
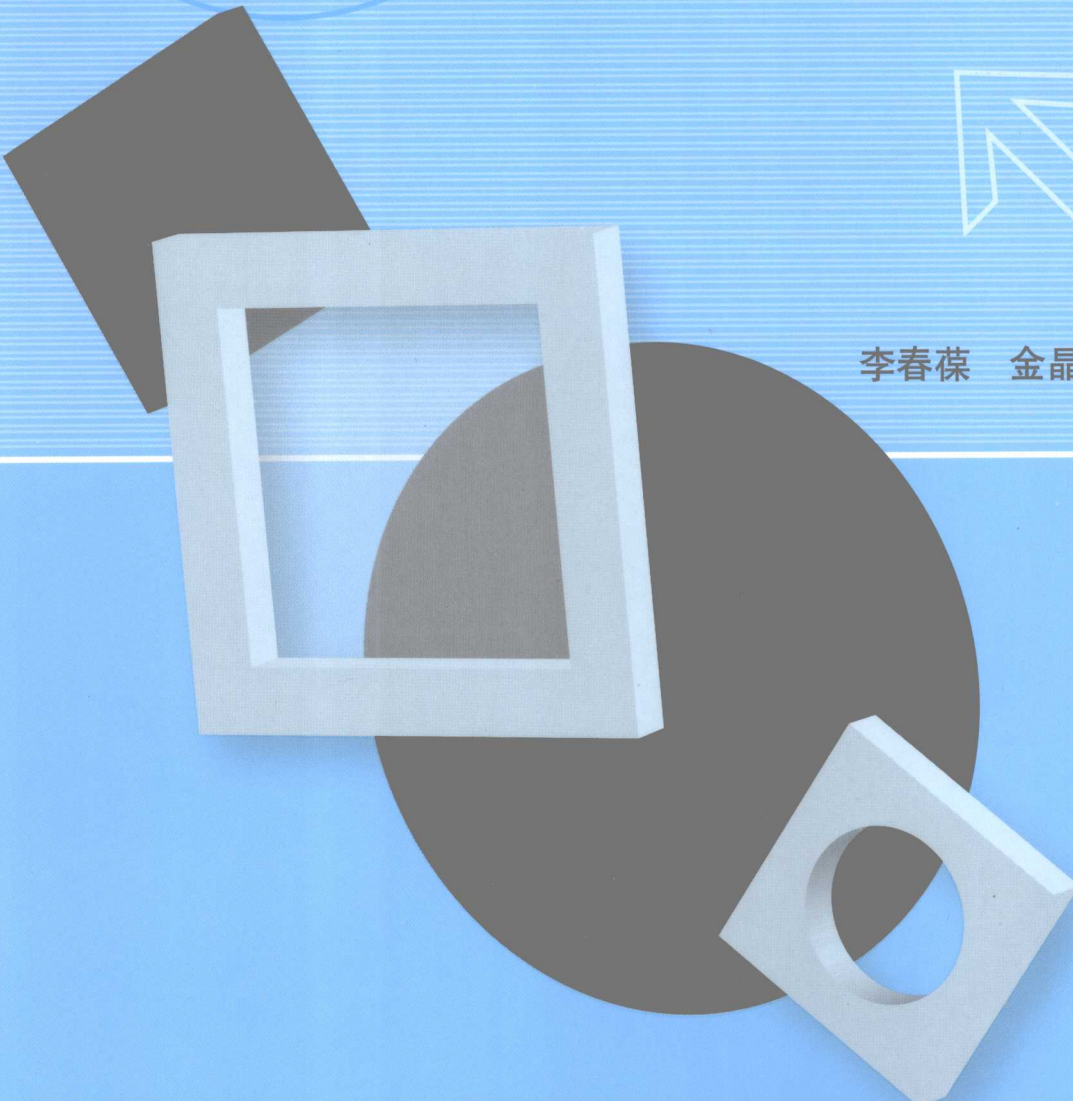



计算机与信息技术专业应用教材

数据库原理与应用

——基于SQL Server 2005



李春葆 金晶 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 2005

李春葆 金晶 曾平 编著

清华大学出版社

北京

内 容 简 介

本书介绍数据库基本原理,并以 SQL Server 2005 为平台介绍数据库管理系统的应用。全书分为 3 部分,第 1~5 章介绍数据库的一般原理,第 6~18 章介绍 SQL Server 2005 的数据管理功能,第 19~20 章介绍以 VB 作为前端设计工具、以 SQL Server 2005 作为数据库平台开发数据库应用系统的技术。本书最后的附录还给出了一个数据库系统设计过程。

本书由浅入深、循序渐进地介绍各个知识点。书中提供了大量例题,有助于读者理解概念和巩固知识;此外,还提供了一定数量的练习题和上机实验题,便于学生训练和上机实习。

为了便于学习,我们提供了《数据库原理与应用——基于 SQL Server 2005 学习与上机实验指导》,对本书中所有练习题和上机实验题进行了详细分析解答,请登录 <http://www.khp.com.cn> 免费下载。

本书可以作为各类院校相关专业及培训班的“数据库原理与应用”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。

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。

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

图书在版编目(CIP)数据

数据库原理与应用:基于 SQL Server 2005 /李春葆,金晶,曾平编著.

—北京:清华大学出版社,2008

ISBN 978-7-302-19042-4

I. 数… II. ①李… ②金… ③曾… III. 关系数据库—数据库管理系统, SQL Server 2005 IV. TP311.138

中国版本图书馆 CIP 数据核字(2008)第 190846 号

责任编辑:陈洁

封面设计:林陶

出版者:清华大学出版社

地 址:北京清华大学学研大厦 A 座

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

邮 编:100084

社总机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印装者:北京市科普瑞印刷有限责任公司

经 销:全国新华书店

开 本:185×260 印 张:26.25

字 数:638 千字

版 次:2009 年 1 月第 1 版

印 次:2009 年 1 月第 1 次印刷

印 数:1~4 000

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 转 8501 产品编号:030442-01

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，激发他们的学习的热情和创新欲望。

