



普通高等教育“十二五”规划教材

数据库技术与应用 (SQL Server 2008版)

主 编 王小玲 安剑奇
副主编 严 晖 周肆清
主 审 施荣华



中国水利水电出版社
www.waterpub.com.cn

普通高等教育“十二五”规划教材

数据库技术与应用

(SQL Server 2008 版)

主 编 王小玲 安剑奇

副主编 严 晖 周肆清

主 审 施荣华

内 容 提 要

本书根据教育部高等计算机基础课程教学指导委员会 2011 年 10 月出版的《高等学校计算机基础核心课程教学实施方案》(新白皮书)中关于“数据库技术及应用”课程实施方案的精神,并以教指委提出的“普及计算机文化,训练计算思维,培养信息应用能力”为总体目标进行编写的。

本书以 SQL Server 2008 为蓝本,以 Visual Basic 6.0 和 Delphi 7.0 作为开发工具,从数据库技术与应用系统开发的角度介绍数据库系统的基本概念及应用。全书共 10 章,内容包括数据库技术概论、数据库的管理与使用、数据表的管理与维护、数据查询、索引与视图、存储过程与触发器、数据库维护、数据库的安全管理、数据库系统开发工具 Visual Basic、Delphi 的数据访问方法。

本书既可作为高等院校数据库技术与应用课程的教材,又可供社会各类计算机应用人员阅读参考。

本书配套有《数据库技术与应用(SQL Server 2008 版)实践教程》,并提供电子教案,读者可以从中国水利水电出版社网站和万水书苑上下载,网址为: <http://www.waterpub.com.cn/softdown/>和 <http://www.wsbookshow.com>。

图书在版编目(CIP)数据

数据库技术与应用:SQL Server 2008版/王小玲,
安剑奇主编. — 北京:中国水利水电出版社,2014.4
普通高等教育“十二五”规划教材
ISBN 978-7-5170-1892-6

I. ①数… II. ①王… ②安… III. ①关系数据库系
统一高等学校—教材 IV. ①TP311.138

中国版本图书馆CIP数据核字(2014)第070303号

策划编辑:雷顺加 责任编辑:李 炎 加工编辑:刘晶平 封面设计:李 佳

书 名	普通高等教育“十二五”规划教材 数据库技术与应用(SQL Server 2008 版)
作 者	主 编 王小玲 安剑奇 副主编 严 晖 周肆清 主 审 施荣华
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路1号D座 100038) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn
经 售	电话: (010) 68367658 (发行部)、82562819 (万水) 北京科水图书销售中心(零售) 电话: (010) 88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	北京万水电子信息有限公司
印 刷	三河市铭浩彩色印装有限公司
规 格	184mm×260mm 16 开本 18.75 印张 474 千字
版 次	2014 年 4 月第 1 版 2014 年 4 月第 1 次印刷
印 数	0001—3000 册
定 价	35.00 元

凡购买我社图书,如有缺页、倒页、脱页的,本社发行部负责调换
版权所有·侵权必究

前 言

本书根据教育部高等计算机基础课程教学指导委员会 2011 年 10 月出版的《高等学校计算机基础核心课程教学实施方案》(新白皮书)中关于“数据库技术及应用”课程实施方案的精神,并以教指委提出的“普及计算机文化,训练计算思维,培养信息应用能力”为总体目标进行编写。

全书将知识传授与能力培养融为一体,以应用为目的,以操作案例为驱动,构建完整的数据库知识体系。用一个具有代表性的实例数据库——“学生信息数据库”贯穿全书,并设计了 100 多个在工作和学习中遇到的数据库问题,指导读者循序渐进地寻找答案。每章配有精心设计的思考题,引导读者在解决问题的过程中加深对知识的理解、巩固,在实际运用中拓展思维能力。

本书以流行的 SQL Server 2008 数据库管理系统作为实验平台,介绍 SQL Server 2008 的主要功能和数据库的基本操作方法,其中 SQL 语法均用实例验证,大部分例题配有图片说明。系统开发平台使用 Windows 7 环境下的 Visual Basic 6.0 和 Delphi 7.0,书中全部例题均在系统环境中运行通过,图片均为 SQL Server 2008 系统运行界面、Visual Basic 6.0 和 Delphi 7.0 界面截图,直观、清晰,方便读者对照学习。

