序和数据之间没有文件系统这一层，那么，程序必然随着数据存储位置的改变而改变。由文件系统负责数据物理结构和逻辑结构之间的转换并且提供存取方法（读/写），文件系统只需把映射表里的地址改动，应用程序就不用改变了。程序员根据数据的逻辑视图编写程序，按名存取数据，不必考虑数据存储的物理位置。用计算机专业术语说，程序与数据有一定的独立性，但这种独立性只是初级的、不彻底的。程序员观点下的数据逻辑结构还不是彻底地独立于外存的物理结构。在数据物理结构修改时，仍然需要修改应用程序。