

计算机与信息技术专业应用教材

数据库原理与应用

——基于SQL Server 2000

李春葆 曾平 编著



清华大学出版社

► 计算机与信息技术专业应用教材

数据库原理与应用

——基于 SQL Server 2000

李春葆 曾 平 编著

清华大学出版社

北京

内 容 简 介

数据库技术是目前 IT 行业中发展最快的领域之一,已经被广泛应用于各种类型的数据处理系统之中。了解并掌握数据库知识已经成为各类科技人员和管理人员的基本要求。

本书基于 SQL Server 2000 系统讨论数据库的原理和应用方法。全书分为 3 部分:第 1 章~第 5 章介绍数据库的一般原理;第 6 章~第 18 章介绍 SQL Server 2000 数据库管理方法;第 19 章~第 21 章介绍 VB 6.0 环境下开发 SQL Server 2000 数据库应用系统的相关技术。每章都配有练习题,书中还安排了 16 个上机实验题供选做。

本书内容由浅入深,循序渐进,通俗易懂,适合自学,可作为各类院校相关专业及其他培训班的“数据库原理与应用”或“SQL Server 2000 应用系统开发”课程的教学用书,对于计算机应用人员和计算机爱好者本书也是一本实用的自学参考书。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

本书防伪标签采用特殊防伪技术,用户可通过在图案表面涂抹清水,图案消失,水干后图案复现;或将表面膜揭下,放在白纸上用彩笔涂抹,图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

数据库原理与应用:基于 SQL Server 2000/李春葆,曾平编著.

—北京:清华大学出版社,2005.10

计算机与信息技术专业应用教材

ISBN 7-302-12063-3

I. 数... II. ①李... ②曾... III. 关系数据库—数据库管理系统,

SQL Server 2000 IV. TP311.138

中国版本图书馆 CIP 数据核字(2005)第 127515 号

出版者:清华大学出版社

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

社总机:010-62770175

地 址:北京清华大学学研大厦

邮 编:100084

客户服务:010-82896445

组稿编辑:夏非彼

文稿编辑:洪英 潘秀燕 王红华

封面设计:付剑飞

版式设计:科海

印刷者:北京市耀华印刷有限公司

发 行 者:新华书店总店北京发行所

开 本:787×1092 1/16 印张:30.25 字数:736 千字

版 次:2006 年 1 月第 1 版 2006 年 4 月第 2 次印刷

书 号:ISBN 7-302-12063-3/TP·7809

印 数:4001~7000

定 价:39.80 元

本书如存在文字不清、漏印以及缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010) 82896445

丛书序

为适应信息社会高速发展的需求，目前全国各类高等院校都在进行计算机教学的全方位改革，目的是规划出一整套面向计算机与信息技术专业、具有中国高校计算机教育特色的课程计划和教材体系，本丛书就是在这一背景下应运而生的。我们组织了由全国高校计算机专业的专家教授组成的“计算机与信息技术专业应用教材”课题研究组，通过对计算机和信息技术专业全方位的研讨，并结合我国当前的实际情况，编写了这套系统性、科学性和实践性都很强的丛书。

丛书特色

☑ 先进性：力求介绍最新的技术和方法

先进性和时代性是教材的生命，计算机与信息技术专业的教学具有更新快、内容多的特点，本丛书在体例安排和实际讲述过程中都力求介绍最新的技术和方法，并注重拓宽学生的知识面，激发他们学习的热情和创新的欲望。

☑ 理论与实践并重：阐明基础理论，强调实践应用

理论是实践的基础，实践是理论的升华；不能有效指导实践的理论是空头理论，没有理论指导的实践是盲目的实践。对于时代呼唤的信息化人才而言，二者缺一不可。本丛书以知识点为主线，穿插演示性案例于理论讲解之中，使枯燥的理论变得更易于理解、易于接受；此外，还在每一章的末尾提供大量的实习题和综合练习题，目的是提高学生综合利用所学知识解决实际问题的能力。

☑ 易教易学：创新体例，合理布局，通俗易懂

本丛书结构清晰，内容系统详实，布局合理，体例较好；力求把握各门课程的核心，通俗易懂，便于教学的展开，也便于学生学习。

丛书组成

首批推出的计算机与信息技术专业应用教材涵盖计算机基础、程序设计和数据库三大领域，共 16 本：

- 操作系统教程
- 计算机系统结构教程
- 数据结构与算法教程
- Java 语言程序设计
- Access 数据库程序设计

- C 程序设计教程（基于 Visual C++ 平台）
- C 程序设计教程学习与上机指导（基于 Visual C++ 平台）
- C++ 程序设计
- C++ 程序设计学习与上机实验指导
- Visual FoxPro 程序设计
- Visual Basic 程序设计
- SQL Server 2000 应用系统开发教程
- SQL Server 2000 学习与上机实验指导
- 数据库原理与应用——基于 Access
- 数据库原理与应用——基于 Visual FoxPro
- 数据库原理与应用——基于 SQL Server 2000