☑ 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

本次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 22 本：

- 计算机系统结构教程
- Java 语言程序设计
- C++ 程序设计教程（第 2 版）
- C++ 程序设计学习与上机实验指导
- 数据结构与算法教程（第 2 版）
- 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 程序设计教程（基于 Turbo C 平台）
- Access 数据库程序设计
- Access 2003 程序设计教程（第 2 版）
- 数据库原理与应用——基于 Visual FoxPro（第 2 版）
- 数据库原理与应用——基于 SQL Server 2000（第 2 版）
- Visual FoxPro 程序设计（第 2 版）
- Visual Basic 程序设计（第 2 版）
- 操作系统教程（第 2 版）
- SQL Server 2000 应用系统开发教程（第 2 版）
- 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 Access 2003（第 2 版）
- 数据库系统开发教程——基于 SQL Server 2005+VB
- 数据库系统开发教程——基于 SQL Server 2005+VB 学习与上机实验指导
- 数据库原理与应用——基于 SQL Server 2005

服务之窗

本丛书的出版者和作者竭诚为读者提供服务。

本套丛书免费为教师提供 PowerPoint 演示文档，该文档可将书中的内容及图片以幻灯片的形式呈现在学生面前，在很大程度上减轻了教师的备课负担，所以深受广大教师的欢迎。请致电：010-82896438 或发 E-mail：feedback@khp.com.cn 获取电子教案。

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：licb1964@126.com

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对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、专科学生的一门重要的专业课程，该课程既具有较强的理论性，又具有较强的实践性。

本书基于SQL Server 2005讨论数据库的原理和应用方法。

全书分为3部分，第1~5章介绍数据库的一般原理，第6~18章介绍SQL Server 2005的数据管理功能，第19~20章介绍以VB作为前端设计工具、以SQL Server 2005作为数据库平台开发数据库应用系统的技术。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 2005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T-SQL基础，第10章为SQL高级应用，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章为数据备份与恢复，第18章为数据转换，第19章为VB与数据库，第20章为使用ADO操作SQL Server数据库。