为了方便教学和读者上机操作练习,作者还组织编写了《数据库技术与应用(SQL Server 2008 版)实践教程》一书,作为与本书配套的实验和课程设计教材。另外,还有与本书配套的教学课件,供教师教学参考。

本书由王小玲、安剑奇任主编,严晖、周肆清任副主编,施荣华任主审。全书由王小玲、安剑奇负责统稿和整理。另外,参加编写工作的还有刘卫国、杨长兴、童键、田琪、邵自然、温国海、孙岱、奎晓燕、韩华、董密、蒋朝辉等。在本书编写过程中,得到了作者所在学校信息科学与工程学院相关领导和教学管理人员、计算机基础教学实验中心全体教师和自动化系部分教师的大力支持和指导,在此表示衷心的感谢!

由于本书的编写人员都是奋战在本课程教学一线的教师,教学、教改和科研任务繁重,书中不当或错误之处在所难免,恳请广大读者批评指正,读者可通过邮箱 wxling@csu.edu.cn 与作者联系。

编 者

2014 年 2 月

目 录

前言

第1章 数据库技术概论	1	2.2.2 使用 T-SQL 创建数据库	49
1.1 数据库技术的产生与发展	1	2.3 数据库的修改	54
1.2 数据库系统	4	2.3.1 使用对象资源管理器修改数据库	54
1.2.1 数据库系统的组成	4	2.3.2 使用 T-SQL 修改数据库	56
1.2.2 数据库的结构体系	5	2.4 数据库的删除	59
1.2.3 数据库系统的特点	6	2.4.1 使用图形界面方式删除数据库	59
1.3 数据模型	7	2.4.2 使用 T-SQL 删除数据库	61
1.3.1 数据模型的组成要素	7	2.5 数据库的分离和附加	61
1.3.2 数据抽象的过程	8	2.5.1 数据库的分离	61
1.3.3 概念模型	9	2.5.2 数据库的附加	63
1.3.4 逻辑模型	11	2.6 数据库的扩大和收缩	66
1.4 关系数据库	12	2.6.1 数据库的扩大	66
1.4.1 关系数据库的基本概念	12	2.6.2 数据库的收缩	67
1.4.2 关系运算	14	习题 2	70
1.4.3 关系的完整性约束	17	第3章 数据表的管理与维护	72
1.4.4 关系数据库设计实例	18	3.1 数据表的创建和管理	72
1.5 SQL Server 2008 数据库概述	19	3.1.1 使用对象资源管理器创建数据表	72
1.5.1 SQL Server 的初步认识	19	3.1.2 使用 T-SQL 创建数据表	74
1.5.2 SQL Server 2008 的服务器组件	23	3.1.3 使用对象资源管理器对数据表进行管理	77
1.5.3 SQL Server 2008 的常用管理工具	24	3.1.4 使用 T-SQL 对数据表进行管理	79
1.5.4 SQL Server 数据类型	29	3.2 表数据的管理	81
1.6 Transact-SQL 语言简介	32	3.2.1 使用对象资源管理器管理表数据	81
1.6.1 SQL 与 Transact-SQL	32	3.2.2 使用 T-SQL 管理表数据	85
1.6.2 运算符与表达式	33	3.3 数据库完整性管理	87
1.6.3 语句块和注释	38	3.3.1 数据库完整性概述	87
1.6.4 流程控制语句	38	3.3.2 数据库完整性的类型	88
习题 1	41	3.3.3 使用对象资源管理器实现数据库完整性的设置	90
第2章 数据库的管理与使用	43	习题 3	94
2.1 SQL Server 数据库的存储结构	43	第4章 数据库查询	96
2.1.1 逻辑存储结构	43	4.1 查询概述	96
2.1.2 物理存储结构	45	4.1.1 图形界面的菜单方式	96
2.2 数据库的创建	46		
2.2.1 使用对象资源管理器创建数据库	47		

