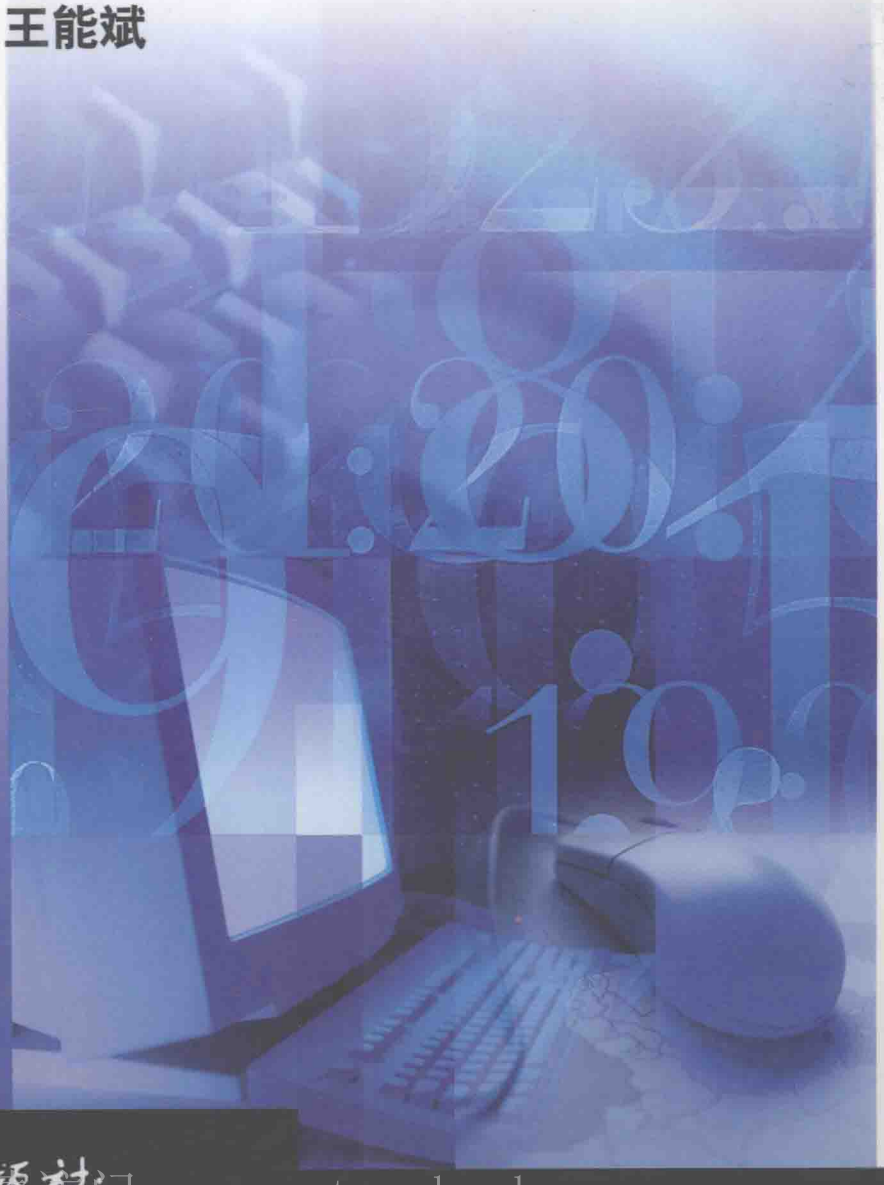
新世纪高等学校计算机系列教材

数据库系统教程

上册

王能斌



電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.

