



工业和信息化部普通高等教育“十二五”规划教材

21世纪高等教育计算机规划教材

COMPUTER

数据库原理与设计

(第2版)

Database Principle and Design

- 杨海霞 主编
- 相洁 南志红 副主编
- 张永奎 审

- 用具体软件的操作实例诠释数据库理论
- 用建构主义的知识结构图理清知识脉络
- 用任务驱动法分散难点尽享探究的乐趣



NLIC2970937174



 人民邮电出版社
POSTS & TELECOM PRESS



工业和信息化普通高等教育“十二五”规划教材
21世纪高等教育计算机规划教材

COMPUTER

数据库原理与设计

(第2版)

Database Principle and Design

- 杨海霞 主编
- 相洁 南志红 副主编
- 张永奎 审



NLIC2970937174



人民邮电出版社

北京

图书在版编目 (CIP) 数据

数据库原理与设计 / 杨海霞主编. -- 2版. -- 北京:
人民邮电出版社, 2013.9
21世纪高等教育计算机规划教材
ISBN 978-7-115-32565-5

I. ①数… II. ①杨… III. ①关系数据库系统—高等
学校—教材 IV. ①TP311.138

中国版本图书馆CIP数据核字(2013)第174586号

内 容 提 要

本书系统地讲述了数据库的基本理论和开发技术。全书分7章, 主要内容包括: 数据库系统的组成、表及查询、SQL、关系规范化理论、数据库的设计方法、SQL 程序设计、管理数据库以及数据库设计实例等, 每章后均附有习题。

本书语言通俗易懂, 既有理论的概括与探讨, 又有实际应用中经验方法的总结, 本书可作为高等学校计算机、信息及相关专业的数据库教材, 也适合从事数据库开发的人员参考阅读。

-
- ◆ 主 编 杨海霞
 - 副 主 编 相 洁 南志红
 - 审 张永奎
 - 责任编辑 邹文波
 - 责任印制 彭志环 杨林杰
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号
邮编 100061 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京天宇星印刷厂印刷
 - ◆ 开本: 787×1092 1/16
印张: 16.25
字数: 425千字
-

定价: 36.00 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223

反盗版热线: (010)67171154

广告经营许可证: 京崇工商广字第 0021 号

第2版前言

数据库技术一直是计算机科学技术中发展最快和应用最广的技术之一，但长期以来，数据库作为计算机及其相关专业的一门课程，其教学一直处于以理论为主的状态，致使教师感觉抽象难教，学生认为枯燥难学。究其原因，主要是由于单纯地讲解理论，不能把数据库的理论知识依托于一个具体的平台来就事论理。只有把数据库原理与实际的应用开发有机结合，增强学生实践动手能力，才能培养出真正满足社会需求的数据库技术人才。正是在此动机推动下，笔者根据多年的教学及项目开发的实际经验，有感而发，写成此书。

本教材于2002年开始以讲义的形式在校内使用，2006年与该教材配套的《数据库实验指导》被评为国家级“十一五”规划教材，2007年该套书正式出版，经过6年的实际使用，在征求教师建议和听取学生意见的基础上，考虑到数据库技术的发展与进步，现已完成《数据库原理与设计（第2版）》的修订。

《数据库原理与设计（第2版）》在第1版的基础上，做了如下补充与修改。

(1) 更新了环境。数据库管理系统更新为SQL Server 2008版；数据库设计工具采用Power Designer 15.2版。

(2) 增加了内容。在每一章前面都加了“知识结构”和“任务驱动”两个栏目。知识结构图让读者从理论上理清知识脉络，任务驱动通过设计完成具体的任务让读者享受探究的乐趣。

(3) 调整了编写思路。采用即学即用的思想，每个知识点先做理论介绍，接着就是具体操作案例，使理论教学与实验教学环环紧扣，浑然一体，增强了教材的实用性。

(4) 整合了内容。为了与《数据库实验指导（第2版）》（第1版被评为普通高等教育“十一五”国家级规划教材）一书配套，采用循序渐进，步步深入的方法，从数据库→表→表查询→设计数据库→管理数据库→应用系统开发，由点及面，加强了教材的系统性。