4.1.2 查询语句 SELECT	97	第 7 章 数据库维护	156
4.2 基本查询	98	7.1 数据备份和还原	156
4.2.1 简单查询	98	7.1.1 数据备份	156
4.2.2 条件查询	101	7.1.2 数据还原	158
4.2.3 查询结果处理	105	7.1.3 数据备份和还原操作	159
4.3 嵌套查询	109	7.2 导入导出数据	166
4.3.1 单值嵌套查询	109	7.2.1 导入数据表	166
4.3.2 多值嵌套查询	110	7.2.2 导入其他数据源的数据	171
4.4 连接查询	111	7.2.3 导出 SQL Server 数据表	177
4.4.1 自连接	111	7.3 生成与执行 SQL 脚本	179
4.4.2 内连接	112	7.3.1 将数据库生成 SQL 脚本	179
4.4.3 外连接	114	7.3.2 将数据表生成 SQL 脚本	179
4.4.4 交叉连接	116	7.3.3 执行 SQL 脚本	181
习题 4	117	习题 7	182
第 5 章 索引与视图	120	第 8 章 数据库安全的管理	184
5.1 索引	120	8.1 SQL Server 2008 的安全机制	184
5.1.1 索引的概念	120	8.1.1 身份验证	184
5.1.2 索引的分类	121	8.1.2 身份验证模式的设置	186
5.1.3 索引的管理	122	8.2 SQL Server 安全管理	187
5.2 视图	125	8.2.1 登录管理	187
5.2.1 视图的概念	125	8.2.2 数据库用户管理	192
5.2.2 视图的创建	126	8.2.3 角色管理	193
5.2.3 视图的查询	129	8.2.4 权限管理	199
5.2.4 视图的修改	129	习题 8	202
5.2.5 视图的删除	130	第 9 章 数据库系统开发工具 VB	204
习题 5	131	9.1 数据库系统开发工具概述	204
第 6 章 存储过程与触发器	133	9.2 VB 概述	205
6.1 存储过程概述	133	9.2.1 VB 6.0 集成开发环境	205
6.1.1 存储过程的特点和类型	133	9.2.2 创建简单的 VB 应用程序	206
6.1.2 存储过程的创建和执行	134	9.2.3 VB 程序的特点	207
6.1.3 存储过程参数和执行状态	139	9.3 VB 语言基础	208
6.1.4 存储过程的查看和修改	143	9.3.1 基本数据类型	209
6.1.5 存储过程的删除	145	9.3.2 变量和常量	209
6.2 触发器概述	146	9.3.3 运算符与表达式	212
6.2.1 触发器的特点和类型	146	9.3.4 数组与自定义类型	213
6.2.2 触发器的创建	147	9.4 程序控制结构	216
6.2.3 触发器的查看和修改	151	9.4.1 选择结构	216
6.2.4 触发器的删除	153	9.4.2 循环控制结构	218
习题 6	153	9.5 控件	222

9.5.1 标签	223	9.9.2 数据查询	258
9.5.2 文本框	225	9.10 数据库应用系统开发	261
9.5.3 图片框与图像框	228	习题 9	263
9.5.4 菜单	230	第 10 章 Delphi 的数据访问方法	265
9.5.5 单选按钮与复选框	231	10.1 Delphi 7.0 的 BDE 组件	265
9.5.6 列表框与组合框	233	10.1.1 BDE 组件	266
9.5.7 滚动条与定时器	236	10.1.2 TDatabase 组件	266
9.6 过程	238	10.1.3 TTable 组件	268
9.6.1 子过程	238	10.1.4 TQuery 组件	272
9.6.2 函数过程	241	10.2 Delphi 7.0 的 ADO 组件	273
9.6.3 变量的作用域和生存期	242	10.2.1 ADO 组件	274
9.7 数据访问方法	244	10.2.2 TADOConnection 组件	274
9.7.1 VB 访问的数据库类型	244	10.2.3 TADOCommand 组件	276
9.7.2 VB 访问数据的接口	245	10.2.4 TADODataSet 和 TADOQuery 组件	277
9.7.3 VB 数据库的访问过程	245	10.3 数据库应用系统开发案例	277
9.8 使用数据控件访问数据库	246	习题 10	285
9.8.1 连接数据库	246	附录 1 SQL Server 2008 常用函数	287
9.8.2 数据绑定	249	附录 2 Visual Basic 常用函数	289
9.9 数据库操作	253	附录 3 Visual Basic 常用方法	291
9.9.1 数据库编辑操作	253	参考文献	293

第1章 数据库技术概论



- **了解：**数据与数据处理的概念、数据库技术的产生背景与发展概况、SQL Server 2008 的特点、常用管理工具、Transact SQL 语言的功能。
- **理解：**数据库系统的组成与特点、数据独立性的概念、数据模型的概念。
- **掌握：**关系模型的基本知识、关系数据库的设计方法、SQL Server 的数据类型及各种运算符和语句。