需要全本请访问 www.ertongbook.com

新世纪高等学校计算机系列教材

数据库系统教程

上册

王能斌

電子工業出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

本书是一部全面、系统、深入阐述数据库系统原理的教程和科技参考书。全书共有六篇23章，分上、下两册。第一、二篇分在上册，共有12章；第三至六篇分在下册，共有11章。

第一篇是序篇，介绍数据库系统的基本概念和术语、数据模型。第二篇全面介绍关系数据库系统，包括关系数据库语言、DBMS结构和数据目录、数据库的存储结构、查询处理和优化、事务管理、数据库安全和完整性约束、触发子和主动数据库、数据依赖和关系模式规范化、数据库设计、数据库管理。SQL:1999标准的新增内容在有关章节都有所反映。各章附有习题，书后附有参考文献。

本书取材新颖，内容丰富，文字精炼、可读性好，可以作为计算机专业或其他专业的"数据库概论"、"数据库系统"、"数据库管理系统"的教材。其中的部分内容可选做研究生数据库课程的教材。本书便于自学，也可供广大科技人员学习、参考。

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

数据库系统教程. 上册/王能斌. —北京：电子工业出版社，2002.8
新世纪高等学校计算机系列教材
ISBN 7-5053-7827-9

Ⅰ. 数… Ⅱ. 王… 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(2002)第051732号

责任编辑：张荣琴　　特约编辑：吴明卒
印　　刷：北京天竺颖华印刷厂
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：787×1 092　1/16　印张：17　字数：435千字
版　　次：2002年8月第1版　2002年8月第1次印刷
印　　数：6 000册　　定价：22.00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

《新世纪高等学校计算机系列教材》组织结构

顾问(以汉语拼音为序)

顾冠群(中国工程院院士,东南大学校长)

李　未(中国科学院院士,国务院学位委员会委员,北京航空航天大学教授)

陆汝钤(中国科学院院士,中国科学院软件所研究员)

潘云鹤(中国工程院院士,浙江大学校长)

孙钟秀(中国科学院院士,南京大学教授)

王　珉(教授,博导,江苏省人民政府副省长)

王　荣(研究员,江苏省教育厅厅长)

杨芙清(中国科学院院士,北京大学教授)

周兴铭(中国科学院院士,国防科技大学教授)

编辑委员会

主　任　徐宝文

副主任　史忠植　　陈道蓄　　文宏武

委　员　（以汉语拼音为序）

蔡家楣　陈　纯　陈道蓄　程自强　胡学龙

付育熙　何炎祥　怀进鹏　金远平　鞠时光

刘方鑫　吕　建　梅　宏　钱培德　秦小麟

瞿裕忠　史忠植　宋方敏　王　茜　王怀民

王汝传　王士同　王绍棣　王志坚　文宏武

夏士雄　徐宝文　徐汀荣　徐福培　须文波

张　宏　张　明　张茂坤　张荣琴　张岳新

张在跃　周傲英

需要全本请访问 www.ertongbook.com

总　序

为了认真贯彻《中国教育改革和发展纲要》和教育部“面向21世纪教育振兴行动计划”精神，适应现代社会、经济、科技、文化，特别是教育的发展方向，适应培养新世纪计算机人才的需要，根据计算机科学技术学科发展和新世纪高等学校教学内容和课程体系改革的要求，我们决定编写这套《新世纪高等学校计算机系列教材》。

教学改革和教学质量是高等学校的头等大事。教学改革的关键在于教材的改革，我们组织编写这套系列教材的指导思想是：立足于新世纪培养创造型开拓型人才的培养目标，立足于高速发展的计算机科学技术，科学地预测本学科前沿发展趋势；注重教材内容的思想性、科学性、先进性、系统性和广泛的适应性，继承与改革现有教材内容和体系；注重加强能力的培养，切实提高学生的综合素质。同时，为了适合我国国情，适应绝大多数高等学校计算机科学与技术专业类的教师知识结构和课程改革的实际情况，一方面，我们尽力拓宽教材内容，充实和加强有关理论，引用新的科技成果；另一方面，我们兼顾各门课程的历史、现状与发展，在具体章节的内容编排上力求循序渐进，顺理成章。在保证学科系统的前提下，充分考虑到学生的接受能力，使教材体系尽量符合学生的认知规律，便于学生接受、理解、掌握和巩固所学知识，促使学生学以致用，理论密切联系实际。