服务之窗

本丛书的出版者和作者竭诚为读者提供服务。

本丛书的网络支持与服务网址为 <http://www.khp.com.cn/xxjsjc/>。在这里提供了实用的相关资源与最新信息，读者可以方便地下载本丛书的实例源代码，便捷地参与讨论，并可同作者进行交流，得到作者和专家的帮助。我们将努力有效、快捷地解决读者在图书使用和学习中遇到的疑难问题，并致力于提供更多的增值服务。

丛书编委会

主任委员：	李春葆
副主任委员：	苏光奎 朱福喜
委员：	尹为民 尹朝庆 李春葆 伍春香 朱福喜
	苏光奎 胡新启 徐爱萍 曾平 曾慧

编者寄语

如果说科学技术的飞速发展是 21 世纪的一个重要特征的话，那么教学改革将是 21 世纪教育工作不变的主题。要紧跟教学改革，不断创新，真正编写出满足新形势下教学需求的教材，还需要我们不断地努力实践、探索和完善。本丛书虽然经过细致的编写与校订，仍难免有疏漏和不足，需要不断地补充、修订和完善。我们热情欢迎使用本丛书的教师、学生和读者朋友提出宝贵意见和建议，使之更臻成熟。

本丛书作者的电子邮件：licb@public.wh.hb.cn

本丛书出版者的电子邮件：feedback@khp.com.cn

前 言

数据库技术是目前 IT 行业中发展最快的领域之一，已经广泛应用于各种类型的数据处理系统之中。了解并掌握数据库知识已经成为各类科技人员和管理人员的基本要求。“数据库原理与应用”课程已逐渐成为普通高校各个专业本、专科学生的必修课程。

本书基于 SQL Server 2000 讨论数据库的原理和应用方法。全书分为 3 部分，第 1 章～第 5 章介绍数据库的一般原理；第 6 章～第 18 章介绍 SQL Server 2000 数据库管理方法；第 19 章～第 21 章介绍 VB 6.0 环境下开发 SQL Server 2000 数据库应用系统的相关技术。

第 1 章为数据库系统概述，第 2 章为数据模型，第 3 章为关系数据库，第 4 章为关系数据库规范化理论，第 5 章为数据库设计，第 6 章为 SQL Server 2000 系统概述，第 7 章为创建和使用数据库，第 8 章为创建和使用表，第 9 章为 Transact-SQL，第 10 章为 SQL 高级使用，第 11 章为索引，第 12 章为视图，第 13 章为数据库完整性，第 14 章为存储过程，第 15 章为触发器，第 16 章为 SQL Server 的安全管理，第 17 章为数据备份与还原，第 18 章为数据转换，第 19 章为 VB 与数据库，第 20 章为使用 DAO 操作数据库，第 21 章为使用 ADO 操作数据库。每一章后都给出相应的练习题，书中还安排了 16 个上机实验题供读者选做。

本书内容由浅入深，循序渐进，通俗易懂，适合自学。书中提供了大量例题，有助于读者理解概念、巩固知识、掌握要点、攻克难点。

本书可以作为各类院校相关专业及其他培训班的“数据库原理与应用”或“SQL Server 2000 应用系统开发”课程的教学用书，对于计算机应用人员和计算机爱好者本书也是一本实用的自学参考书。

由于时间仓促，编者水平有限，书中难免出现错误和疏漏之处，敬请广大读者指正。

编 者
2005 年 11 月

目 录

第 1 章 数据库系统概述	1	第 3 章 关系数据库	21
1.1 信息、数据和数据处理	1	3.1 关系模型的基本概念	21
1.1.1 信息与数据	1	3.2 关系的数学定义	22
1.1.2 数据处理	2	3.3 关系代数	24
1.2 数据管理技术的发展	2	3.3.1 传统的集合运算	24
1.2.1 人工管理阶段 (20 世纪 50 年代)	2	3.3.2 专门的关系运算	24
.....	2	练习题 3	26
1.2.2 文件系统阶段 (20 世纪 60 年代)	2	第 4 章 关系数据库规范化理论	28
.....	2	4.1 问题的提出	28
1.2.3 数据库系统阶段 (20 世纪 60 年	3	4.2 函数依赖	29
代后期)	3	4.2.1 函数依赖的定义	29
1.3 数据库系统的组成与结构	4	4.2.2 函数依赖与属性关系	30
1.3.1 数据库系统的组成	4	4.2.3 Armstrong 公理	31
1.3.2 数据库系统体系结构	5	4.2.4 闭包及其计算	32
1.4 数据库管理系统	7	4.3 范式和规范化	33
1.4.1 DBMS 的主要功能	7	4.3.1 什么叫范式	33
1.4.2 DBMS 的组成	8	4.3.2 范式的判定条件与规范化	33
练习题 1	9	4.4 关系模式的分解	35
第 2 章 数据模型	10	4.4.1 模式分解中存在的问题	35
2.1 什么是数据模型	10	4.4.2 无损分解的定义和性质	35
2.1.1 数据的描述	11	4.4.3 无损分解的测试方法	36
2.1.2 数据间联系的描述	11	4.4.4 保持函数依赖的分解	36
2.2 概念模型	11	练习题 4	37
2.2.1 信息世界中的基本概念	11	第 5 章 数据库设计	39
2.2.2 实体间的联系方式	12	5.1 数据库设计概述	39
2.2.3 实体联系表示法 (E-R 方法)	13	5.2 需求分析	40
2.2.4 怎样设计 E-R 图	15	5.2.1 需求分析的步骤	40
2.3 数据库类型	16	5.2.2 需求分析的方法	42
2.3.1 层次模型	17	5.3 概念结构设计	45
2.3.2 网状模型	18	5.3.1 局部应用 E-R 模型设计	45
2.3.3 关系模型	19	5.3.2 总体概念 E-R 模型设计	47
练习题 2	20		