1.1 数据库技术的产生与发展

数据库技术是一门研究如何存储、使用和管理数据的技术，是计算机数据管理技术的最新发展阶段，它能把大量的数据按照一定的结构存储起来，在数据库管理系统的集中管理下实现数据共享。

人类在长期的社会生产实践中会产生大量数据，如何对数据进行分类、组织、存储、检索和维护成为迫切的实际需要，在计算机成为数据处理的工具之后，数据处理现代化成为可能。数据库系统的核心任务是数据管理，但并不是一开始就有数据库技术，它是随着数据管理技术的不断发展而逐步产生与发展的。

1. 人工管理阶段

20 世纪 50 年代中期以前，计算机主要应用于科学计算，虽然此时也有数据管理的问题，但这时的数据管理是以人工管理方式进行的。在硬件方面，外存储器只有磁带、卡片和纸带等，没有磁盘等可以直接存取的外存储器。在软件方面，只有汇编语言，没有操作系统，没有对数据进行管理的软件。数据处理方式基本上是批处理。在此阶段，数据管理的特点如下：

(1) 数据不保存。此阶段处理的数据量较少，一般不需要将数据长期保存，只是在计算时将数据随程序一起输入，计算完后将结果输出，而数据和程序则一起从内存中被释放。若再计算，则需重新输入数据和程序。

(2) 由应用程序管理数据。系统没有专门的软件对数据进行管理，数据需要由应用程序自行管理。每个应用程序不仅要规定数据的逻辑结构，而且要设计数据的存储结构及输入/输出方法等，程序设计任务繁重。

(3) 数据有冗余，无法实现共享。程序与数据是一个整体，一个程序中的数据无法被其他程序使用，因此程序与程序之间存在大量的重复数据，数据无法实现共享。

(4) 数据对程序不具有独立性。由于程序对数据的依赖性，数据的逻辑结构或存储结构一旦有所改变，则必须修改相应的程序，这就进一步加重了程序设计的负担。

2. 文件管理阶段

20 世纪 50 年代后期至 60 年代后期, 计算机开始大量用于数据管理。硬件上出现了可以直接存取的大容量外存储器, 如磁盘、磁鼓等, 这为计算机数据管理提供了物质基础。软件上出现了高级语言和操作系统。操作系统中的文件系统专门用于管理数据, 这又为数据管理提供了技术支持。数据处理方式不仅有批处理, 而且有联机实时处理。

数据处理应用程序利用操作系统的文件管理功能将相关数据按一定的规则构成文件, 通过文件系统对文件中的数据进行存取和管理, 实现数据的文件管理方式。其特点如下:

(1) 数据可以长期保存。文件系统为程序和数据之间提供了一个公共接口, 使应用程序采用统一的存取方法来存取和操作数据。数据可以组织成文件, 能够长期保存、反复使用。

(2) 数据对程序有一定的独立性。程序和数据不再是一个整体, 而是通过文件系统把数据组织成一个独立的数据文件, 由文件系统对数据的存取进行管理, 程序员只需通过文件名来访问数据文件, 不必过多考虑数据的物理存储细节, 因此程序员可集中精力进行算法设计, 从而大大减少了程序维护的工作量。

文件管理使计算机在数据管理方面有了长足的进步。时至今日, 文件系统仍是一般高级语言普遍采用的数据管理方式。

3. 数据库管理阶段

20 世纪 60 年代后期, 计算机用于数据管理的规模更加庞大, 数据量急剧增加, 数据共享性要求更加强烈。同时, 计算机硬件价格下降, 软件价格上升, 编制和维护软件所需成本相对增加, 其中维护成本更高。这些都成为数据管理技术在文件管理的基础上发展到数据库管理的原动力。

数据库 (Data Base, DB) 是按照一定的组织方式存储起来的、相互关联的数据集合。在数据库管理阶段, 由一种叫做数据库管理系统 (Data Base Management System, DBMS) 的系统软件来对数据进行统一的控制和管理, 把所有应用程序中使用的相关数据汇集起来, 按统一的数据模型存储在数据库中, 为各个应用程序所使用。在应用程序和数据库之间保持较高的独立性, 数据具有完整性、一致性和安全性高等特点, 并且具有充分的共享性, 有效地减少了数据冗余。