本套系列教材的组织编写得到了中国计算机学会、江苏省计算机学会的大力支持和帮助，凝聚了参加编审出版工作的许多教师和编辑的心血和汗水，在此一并对他们表示衷心的感谢。

限于我们的水平、能力和经验，本套系列教材的编审与出版工作一定还存在着很多错漏和其他不如意之处，我们真诚地希望使用这套教材的高校教师、学生和广大读者积极提出宝贵意见，以鞭策我们不断提高教材的编写、修订和出版质量，为我国的计算机科学技术专业教材建设服务。

《新世纪高等学校计算机系列教材》编辑委员会

前　言

本书是一部面向本科生和研究生的数据库系统教程。全书共有六篇23章，分为上、下两册。

第一篇为序篇，其中第1章概括地介绍数据管理的发展，数据库系统的构成以及基本的术语和概念；第2章阐述有代表性的数据模型，分别介绍其概念、结构、约束和操作。

第二篇共分为10章，全面阐述关系数据库系统，包括关系数据库语言、DBMS原理及实现技术、数据库设计以及数据库的运行、维护和管理。关系数据库是当前应用最广、技术最成熟的一种数据库。对于数据库的初学者，先通过关系数据库全面掌握数据库的基本概念、理论、技术和方法，再在此基础上学习关系数据库的各种扩充或非关系数据库，不失为一种循序渐进、符合认识规律的学习方法。如果学时有限，作为数据库的入门课，也可以只学关系数据库。本书把关系数据库的内容全部集中在第二篇中，并列为上册的主要内容，其目的在此。

第三篇论述对象和对象-关系数据库系统。对象数据模型是继关系数据模型之后的最重要的数据模型，也是支持各种非传统数据库应用的基础。本篇的第13章介绍对象数据模型和对象数据库语言；第14章介绍对象DBMS的实现技术。

第四篇论述非传统数据的管理，这是近20多年来数据库的重要成就之一，其应用愈来愈多。本篇按空间数据库、时态数据库和多媒体数据库分别阐述其所需的特殊功能及实现方法。

第五篇论述分布环境下的数据访问与管理。除系统地介绍分布式数据库管理系统外，还介绍并行数据库系统以及用得比较多的各种异构多数据源访问技术。XML已成为描述半结构化数据的标准语言；XML文档管理也将成为数据库的标准功能之一。第21章专门阐述此问题。

第六篇从实现技术角度介绍数据库所支持的一些新应用。数据库对决策的支持是其中用得较广的一种，第22章介绍了数据仓库和OLAP。第23章介绍数据挖掘、工作流和电子商务等三个以数据库为基础的应用，并结合工作流扼要地介绍非传统事务模型及其应用，结合电子商务简单介绍分布环境下的安全问题及解决途径。

历经40年的发展，数据库在理论和技术上都有丰富的积累，并正在蓬勃地发展。国际上新出版的数据库教材大部分都在千页左右。为了使读者能系统地掌握数据库内容，不致在读后有支离破碎之感，作者初次尝试按上述系统组织本教程。尚祈海内外同行和广大读者，不吝赐教，以资改进。

本书着重基本原理、基本概念和基本技术的阐述，力图使读者对数据库系统有全面、深入、系统的领会。各种DBMS产品，除用于举例外，不做专门介绍。具体技术很容易过时，而基本原理、概念和技术可起相对持久的作用。在掌握了基本原理、概念和技术后，读者不难通过阅读手册或培训材料较快地熟悉产品，并以宽阔的视野分析和评价产品。

教材不同于文献综述，教材应主要阐述相对成熟的内容。对于那些尚在探讨中的新内容，本书除在展望中偶尔提及外，不做详细介绍。如果读者因研究需要，须深入了解某一研究方向的发展和动态，可查阅有关文献。

