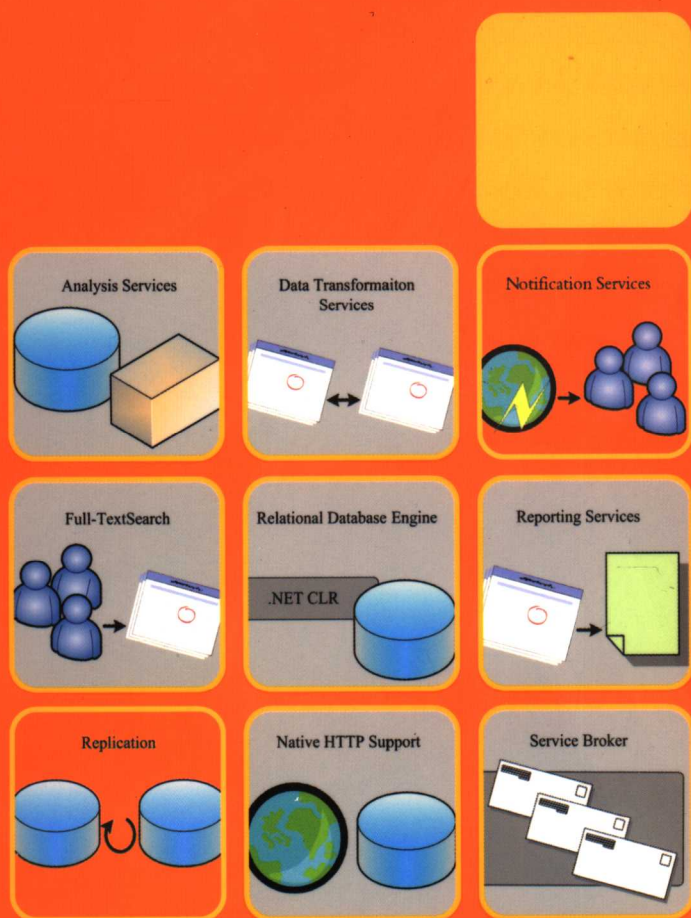


计算机应用能力培养丛书

SQL Server 2005

数据库原理与应用简明教程

李维杰 孙乾君 编著



清华大学出版社

计算机应用能力培养丛书

SQL Server 2005 数据库原理 与应用简明教程

李维杰 孙乾君 编著

清华大学出版社

北 京

内 容 简 介

本书通过大量翔实的示例,向读者展示了 SQL Server 2005 数据库的原理、开发、应用及管理的相关技术和技巧。全书 15 章,可分为 3 大部分:第 1 部分(第 1、2 章)概述了数据库的基本原理,包括数据处理的基础知识、数据库的基本体系结构、开发过程、关系数据库的设计理论等内容;第 2 部分(第 3~9 章)是本书的重点,也是读者深入学习 SQL Server 2005 的基础,详细介绍了 SQL Server 2005 的安装和配置过程,如何创建和管理数据库、数据表,以及如何对数据进行查询、修改、输出等内容;第 3 部分(第 10~14 章)详细介绍了 SQL Server 2005 的一些高级话题,如存储过程、触发器、约束、同义词,以及数据库的备份和还原、安全管理技术,数据的分析与挖掘等内容。为提高读者的应用开发能力,本书第 15 章提供了电子相册系统的具体开发过程,以供学习和参考。

本书内容丰富,结构清晰,核心概念和关键技术讲解清楚,同时提供了丰富的示例以展示具体应用。本书具有很强的操作性和实用性,可作为高等院校、高职学校“SQL Server 数据库原理与应用”课程的教材,或社会培训班“SQL Server 2005”的基础入门教材,同时对从事数据库开发和管理人员也具有较高参考价值。

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,侵权必究。侵权举报电话:010-62782989 13501256678 13801310933

图书在版编目(CIP)数据

SQL Server 2005 数据库原理与应用简明教程/李维杰,孙乾君编著. —北京:清华大学出版社,2007.5

(计算机应用能力培养丛书)

ISBN 978-7-302-15144-9

I. S… II. ①李…②孙… III. 关系数据库—数据库管理系统, SQL Server 2005—教材 IV. TP311.138

中国版本图书馆 CIP 数据核字(2007)第 061859 号

责任编辑:王 军 杜一民

装帧设计:康 博

责任校对:胡雁翎

责任印制:何 芊

出版发行: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

c-service@tup.tsinghua.edu.cn

社 总 机:010-62770175 邮购热线:010-62786544

投稿咨询:010-62772015 客户服务:010-62776969

印 刷 者:北京市人民文学印刷厂

装 订 者:三河市金元印装有限公司

经 销:全国新华书店

开 本:185×260 印 张:22.25 字 数:541 千字

版 次:2007 年 5 月第 1 版 印 次:2007 年 5 月第 1 次印刷

印 数:1~5000

定 价:29.00 元

本书如存在文字不清、漏印、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)62770177 转 3103 产品编号:019408-01

读者意见反馈卡

亲爱的读者：

您好！感谢您购买本书。为了今后能为您提供更优秀的图书，请您抽出宝贵的时间填写下面的意见反馈卡，然后剪下寄到：北京清华大学出版社第五事业部(邮编100084)。您也可以把意见反馈到bookservice@263.net。我们将充分考虑您的意见和建议，并尽可能地给您满意的答复。谢谢！