本书最后的附录还给出了一个数据库系统设计过程。

本书内容由浅入深，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。每一章后都给出相应的练习题，多数章后给出了一定数量的上机实验题。本书可以作为各类院校相关专业及培训班的“数据库原理与应用”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。

为了便于学习，我们提供了《数据库原理与应用——基于SQL Server 2005学习与上机实验指导》，对本书中所有练习题和上机实验题进行了详细分析解答，请登录<http://khp.com.cn>免费下载。

由于编者水平所限，书中难免存在不足之处，敬请广大读者指正。编者的E-mail为licb1964@126.com。

编者
2009年1月

目 录

第1章 数据库系统概述	1	3.2 关系的数学定义	24
1.1 信息、数据和数据处理	1	3.3 关系代数	26
1.1.1 信息与数据	1	3.3.1 传统的集合运算	26
1.1.2 数据处理	2	3.3.2 专门的关系运算	27
1.2 数据管理技术的发展	2	练习题3	29
1.2.1 人工管理阶段	2	第4章 关系数据库规范化理论	30
1.2.2 文件系统阶段	3	4.1 问题的提出	30
1.2.3 数据库系统阶段	3	4.2 函数依赖	31
1.3 数据库系统的组成与结构	4	4.2.1 函数依赖的定义	31
1.3.1 数据库系统的组成	4	4.2.2 函数依赖与属性关系	32
1.3.2 数据库系统体系结构	6	4.2.3 Armstrong公理	33
1.4 数据库管理系统	8	4.2.4 闭包及其计算	34
1.4.1 DBMS的主要功能	9	4.3 范式和规范化	35
1.4.2 DBMS的组成	10	4.3.1 什么叫范式	35
练习题1	11	4.3.2 范式的判定条件与规范化	35
第2章 数据模型	12	4.4 关系模式的分解	37
2.1 什么是数据模型	12	4.4.1 模式分解中存在的问题	37
2.1.1 数据的描述	13	4.4.2 无损分解的定义和性质	37
2.1.2 数据间联系的描述	13	4.4.3 无损分解的测试方法	38
2.2 概念模型	13	4.4.4 保持函数依赖的分解	38
2.2.1 信息世界中的基本概念	13	练习题4	39
2.2.2 实体间的联系方式	15	第5章 数据库设计	41
2.2.3 实体联系表示法	15	5.1 数据库设计概述	41
2.2.4 设计E-R图的步骤	17	5.2 需求分析	42
2.3 数据库类型	18	5.2.1 需求分析的步骤	42
2.3.1 层次模型	18	5.2.2 需求分析的方法	43
2.3.2 网状模型	19	5.3 概念结构设计	47
2.3.3 关系模型	21	5.3.1 局部应用E-R模型设计	47
练习题2	22	5.3.2 总体概念E-R模型设计	49
第3章 关系数据库	23	5.4 逻辑结构设计	51
3.1 关系模型的基本概念	23	5.5 物理结构设计	52
		5.6 数据库的实施和维护	53