标准是技术发展的反映或导向。虽然其中难免有折衷平衡的成分，但总比个别产品全面、

系统、周到些。本书尽可能向现有标准靠拢，但不介绍标准本身。在数据库的诸多标准中，SQL:1999 也许是最重要的。在本书中，介绍了递归查询、存储过程、按角色授权、触发子、对象-关系数据库、SQL/CLI 等 SQL:1999 新增的内容。

本书可满足各种课程设置的需要。作为入门的数据库概论或 DBMS 概论课程，可选用上册的内容。下册的四篇可作为本科生高年级选修课或研究生课程的教材。这四篇的内容相对独立，可以根据学时及研究方向选学其中的一部分或全部。上册也可用做非计算机专业研究生课程教材。目前学生看参考书比较少，无适合的参考书是原因之一。本书的内容超出一般课堂教学的需要，可以选择其中的一些内容作为学生的参考材料或自学材料。本书便于自学，可供广大科技人员学习、参考。

本书是由徐宝文教授倡议编写的；Michigan 大学朱强教授和 Nova Southeastern 大学计算机及信息研究生院孙俊平教授赠送美国新出版的数据库教材多本，这都是对作者宝贵的支持，作者在此表示诚挚的谢意。本书是在作者过去的科研、教学工作和编写教材实践的基础上写成的，对于曾在这方面提供过帮助的谢希仁教授、徐洁磐教授以及本校教师董逸生、孙志挥、徐宏炳、金远平、陈钢、徐立臻和作者所指导过的历届研究生表示感谢。对于曾支持过作者研究经费的国家自然科学基金委员会、863 高技术计划、原国防科工委、原电子工业部等部门表示感谢。电子工业出版社吴浩源先生多年来给予作者很大的帮助。夫人包桂林女士长期以来给予作者充分的理解、支持和慰勉，使作者能在愉快的氛围中笔耕不辍，终成此书。对此作者深表感谢。

由于编者水平有限，书中难免存在一些缺点和错误，殷切希望广大读者批评指正。对本书的意见请按电子邮件地址 nbwang@seu.edu.cn 反馈给作者，谢谢。

王能斌
于玄武湖畔
2002 年 3 月 26 日

目　　录

上　　册

第一篇　序　　篇

第 1 章　数据库系统引论

1.1　数据管理的发展 …… (2)
1.2　数据库系统 …… (5)
1.3　数据、数据模型和数据模式 …… (7)
　1.3.1　数据 …… (7)
　1.3.2　数据模型 …… (7)
　1.3.3　数据模式 …… (8)
1.4　数据库的生存周期 …… (9)
习题 …… (10)

第 2 章　数 据 模 型

2.1　层次数据模型 …… (12)
　2.1.1　基本概念和结构 …… (12)
　2.1.2　约束 …… (16)
　2.1.3　操作 …… (16)
2.2　网状数据模型 …… (17)
　2.2.1　基本概念和结构 …… (17)
　2.2.2　约束 …… (20)
　2.2.3　操作 …… (21)
2.3　关系数据模型 …… (22)
　2.3.1　基本概念和定义 …… (22)
　2.3.2　约束 …… (24)
　2.3.3　操作 …… (26)
　2.3.4　关系演算 …… (30)
2.4　对传统数据模型的评价 …… (31)
2.5　E-R 数据模型 …… (33)
　2.5.1　基本概念 …… (33)
　2.5.2　E-R 图 …… (36)
　2.5.3　扩充 E-R 数据模型 …… (37)
2.6　面向对象数据模型 …… (40)
　2.6.1　对象 …… (40)
　2.6.2　类和实例 …… (41)

2.6.3 类层次结构和继承 …… (42)
2.6.4 对象的标识 …… (44)
2.7 基于逻辑的数据模型 …… (45)
2.7.1 一阶谓词逻辑作为数据模型 …… (45)
2.7.2 有关基于逻辑的数据模型的语法问题 …… (46)
2.7.3 有关基于逻辑的数据模型的语义问题 …… (49)
习题 …… (54)