本系列图书订购咨询电话：010-62794504；邮购咨询电话：010-62770175/77 转 3505。

图书名称：SQL Server 2005 数据库原理与应用简明教程

读者资料卡

姓 名：_____ 性 别：☐男 ☐女 年 龄：_____ 文化程度：_____

职 业：☐教师 ☐学生 ☐其他 院校专业：_____

联系电话：_____ E-mail：_____

您使用本书是作为：☐指定教材 ☐教辅用书 ☐个人自学 ☐参加认证考试

您对本书的满意程度：

语言文字：☐很满意 ☐比较满意 ☐一般 ☐较不满意 ☐不满意 建议_____

技术含量：☐很满意 ☐比较满意 ☐一般 ☐较不满意 ☐不满意 建议_____

封面设计：☐很满意 ☐比较满意 ☐一般 ☐较不满意 ☐不满意 建议_____

印刷质量：☐很满意 ☐比较满意 ☐一般 ☐较不满意 ☐不满意 建议_____

您对本书的总体满意度：

☐很满意 ☐比较满意 ☐一般 ☐较不满意 ☐不满意

您希望本书在哪些方面进行修改或改进？（可附页）_____

影响您购买图书的因素：

☐书名 ☐作者名声 ☐出版机构 ☐封面封底 ☐装帧设计 ☐内容提要、前言或目录

☐价格（☐20 元内 ☐30 元内 ☐50 元内 ☐100 元内） ☐书店宣传 ☐网络宣传

☐知名专家学者的书评或推荐 ☐其他：_____

您是如何获得图书信息的：

☐朋友推荐 ☐出版社图书目录 ☐网站 ☐书店 ☐杂志、报纸 ☐其他_____

您对我们的建议：（可附页）_____

电子教案支持

敬爱的教师：

为了配合本课程的教学需要，本系列教材配有相应的电子教案，有需求的教师可以与我们联系，我们将向使用本教材进行教学的教师免费赠送电子教案，以方便教学工作的开展。有需求的教师，请拨打电话010-62794504 或发电子邮件至 wkservice@tup.tsinghua.edu.cn 进行咨询，也可以直接填写读者意见反馈卡并在右下角进行标注，然后加盖院系公章后，按反馈卡地址寄给我们。

☐ 索取电子教案

前言

高职高专教育以就业为导向,以技术应用型人才为培养目标,担负着为国家经济高速发展输送一线高素质技术应用人才的重任。近年来,随着我国高等职业教育的发展,高职院校数量和在校生人数均有了大幅激增,已经成为我国高等教育的重要组成部分。

根据目前我国高级应用型人才的紧缺情况,教育部联合六部委推出“国家技能型紧缺人才培养培训项目”,并从2004年秋季起,在全国两百多所学校的计算机应用与软件技术、数控项目、汽车维修与护理等专业推行两年制和三年制改革。

为了配合高职高专院校的学制改革和教材建设,清华大学出版社在主管部门的指导下,组织了一批工作在高等职业教育第一线的资深教师和相关行业的优秀工程师,编写了适应新教学要求的计算机系列高职高专教材——《计算机应用能力培养丛书》。该丛书主要面向高等职业教育,遵循“以就业为导向”的原则,根据企业的实际需求来进行课程体系设置和教材内容选取。根据教材所对应的专业,以“实用”为基础,以“必需”为尺度,为教材选取理论知识;注重和提高案例教学的比重,突出培养人才的应用能力和实际问题解决能力,满足高等职业教育“学校评估”和“社会评估”的双重教学特征。

每本教材的内容均由“授课”和“实训”两个互为联系和支持的部分组成,“授课”部分介绍在相应课程中,学生必须掌握或了解的基础知识,每章都设有“学习目标”、“实际问题解答”、“小结”、“习题”等特色段落;“实训”部分设置了一组源于实际应用的上机实例,用于强化学生的计算机操作使用能力和解决实际问题的能力。每本教材配套的习题答案、电子教案和一些教学课件均可在该丛书的信息支持网站(<http://www.tupwk.com.cn/GZGZ>)上下载或通过 Email(wkservice@tup.tsinghua.edu.cn)索取,读者在使用过程中遇到了疑惑或困难可以在支持网站的互动论坛上留言,本丛书的作者或技术编辑会提供相应的技术支持。

作为当前使用最流行的关系型数据库,SQL Server 在大型数据库开发和管理、企业业务数据处理等方面具有广泛应用,SQL Server 2005 相对于 2000 及以前版本,功能和安全等方面性能有了很大提升,已经从单纯的数据库系统发展为数据处理平台,包括数据库引擎、Analysis Services、Reporting Services、Integration Services、Notification Services 等服务,提供了更加专业级的商务智能解决方案,成为企业处理和管理业务数据的首选商业数据平台。

本书依据教育部《高职高专教育计算机公共基础课程教学基本要求》编写而成,基于 SQL Server 2005 数据库,在注重讲解基础知识和开发、管理技术的同时,注重体现其最新特色,并提供精彩的示例,展示 SQL Server 2005 的开发和管理过程。