5.4 逻辑结构设计	49	7.1.5 查看存储过程	88
5.5 物理结构设计	51	7.1.6 查看用户和角色	89
5.6 数据库的实施和维护	51	7.1.7 数据库的其他组成部分	91
练习题 5	52	7.2 SQL Server 数据库和文件	91
第 6 章 SQL Server 2000 系统概述	54	7.2.1 文件和文件组	91
6.1 SQL Server 2000 系统简介	54	7.2.2 数据库存储结构	93
6.1.1 概况	54	7.2.3 事务日志	94
6.1.2 特性	55	7.3 创建数据库	95
6.1.3 新增或增强功能	56	7.3.1 直接建立数据库	95
6.2 系统需求	58	7.3.2 使用向导建立数据库	98
6.2.1 硬件需求	58	7.3.3 数据库的属性设置	99
6.2.2 软件需求	59	7.3.4 数据库大小估算和收缩数据库	100
6.3 SQL Server 2000 的安装	60	7.4 数据库更名	102
6.3.1 SQL Server 2000 的配置选项	60	7.5 删除数据库	103
6.3.2 安装 SQL Server 2000	66	练习题 7	104
6.3.3 其他安装方式	69	上机实验题 2	104
6.3.4 从其他版本升级到 SQL Server 2000	70	第 8 章 创建和使用表	105
6.4 客户/服务器体系结构	71	8.1 建立表	105
6.5 SQL Server 服务器的管理	73	8.2 修改表的结构	108
6.6 SQL Server 2000 的主要组件	75	8.3 建立表间的关联	108
6.6.1 联机丛书	76	8.4 删除表	111
6.6.2 企业管理器	76	8.5 记录的新增和修改	112
6.6.3 服务器网络实用工具和客户网络实用工具	77	练习题 8	113
6.6.4 查询分析器	78	上机实验题 3	114
6.6.5 导入和导出数据	78	第 9 章 Transact-SQL	116
6.6.6 服务管理器	79	9.1 SQL 语言	116
6.6.7 事件探查器	79	9.1.1 概述	116
练习题 6	79	9.1.2 SQL 语言的分类	117
上机实验题 1	79	9.2 Transact-SQL 基础	118
第 7 章 创建和使用数据库	80	9.2.1 查询分析器	118
7.1 查看数据库	80	9.2.2 数据库的操作语句	120
7.1.1 查看数据库	80	9.2.3 表的操作语句	124
7.1.2 查看关系图	83	9.2.4 数据查询	127
7.1.3 查看表的结构和内容	85	9.3 Transact-SQL 程序设计基础	141
7.1.4 查看视图	86	9.3.1 标识符	142
		9.3.2 数据类型	144