第二篇 关系数据库系统

第3章 关系数据库语言

3.1 数据库的用户接口 …… (58)
3.2 SQL语言概况 …… (59)
3.3 SQL数据定义语言 …… (60)
3.3.1 术语和数据类型 …… (60)
3.3.2 基表模式的定义 …… (60)
3.3.3 基表模式的修改 …… (62)
3.3.4 索引的建立和撤销 …… (63)
3.4 SQL查询语言 …… (64)
3.4.1 基本SQL查询语言 …… (64)
3.4.2 查询条件比较复杂的SQL查询语句 …… (66)
3.4.3 GROUP BY 和 ORDER BY 子句的应用 …… (70)
3.3.4 包含UNION的查询 …… (71)
3.5 SQL数据操纵语言 …… (71)
3.5.1 INSERT语句 …… (71)
3.5.2 DELETE语句 …… (72)
3.5.3 UPDATE语句 …… (73)
3.6 SQL中的视图 …… (73)
3.6.1 普通视图 …… (73)
3.6.2 临时视图和递归查询 …… (75)
3.7 嵌入式SQL …… (77)
3.7.1 嵌入式SQL的介绍 …… (77)
3.7.2 嵌入式SQL的说明部分 …… (77)
3.7.3 嵌入式SQL的可执行语句 …… (78)
3.7.4 嵌入式SQL的处理过程 …… (80)
3.8 动态SQL …… (81)
3.8.1 直接执行的动态SQL …… (81)
3.8.2 带动态参数的动态SQL …… (82)
3.8.3 查询类动态SQL …… (82)
3.9 SQL的存储过程 …… (83)
3.10 QBE数据库语言 …… (84)
3.10.1 QBE的查询操作 …… (85)

3.10.2 QBE 的更新操作 …… (87)
3.10.3 QBE 的其他功能 …… (87)
习题 …… (88)

第 4 章 数据库管理系统引论

4.1 数据库管理系统结构简介 …… (90)
4.2 事务 …… (91)
4.3 DBMS 的进程结构 …… (92)
4.4 DBMS 的系统结构 …… (95)
4.5 数据目录 …… (97)
习题 …… (99)

第 5 章 数据库的存储结构

5.1 数据库存储介质的特点 …… (100)
5.2 记录的存储结构 …… (101)
5.2.1 记录的物理表示 …… (101)
5.2.2 记录在物理块上的分配 …… (102)
5.2.3 物理块在磁盘上的分配 …… (103)
5.2.4 数据压缩技术 …… (104)
5.3 文件结构和存取路径 …… (105)
5.3.1 访问文件的方式 …… (105)
5.3.2 数据库对文件的要求 …… (106)
5.3.3 文件的基本类型 …… (106)
5.4 动态索引 …… (111)
5.5 动态散列 …… (115)
5.6 多键查询 …… (117)
5.7 存储系统的发展 …… (118)
习题 …… (121)

第 6 章 查询处理和优化

6.1 引言 …… (122)
6.2 代数优化 …… (123)
6.3 依赖于存取路径的规则优化 …… (127)
6.3.1 选择操作的实现和优化 …… (127)
6.3.2 连接操作的实现和优化 …… (129)
6.3.3 投影操作的实现 …… (132)
6.3.4 集合操作的实现 …… (132)
6.3.5 组合操作 …… (134)
6.4 代价估算优化 …… (134)
6.4.1 查询执行代价的组成和代价模型 …… (134)
6.4.2 选择操作的代价估算 …… (135)

需要全本请访问 www.ertongbook.com

6.4.3 连接操作的代价估算 …… (138)
6.5 结束语 …… (140)
习题 …… (140)