本书结构紧凑,内容承上启下,共 15 章,可分为 3 大部分:第 1 部分(第 1、2 章)概述了数据库的基本原理,包括数据处理的基础知识、数据库的基本体系结构、开发过程,



以及关系数据库的设计理论等内容；第 2 部分(第 3~9 章)是本书的重点，也是读者深入学习 SQL Server 2005 的基础，详细介绍了 SQL Server 2005 的安装和配置过程，如何创建和管理数据库、数据表，以及如何对数据进行查询、修改、输出等内容；第 3 部分(第 10~14 章)详细介绍了 SQL Server 2005 的一些高级话题，如存储过程、触发器、约束、同义词，以及数据库的备份和还原、安全管理技术、数据的分析与挖掘等内容。为提高读者的应用开发能力，本书第 15 章提供了电子相册系统的具体开发过程，以供学习和参考。

由于计算机科学技术发展迅速，再者受自身水平和编写时间所限，书中如有错误或不足之处，欢迎广大读者对我们提出意见或建议。

作 者

2006 年 11 月

目 录

第1章 数据库基础知识	1
1.1 数据库概述	1
1.1.1 基本概念	1
1.1.2 数据管理技术的产生和发展	3
1.1.3 数据库系统的应用	6
1.1.4 常见的数据库管理系统	7
1.2 数据库系统的结构	8
1.2.1 模式(Schema)的概念	8
1.2.2 数据库的三级模式结构	8
1.3 数据库系统的组成	9
1.4 数据库的设计流程	10
1.4.1 需求分析	11
1.4.2 概念结构设计	11
1.4.3 逻辑结构设计	11
1.4.4 物理结构设计	12
1.4.5 数据库的实施	12
1.4.6 数据库的运行和维护	12
本章小结	12
习题	13
第2章 关系数据库及其设计理论	14
2.1 关系模型	14
2.1.1 数据模型	14
2.1.2 关系模型的数据结构	16
2.1.3 关系模型的特点	17
2.2 关系数据库	17
2.2.1 关系操作	17
2.2.2 关系的完整性	20
2.3 关系数据库的设计理论	21
2.3.1 函数依赖	21
2.3.2 关系模式的范式(NF, Normal Form)	22

2.3.3 关系模式的规范化	23
2.3.4 范式应用——设计论坛数据库	24
2.4 数据库的恢复与并发控制	25
2.4.1 事务的基础知识	25
2.4.2 数据库的恢复	26
2.4.3 数据库的并发控制	27
本章小结	31
习题	31
第3章 安装与配置 SQL Server 2005	33
3.1 初识 SQL Server 2005	33
3.1.1 OLTP 和 OLAP	33
3.1.2 客户机/服务器(C/S)架构	34
3.1.3 SQL Server 2005 数据平台	35
3.2 安装 SQL Server 2005	37
3.2.1 安装前的准备工作	37
3.2.2 SQL Server 2005 的安装过程	39
3.2.3 验证安装是否成功	43
3.3 注册、配置和管理 SQL Server 2005 服务器	43
3.3.1 Management Studio 概述	44
3.3.2 注册服务器	45
3.3.3 配置 SQL Server 2005 服务器	47
3.3.4 管理 SQL Server 2005 服务器	54
3.4 配置 SQL Server 2005 网络环境	56
3.4.1 配置本地共享内存的 SQL Server 2005 网络环境	56
3.4.2 配置 TCP/IP 协议的 SQL Server 2005 网络环境	57
3.4.3 配置命名管道的 SQL Server 2005 网络环境	59
本章小结	60
习题	60

**第4章 创建和管理 SQL Server 2005**

数据库	62
4.1 适合数据库存储的数据类型	62
4.2 重要的数据对象	63
4.3 设计与创建数据库	66
4.3.1 规划数据库	66
4.3.2 确定数据库类型	66
4.3.3 估计数据库规模	67
4.3.4 使用 Management Studio 创建数据库	67
4.4 管理数据库	70
4.4.1 查看数据库信息	70
4.4.2 增大数据库容量	71
4.4.3 压缩数据库容量	73
4.4.4 分离和附加数据库	74
4.4.5 数据库的脱机和联机使用	76
4.4.6 数据库的并发访问	76
4.4.7 设置数据库故障恢复模型	76
4.4.8 设置数据库排序规则	77
4.4.9 重命名和删除数据库	77
本章小结	78
习题	78

第5章 创建和管理 SQL Server 2005

数据表	79
5.1 准备工作	79
5.1.1 数据表的常见类型	79
5.1.2 数据表的特点	80
5.1.3 规划表结构	80
5.2 创建和修改数据表	81
5.2.1 创建数据表	81
5.2.2 查看数据表属性	83
5.2.3 修改数据表结构	84
5.3 添加和修改表数据	85
5.3.1 添加数据表数据	85
5.3.2 查看数据表记录	86
5.3.3 插入一行记录	86
5.3.4 更新数据	87