9.3.3 运算符.....	158	11.2 索引类型.....	229
9.3.4 变量.....	163	11.2.1 B-树索引结构.....	229
9.3.5 批处理.....	165	11.2.2 聚集索引和非聚集索引.....	229
9.3.6 注释.....	166	11.2.3 惟一索引和组合索引.....	232
9.3.7 控制流语句.....	167	11.3 创建索引.....	233
9.3.8 函数.....	174	11.3.1 通过企业管理器创建索引.....	233
练习题 9.....	177	11.3.2 通过向导创建索引.....	236
上机实验题 4.....	177	11.3.3 使用 SQL 语言创建索引.....	238
第 10 章 SQL 高级应用.....	179	11.3.4 创建索引的选项设置.....	240
10.1 SELECT 高级查询.....	179	11.3.5 创建索引的空间考虑.....	242
10.1.1 数据汇总.....	179	11.3.6 在视图和计算列上创建索引 ...	242
10.1.2 联接查询.....	184	11.4 查看和删除索引.....	242
10.1.3 子查询.....	188	11.4.1 使用企业管理器.....	243
10.1.4 在查询的基础上创建新表.....	193	11.4.2 使用 SQL 语言.....	243
10.2 管理 ntext、text 和 image 数据.....	194	练习题 11.....	244
10.2.1 检索 ntext、text 或 image 值 ...	195	上机实验题 6.....	244
10.2.2 修改 ntext、text 或 image 值 ...	197	第 12 章 视图.....	245
10.3 事务处理.....	199	12.1 视图概述.....	245
10.3.1 事务分类.....	200	12.2 创建视图.....	246
10.3.2 显式事务.....	200	12.2.1 使用企业管理器创建视图.....	247
10.3.3 自动提交事务.....	204	12.2.2 使用向导创建视图.....	249
10.3.4 隐式事务.....	205	12.2.3 使用 SQL 语句创建视图.....	251
10.4 数据的锁定.....	206	12.3 使用视图.....	252
10.4.1 并发问题.....	206	12.3.1 使用视图进行数据检索.....	252
10.4.2 事务的隔离级别.....	208	12.3.2 通过视图修改数据.....	253
10.4.3 SQL Server 中的锁定.....	209	12.4 视图的修改.....	255
10.4.4 自定义锁.....	212	12.4.1 修改视图.....	255
10.5 使用游标.....	217	12.4.2 重命名视图.....	256
10.5.1 游标的概念.....	218	12.5 视图信息的查询.....	258
10.5.2 使用游标.....	218	12.5.1 使用企业管理器查询视图信息.....	258
10.5.3 游标类型.....	223	12.5.2 使用 sp_helptext 存储过程查询 视图信息.....	258
练习题 10.....	224	12.6 视图的删除.....	259
上机实验题 5.....	224	12.6.1 使用企业管理器删除视图.....	260
第 11 章 索引.....	226	12.6.2 使用 Transact-SQL 删除视图... ..	261
11.1 索引简介.....	226	练习题 12.....	261
11.1.1 创建索引的原因.....	226	上机实验题 7.....	261
11.1.2 建立索引应该考虑的问题.....	227		

第 13 章 数据库完整性	262	16.1.1 操作系统的安全性.....	300
13.1 约束.....	262	16.1.2 SQL Server 的安全性.....	301
13.1.1 PRIMARY KEY 约束.....	262	16.1.3 数据库的安全性.....	301
13.1.2 FOREIGN KEY 约束.....	263	16.1.4 SQL Server 数据库对象的 安全性.....	302
13.1.3 UNIQUE 约束.....	264	16.2 SQL Server 的验证模式.....	302
13.1.4 CHECK 约束.....	265	16.2.1 NT 验证模式.....	302
13.1.5 列约束和表约束.....	265	16.2.2 混合验证模式.....	303
13.2 默认值.....	266	16.2.3 NT 验证模式对 SQL Server 的 影响.....	303
13.2.1 在创建表时指定默认值.....	267	16.2.4 设置验证模式.....	304
13.2.2 使用默认对象.....	268	16.3 账号和角色.....	305
13.3 规则.....	272	16.3.1 服务器的登录账号.....	305
13.3.1 创建规则.....	273	16.3.2 数据库的用户.....	308
13.3.2 绑定规则.....	274	16.3.3 数据库角色.....	312
13.3.3 解除和删除规则.....	275	16.3.4 用户和角色的权限问题.....	316
练习题 13.....	275	练习题 16.....	317
上机实验题 8.....	276	上机实验题 11.....	317
第 14 章 存储过程	277	第 17 章 数据备份与还原	318
14.1 概述.....	277	17.1 备份和还原构架.....	318
14.2 创建存储过程.....	277	17.2 备份数据库.....	321
14.3 执行存储过程.....	282	17.2.1 备份类型和常用备份策略.....	321
14.4 存储过程的参数.....	283	17.2.2 备份数据库中各种对象.....	322
14.5 存储过程的查看、修改和删除.....	286	17.3 数据库还原.....	334
练习题 14.....	288	17.3.1 数据库恢复模型.....	334
上机实验题 9.....	288	17.3.2 还原数据库备份.....	337
第 15 章 触发器	289	17.3.3 还原事务日志备份.....	340
15.1 概述.....	289	练习题 17.....	341
15.2 创建触发器.....	289	上机实验题 12.....	341
15.3 inserted 表和 deleted 表.....	294	第 18 章 数据转换	342
15.4 使用触发器.....	295	18.1 DTS 的基本概念.....	342
15.5 修改触发器.....	297	18.2 数据的导入.....	343
15.6 删除触发器.....	298	18.3 数据的导出.....	346
15.7 嵌套触发器.....	298	练习题 18.....	350
练习题 15.....	299	上机实验题 13.....	350
上机实验题 10.....	299		
第 16 章 SQL Server 的安全管理	300		
16.1 SQL Server 安全体系结构.....	300		