第7章 事务管理

7.1 恢复引论 …… (142)
7.2 运行记录的结构 …… (144)
7.3 更新事务的执行与恢复 …… (146)
7.4 易地更新恢复技术 …… (148)
7.5 消息的处理 …… (149)
7.6 失效的类型及恢复的对策 …… (150)
7.7 并发控制引论 …… (151)
7.7.1 数据库系统中的并发 …… (151)
7.7.2 并发的目的 …… (152)
7.7.3 并发所引起的问题 …… (152)
7.7.4 并发控制的正确性准则 …… (153)
7.8 加锁协议 …… (156)
7.8.1 X锁 …… (156)
7.8.2 两段封锁协议 …… (157)
7.8.3 (S,X)锁 …… (158)
7.8.4 (S,U,X)锁 …… (159)
7.9 死锁的检测、处理和防止 …… (159)
7.9.1 死锁的检测和处理 …… (159)
7.9.2 死锁的防止 …… (161)
7.10 多粒度封锁 …… (162)
7.11 索引的并发控制 …… (164)
7.12 幽灵及其防止 …… (165)
7.13 事务的隔离等级 …… (167)
7.14 基于时间标记的并发控制技术 …… (168)
7.14.1 基本的时间标记协议 …… (168)
7.14.2 多版本并发控制技术 …… (170)
7.15 乐观并发控制技术 …… (170)
习题 …… (172)

第8章 数据库的安全和完整性约束

8.1 数据库的安全 …… (174)
8.1.1 视图定义和查询修改 …… (174)
8.1.2 访问控制 …… (175)
8.1.3 数据加密 …… (179)
8.1.4 跟踪审查 …… (179)
8.2 统计数据库的安全 …… (179)

8.3 完整性约束 …… (183)
8.3.1 完整性约束的类型 …… (183)
8.3.2 完整性约束的说明 …… (184)
8.3.3 完整性约束的实施 …… (185)
习题 …… (185)

第 9 章 触发子和主动数据库系统

9.1 主动数据库系统引论 …… (187)
9.2 规则的表示 …… (188)
9.3 规则的执行 …… (190)
9.4 规则的实现 …… (192)
9.5 规则的应用 …… (194)
习题 …… (200)

第 10 章 数据依赖和关系模式的规范化

10.1 关系模式设计中的一些数据语义问题 …… (201)
10.2 函数依赖 …… (203)
10.3 多值依赖 …… (207)
10.4 连接依赖 …… (210)
10.5 关系模式的分解及其问题 …… (211)
10.6 关系模式的规范化 …… (218)
习题 …… (225)

第 11 章 数据库设计

11.1 数据库设计引论 …… (227)
11.2 数据库的概念设计 …… (229)
11.2.1 数据库概念设计的基本方法 …… (229)
11.2.2 视图设计 …… (230)
11.2.3 视图集成 …… (231)
11.3 数据库的逻辑设计 …… (233)
11.3.1 E-R 图到关系模式的转换 …… (233)
11.3.2 逻辑模式的规范化、调整和实现 …… (239)
11.3.3 外模式的设计 …… (240)
11.4 数据库的物理设计 …… (241)
11.4.1 簇集设计 …… (241)
11.4.2 索引的选择 …… (244)
11.4.3 分区设计 …… (246)
11.5 计算机辅助数据库设计工具 …… (246)
习题 …… (247)

第 12 章 数据库管理

12.1 数据库的建立 …… (249)

12.2 数据库的调整 …… (250)
12.3 数据库的重组 …… (251)
12.4 数据库的重构 …… (253)
12.5 数据库的安全控制 …… (253)
12.6 数据完整性控制 …… (254)
12.7 DBA 的作用和职责 …… (254)
习题 …… (255)

参考文献 …… (256)

第一篇　序　　篇

第 1 章　数据库系统引论

第 2 章　数据模型

读书是个认识过程。见林不见树，只知其大概，无从深入；见树不见林，虽然学了不少片断的知识，难窥其全貌。只有全局在胸，既能深入研讨具体问题，又不致在知识海洋中迷航，这才是系统、深入掌握一门知识的正确之道。