5.3.5 删除数据	88
本章小结	89
习题	89

第6章 数据查询

6.1 选择查询	90
6.2 条件查询	93
6.2.1 带 not 的条件查询	94
6.2.2 带搜索范围的条件查询	94
6.2.3 带搜索列表的条件查询	94
6.2.4 带模糊匹配的条件查询	94
6.2.5 多重条件查询	95
6.3 连接查询	95
6.3.1 等值与非等值连接查询	95
6.3.2 自然连接查询	96
6.3.3 自连接查询	96
6.3.4 外连接查询	97
6.3.5 复合条件连接查询	98
6.4 嵌套查询	98
6.4.1 带有 in 的嵌套查询	98
6.4.2 带有比较运算符的嵌套查询	99
6.4.3 带有 any 或 all 的嵌套查询	100
6.4.4 带有 exists 的嵌套查询	100
6.5 集合查询	101
6.5.1 union(并)操作	101
6.5.2 intersect(交)操作	101
6.5.3 except(差)操作	101
6.6 数据的排序	102
6.7 数据的成组	103
6.8 利用可视化编辑器设计查询	104
6.9 嵌入式 Transact-SQL 与动态 查询技术	105
6.9.1 嵌入式 Transact-SQL 的一般 形式	105
6.9.2 动态查询技术	106
本章小结	107
习题	107

第 7 章 查询的优化——索引和视图..... 108**7.1 索引(Index)..... 108**

7.1.1 索引的结构和基本类型..... 109

7.1.2 创建索引时的考虑因素和
基本原则..... 1127.1.3 在 Management Studio 中
创建索引..... 113

7.1.4 检测与处理索引碎片..... 117

7.1.5 查看与维护索引统计信息..... 121

7.1.6 重命名和删除索引..... 122

7.2 视图(View)..... 122

7.2.1 视图的作用和基本类型..... 123

7.2.2 创建视图的基本原则..... 124

7.2.3 在 Management Studio 中
创建标准视图..... 124

7.2.4 通过视图修改数据..... 125

7.2.5 修改和删除视图..... 128

本章小结..... 129

习题..... 129

第 8 章 数据的修改与计算——事务**和函数..... 131****8.1 数据修改..... 131**

8.1.1 数据修改与事务..... 131

8.1.2 自动提交事务..... 132

8.1.3 显式事务..... 132

8.1.4 隐式事务..... 134

8.1.5 如何避免阻塞和死锁..... 135

8.2 计算和汇总数据..... 135

8.2.1 GUID(全局唯一性标识符)..... 136

8.2.2 聚合函数..... 136

8.2.3 字符串函数..... 136

8.2.4 日期与时间函数..... 137

8.2.5 数学函数..... 137

8.2.6 系统与元数据函数..... 138

本章小结..... 138

习题..... 139

第 9 章 数据输出——Reporting**Services..... 140****9.1 Reporting Services 概述..... 140**9.1.1 Reporting Services 的
需求分析..... 1409.1.2 Reporting Services 的
基本概念..... 1419.1.3 Reporting Services 的核心
组件和工具..... 143

9.1.4 报表的设计流程..... 146

9.2 创建和修改报表..... 146

9.2.1 使用报表生成器创建报表..... 146

9.2.2 使用报表设计器创建报表..... 151

9.2.3 修改报表..... 159

9.3 发布和管理报表..... 160

9.3.1 发布报表..... 160

9.3.2 使用报表管理器管理发布
报表..... 1619.3.3 使用 Management Studio
管理报表..... 169**9.4 报表的传递技术..... 171**