4. 新型数据库系统

数据库技术的发展先后经历了层次数据库、网状数据库和关系数据库阶段。层次数据库和网状数据库可以看做是第一代数据库系统, 关系数据库可以看做是第二代数据库系统。自 20 世纪 70 年代提出关系数据模型和关系数据库后, 数据库技术得到了蓬勃发展, 应用也越来越广泛。但随着应用的不断深入, 占主导地位的关系数据库系统已不能满足新应用领域的需求。例如, 在实际应用中, 除了需要处理数字、字符数据的简单应用外, 还需要存储并检索复杂的复合数据 (如集合、数组、结构)、多媒体数据、计算机辅助设计绘制的工程图纸和 GIS (地理信息系统) 提供的空间数据等。对于这些复杂数据, 关系数据库无法实现对它们的管理。正是这些实际应用中涌现出的问题促使数据库技术不断向前发展, 出现了许多不同类型的新型数据库系统。下面简要地做一些介绍。

(1) 分布式数据库系统。

分布式数据库系统 (Distributed Data Base System, DDBS) 是数据库技术与计算机网络技术、分布式处理技术相结合的产物。分布式数据库系统是将系统中的数据地理上分布在计算机网络的不同节点, 但逻辑上属于一个整体的数据库系统, 它不同于将数据存储在服务

户共享存取的网络数据库系统, 分布式数据库系统不仅能支持局部应用(访问本地数据库), 而且能支持全局应用(访问异地数据库)。分布式数据库系统主要应用于航空、铁路、旅游订票系统, 银行通存通兑系统, 水陆空联运系统, 跨国公司管理系统, 连锁配送管理系统等。

(2) 面向对象数据库系统。

面向对象数据库系统(Object-Oriented Data Base System, OODBS)是将面向对象的模型、方法和机制与先进的数据库技术有机结合而形成的新型数据库系统。它从关系模型中脱离出来, 强调在数据库框架中发展类型、数据抽象、继承和持久性。它的基本设计思想是: 把面向对象语言向数据库方向扩展, 使应用程序能够存取并处理对象; 扩展数据库系统, 使其具有面向对象的特征, 提供一种综合的语义数据建模概念集, 以便对现实世界中复杂应用的实体和联系建模。因此, 面向对象数据库系统首先是一个数据库系统, 具备数据库系统的基本功能; 其次是一个面向对象的系统, 针对面向对象的程序设计语言的永久性对象存储管理而设计, 充分支持完整的面向对象概念和机制。面向对象数据库系统对一些特定应用领域(如 CAD 等), 能较好地满足其应用需求。

(3) 多媒体数据库系统。

多媒体数据库系统(Multi-media Data Base System, MDBS)是数据库技术与多媒体技术相结合的产物。随着信息技术的发展, 数据库应用从传统的企业信息管理扩展到计算机辅助设计(CAD)、计算机辅助制造(CAM)、办公自动化(OA)、人工智能(AI)等多个应用领域。这些领域中处理的数据不仅包括传统的数字、字符等格式化数据, 还包括大量多媒体形式的非格式化数据, 如图形、图像、声音等。这种能存储和管理多媒体数据的数据库称为多媒体数据库。多媒体数据库系统主要应用于军事、医学病例管理、航天测控、商标管理、地理信息系统、数字图书馆和期刊出版系统等。

(4) 数据仓库技术。

随着信息技术的高速发展, 数据库应用的规模、范围和深度不断扩大, 一般的事务处理已不能满足应用的需要, 企业需要在大量数据基础上的决策支持, 数据仓库(Data Warehouse, DW)技术的兴起满足了这一需求。数据仓库作为决策支持系统(Decision Support System, DSS)的有效解决方案, 涉及3个方面的技术内容: 数据仓库技术、联机分析处理(On-Line Analytical Processing, OLAP)技术和数据挖掘(Data Mining, DM)技术。

数据仓库、OLAP 和数据挖掘是作为3种独立的数据处理技术出现的。数据仓库用于数据的存储和组织, OLAP 集中于数据的分析, 数据挖掘则致力于知识的自动发现。它们都可以分别应用到信息系统的设计和实现中, 以提高相应部分的处理能力。但是, 由于这3种技术内在的联系性和互补性, 将它们结合起来即是一种新的 DSS 架构。这一架构以数据库中的大量数据为基础, 系统则由数据驱动。数据仓库技术应用遍及通信、零售业、金融及制造业等领域。