第 1 章对数据管理做了系统、概括的介绍，并且阐明一些基本术语、概念和观点，可为读者提供本课程的一个鸟瞰，这对初学者特别重要。

数据模型是数据的抽象。没有这种抽象，就没有数据管理这门学科。它是研究数据管理的出发点和基础。数据模型随着应用的需求和技术的发展而不断地演化。从这个意义来说，数据模型的发展轨迹反映数据管理的历史。第 2 章对有代表性的数据模型做了较全面的介绍，这是学习其他各章的预备和基础。

第 1 章　数据库系统引论

1.1　数据管理的发展

在众多的计算机应用中，有一类重要的计算机应用，称为数据密集型应用(data intensive applications)。它有下列 3 个特点：

(1) 涉及的数据量大，一般需存放在辅助存储器中，内存中只能暂存其中很小的一部分；

(2) 数据不随程序的结束而消失，而需长期保留在计算机系统中，这种数据称为持久数据(persistent data)；

(3) 数据为多个应用程序所共享，甚至在一个单位或更大范围内共享。

这是最大的计算机应用领域，管理信息系统、办公信息系统、银行信息系统、民航订票系统、情报检索系统等都属于这一类。管理这种大量的、持久的、共享的数据是这类计算机应用面临的共同问题。从 20 世纪 50 年代末以来，数据管理一直是计算机科学技术领域中的一门重要的技术和研究课题。

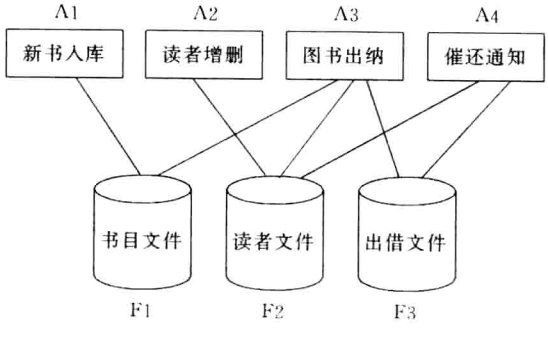


图 1-1　图书馆文件系统

早期的数据管理都采用文件系统(file system)。在文件系统中，数据按其内容、结构和用途组成若干命名的文件。文件一般为某个用户或用户组所有，但可供指定的其他用户共享。用户可以通过操作系统对文件进行打开、读、写和关闭等操作。图 1-1 是一个图书馆文件系统的例子，其中共有 F1，F2 和 F3 三个文件，A1，A2，A3 和 A4 是建立在这三个文件上的四个应用程序，图中的实线表示每个应用程序所访问的文件。例如，图书出纳应用程序若要处理一个借书申请，须访问 F1，查询有无此书；访问 F2，查询借阅者是否为本馆的读者；访问 F3，查询此书是否借出。这样的文件系统明显地有下列 5 个缺点。

(1) 编写应用程序很不方便。应用程序的设计者必须对所用的文件的逻辑及物理结构有清楚的了解。操作系统只提供打开、关闭、读、写等几个低级的文件操作命令，对文件的查询、修改等处理都须在应用程序内解决。应用程序还不可避免地在功能上有所重复，例如 A2，A3，A4 三个应用程序都要有查询 F2 的功能。因此，在文件系统上编写应用程序的生产率不高；如果有新的要求，例如要求提供某一专题书目，则必须请熟悉文件的程序员编写相应的应用程序，一般需要较长的时间，而不能由用户直接向系统查询。

(2) 文件的设计很难满足多种应用程序的不同要求，数据冗余往往是不可避免的。例如，A2 要求 F2 包含读者的所有属性，而 A3 只要从 F2 查询借阅者是否为合法读者，A4 只要从