练习题5.....	54	7.5.1 添加和删除数据文件、 日志文件.....	78
第6章 SQL Server 2005系统概述.....	55	7.5.2 数据库选项设置.....	82
6.1 SQL Server 2005系统简介.....	55	7.5.3 更改数据库所有者.....	83
6.1.1 SQL Server的发展历史.....	55	7.5.4 数据库重命名.....	84
6.1.2 SQL Server 2005的各种版本.....	55	7.6 删除数据库.....	85
6.1.3 SQL Server 2005的组成部分.....	56	练习题7.....	85
6.1.4 SQL Server 2005组件的分类.....	57	上机实验题2.....	86
6.2 系统需求.....	59	第8章 创建和使用表.....	87
6.2.1 硬件需求.....	59	8.1 表的概念.....	87
6.2.2 软件需求.....	60	8.1.1 什么是表.....	87
6.2.3 SQL Server 2005的网络环境 需求.....	61	8.1.2 表中数据的完整性.....	88
6.2.4 SQL Server 2005的其他需求.....	61	8.2 创建表.....	89
6.2.5 SQL Server 2005安装的 注意事项.....	61	8.3 修改表的结构.....	92
6.3 SQL Server 2005的安装.....	62	8.4 建立数据库关系图.....	93
6.4 SQL Server 2005的工具和实用程序....	65	8.5 更改表名.....	95
6.4.1 SQL Server Management Studio... 66		8.6 删除表.....	96
6.4.2 SQL Server Business Intelligence Development Studio.....	67	8.7 记录的新增和修改.....	97
6.4.3 数据库引擎优化顾问.....	68	练习题8.....	98
6.4.4 Analysis Services.....	69	上机实验题3.....	99
6.4.5 SQL Server Configuration Manager.....	69	第9章 T-SQL基础.....	101
6.4.6 文档和教程.....	70	9.1 SQL.....	101
练习题6.....	70	9.1.1 SQL概述.....	101
上机实验题1.....	70	9.1.2 SQL的分类.....	101
第7章 创建和使用数据库.....	71	9.2 T-SQL语句的执行.....	102
7.1 数据库对象.....	71	9.3 数据定义语言.....	103
7.2 系统数据库.....	72	9.3.1 数据库的操作语句.....	103
7.3 SQL Server数据库的存储结构.....	73	9.3.2 表的操作语句.....	108
7.3.1 文件和文件组.....	73	9.4 数据操纵语言.....	111
7.3.2 数据库的存储结构.....	74	9.4.1 INSERT语句.....	111
7.3.3 事务日志.....	75	9.4.2 UPDATE语句.....	112
7.4 创建数据库.....	76	9.4.3 DELETE语句.....	112
7.5 配置数据库.....	78	9.5 数据查询语言.....	112
		9.5.1 投影查询.....	113
		9.5.2 选择查询.....	114
		9.5.3 排序查询.....	115

9.5.4 使用聚合函数	115	第11章 索引	193
9.5.5 表的连接查询	117	11.1 什么是索引	193
9.5.6 子查询	119	11.2 索引类型	193
9.5.7 相关子查询	121	11.2.1 聚集索引	194
9.5.8 带EXISTS测试的子查询	122	11.2.2 非聚集索引	194
9.5.9 空值及其处理	123	11.2.3 唯一性索引	195
9.6 T-SQL程序设计基础	124	11.3 创建索引	195
9.6.1 标识符	124	11.3.1 使用SQL Server控制管理器 创建索引	195
9.6.2 数据类型	125	11.3.2 使用CREATE INDEX语句 创建索引	199
9.6.3 变量	135	11.3.3 使用CREATE TABLE语句 创建索引	202
9.6.4 运算符	139	11.4 查看和修改索引属性	202
9.6.5 批处理	143	11.4.1 使用SQL Server控制管理器 查看和修改索引属性	202
9.6.6 注释	144	11.4.2 使用T-SQL语句查看和修改 索引属性	203
9.6.7 控制流语句	145	11.5 删除索引	204
9.6.8 函数	150	11.5.1 使用SQL Server控制管理器 删除索引	204
练习题9	158	11.5.2 使用T-SQL语句删除索引	205
上机实验题4	158	练习题11	205
第10章 SQL高级应用	160	上机实验题6	205
10.1 SELECT高级查询	160	第12章 视图	206
10.1.1 数据汇总	160	12.1 视图概述	206
10.1.2 连接查询	163	12.2 创建视图	206
10.1.3 子查询	166	12.2.1 使用SQL Server管理控制器 创建视图	207
10.1.4 数据来源是一个查询的结果	169	12.2.2 使用SQL语句创建视图	210
10.2 事务处理	170	12.3 使用视图	211
10.2.1 事务分类	170	12.3.1 使用视图进行数据查询	212
10.2.2 显式事务	171	12.3.2 通过视图向基表中插入数据	213
10.2.3 自动提交事务	174	12.3.3 通过视图修改基表中的数据	214
10.2.4 隐性事务	175	12.3.4 通过视图删除基表中的数据	215
10.3 数据的锁定	176	12.4 视图定义的修改	215
10.3.1 SQL Server中的锁定	176		
10.3.2 自定义锁	179		
10.4 使用游标	184		
10.4.1 游标的概念	185		
10.4.2 游标的基本操作	185		
10.4.3 使用游标	188		
练习题10	191		
上机实验题5	192		