(5) 删减了部分内容。考虑到与《数据库实验指导（第2版）》的配套使用，为节省篇幅，删去了“数据库建模工具”一节，但在实验指导书中仍保留该节的详细内容。

本教材分7章来完成数据库课程的教学内容，具体教学安排情况见下表。其中，第7章及配套实验教材的训练篇和实例篇，教师可根据实际教学具体情况予以调整。

章	理论要求	能力要求	与《数据库实验指导》配合使用情况
1	熟悉数据库系统的基本概念和基本理论	能够正确创建及维护数据库	基础篇——实验1、2、3、6
2	掌握关系模型的基本概念及理论规范	能够正确创建表及维护表	基础篇——实验4、5、6、7
3	掌握关系代数理论及各种查询方法	能够灵活运用SQL语句实现各类查询,通过视图与索引优化查询	基础篇——实验8、9、10
4	熟练掌握数据库设计从需求分析到运行维护的各个过程	能正确描述数据流图与数据字典,能使用PowerDesigner软件设计CDM模型、PDM模型、创建脚本	设计篇——设计1、2、3、4、5
5	熟悉SQL程序的基本成分,掌握SQL编程规范与程序调试方法	能正确使用SQL语句编写游标、自定义函数、存储过程、触发器等批处理程序	提高篇——实验2、3、4、5、6 设计篇——设计6、7、8、9
6	了解数据库安全性、完整性、数据恢复、并发控制等数据库保护技术,掌握其使用方法	能够创建用户、定义角色并授权;能够进行数据库安全认证;能够正确运用DBMS提供的各种完整性约束机制;能够对数据库进行复制与转储等	提高篇——实验1、7、8、9
7	了解应用程序设计并实现的整个流程	能够正确连接数据库,配置应用程序开发环境,对数据库的开发与应用有整体认识与感悟,加强对所学知识的驾驭能力	提高篇——实验10 训练篇——综合训练1、2、3、4 实例篇

与其他同类教材相比,本书具有以下特点。

(1) 理论依托于具体软件,以感性认识固化理性认识。

本书以SQL Server 2008作为数据库管理系统,以Power Designer为设计工具,把数据库的基本概念、基本理论和基本技术及数据库的设计贯穿于一个学籍管理实例中,通过详尽的分析,使学生感性认识数据库,从而加深对数据库理论知识的理解。

(2) 循序渐进,步步深入,形成由点及面到体的体系结构。

本书从数据库的基本知识开始,首先讲解每个数据库对象的基本功能和基本操作,然后应用前面的知识进行学籍管理数据库的设计,从中领会数据库的设计技巧,最后介绍数据库的访问技术和维护技术,完成学籍管理实例,形成了由点及面到体的体系结构。

(3) 注重应用,丰富实例,精编实验指导教程。

为加强学生的实际操作能力,精心编写了配套的实验指导书,通过基础篇、提高篇、设计篇、训练篇和实例篇五部分的训练,由点及面,层次分明,使理论教学与实验教学环环相扣,既方便教师教学,又有助于学生学习,具有很强的实用性。

(4) 设置知识结构图和任务驱动,强化知识脉络。

用知识结构图的形式呈现每章知识点,将孤立的知识点连成线,引导学生将所学知识系统化,从而理清知识脉络;通过任务驱动,把每个知识点又分解到一个具体软件的应用操作中,让学生通过自主练习,享受探究知识的乐趣。

下面是作者的几点建议,仅供安排与组织教学时参考。

(1) 应该先开设程序设计类课程和简单的网页制作课程,尤其是面向对象程序设计语言。建议将《数据库原理与设计》(第2版)和《数据库实验指导》(第2版)配合使用。

(2) 建议该课程的教学环境为多媒体环境,系统为 Windows XP、SQL Server 2008 以上版本。

(3) 有条件的学校可安排课程设计,学生可以分小组练习实训篇的内容,并将此成绩作为本课程成绩的一部分。

本书语言通俗易懂,既有理论的概括与探讨,又有实际的经验方法的总结,本书可作为高等学校计算机、信息及相关专业的数据库课程教科书。为便于教师进行教学,书中的源程序代码和教学用的 PPT 文件,请登录人民邮电出版社教学服务与资源网:<http://www.ptpedu.com.cn>,从网上下载。

本书由杨海霞、相洁、南志红、任菊香、殷鹏、杨潞霞、赵宇兰编写。其中,第1章由杨潞霞编写,第2章由赵宇兰编写,第3章由殷鹏编写,第4章由杨海霞编写,第5章、附录A和附录B由相洁编写,第6章由任菊香编写,第7章由南志红编写。全书由杨海霞统稿,张永奎审阅。在编写过程中得到了徐仲安教授、杨继平教授、石冰教授、樊东燕教授、王建珍教授的支持与帮助,同时,本书的出版还得到了山西财经大学、太原理工大学等高校教师的大力支持,在此一并致以衷心的感谢。