第 19 章 VB 与数据库	351
19.1 概述	351
19.1.1 数据访问对象模型	351
19.1.2 数据库的分类	351
19.1.3 ODBC 和数据源	353
19.2 常用的数据访问控件	355
19.2.1 Data 控件	355
19.2.2 ADO Data 控件	360
19.2.3 DataCombo 控件	365
19.2.4 DataGrid 控件	368
19.3 常用的高级用户界面控件	371
19.3.1 图像列表控件	371
19.3.2 TreeView 控件	374
19.3.3 ListView 控件	380
19.3.4 公共对话框控件	387
练习题 19	392
上机实验题 14	392
第 20 章 使用 DAO 操作数据库	393
20.1 什么是 DAO	393
20.1.1 DAO 操作数据库的步骤	394
20.1.2 DBEngine 初始化数据库环境	396
20.1.3 打开数据库	396
20.1.4 打开数据集	396
20.2 DBEngine 对象	396
20.2.1 DBEngine 对象的属性	396
20.2.2 DBEngine 对象的方法	397
20.3 Workspace 对象	398
20.3.1 Workspace 对象的属性	399
20.3.2 Workspace 对象的方法	399
20.4 Database 对象	401
20.4.1 Database 对象的属性	401
20.4.2 Database 对象的方法	402
20.4.3 TableDef 对象	403
20.4.4 Recordset 对象	405
20.4.5 QueryDef 对象	409
20.5 Fields 集合和 Field 对象	410
20.5.1 Field 对象的属性	410
20.5.2 Field 对象的方法	411
20.5.3 Fields 集合的属性	411
20.5.4 Fields 集合的方法	411
20.6 Index 对象	412
20.6.1 Index 对象的属性	412
20.6.2 Index 对象的方法	412
20.7 DAO 编程实例	413
练习题 20	418
上机实验题 15	418
第 21 章 使用 ADO 操作数据库	419
21.1 ADO 概述	419
21.2 Connection 对象	423
21.2.1 Connection 对象的属性	424
21.2.2 Connection 对象的方法	426
21.2.3 Connection 对象的事件	428
21.3 Recordset 对象	429
21.3.1 Recordset 对象的属性	430
21.3.2 Recordset 对象的方法	436
21.3.3 Recordset 对象的事件	447
21.4 Field 对象	451
21.4.1 Field 对象的属性	451
21.4.2 Field 对象的方法	454
21.4.3 Fields 集合	454
21.5 Command 对象	459
21.5.1 Command 对象的属性	460
21.5.2 Command 对象的方法	461
21.5.3 Parameters 集合和 Parameter 对象	463
21.6 ADO 编程综合实例	468
练习题 21	471
上机实验题 16	471
参考文献	472

第 1 章

数据库系统概述

数据库是一门研究数据管理的技术，始于 20 世纪 60 年代末，经过 30 多年的发展，已形成理论体系，成为计算机软件的一个重要分支。数据库技术主要研究如何存储、使用和管理数据，是计算机数据管理技术发展的最新阶段。在本章中，主要介绍数据管理技术的发展、数据模型和数据库系统的基本概念等，为后面各章的学习奠定基础。

1.1 信息、数据和数据处理

计算机的出现，开辟了数据处理的新纪元。数据处理的基本问题是数据的组织、存储、检索、维护和加工利用，这些正是数据库系统所要解决的问题。

数据是数据库系统研究和处理的对象。但是数据与信息是分不开的，它们既有联系又有区别，因此首先要搞清数据与信息在概念上的不同。

1.1.1 信息与数据

1. 信息

“信息”是对现实世界事物存在方式或运动状态的反映。具体地说，信息是一种已经被加工为特定形式的数据，这种数据形式对接收者来说是有意义的，而且对当前和将来的决策具有明显的或实际的价值。

信息有如下一些重要特征：

- 信息传递需要物质载体，信息的获取和传递要消耗能量。
- 信息是可以感知的。不同的信息源有不同的感知方式（如感觉器官、仪器或传感器等）。
- 信息是可以存储、压缩、加工、传递、共享、扩散、再生和增值的。

信息是资源，人类进行各项社会活动，不仅要考虑物质条件，而且要认真研究信息和利用信息。正因为如此，人们才将能源、物质和信息并列为人类社会活动的三大要素。

2. 数据

数据本质上是对信息的一种符号化表示，即用一定的符号表示信息。采用什么符号，完全是人为规定。为了用计算机进行信息处理，就得把信息转换为计算机能够识别的符号，

即用 0 和 1 两个符号编码来表示各种各样的信息。从这个意义上说，数据是用来载荷信息的。

3. 数据与信息的联系

信息与数据是两个既有联系、又有区别的概念。数据是信息的载体，而信息是数据的内涵。同一信息可以有不同的数据表示形式；而同一数据也可能有不同的解释。例如，“李明同学“数据库系统与应用”课程考试分数为 95 分”。这段文字（数据）提供了李明考试成绩优秀的信息。可见数据与信息是密切相关的。因此，在许多场合下，对它们不做严格的区分，可互换使用。例如通常说的，“信息处理”与“数据处理”具有同义性。