12.4.1 使用SQL Server管理控制器 修改视图定义	216	14.2.2 使用CREATE PROCEDURE 语句创建存储过程.....	238
12.4.2 重命名视图	218	14.3 执行存储过程	239
12.5 查看视图的信息	219	14.4 存储过程的参数	241
12.5.1 使用SQL Server管理控制器 查看视图信息	219	14.4.1 在存储过程中使用参数.....	241
12.5.2 使用sp_helptext存储过程 查看视图信息	220	14.4.2 在存储过程中使用默认参数	242
12.6 视图的删除	220	14.4.3 在存储过程中使用返回参数	243
12.6.1 使用SQL Server管理控制器 删除视图	220	14.4.4 存储过程的返回值.....	244
12.6.2 使用T-SQL删除视图	221	14.5 存储过程的管理	245
练习题12.....	222	14.5.1 查看存储过程.....	245
上机实验题7.....	222	14.5.2 修改存储过程.....	247
第13章 数据库完整性	223	14.5.3 重命名存储过程.....	248
13.1 约束.....	223	14.5.4 删除存储过程.....	249
13.1.1 PRIMARY KEY约束	223	练习题14.....	250
13.1.2 FOREIGN KEY约束	224	上机实验题9.....	250
13.1.3 UNIQUE约束	225	第15章 触发器	251
13.1.4 CHECK约束	226	15.1 概述	251
13.1.5 列约束和表约束.....	227	15.2 创建DML触发器	252
13.2 默认值.....	227	15.2.1 使用SQL Server管理控制器 创建DML触发器.....	252
13.2.1 在创建表时指定默认值.....	228	15.2.2 使用T-SQL语句创建DML 触发器.....	253
13.2.2 使用默认对象	229	15.2.3 创建DML触发器的注意事项... ..	256
13.3 规则.....	232	15.3 inserted表和deleted表	256
13.3.1 创建规则	233	15.4 使用DML触发器	257
13.3.2 绑定规则	234	15.4.1 使用INSERT触发器	258
13.3.3 解除和删除规则.....	234	15.4.2 使用UPDATE触发器	259
练习题13.....	235	15.4.3 使用DELETE触发器.....	261
上机实验题8.....	235	15.4.4 使用INSTEAD OF触发器.....	262
第14章 存储过程	236	15.5 创建和使用DDL触发器	263
14.1 概述.....	236	15.5.1 创建DDL触发器.....	263
14.2 创建存储过程.....	237	15.5.2 DDL触发器的应用.....	264
14.2.1 使用SQL Server管理控制器 创建存储过程	237	15.6 触发器的管理	265
		15.6.1 查看触发器.....	265
		15.6.2 修改触发器.....	267
		15.6.3 删除触发器.....	268
		15.6.4 启用或禁用触发器.....	268