由于作者水平有限,书中难免存在疏漏、欠妥之处,敬请读者批评指正。

编者

2013年7月

目 录

第 1 章 认识数据库	1
1.1 数据管理技术的发展.....	2
1.1.1 数据处理技术.....	2
1.1.2 数据库技术的发展阶段.....	2
1.1.3 数据库技术的新进展.....	4
1.2 数据库系统.....	6
1.2.1 数据库系统的组成.....	6
1.2.2 数据库系统的模式结构.....	6
1.2.3 数据库系统的体系结构.....	9
1.3 SQL Server 数据库管理系统.....	11
1.3.1 SQL Server 2008 简介.....	11
1.3.2 配置 SQL Server 2008.....	13
1.3.3 SQL Server 2008 主要管理工具.....	15
1.3.4 SQL 概述.....	18
1.4 数据库的基本操作.....	22
1.4.1 SQL Sever 数据库的存储结构.....	22
1.4.2 SQL Server 系统数据库.....	24
1.4.3 创建数据库.....	25
1.4.4 修改数据库.....	28
1.4.5 删除数据库.....	30
1.4.6 数据备份与还原.....	31
1.5 数据库的其他操作.....	35
1.5.1 分离和附加数据库.....	35
1.5.2 脱机和联机数据库.....	37
1.5.3 收缩数据库.....	37
思考与练习.....	38
第 2 章 管理与维护表	39
2.1 关系模型.....	39
2.1.1 关系模型的基本术语.....	39
2.1.2 关系的键.....	40
2.1.3 关系的完整性.....	41
2.2 表结构与数据类型.....	43
2.2.1 数据类型.....	43
2.2.2 约束.....	47
2.2.3 表结构设计.....	49
2.3 表的创建、修改和删除.....	50
2.3.1 表的创建.....	50
2.3.2 表的修改.....	51
2.3.3 表的删除.....	52
2.4 表的数据操作.....	53
2.4.1 插入数据.....	53
2.4.2 修改数据.....	54
2.4.3 删除数据.....	54
2.4.4 更新操作与数据库的一致性.....	55
2.5 实例——学籍管理数据库中表的创建.....	57
思考与练习.....	59
第 3 章 查询、视图和索引	61
3.1 关系代数.....	61
3.1.1 传统的集合运算.....	64
3.1.2 专门的关系运算.....	65
3.1.3 关系代数查询综合举例.....	70
3.2 简单查询.....	71
3.2.1 选择表中的若干列.....	71
3.2.2 选择表中的若干元组.....	74
3.2.3 对查询结果排序.....	79
3.3 连接查询.....	80
3.3.1 等值与非等值连接查询.....	80
3.3.2 自身连接.....	83
3.3.3 外连接.....	83
3.4 分组及汇总查询.....	85
3.4.1 聚合函数.....	85
3.4.2 分组聚合.....	86
3.5 嵌套查询.....	88
3.5.1 使用 IN 的子查询.....	89
3.5.2 使用比较运算符的子查询.....	91
3.5.3 使用 ANY 或 ALL 的子查询.....	92
3.5.4 使用存在量词 EXISTS 的子查询.....	93
3.6 集合查询.....	95
3.7 SQL 查询的一般格式.....	96
3.8 视图.....	97
3.8.1 创建视图.....	97
3.8.2 管理视图.....	100
3.8.3 查询视图.....	101
3.8.4 更新视图.....	102
3.8.5 视图的作用.....	105
3.9 索引.....	106

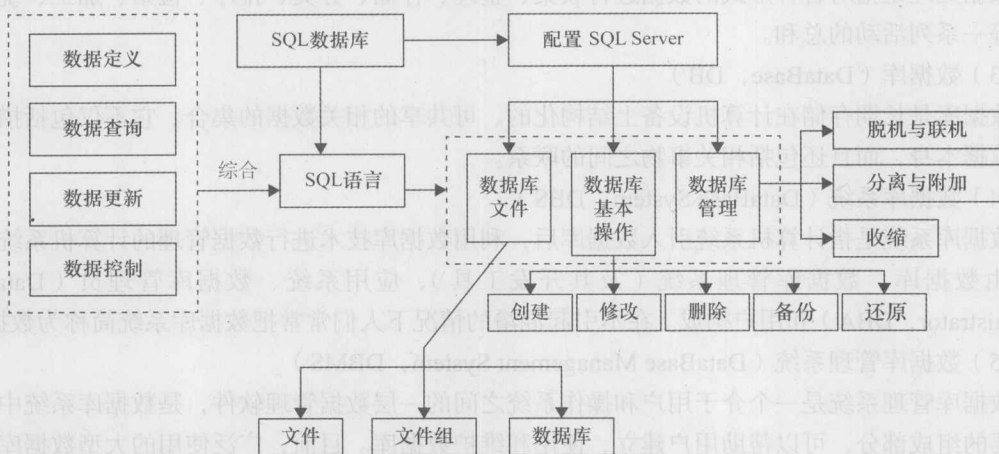
3.9.1 索引的概念.....	106	5.1.1 批处理.....	159
3.9.2 索引的分类.....	107	5.1.2 脚本.....	160
3.9.3 创建索引.....	107	5.2 SQL 程序设计基础.....	160
3.9.4 管理索引.....	109	5.2.1 SQL 程序基本成分.....	161
思考与练习.....	110	5.2.2 SQL 程序编写规范.....	163
第4章 数据库设计.....	112	5.3 流程控制语句.....	164
4.1 数据库设计概述.....	112	5.3.1 语句块: BEGIN...END.....	165
4.1.1 数据库设计任务及方法.....	112	5.3.2 条件执行: IF...ELSE 语句.....	165
4.1.2 数据库设计阶段划分.....	114	5.3.3 多分支 CASE 表达式.....	166
4.2 需求分析.....	114	5.3.4 循环: WHILE 语句.....	167
4.2.1 需求分析方法.....	115	5.3.5 非条件执行: GOTO 语句.....	168
4.2.2 数据流图与数据字典.....	116	5.3.6 调度执行: WAITFOR.....	168
4.2.3 实例——学籍管理需求分析.....	119	5.4 游标.....	168
4.3 概念结构设计.....	123	5.4.1 游标.....	168
4.3.1 概念结构设计的特点和方法.....	123	5.4.2 声明游标.....	169
4.3.2 概念模型.....	123	5.4.3 打开游标.....	170
4.3.3 数据抽象与局部视图设计.....	125	5.4.4 读取数据.....	170
4.3.4 视图的集成.....	126	5.4.5 关闭游标.....	172
4.3.5 实例——学籍管理概念 结构设计.....	127	5.4.6 释放游标.....	172
4.4 逻辑结构设计.....	129	5.5 SQL 程序的调试与错误处理.....	174
4.4.1 E-R 图向关系数据模型的转换.....	129	5.5.1 SQL 程序的错误类型.....	174
4.4.2 函数依赖.....	133	5.5.2 SQL 程序的错误处理.....	174
4.4.3 范式.....	136	5.6 SQL 程序实例.....	175
4.4.4 关系模式的规范化.....	145	5.6.1 自定义函数.....	175
4.4.5 设计用户子模式.....	148	5.6.2 存储过程.....	176
4.4.6 实例——学籍管理 逻辑结构设计.....	149	5.6.3 触发器.....	179
4.5 物理结构设计.....	150	思考与练习.....	180
4.5.1 分析影响数据库物理 设计的因素.....	151	第6章 管理数据库.....	182
4.5.2 关系模式存取方法选择.....	152	6.1 数据库的安全管理.....	183
4.5.3 确定系统配置.....	153	6.1.1 安全性概述.....	183
4.5.4 评价物理结构.....	153	6.1.2 数据库管理系统的 身份识别机制.....	183
4.6 数据库实施与维护.....	153	6.1.3 SQL Server 的用户和角色管理.....	184
4.6.1 创建数据库.....	153	6.1.4 SQL Server 的安全认证模式.....	192
4.6.2 组织数据入库.....	154	6.2 数据库的完整性管理.....	193
4.6.3 数据库试运行.....	154	6.2.1 完整性约束.....	193
4.6.4 数据库运行与维护.....	154	6.2.2 完整性控制.....	195
4.7 数据库建库工具.....	155	6.2.3 SQL Server 数据完整性的 实现方法.....	197
思考与练习.....	157	6.3 SQL Server 数据库恢复.....	200
第5章 SQL 程序设计与开发.....	158	6.3.1 恢复的原理.....	201
5.1 批处理与脚本.....	159	6.3.2 恢复的实现.....	202
		6.3.3 数据库镜像.....	207
		6.4 并发控制.....	208