9.4.1 报表的订阅功能..... 171

9.4.2 共享计划..... 175

本章小结..... 176

习题..... 176

第 10 章 存储过程和触发器..... 178**10.1 存储过程..... 178**10.1.1 存储过程的特点和基本
类型..... 178

10.1.2 创建存储过程的基本原则..... 179

10.1.3 创建和调用存储过程..... 180

10.1.4 修改和删除存储过程..... 184

10.1.5 存储过程的执行过程和
重编译..... 185

10.1.6 加密存储过程..... 186

10.1.7 使用系统存储过程和
扩展存储过程..... 187



10.2 触发器	192	12.3 管理服务器安全	230
10.2.1 触发器的特点和基本类型 ..	192	12.3.1 SQL Server 2005 服务器 的登录模式	230
10.2.2 创建触发器的基本原则	193	12.3.2 创建和管理登录名	231
10.2.3 创建和激活 DML 触发器 ..	193	12.3.3 服务器角色管理	233
10.2.4 创建和激活 DDL 触发器 ..	201	12.4 管理数据库安全	235
10.2.5 使用 RAISEERROR() 命令 报告错误	202	12.4.1 数据库架构管理	235
10.2.6 禁用和重新激活触发器	203	12.4.2 数据库用户名管理	237
10.2.7 修改和删除触发器	204	12.4.3 数据库角色管理	238
本章小结	205	12.5 SQL Server 2005 的安全 层次机制	243
习题	206	本章小结	245
第 11 章 约束、同义词和自定义函数 ..	208	习题	246
11.1 约束	208	第 13 章 数据库的备份和恢复	247
11.1.1 创建 DEFAULT 约束	208	13.1 数据库备份概述	247
11.1.2 创建 CHECK 约束	210	13.1.1 数据库常见故障	247
11.1.3 创建主键约束	211	13.1.2 备份及其工作原理	248
11.1.4 创建唯一约束	212	13.1.3 数据恢复模型	248
11.1.5 创建外键约束	213	13.2 备份数据库	250
11.1.6 管理约束	213	13.2.1 创建备份设备	250
11.2 同义词	214	13.2.2 完整备份	251
11.2.1 同义词简介	215	13.2.3 差异备份	253
11.2.2 创建和管理同义词	215	13.2.4 事务日志备份	254
11.3 用户自定义函数	218	13.2.5 文件/文件组备份	255
11.3.1 含义和特点	218	13.3 恢复数据库	257
11.3.2 组成	219	13.3.1 标准恢复	257
11.3.3 类型	219	13.3.2 文件/文件组恢复	259
11.3.4 设计原则	219	13.3.3 时间点恢复	260
11.3.5 创建和调用表值函数	220	13.4 选择备份和恢复策略	261
11.3.6 创建和调用标量值函数	223	13.4.1 确定恢复目标和要求	261
11.3.7 管理用户自定义函数	224	13.4.2 纯完整备份和恢复策略	262
11.3.8 在用户自定义函数中语句 的使用限制	226	13.4.3 完整兼差异备份和恢复 策略	262
本章小结	226	13.4.4 完整兼事务日志备份和 恢复策略	262
习题	227	13.4.5 完整、差异兼事务日志 备份和恢复策略	262
第 12 章 数据库安全与角色管理	228		
12.1 SQL Server 2005 的安全机制 ..	228		
12.2 基本概念	229		

13.4.6 文件组备份和恢复策略	263
本章小结	263
习题	263

第 14 章 数据的分析与

挖掘——Analysis Services..... 265

14.1 Analysis Services 概述..... 265

14.1.1 Analysis Services 的 基本概念	265
--	-----

14.1.2 Analysis Services 的 基本对象	268
--	-----

14.1.3 Analysis Services 的 体系结构	271
--	-----

14.2 使用 Analysis Services 进行 数据分析 274 |

14.2.1 创建 Analysis Services 项目	274
---	-----

14.2.2 定义多维数据集	279
----------------------	-----

14.2.3 部署 Analysis Services 项目	281
---	-----

14.2.4 查看和分析数据	282
----------------------	-----

14.3 使用数据挖掘 285 |

14.3.1 准备 Analysis Services 数据库	286
--	-----

14.3.2 生成目标邮件方案	287
-----------------------	-----

14.3.3 生成预测方案	292
---------------------	-----

14.4 Integration Services 和 Notification Services 296 |

14.4.1 Integration Services (集成服务)	296
---	-----

14.4.2 Notification Services (通知服务)	296
--	-----

本章小结	297
------------	-----

习题	297
----------	-----

第 15 章 实训——开发“电子相册”

数据库系统 299 |

15.1 系统设计 299 |

15.1.1 需求分析	299
-------------------	-----

15.1.2 功能模块划分	300
---------------------	-----

15.1.3 流程分析	300
-------------------	-----

15.2 设计数据库 300 |

15.3 项目的创建和设计 302 |

15.3.1 创建项目并添加数据源	303
-------------------------	-----

15.3.2 用户登录模块设计	304
-----------------------	-----

15.3.3 照片列表模块设计	309
-----------------------	-----

15.3.4 照片查询模块设计	317
-----------------------	-----

15.3.5 照片浏览模块设计	321
-----------------------	-----

15.3.6 照片添加模块设计	323
-----------------------	-----

15.3.7 照片修改模块设计	325
-----------------------	-----

15.3.8 照片删除模块设计	327
-----------------------	-----

15.4 调试和运行系统 328 |

15.5 安装和部署项目 330 |

附录 A Transact-SQL 语言 333 |

A.1 Transact-SQL 语言的分类和 语法规则	333
---------------------------------------	-----

A.2 Transact-SQL 的数据对象 命名方法	334
--------------------------------------	-----

A.3 Transact-SQL 变量和常量	334
------------------------------	-----

A.4 Transact-SQL 表达式	337
----------------------------	-----

A.5 Transact-SQL 数据定义语言	337
-------------------------------	-----

A.6 Transact-SQL 数据操纵语言	338
-------------------------------	-----

A.7 Transact-SQL 控制语句	339
-----------------------------	-----

A.8 Transact-SQL 流程控制语言	340
-------------------------------	-----

A.9 Transact-SQL 函数	340
---------------------------	-----

A.10 Transact-SQL 注释	342
----------------------------	-----

附录 B 安装 AdventureWorksDW 示例

数据库 343 |

第 1 章

数据库基础知识

本章介绍数据、信息、DB、DBS、DBMS 等数据库基本概念，数据库管理技术的产生和发展，数据库系统的应用领域，常见的数据库系统，数据库系统的结构和组成，以及数据库的开发过程等内容。通过本章的学习，应完成以下学习目标：

- ☑ 掌握数据、信息、DB、DBS、DBMS 等数据库基本概念
- ☑ 了解数据库技术的演变以及数据仓库、数据挖掘等当前流行的数据库技术
- ☑ 了解数据库系统的作用和应用领域
- ☑ 了解常见的数据库系统
- ☑ 掌握数据库系统的结构和组成
- ☑ 了解数据库的开发过程