F2 查询读者的地址；对文件的划分也有不同的要求，例如，A4 希望把读者的地址列入 F3 中，可以在处理时少打开一个文件，但这又与 F2 的内容重复，成为数据冗余(redundancy)，而且 A3 又无此要求。为了兼顾各种应用程序的要求，在设计文件系统时，往往不得不增加数据冗余，有时甚至会出现大量的数据冗余。数据冗余不仅浪费存储空间，而且会带来数据的不一致(inconsistency)。例如，在 F2 和 F3 中都列入读者地址，如果读者搬家了，只修改 F2，不修改 F3，则会造成数据的不一致。在文件系统中，无维护数据一致性的监控机制，数据的一致性完全由用户负责维护。这在简单系统中还可勉强应付，在复杂的系统中，保证数据的一致性几乎是不可能的。

(3) 文件结构的每个修改将导致应用程序的修改，应用程序的维护工作量很大。随着应用的环境和需求的变化，修改文件的结构是常有的事，例如在某些文件的记录中增加一些字段，扩充某些字段的长度，改变某些字段的计量单位或表示格式，把书目文件按学科分为几个文件等，都会引起应用程序的一连串的修改。众所周知，修改程序是很麻烦的事，首先要熟悉原有的程序，修改后还要对程序进行测试，以免因修改而引起不希望的副作用。这些问题之所以产生，主要是由于应用程序对文件的过分依赖；或者说，文件系统的数据独立性(data independence)不好。

(4) 文件系统一般不支持对文件的并发访问(concurrent access)。在现代计算机系统中，为了有效地利用计算机的资源，一般允许多个应用程序并发地运行。文件当初就是作为某个程序的附属数据出现的。文件系统一般不支持多个应用程序对同一文件的并发访问，例如在图 1-1 中，如果在 A2 修改 F2 的同时，A3 或 A4 并发地查询 F2，则 A3 或 A4 读到的数据可能是正在修改中的数据，有可能不一致，甚至错误。为了避免这个问题的发生，可限制对文件的并发访问或者为每个应用程序准备一个文件的复本。前一种办法降低了程序的并发程度；后一种办法增加了数据的冗余，从而会引起数据的不一致。

(5) 由于数据缺少统一管理，在数据的结构、编码、表示格式、命名以及输出格式等方面不容易做到规范化、标准化；在数据的安全和保密方面，也难以采取有效的措施。

针对文件系统的上述缺点，人们逐步发展了以统一管理和共享数据为主要特征的数据库系统(database system)。在数据库系统中，数据不再仅仅服务于某个程序或用户，而是看成一个单位的共享资源，由一个叫做数据库管理系统(DataBase Management System，简称 DBMS)的软件统一管理。由于有 DBMS 的统一管理，应用程序不必直接介入诸如打开、关闭、读、写文件等低级操作，而由 DBMS 代办。用户也不必关心数据存储和其他实现的细节，可在更高的抽象级别上观察和访问数据。文件结构的一些修改也可以由 DBMS 屏蔽，使用户看不到这些修改，从而减少应用程序的维护工作量，提高数据的独立性。由于数据的统一管理，人们可以从全单位着眼，合理组织数据，减少数据冗余；还可以更好地贯彻规范化和标准化，从而有利于数据的转移和更大范围内的共享。由于 DBMS 不是为某个应用程序服务，而是为整个单位服务的，DBMS 做得复杂一些也是可以接受的。许多在文件系统中难以实现的功能，在 DBMS 中都一一实现了。例如，适合不同类型用户的多种用户界面，保证并发访问时的数据一致性的并发控制(concurrent control)，增进数据安全性(security)的访问控制(access control)，在故障情况下保证数据一致性的恢复(recovery)功能，保证数据在语义上的一致性的完整性约束(integrity constraints)检查功能等。随着计算机应用的发展，DBMS 的功能愈来愈强，规模愈来愈大，复杂性和开销也随之增加。目前，在一些功能非常明确且无数据共享的简单应用系统中，为了减少开销、提高性能，有时仍采用文件系统；不过在本节开头所列举的数据密集型

需要全本请访问 www.ertongbook.com