6.4.1 并发控制概述.....	209	7.4.3 ADO.NET 访问数据库的步骤	221
6.4.2 并发调度的可串行性.....	211	7.5 系统实现	222
6.4.3 封锁协议.....	212	7.5.1 系统主框架布局设计.....	223
6.4.4 死锁和活锁.....	214	7.5.2 系统页面模板设计.....	223
思考与练习	215	7.5.3 设置数据库连接信息.....	223
第 7 章 应用实例——		7.5.4 系统公共类.....	224
学籍管理系统.....	217	7.5.5 系统部分存储过程、自定义函数、 视图.....	230
7.1 系统开发环境	217	7.5.6 系统部分页面.....	233
7.2 系统功能结构设计	218	附录 A SQL Server 支持的运算符	
7.3 数据库设计	218	与函数	241
7.4 ADO.NET 数据库访问		附录 B 学籍管理数据表结构	245
SQL Server 技术.....	219	参考文献	248
7.4.1 ADO.NET 的数据处理类	219		
7.4.2 ADO.NET 的数据处理类的 常用属性和方法.....	220		

第 1 章

认识数据库

► 知识结构·理清知识脉络



► 任务驱动·享受探究乐趣

任务 1 理解数据库相关概念及数据库的作用。

任务 2 了解数据管理技术的发展阶段及各阶段的特点。

任务 3 理解并掌握数据库体系的三级模式结构、两级映像及其作用。

任务 4 理解 SQL 语言并掌握 SQL 语言的语句结构。

任务 5 创建学籍管理数据库，并会对其进行修改、删除、备份与还原操作。

任务 6 使用 SSMS 管理数据库，包括数据库的附加与分离、脱机与联机、收缩等操作。

数据库技术是现代信息科学与技术的重要组成部分，是计算机数据处理与信息管理的核心。随着数据库技术的发展，其应用已深入到工农业生产、商业、金融、行政管理、科学研究、工程技术、国防和军事等各个领域，成为信息管理、办公自动化、计算机辅助决策、计算机辅助设计、智能信息处理等计算机应用系统的核心部分。作为计算机软件学科中的一个十分活跃而重要的独立分支，数据库技术已形成了一整套的理论与技术体系。

本章将从数据管理技术的发展着手，比较详细地介绍数据库系统的基本知识。

1.1 数据管理技术的发展

1.1.1 数据处理技术

1. 数据库的基础概念

(1) 数据 (Data)

数据就是对客观事物反映的一种物理符号,可以是数字、文字、图形、图像、声音、语言和视频等。这些多种形式的数 据都需经过数字化后才能存入计算机。

(2) 数据处理 (Data Processing)