练习题15.....	269	17.4.3 完整+差异+日志数据库 备份和恢复.....	306
上机实验题10.....	269	练习题17.....	307
第16章 SQL Server的安全管理	271	上机实验题12.....	307
16.1 SQL Server安全体系结构	271	第18章 数据转换.....	308
16.1.1 操作系统的安全性.....	272	18.1 数据的导出和导入	308
16.1.2 SQL Server的安全性.....	272	18.1.1 数据的导出.....	308
16.1.3 数据库的安全性.....	272	18.1.2 数据的导入.....	311
16.1.4 SQL Server数据库对象的 安全性	272	18.2 分离和附加用户数据库	311
16.2 SQL Server的身份验证模式	273	18.2.1 分离用户数据库.....	311
16.2.1 Windows身份验证模式	273	18.2.2 附加用户数据库.....	313
16.2.2 混合身份验证模式.....	274	18.3 脚本	315
16.2.3 设置身份验证模式.....	274	18.3.1 将数据库生成脚本.....	315
16.3 SQL Server账号管理	275	18.3.2 将数据表生成脚本.....	317
16.3.1 SQL Server服务器登录 账号管理	275	18.3.3 执行脚本.....	318
16.3.2 SQL Server数据库用户 账号管理	279	练习题18.....	318
16.4 权限和角色	283	上机实验题13	318
16.4.1 什么是角色	283	第19章 VB与数据库	319
16.4.2 权限管理	285	19.1 概述	319
16.4.3 创建数据库角色.....	289	19.1.1 数据访问对象模型.....	319
16.4.4 数据库角色管理.....	290	19.1.2 ODBC和数据源.....	319
练习题16.....	291	19.2 常用的数据访问控件	323
上机实验题11	291	19.2.1 ADO Data控件	323
第17章 数据备份与恢复.....	292	19.2.2 DataCombo控件	330
17.1 数据备份和恢复概述	292	19.2.3 DataGrid控件.....	332
17.1.1 备份类型	292	19.3 常用的高级用户界面控件.....	337
17.1.2 恢复类型	293	19.3.1 图像列表控件.....	337
17.2 备份设备.....	294	19.3.2 TreeView控件.....	340
17.2.1 创建数据库备份设备.....	294	19.3.3 ListView控件.....	345
17.2.2 删除数据库备份设备.....	296	19.4 综合数据库操作设计实例.....	352
17.3 选择数据库恢复类型	297	练习题19	357
17.4 数据库备份和恢复过程	298	上机实验题14	358
17.4.1 完整数据库备份和恢复.....	298	第20章 使用ADO操作SQL Server 数据库.....	359
17.4.2 完整+日志数据库备份和恢复... 303		20.1 ADO概述.....	359
		20.2 Connection对象.....	363

20.2.1	使用Connection对象的操作 ...	363
20.2.2	ADO连接数据库的方式	363
20.2.3	Connection对象的属性	364
20.2.4	Connection对象的方法	366
20.2.5	Connection对象的事件	368
20.3	Recordset对象	369
20.3.1	ADO的游标类型	370
20.3.2	使用Recordset对象的操作	370
20.3.3	Recordset对象的属性	370
20.3.4	Recordset对象的方法	374
20.3.5	Recordset对象的事件	379
20.4	Field对象	382
20.4.1	Field对象的操作	382
20.4.2	Field对象的属性	383
20.4.3	Field对象的方法	386
20.4.4	Fields集合	386
20.5	Command对象	390
20.5.1	Command对象的操作	390
20.5.2	Command对象的属性	391
20.5.3	Command对象的方法	392
20.5.4	Parameters集合和Parameter 对象	394
20.6	ADO编程综合实例	398
	练习题20	404
	上机实验题15	405
	参考文献	406
	附录 数据库应用系统开发综合 实习题	407

第 1 章

数据库系统概述

CHAPTER 01

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。本章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。

3. 数据与信息的关系

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

人工管理阶段（20 世纪 50 年代），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

文件系统阶段（20世纪60年代），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

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，文件系统中的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，所以叫做逻辑文件。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。

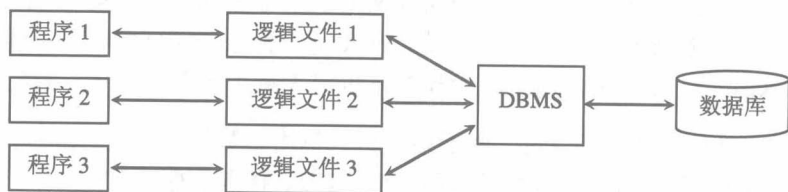


图 1.1 应用程序使用从数据库中导出的逻辑文件

- 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。
- 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