1.1.2 数据处理

当把信息表示成数据后，这些数据便被人们赋予了特定的含义，反映了现实世界事物的存在特性和变化状态。由于现实世界事物往往是相互关联的，基于这一事实，可以从已知数据出发，参照相关数据，进行加工计算，产生出一些新的数据。这些新的数据又表示了新的信息，可以作为某种决策的依据。上述的整个过程，就叫做数据处理。

在数据处理的一系列活动中，数据收集、存储、分类、传输等操作为基本操作，这些基本操作环节称为数据管理，而加工、计算、输出等操作是千变万化的，不同业务有不同的处理。数据管理技术是解决上述基本环节的，而其他环节是由应用程序实现的。

1.2 数据管理技术的发展

随着计算机软硬件技术的发展，数据管理技术的发展大致经历了人工管理、文件系统和数据库系统 3 个阶段。

1.2.1 人工管理阶段（20 世纪 50 年代）

这一时期，没有磁盘，没有专门的数据管理软件。计算机主要用于科学计算，数据量不大。人工管理方式的特点是：

- 数据不保存。
- 程序与数据合在一起，因而数据没有独立性，要修改数据必须修改程序。
- 编写程序时要安排数据的物理存储。一旦数据的物理存储改变，必须要重新编程，程序员的工作量大、繁琐，程序难以维护。
- 数据面向应用，这意味着即使多个不同程序用到相同数据，也得各自定义，数据不仅高度冗余，而且不能共享。

1.2.2 文件系统阶段（20 世纪 60 年代）

这一时期，计算机外存已有了磁鼓、磁盘等存储设备，软件有了操作系统。人们在操作系统的支持下，设计开发了一种专门管理数据的计算机软件，称之为文件系统。这时，计算机不仅用于科学计算，也已大量用于数据处理，其特点是：

- 数据以文件的形式长期保存。由于计算机大量用于数据处理，数据需要长期保留在

外存上反复处置,即经常对其进行查询、修改、插入和删除等操作。因此,在文件系统中,按一定的规则将数据组织为一个文件,存放在外存储器中长期保存。

- 数据的物理结构与逻辑结构有了区别,但比较简单。程序员只需用文件名与数据打交道,不必关心数据的物理位置,可由文件系统提供的读写方法去读/写数据。
- 文件形式多样化。为了方便数据的存储和查找,人们研究了许多文件类型,如索引文件、链接文件、顺序文件和倒排文件等。数据的存取基本上是以记录为单位的。
- 程序与数据之间有一定的独立性。应用程序通过文件系统对数据文件中的数据进行存取和加工,因此,处理数据时,程序不必过多地考虑数据的物理存储的细节,文件系统充当应用程序和数据之间的一种接口,可使应用程序和数据都具有一定的独立性。这样,程序员可以集中精力于算法,而不必过多地考虑物理细节。并且,数据在存储上的改变不一定反映在程序上,这可以大大节省维护程序的工作量。

尽管文件系统有上述优点,但是,这些数据在数据文件中只是简单地存放,文件中的数据没有结构,文件之间并没有有机的联系,仍不能表示复杂的数据结构;数据的存放仍依赖于应用程序的使用方法,基本上是一个数据文件对应于一个或几个应用程序;数据面向应用,独立性较差,仍然出现数据重复存储、冗余度大、一致性差(同一数据在不同文件中的值不一样)等问题。

1.2.3 数据库系统阶段(20世纪60年代后期)

随着计算机软硬件的发展、数据处理规模的扩大,20世纪60年代后期出现了数据库技术。关于什么是数据库,从不同的角度去定义可能差别较大,但是对数据库所应具有的特点,其认识大体上是一致的。下面指出数据库技术的若干特点:

- **数据结构化** 数据库是存储在磁盘等外部直接存取设备上的数据集合,按一定的数据结构组织起来。与文件系统相比,文件系统文件之间不存在联系,因而从总体上看数据是没有结构的;而数据库中的文件是相互联系着的,并在总体上遵从一定的结构形式。这是文件系统与数据库系统的最大区别。数据库正是通过文件之间的联系反映现实世界事物间的自然联系。
- **数据共享** 数据库中的数据是考虑所有用户的数据需求、面向整个系统组织的。因此数据库中包含了所有用户的数据成分,但每个用户通常只用到其中一部分数据。不同用户所使用的数据可以重叠,同一部分数据也可为多个用户共享。
- **减少了数据冗余** 在数据库方式下,用户不是自建文件,而是取自数据库中的某个子集,它并非独立存在,而是靠数据库管理系统(DataBase Management System, DBMS)从数据库中映象出来的,所以叫作逻辑文件。如图 1.1 所示,用户使用的是逻辑文件,因此尽管一个数据可能出现在不同的逻辑文件中,但实际上的物理存储只可能出现一次,这就减少了数据冗余。
- **有较高的数据独立性** 数据独立的好处是数据存储方式的改变不会影响到应用程序。数据独立又有两个含义,即物理数据独立性和逻辑数据独立性。所谓物理数据独立性是指数据库物理结构(包括数据的组织和存储、存取方法以及外部存储设备等)发生改变时,不会影响到逻辑结构,而用户使用的是逻辑数据,所以不必改动

程序；所谓逻辑数据独立性是指数据库全局逻辑发生改变时，用户也不需改动程序，就像数据库并没发生变化一样。这是因为用户仅使用数据库的一个子集，全局变化与否与具体用户无关，只要能从数据库中导出所用到的数据就行。

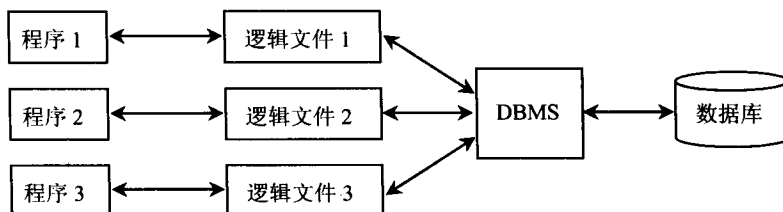


图 1.1 应用程序使用从数据库中导出的逻辑文件

- 用户接口 在数据库系统中，数据库管理系统作为用户与数据库的接口，提供了数据库定义、数据库运行、数据库维护和数据安全性、完整性等控制功能。此外，还支持某种程序设计语言，并设有专门的数据操纵语言，为用户编程提供了方便。

从文件系统管理发展到数据库系统管理是信息处理领域的重大变化，人们由传统的关注系统功能设计（因为程序设计处于主导地位，数据服从于程序）转向关注数据的结构设计，数据的结构设计成为信息系统首要关心的中心问题。

1.3 数据库系统的组成与结构

通常把引进了数据库技术的计算机系统称为数据库系统，它的目的是存储和产生所需要的有用信息。这些有用的信息可以是使用该系统的个人或组织的有意义的任何事情，换句话说，是对某个人或组织辅助决策过程中不可少的事情。

1.3.1 数据库系统的组成

数据库系统由数据库、支持数据库运行的软硬件、数据库管理系统、应用程序和人员等部分组成。

- 数据库 数据库是存储在外存上的若干个设计合理、满足应用需要的结构化的数据集合。
- 硬件 硬件是数据库赖以存在的物理设备，包括 CPU、存储器和其他外部设备等。数据库系统需要有足够大的内存和外存，用来运行操作系统、数据库管理系统核心模块和应用程序，以及存储数据库。
- 数据库管理系统（DBMS） 这是帮助用户创建、维护和使用数据库的软件系统，是数据库系统的核心。较流行的微机数据库管理系统有 FoxPro、Visual FoxPro、SQL Server 等。
- 相关软件 包括操作系统、编译系统、应用开发工具软件和计算机网络软件等。
- 应用程序 数据库是多用户共享的，不同用户的数据视图已由设计者组织在数据库中，但是如何使用是用户自己的事，可以在远程终端上查询数据，也可以编程处理自己的业务，其操作权限仅是数据库的一个子集。

- 人员 包括数据库管理员 (Data Base Administrator, DBA) 和用户。在大型数据库系统中, 需要有专人负责数据库系统的建立、维护和管理的工作, 承担该任务的人员称为数据库管理员。用户分为两类: 专业用户和最终用户。专业用户侧重设计数据库、开发应用系统程序, 为最终用户提供友好的用户界面。最终用户侧重对数据库的使用, 主要是通过数据库进行联机查询, 或者通过数据库应用系统提供的界面使用数据库。

1.3.2 数据库系统体系结构

数据库系统有着严谨的体系结构。目前世界上有大量的数据库在运行中, 其类型和规模可能相差很大, 但是就其体系结构而言却是大体相同的。

1. 数据库系统的三级模式结构

美国国家标准学会 (ANSI) 所属标准计划和要求委员会在 1975 年公布了一个关于数据库标准的报告, 提出了数据库的三级模式结构, 这就是有名的 SPARC 分级结构。三级结构对数据库的组织从内到外分 3 个层次描述, 分别称为内模式、概念模式和外模式, 如图 1.2 所示。