数据处理是指对各种形式的数 据进行收集、整理、存储、分类、排序、检索、加工、统计和传输等一系列活动的总和。

(3) 数据库 (DataBase, DB)

数据库是长期存储在计算机设备上结构化的、可共享的相关数据的集合。它不仅包括描述事物的数据本身,而且还包括相关事物之间的联系。

(4) 数据库系统 (DataBase System, DBS)

数据库系统是指计算机系统引入数据库后,利用数据库技术进行数据管理的计算机系统。它一般由数据库、数据库管理系统 (及其开发工具)、应用系统、数据库管理员 (DataBase Administrator, DBA) 和用户构成。在不引起混淆的情况下人们常常把数据库系统简称为数据库。

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

数据库管理系统是一个介于用户和操作系统之间的一层数据管理软件,是数据库系统中的一个重要的组成部分,可以帮助用户建立、使用和维护数据库。目前,广泛使用的大型数据库管理系统有 Oracle、Sybase 等,小型数据库管理系统有 SQL Server、Visual FoxPro、Access 等。

2. 数据处理技术

数据处理技术一般包括两部分内容:数据管理和数据处理。

数据管理是指对数据的收集、整理、组织、存储、维护、检索、传送等操作过程。数据处理是指对数据进行收集、存储、加工、传播等一系列活动的总和,其基本目的是从大量的、杂乱无章的、难以理解的数据中抽取并导出对于某些特定的应用来说是有价值的、有意义的数 据,借以作为决策的依据。可见,数据管理是任何数据处理业务中必不可少的部分,是数据处理的中心问题。

从根本上说,数据处理是借助于计算机科学地保存和管理大量复杂的数据,使人们充分而高效地利用这些宝贵的信息资源。

1.1.2 数据库技术的发展阶段

在数据的处理过程中,对于不同的业务部门其对数据的加工、计算、打印报表等操作可能有不同的内容需求。数据库技术正是应数据管理任务的需求而产生的。数据库技术研究的问题就是如何科学地组织和存储数据,如何高效地获取和处理数据。

随着计算机技术及应用的发展,人们对数据管理技术也不断提出了更高的要求,其先后经历了从低级到高级的 3 个阶段,即人工管理阶段、文件系统阶段和数据库系统阶段。

1. 人工管理阶段

20世纪50年代中期前,从计算机硬件看,外存储器只有磁带、卡片和纸带,没有磁盘;从软件看,没有操作系统,该阶段的计算机系统还没有支持管理数据的软件,主要应用于科学计算。人工管理阶段的特点如下。

- (1) 数据不长期保存在计算机里,用完就删除。
- (2) 使用应用程序管理数据。
- (3) 数据不共享,一个程序对应一组数据,存在大量的重复数据。
- (4) 数据不具有独立性,应用程序和数据之间的关系如图1-1所示。

2. 文件系统阶段

20世纪50年代后期至60年代中期,计算机不仅用于科学计算,而且开始用于数据处理工作。从硬件来看,外存储器有了磁盘、磁鼓等直接存取的存储设备,软件方面有了操作系统,在操作系统中已经有了专门的管理数据的软件——文件系统。这一阶段的特点如下。

- (1) 数据能够以文件的形式长期保留在外存上,可以反复对其进行查询、修改、插入和删除等操作。
- (2) 由文件系统管理数据的输入和输出,程序和数据能够分离,有了一定的独立性,应用程序和数据之间的关系如图1-2所示。
- (3) 数据能够共享,但数据文件之间彼此孤立,数据之间的联系弱,数据大量冗余。
- (4) 数据没有独立性。

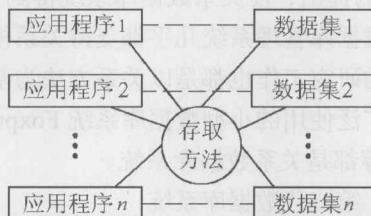


图 1-1 人工管理阶段应用程序与数据之间的对应关系 图 1-2 文件系统阶段应用程序与数据之间的对应关系

3. 数据库系统阶段

20世纪60年代后期,计算机硬件和软件技术有了较大发展,出现了大容量直接存取设备。计算机在企业管理中的广泛应用,对数据管理提出了更高的要求。为了实现数据的统一管理,达到数据共享的目的,产生了数据库技术。其满足了多用户、多应用共享数据的需求,与人工管理阶段、文件系统阶段相比,这一阶段的特点如下。

- (1) 数据结构化。
- (2) 数据的共享性高,冗余度低、易于扩充。
- (3) 数据的独立性高。
- (4) 数据由 DBMS 统一管理和控制,应用程序和数据之间的关系如图1-3所示。