1.1 数据库概述

数据库技术是计算机技术中发展最为迅速的领域之一，已经成为人们存储数据、管理信息和共享资源的最常用、最先进技术。在系统介绍数据库之前，先简要介绍一下数据库的一些基本概念和应用。

1.1.1 基本概念

1. 数据(Data)与信息(Information)

提起数据，人们往往首先想到的是数字，如 0、1、22、267 等。其实，数字只是数据的一部分，是狭义和传统意义上的数据。在信息时代，数据是用于记录、说明人类活动的事实或事物的所有符号，包括数字、文字、图形、声音、语言等，虽然数据有多种表现形式，但它们最终都可以数字化为 0、1 代码存储到计算机中。

信息针对某一特定目的的事实或事物，是原始数据经过系统处理过的数据，是原始数据的提炼和升华，为人们决策或判断提供支持。信息与数据的不同主要体现在以下几个方面：

- 信息是经过记录、分类、组织、关联与解释的数据，就某一特定目的而言，它具有意义，而在其他方面则可能没有任何价值，信息具有相对性和时效性；
- 信息是针对某一特定目的，对原始数据进行提炼的结果；
- 信息能够为人们进行决策或判断提供支持。

综上,数据只是事实或事物的记录,没有特定的目的;而信息是针对某一问题来收集数据并进行处理,作为决策或参考的依据。所以两者的本质差别在于是否由决策问题来鉴定信息,信息处理的目的是降低不确定因素,以便选择方案。数据与信息的关系如图 1-1 所示。

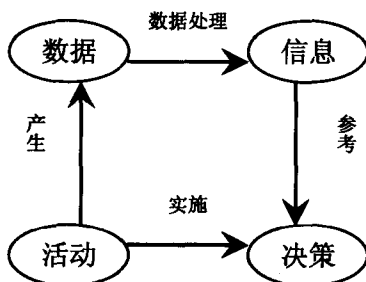


图 1-1 数据与信息的关系

当前,人们利用计算机进行数据处理(对数据进行采集、输入、存储、建库、处理、加工、转换、合并、分类、计算、统计、汇总、传送等),以便获得信息,再通过信息作为未来活动的依据。而根据决策时所制定的准则,进行所有的活动,从活动中又产生新的数据,如此不断的循环。一般来说,数据对人们并无直接的用处,因此,必须经过处理才能产生对人类较有用的信息。

2. 数据库(DB, DataBase)

数据库,简言之,就是存放数据的仓库。人们针对某个具体应用,采集并整理出该应用所需的大量数据,然后利用计算机将它们有组织、有条理地存储到计算机中。因而,从严格意义上讲,数据库是计算机中存储的大量有组织、可共享的数据集合,这些数据按照一定数据模型来组织、描述和存储,具有较小的冗余度,较高的数据独立性和可扩展性,且能为各种用户所共享。

3. 数据库系统(DBS, DataBase System)

数据库系统是指在计算机系统中引入数据库后的系统,是实现有组织地、动态地存储和管理大量关联数据,支持多用户访问的,由各种软、硬件资源构成,并有相关技术人员参与实施和管理的系统,如图 1-2 所示。

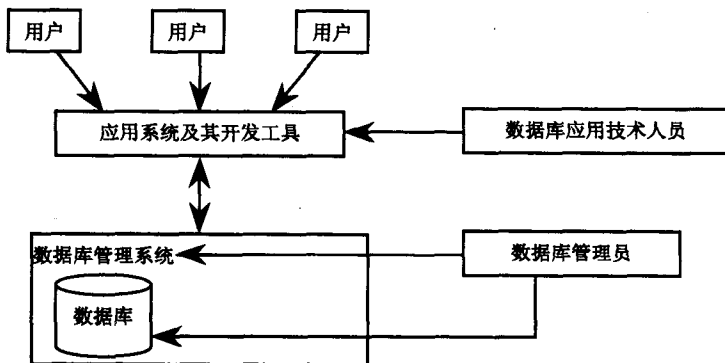


图 1-2 数据库系统

4. 数据库管理系统(DBMS, DataBase Management System)

数据库管理系统是数据库系统的重要组成部分,是位于用户和操作系统之间的一层数据库管理软件,主要解决如何科学地组织和存储数据,如何高效地获取和维护数据等问题。数据库管理系统的主要功能包含以下几个方面。

- 数据定义功能:借助 DBMS 提供的数据定义语言(DDL, Data Definition Language)定义表、索引、视图等数据对象。
- 数据操作功能:借助 DBMS 提供的数据操纵语言(DML, Data Manipulation Language)来实现对数据库的查询、录入、删除、修改等基本操作。
- 数据库的映射功能:实现外模式——模式、模式——内模式之间的映射。
- 数据库的运行管理功能:数据库在建立、运行和维护时由 DBMS 统一管理和控制,以保证数据的安全性、完整性、并发控制以及系统的备份和发生故障后的恢复。
- 数据库的建立和维护功能:DBMS 提供一些实用程序,完成数据库初始数据的输入、转换,数据库的转储、恢复、重组织、性能监视和分析等。