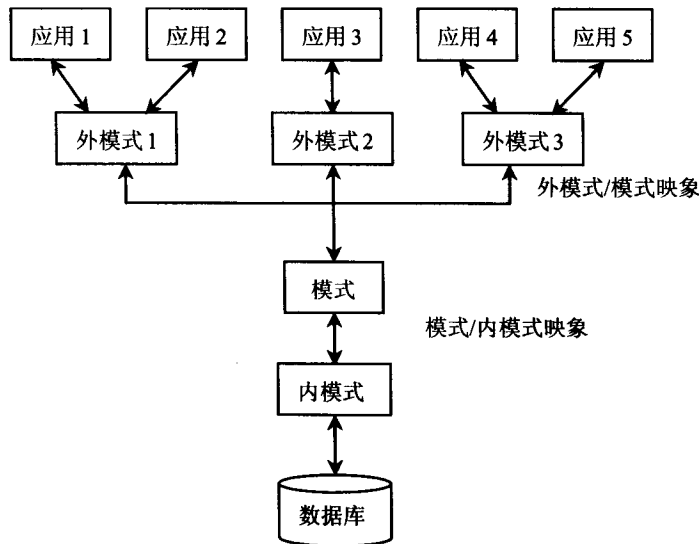


图 1.2 SPARC 分级结构

- 概念模式 简称模式, 是对数据库的整体逻辑结构和特征的描述, 并不涉及数据的物理存储细节和硬件环境, 与具体的应用程序和使用的应用开发工具无关。
- 内模式 又称存储模式, 具体描述了数据如何组织存储在存储介质上。内模式是系统程序员用一定的文件形式组织起来的一个个存储文件和联系手段; 也是由他们编制存取程序, 实现数据存取的。一个数据库只有一个内模式。
- 外模式 外模式通常是模式的一个子集, 故又称外模式为子模式。外模式面向用户, 它是数据库用户能够看到和使用的局部数据的逻辑结构和特征的描述, 是与某一应用有关的数据的逻辑表示。

综上所述，模式是内模式的逻辑表示，内模式是模式的物理实现，外模式则是模式的部分抽取。3 个模式反映了对数据库的 3 种不同观点：模式表示了概念级数据库，体现了对数据库的总体观；内模式表示了物理级数据库，体现了对数据库的存储观；外模式表示了用户级数据库，体现了对数据库的用户观。总体观和存储观只有一个，而用户观可能有多个，有一个应用，就有一个用户观。

2. 3 个模式之间的映象

前面谈到的三级模式，只有内模式才是真正存储数据的，而模式和外模式仅是一种逻辑表示数据的方法，但却可以放心大胆地使用它们，这是靠 DBMS 的映象功能实现的。

数据库系统的三级模式是对数据的 3 个抽象级别，它把数据的具体组织留给 DBMS 管理，使用户能逻辑地、抽象地处理数据，而不必关心数据在计算机中的具体表示方式与存储方式。为了能够在内部实现这 3 个抽象层次的联系和转换，数据库管理系统在这三级模式之间提供了两层映象：

- 外模式/模式映象
- 模式/内模式映象

正是这两层映象保证了数据库系统中的数据能够具有较高的逻辑独立性和物理独立性。

(1) 外模式/模式映象

模式描述的是数据的全局逻辑结构，外模式描述的是数据的局部逻辑结构。对应于同一个模式可以有任意多个外模式。对于每一个外模式，数据库系统都有一个外模式/模式映象，它定义了该外模式与模式之间的对应关系。这些映象定义通常包含在各自外模式的描述中。

当模式改变时（例如增加新的关系、新的属性、改变属性的数据类型等），由数据库管理员对各个外模式/模式的映象作相应改变，可以使外模式保持不变。应用程序是依据数据的外模式编写的，从而应用程序不必修改，保证了数据与程序的逻辑独立性，简称数据的逻辑独立性。

(2) 模式/内模式映象

数据库中只有一个模式，也只有一个内模式，所以模式/内模式映象是惟一的，它定义了数据库全局逻辑结构与存储结构之间的对应关系。例如，说明逻辑记录和字段在内部是如何表示的。该映象定义通常包含在模式描述中。当数据库的存储结构改变了（例如选用了另一种存储结构），由数据库管理员对模式/内模式映象作相应改变，可以使模式保持不变，从而应用程序也不必改变。保证了数据与程序的物理独立性，简称数据的物理独立性。

在数据库的三级模式结构中，数据库模式即全局逻辑结构是数据库的中心与关键，它独立于数据库的其他层次。因此设计数据库模式结构时应首先确定数据库的逻辑模式。

数据库的内模式依赖于它的全局逻辑结构，但独立于数据库的用户视图（即外模式），也独立于具体的存储设备。它是将全局逻辑结构中所定义的数据结构及其联系按照一定的物理存储策略进行组织，以达到较好的时间与空间效率。