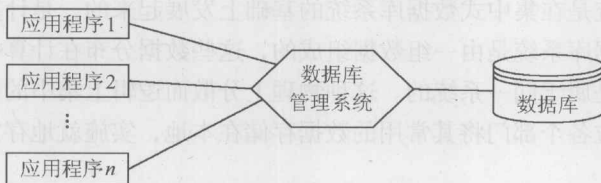


图 1-3 数据库系统阶段应用程序与数据的对应关系

⑤ 为用户提供友好接口。

1.1.3 数据库技术的新进展

数据库技术是计算机科学技术中发展最快的分支之一。20世纪70年代以来,数据库系统从第一代的网状和层次数据库系统发展到第二代的关系数据库系统。目前现代数据库系统正向着第三代的面向对象数据库系统发展,并与网络技术、分布式计算、面向对象程序设计技术相结合。

1. 数据库技术的发展

(1) 第一代数据库系统

第一代数据库系统为网状和层次数据库系统。1969年IBM公司开发了基于层次模型的信息管理系统(Information Management System, IMS)。20世纪60年代末至70年代初,美国数据系统语言委员会(Conference on Data Systems Languages, CODASYL)下属的数据库任务组(DataBase Task Group, DBTG)提出了若干报告(DBTG报告)。该报告确定并建立了网状数据库系统的许多概念、方法和技术。正是基于上述报告,Cullinet Software开发了基于网状模型的产品IDMS(Information Data Management System),IMS和IDMS这两个产品推动了网状和层次数据库系统的发展。

(2) 第二代数据库系统

第二代数据库系统为关系数据库系统。1970年IBM公司研究员E.F.Codd发表的关于关系模型的论文推动了关系数据库系统的研究和开发。尤其是关系数据库标准语言——结构化查询语言(SQL)的提出,使关系数据库系统得到了广泛的应用。20世纪80年代以来,数据库软件厂商新推出的数据库管理系统几乎都支持关系模型,非关系系统的产品也都加上了关系接口。数据库领域当前的研究工作也都是以关系方法为基础。关系数据库已成为目前应用最广泛的数据库系统,如现在广泛使用的小型数据库系统Foxpro、Access,大型数据库系统Oracle、SQL Server、Informix、Sybase等都是关系数据库系统。

(3) 第三代数据库系统

第三代数据库系统为支持面向对象(Object Oriented, OO)数据模型的数据库系统。20世纪90年代,随着数据库系统应用的范围进一步扩大,数据库所处理对象的复杂性和灵活性对数据库系统提出了越来越高的要求。例如多媒体数据、CAD数据、图形图像数据需要更好的数据模型来表达,以便存储、管理和维护。面向对象数据模型能完整地描述现实世界复杂的数据结构,并具有封装性和继承性等面向对象技术的优点。正是在这种形势下,产生了面向对象的数据库系统。

2. 数据库技术的新发展

随着计算机应用领域的不断拓展,数据库的应用越来越广泛,数据库技术的研究取得了重大突破。各种科学技术与数据库技术有机结合,使数据库领域中新内容、新应用、新技术层出不穷,形成了各种新型的数据库系统。如面向对象数据库系统、分布式数据库系统、知识数据库系统、模糊数据库系统、并行数据库系统、多媒体数据库系统等。

(1) 分布式数据库系统(Distributed DataBase System, DDBS)

分布式数据库系统是在集中式数据库系统的基础上发展起来的,是计算机技术和网络技术结合的产物。分布式数据库系统是由一组数据组成的,这些数据分布在计算机网络的不同节点(亦称场地)上,逻辑上是属于同一系统的。这种物理上分散而逻辑上集中的数据库系统适用于单位分散的部门,允许单位各个部门将其常用的数据存储在本地区,实施就地存放本地使用,从而提高响应速度,降低费用。

(2) 并行数据库系统

并行数据库系统是并行技术与数据库技术结合的产物,是在并行机上运行的具有并行处理能力的数据库系统,是新一代高性能的数据库系统,具有数据库操作的时间并行性和空间并行性。并行数据库技术起源于20世纪70年代,主要是研究利用硬件实现关系代数操作的并行化。20世纪80年代后,逐步转向通用并行机的研究。20世纪90年代以后,迅猛发展的存储技术、网络技术和通用并行计算机技术,为并行数据库技术的研究奠定了基础。最近几年,以通用并行计算系统为基础的并行数据库系统研究出现了高潮。很多研究者认为,在具有数百个甚至上千个处理机的并行计算机上建立并行数据库系统是可行的。

(3) 主动数据库系统

主动数据库系统是相对于传统数据库的被动性而言的。它是在传统数据库的基础上,结合人工智能技术研制和开发的主动数据库。主动数据库除了完成一切传统数据库的服务外,还具有各种主动服务功能的数据库系统。

(4) 模糊数据库系统

模糊性是客观世界的一个重要属性,传统的数据库系统描述和处理的是精确的或确定的客观事物,不能描述许多模糊性和不具完全性的事情。

模糊数据库能够存储模糊数据,而且数据结构、数据联系也是模糊的,数据上的运算和操作、对数据的约束(包括完整性和安全性)、数据的一致性、无冗余性的定义等都是模糊的。

模糊数据库研究主要有两方面,首先是如何在数据库中存放模糊数据;其次是定义各种运算,建立模糊数据上的函数。模糊数的表示主要有模糊区间数、模糊中心数、模糊集合数和隶属函数等。由于理论和实现技术上的困难,模糊数据库技术近年来发展不是很理想,但已在模式识别、过程控制、案情侦破、医疗诊断、工程设计、营养咨询、公共服务以及专家系统等领域得到较好的应用,显示了广阔的应用前景。