1.1.2 数据管理技术的产生和发展

数据管理是指对数据进行分类、组织、编码、存储、检索和维护,是数据处理的中心环节。数据管理技术的发展经历了人工管理、文件管理和数据库管理 3 个阶段。当前,数据库管理技术发展十分迅速,新出现了面向对象数据库、分布式数据库、数据仓库、数据挖掘、多媒体数据库、移动数据库等一系列新技术。

1. 人工管理阶段

早期的计算机主要用于科学计算,数据一般不需要长期保存,只是在具体用到的时候向计算机输入数据,用完就将数据撤走。例如要计算 1~100 之间所有偶数的和,就需要首先编写好程序,然后通过纸带或卡片打孔的方式,将数字输入到计算机中,计算完后数据和程序都不在计算机中保存。

当时计算机的状况是:硬件方面没有磁盘等可以直接存取的存储设备,只有纸带、卡片、磁带等;软件方面没有操作系统,更没有管理数据的软件。这在客观上也决定了计算机不能真正用于数据的管理,数据的管理主要靠人工管理的方式。

人工管理的特点主要如下:

- 数据不保存。
- 由应用程序管理数据。应用程序不仅要规定数据的逻辑结构,而且要设计物理结构,包括数据的存储结构、存取方式、输入方式等,程序员的负担很重。
- 数据不共享。数据是面向应用的,一组数据只能对应一个程序,这导致了数据的大量冗余。
- 数据不具有独立性。如果数据的逻辑结构和物理结构发生改变,就必须对相应的应用程序进行修改,这进一步加重了程序员的负担。

在人工管理阶段,数据和应用程序是一一对应的,如图 1-3 所示。

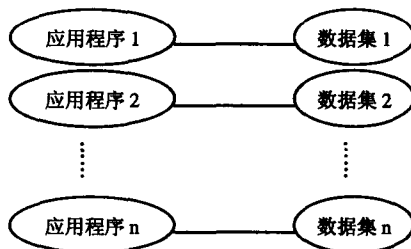


图 1-3 人工管理阶段数据和应用程序间的关系

2. 文件管理阶段

随着计算机技术的发展，硬件方面出现了磁盘、磁鼓等直接存取设备，软件方面出现了操作系统和高级程序设计语言，操作系统中的文件系统可以专门用于处理存储设备上的数据。此时，可以把数据组织成一个个文件存放在计算机中，在需要的时候只要提供文件名，计算机就能够从系统中查找到相应的文件，并将文件中存取的数据提供给用户进行处理。

文件管理的特点主要如下：

- 数据可以长期保存。
- 由文件系统来管理数据，应用程序和数据之间有了初步的独立性，程序员不必过多地考虑数据存储的物理细节。
- 数据共享性差、冗余量大。文件同样是面向应用的，一个应用程序基本上对应一个或一组文件，即使当不同的应用程序使用部分相同的数据时，也必须建立各自的文件，而不能共享相同的数据。因此数据冗余量大，同时由于数据的重复存储，也容易造成数据的不一致。
- 数据的独立性较差，缺乏灵活性。对于特定领域的应用程序，如果需求发生改变，就必须修改程序和文件的结构，而这将耗费大量的时间、人力和物力。

文件管理阶段应用程序和数据之间的关系如图 1-4 所示。

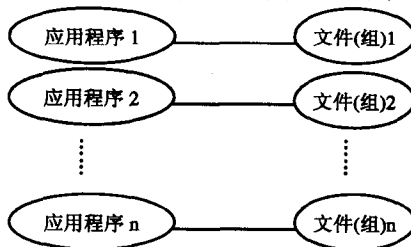


图 1-4 文件管理阶段应用程序和数据间的关系

3. 数据库管理阶段

随着计算机处理的数据量的急剧增长，对数据的管理和分析力度需求加大，同时多种应用、多种语言共享数据的要求也越来越强烈。此时已经产生了大容量磁盘，硬件价格降低了，而编制应用程序和维护软件的成本却相对增加。为解决多用户、多应用共享数据的需求，使数据为尽可能多的应用服务，数据库技术应运而生，出现了统一管理数据的专用软件系统——数据库管理系统。从文件管理到数据库管理，是数据管理技术的重大飞跃。

数据库管理的主要特点如下：

- 数据结构化。数据库系统实现了整体数据的结构化，这是数据库系统的主要特征之

一。所谓“整体”结构化,是指数据库中的数据不再仅仅针对某一个应用,而是面向整个组织,不仅数据内部是结构化的,整体也是结构化的,而且数据的存取方式也更为灵活,可以存取数据库中的某一个数据项、一组数据项、一个记录或一组记录。

- 数据共享性高,冗余量小,易扩充。在数据库系统中,数据不再面向某个具体应用,而是面向整个系统,因而数据可以为多个用户共享和使用。数据共享性高大大降低了数据的重复存储率,减小了数据冗余,极大避免了数据之间的不相容和不一致。同时,由于数据面向整个系统,是结构化的,容易添加新的应用,扩展性很好,能够满足不同应用。
- 数据独立性高。数据的独立性包括物理独立性和逻辑独立性。物理独立性是指用户的应用程序与存储在磁盘上的数据库中的数据之间的独立性;逻辑独立性是指用户的应用程序与数据库的逻辑结构之间的独立性。在数据库系统中,应用程序处理的只是数据的逻辑结构,因此当数据的物理存储发生改变时,应用程序不变;当数据的逻辑结构发生变化时,用户程序不变。应用程序与数据的独立性大大简化了应用程序的编制、修改和维护。
- 数据由 DBMS 统一管理和控制。在数据库系统中,数据是并发(concurrency)共享的, DBMS 提供数据的安全性保护、完整性检查、并发和控制等功能。