通常把引入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不可少的事情。

1.3.1 数据库系统的组成

数据库系统（Database System, DBS）是数据库应用系统的简称。数据库系统是指计算机系统中引入数据库之后组成的系统，是用来组织和存取大量数据的管理系统。数据库系统是由计算机系统、数据库、数据库管理系统、应用程序和用户组成的。数据库系统的组成及其各组件之间的关系如图 1.2 所示。

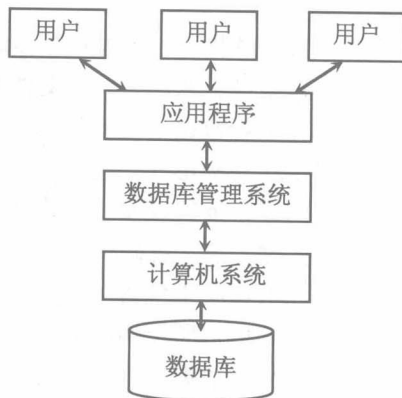


图 1.2 数据库系统组成

1. 计算机系统

计算机系统由硬件和必须的软件组成。

- 硬件。指存储数据库和运行 DBMS（包括操作系统）的硬件资源。它包括物理存储数据库的磁盘、磁鼓、磁带或其他外存储器及其附属设备、控制器、I/O 通道、内存、CPU 及其他外部设备等。
- 必须的软件。指计算机正常运行所需要的操作系统和各种驱动程序等。

2. 数据库

数据库是指数据库系统中集中存储的一批数据的集合。它是数据库系统的工作对象。

为了把输入、输出或中间数据加以区别，通常把数据库数据称为“存储数据”、“工作数据”或“操作数据”。它们是某特定应用环境中进行管理和决策所必须的信息。

特定的应用环境，可以指一个公司、一个银行、一所医院或一所学校等各种各样的应用环境。在这些各种各样的应用环境中，各种不同的应用可通过访问其数据库获得必要的信息，以辅助进行决策，决策完成后，再将决策结果存储在数据库中。

特别需要指出的是，数据库中的存储数据是“集成的”和“共享的”。

所谓“集成”，是指把某特定应用环境中的各种应用相关的数据及其数据之间的联系（联系也是一种数据）全部集中地按照一定的结构形式进行存储。或者说，把数据库看成若干单个性质的数据文件的联合和统一的数据整体，并且在文件之间局部或全部消除了冗余。这使数据库系统具有整体数据结构化和数据冗余小的特点。

所谓“共享”，是指数据库中的一块块数据可为多个不同的用户所共享，即多个不同的用户，使用多种不同的语言，为了不同的应用目的，而同时存取数据库，甚至同时存取同一块数据。共享实际上是基于数据库是“集成的”这一事实的结果。

3. 数据库管理系统

DBMS 用于数据库存取、维护和管理。数据库系统各类用户对数据库的各种操作请求，都是由 DBMS 来完成的，它是数据库系统的核心软件。DBMS 提供一种超出硬件层的对数据库进行观察的功能，并支持用较高的观点来表达用户的操作，使数据库用户不受硬件层细节的影响。DBMS 是在操作系统（OS）支持下工作的。

4. 应用程序

应用程序介于用户和数据库管理系统之间，是指完成用户操作的程序，该程序将用户的操作转换成一序列的命令执行，例如，实现学生平均分统计、打印学生学籍表等。在这些命令中，需要对数据库中的数据进行查询、插入、删除和统计等，应用程序将这些复杂的数据库操作交由数据库管理系统来完成。

5. 用户

用户是指存储、维护和检索数据库中数据的使用人员。数据库系统中主要有 3 类用户：终端用户、应用程序员和数据库管理员。