(5) 空间数据库系统

随着地理信息系统(GIS)的发展,传统的数据模型已经不足以表示空间数据,空间数据库技术逐渐取代原来数据库技术,存储与管理GIS应用中的空间数据。

空间数据库的研究始于20世纪70年代的地图制图与遥感图像处理领域,其目的是为了有效地利用卫星遥感资源迅速绘制出各种经济专题地图。由于传统的关系数据库在空间数据的表示、存储、管理、检索上存在许多缺陷,从而形成了空间数据库这一数据库研究领域。随着计算机技术的发展,目前空间数据库的应用范围已经扩展到了机器人、计算机视觉、图像识别、环境保护、地理信息处理等领域。

(6) 多媒体数据库系统

随着计算机辅助设计、计算机辅助制造等计算机应用技术的不断发展,许多复杂的应用对象中涉及大量的图形、图像、声音、动画等多媒体数据类型。由于数据需求量不断增加,数据量越来越大,数据操作的困难也在不断增加,因此人们有必要寻找到一种较完善的解决办法,在这种形势下,多媒体数据库作为一种强大的多媒体数据处理技术出现了。

多媒体数据库系统应提供比传统数据库管理系统更强的适合非格式化数据查询的搜索功能,允许对非格式化数据做整体和部分搜索,允许通过范围、知识和其他描述符的确定值和模糊值搜索各种媒体数据,允许同时搜索多个数据库中的数据,允许通过对非格式化数据的分析建立图示等索引来搜索数据,允许通过举例查询(Query by Example)和通过主题描述查询使复杂查询简单化。

多媒体数据库系统目前仍然还是处于研究、发展阶段,目前还没有成熟的多媒体数据库管理系统推出。

1.2 数据库系统

数据库系统是对数据进行存储、管理、处理和维护的计算机软件系统。这类系统在硬件上要求有较大的外存容量和 I/O 处理速度等硬件资源, 在软件上要求有数据库管理系统 (DBMS) 的支持, 以及 DBMS 能够正常运行的操作系统。

1.2.1 数据库系统的组成

数据库系统是指带有数据库并利用数据库技术进行数据管理的计算机系统。一般由 5 部分构成: 数据库、数据库管理系统 (及其开发工具)、应用系统、数据库管理员 (DataBase Administrator, DBA) 和用户。

数据库系统可以用图 1-4 表示。其中, 应用系统实际是指用户应用程序, 是由系统开发人员使用开发工具, 利用数据库系统资源开发出来的, 面向某一类实际应用问题的应用软件系统, 例如图书管理系统、人事管理系统等; 数据库管理员是负责管理和维护数据库服务器的人员; 用户则是在数据库系统中具有使用数据库应用系统软件权限的部门或人员。

这里需要注意的是: 数据库、数据库管理系统和数据库系统是 3 个不同的概念。数据库强调的是相互关联的数据; 数据库管理系统强调的是管理数据库的系统软件; 而数据库系统强调的是基于数据库技术的计算机系统, 它包含数据库、数据库管理系统、操作系统、应用软件、硬件系统、数据库管理员、最终用户和开发人员等。

1.2.2 数据库系统的模式结构

1. 数据库系统的三级模式结构

从数据库管理角度看, 数据库系统通常采用三级模式结构。这三级模式结构是指数据库系统是由外模式、模式和内模式三级构成的, 如图 1-5 所示。

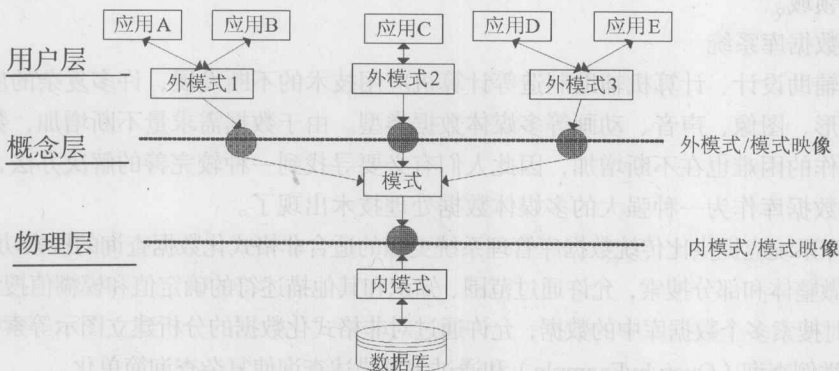


图 1-5 数据库系统的模式结构

(1) 外模式

外模式也称子模式或用户模式，它是对数据库用户（包括应用程序员和最终用户）看见和使用的局部数据的逻辑结构和特征的描述，是数据库用户的数据视图，是与某一应用有关的数据的逻辑表示。