数据库系统的出现使得信息从以加工数据的程序为中心转向以数据共享的数据库为中心的新阶段。既便于数据的集中管理,又有利于应用程序的开发和维护,极大提高了数据的利用率和决策的可靠性。数据库管理阶段应用程序和数据之间的关系如图 1-5 所示。

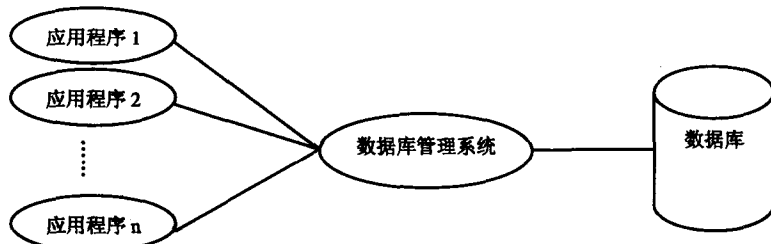


图 1-5 数据库管理阶段应用程序与数据间的关系

4. 数据仓库(Data Warehouse)与数据挖掘

数据仓库是为决策支持系统服务的,主要用于对数据进行多维和历史分析。它将各种数据源收集过来的信息进行存储,为用户提供统一的数据接口,并具备了集数据查询功能和分析工具于一体的完整结构,除了具备传统数据库管理系统的共享性、完整性和数据独立性外,还具有统一性、一致性、面向主题、历史性和只读性等特点。

数据挖掘技术是数据仓库的重要组成部分,它是指运用基于计算机的各种方法和技术,从大量商务、工程等领域的数据中提取出有用知识的整个过程,包含陈述问题和阐明假设、数据收集、数据预处理、模型评估、解释模型和得出结论共 5 个阶段,是一个反复迭代的过程。

5. 分布式数据库

分布式数据库主要是相对集中式数据库而言,是指将数据库管理系统安装在所处地理

位置不同的服务器上,用户的数据也分布在不同的服务器上。但从逻辑上,这些用户的数据仍然是一个整体,构成一个逻辑数据库。在分布式数据库中,本地服务器的用户不仅可以访问本地服务器上数据库中的数据,也可以访问异地服务器上数据库中的数据。

6. 面向对象数据库

面向对象数据库是面向对象技术和数据库技术的结合,在面向对象数据库系统中,文本、图像、视频以及空间数据等都可以存储在数据库中,这与传统的数据库系统不同。在传统数据库中数据仅仅是数据,不包含任何层次结构信息,而面向对象数据库可以将数据视为对象,包含属性和方法,并能体现数据间的继承关系。面向对象数据库具有技术新、效率高的优点,被认为是替代传统数据库系统的理想数据库体系结构。

7. 多媒体数据库

多媒体数据库技术主要用于处理图形、图像、声音、视频等多媒体海量数据,涉及图像处理技术、声音处理技术、视频处理技术、动画处理技术等方面。

8. 移动数据库

随着笔记本电脑和无线网络的发展和应用,移动计算变得越来越重要。移动计算的特点是计算机不在固定的位置,没有固定的网络地址,这样使得数据处理变得十分复杂,为解决这一问题,越来越多的数据会放在由用户管理的计算机上,成为移动数据库。移动数据库技术的核心技术包括移动计算模型、路由技术、查询技术、广播数据技术等。

1.1.3 数据库系统的应用

数据处理涉及到社会生活的方方面面,数据库应用系统已经成为人们生活不可或缺的组成部分,下面介绍一些常见的数据库应用系统。

1. 超市销售系统

目前,在超市购物已经成为人们日常生活的一个重要活动,而数据库技术是超市运营成功的重要技术基础。在超市的销售系统中,主要的信息包括商品信息、供应商信息和销售信息。

当在超市进行购物时,收银员通过使用条形码阅读器来扫描每件货物的条形码而得到货物的信息,而这实际上是从数据库中查询该商品信息,获得商品价格的应用程序,然后将查询到的商品价格显示在收银机上,并求得金额打出清单。

超市销售系统还提供了商品库存情况和销售情况的查询,以确定进货的品种、时间和数量。该系统为超市的经营决策者提供实际数据,进而满足客户的需求并提高管理水平。

2. 铁路售票系统

铁路售票系统的应用大大提高了铁路服务和管理水平,乘客可在售票点通过指定时间、始发站、到达地名来查询可供选择的车次、座位、票价等信息。在确认乘客的选择后,售票员打印出车票,同时在存储车票信息的数据库中将该车票的记录锁住,以防止重复出售。

3. 银行业务系统

银行业务系统是最早使用数据库技术的系统之一。对银行业务的操作可通过各地的银行网点和 ATM(Automated Teller Machine)来进行。该系统为顾客提供账户信息的查询和存