外模式通常是模式的子集。一个数据库通常可以有多个外模式。由于它是各个用户的数据视图，如果不同的用户在应用需求、看待数据的方式、对数据保密的要求等方面存在差异，则他们的外模式描述就是不同的。即使是模式中同一数据，在外模式中的结构、类型、长度、保密级别等都可以不同。另一方面，同一外模式也可以为某一用户的多个应用系统所使用，但一个应用程序只能使用一个外模式。

外模式是保证数据库安全性的一个有力的措施。每个用户只能看见和访问所对应的外模式中的数据，数据库中的其余数据对他们来说是不可见的。

(2) 模式

模式也称逻辑模式，是数据库中全体数据的逻辑结构和特征的描述，是所有用户的公共数据视图。模式是相对稳定的，模式的一个具体的值称为模式的一个实例（Instance）。同一个模式可以有多个实例。实例反映的是数据库某一时刻的状态。模式是数据库系统模式结构的中间层，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、所使用的应用开发工具及高级程序设计语言无关。

模式实际上是数据库中的数据在逻辑上的视图。一个数据库只有一个模式。数据库模式以某一种数据模型为基础，统一综合地考虑了所有用户的需求，并将这些需求有机地结合成一个逻辑整体。定义模式时不仅要定义数据的逻辑结构（例如，数据记录由哪些数据项构成，数据项的名字、类型、取值范围等），而且要定义与数据有关的安全性、完整性要求，定义这些数据之间的联系。

(3) 内模式

内模式也称存储模式，它是数据物理结构和存储结构的描述，是数据在数据库内部的表示方式。例如，记录的存储方式是顺序存储、按照 B 树结构存储还是按散列（hash）方法存储，索引按照什么方式组织；数据是否压缩存储、是否加密，数据的存储记录结构有何规定等。

一个数据库只有一个内模式。内模式是最低层，反映了数据在计算机物理结构中的实际存储形式。

2. 数据库的二级映像功能与数据独立性

数据库系统的三级模式是数据在 3 个层次（用户层、概念层和物理层）的抽象，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。而为了能够在内部实现这三个抽象层次的联系和转换，数据库系统在这三级模式之间提供了两层映像：外模式/模式映像和模式/内模式映像。正是这两层映像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/模式映像

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/模式映像，它定义了该外模式与模式之间的对应关系。这些映像定义通常包含在各自外模式的描述中。当模式改变时（例如，增加新的数据类型、新的数据项、新的关系等），由数据库管理员对各个外模式/模式映像做相应改变，可以使外模式保持不变，从而不必修改应用程序，保证了数据的逻辑独立性。

(2) 模式/内模式映像

数据库中只有一个模式,也只有一个内模式,所以模式/内模式映像唯一的,它定义了数据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,该映像定义通常包含在模式描述中。当数据库的存储结构发生改变时(例如,采用了更先进的存储结构),由数据库管理员对模式/内模式映像做相应改变,可以使模式保持不变,从而保证了数据的物理独立性。

3. 设计三级模式结构的关键问题

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。

数据库的外模式面向具体的应用程序,它定义在逻辑模式之上,但独立于存储模式和存储设备。当应用需求发生较大变化,相应外模式不能满足其视图要求时,该外模式就得做相应改动,所以设计外模式时应充分考虑到应用的扩充性。特定的应用程序是在外模式描述的数据结构上编制的,它依赖于特定的外模式,与数据库的模式和存储结构独立。不同的应用程序有时可以共用同一个外模式。数据库的二级映像保证了数据库外模式的稳定性,从而从底层保证了应用程序的稳定性,除非应用需求本身发生变化,否则应用程序一般不需要修改。

4. 数据库系统所需人员与三级模式结构的匹配

一个完整的数据库系统,除了软、硬件配置以外,还需要各级各类人员进行开发、管理和维护,这些人员分别是:数据库管理员(Data Base Administrator, DBA)、系统分析员、数据库设计人员、应用程序员和最终用户。不同的人员涉及不同的数据抽象级别,具有不同的数据视图,如图1-6所示,其各自的职责如下。

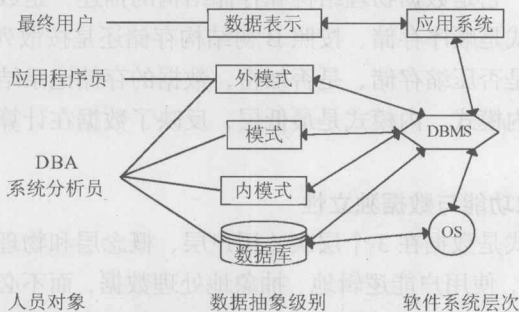


图 1-6 各种人员的数据视图

(1) 数据库管理员

在数据库系统环境下,有两类共享资源。一类是数据库,另一类是数据库管理系统软件。因此需要有专门的管理机构来监督和管理数据库系统。数据库管理员则是这个机构的一个(组)人员,负责全面管理和控制数据库系统。

在数据运行过程中,不断会有数据的插入、删除、修改,时间一长,会影响系统的性能。因此,数据库管理员要定期对数据库进行重组,以提高系统的性能。

当用户的需求增加和改变时,数据库管理员还要对数据库进行较大的改造,例如,修改部分