(5) 内存数据库系统。

内存数据库(Main Memory Data Base, MMDB)系统是实时系统和数据库系统的有机结合。它抛弃了磁盘数据管理的传统方式, 基于全部数据都在内存中这一前提重新设计了体系结构, 并且在数据缓存、快速算法、并行操作方面也进行了相应的改进, 所以数据处理速度比传统数据库的数据处理速度要快很多, 一般都在10倍以上。内存数据库的最大特点是其“主拷贝”或“工作版本”常驻内存, 即活动事务只与实时内存数据库的内存拷贝打交道。内存数据库系统目前广泛应用于航空、军事、电信、电力及工业控制等领域。

1.2 数据库系统

数据库系统 (Data Base System, DBS) 是指基于数据库的计算机应用系统。和一般的应用系统相比, 数据库系统有其自身的特点, 它涉及一些既相互联系又有区别的基本概念。

1.2.1 数据库系统的组成

数据库系统是一个计算机应用系统, 它是把有关计算机硬件、软件、数据和人员组合起来为用户提供信息服务的系统。因此, 数据库系统是由计算机系统、数据库及其描述机制、数据库管理系统和有关人员组成的具有高度组织性的整体。

1. 计算机硬件

计算机硬件系统是数据库系统的物质基础, 是存储数据库及运行数据库管理系统的硬件资源, 主要包括计算机主机、存储设备、输入/输出设备及计算机网络环境。

2. 计算机软件

数据库系统中的软件包括操作系统、数据库管理系统、数据库应用系统等。

数据库管理系统 (DBMS) 是数据库系统的核心软件之一, 它提供数据定义、数据操纵、数据库管理、数据库建立和维护及通信等功能。DBMS 提供对数据库中的数据资源进行统一管理和控制的功能, 将用户、应用程序与数据库数据相互隔离, 是数据库系统的核心, 其功能的强弱是衡量数据库系统性能优劣的主要指标。DBMS 必须运行在相应的系统平台上, 有操作系统和相关系统软件的支持。

DBMS 功能的强弱随系统而异, 大系统功能较强、较全, 小系统功能较弱、较少。目前较流行的数据库管理系统有 Access、Visual FoxPro、SQL Server、Oracle 和 Sybase 等。

数据库应用系统是指系统开发人员利用数据库系统资源开发出来的、面向某一类实际应用的应用软件系统。从实现技术角度而言, 它是以数据库技术为基础的计算机应用系统。

3. 数据库

数据库 (DB) 是指数据库系统中按照一定的方式组织的、存储在外部存储设备上的、能为多个用户共享的、与应用程序相互独立的相关数据集合。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。

数据库中的数据往往不像文件系统那样只面向某一项特定应用, 而是面向多种应用, 可以被多个用户、多个应用程序共享。其数据结构独立于使用数据的程序, 对数据的增加、删除、修改和检索由 DBMS 进行统一管理和控制, 用户对数据库进行的各种操作都是由数据库管理系统实现的。

4. 数据库系统的有关人员

数据库系统的有关人员主要有 3 类: 最终用户、数据库应用系统开发人员和数据库管理员 (Data Base Administrator, DBA)。最终用户指通过应用系统的用户界面使用数据库的人员, 他们一般对数据库知识了解不多。数据库应用系统开发人员包括系统分析员、系统设计员和程序员。系统分析员负责应用系统的分析, 他们和用户、数据库管理员相配合, 参与系统分析; 系统设计员负责应用系统设计和数据库设计; 程序员则根据设计要求进行编码。数据库管理员是数据管理机构的一组人员, 他们负责对整个数据库系统进行总体控制和维护, 以保证数据库系统的正常运行。

综上所述,数据库中包含的数据是存储在存储介质上的数据文件的集合;每个用户均可使用其中的数据,不同用户使用的数据可以重叠,同一组数据可以为多个用户共享;DBMS为用户提供对数据的存储组织、操作管理等功能;用户通过数据库管理系统和应用程序实现对数据库系统的操作与应用。

1.2.2 数据库的结构体系

为了有效地组织、管理数据,提高数据库的逻辑独立性和物理独立性,人们为数据库设计了一个严谨的结构体系,数据库领域公认的标准结构是三级模式及二级映射。三级模式包括外模式、概念模式和内模式;二级映射是概念模式/内模式的映射和外模式/概念模式的映射。这种三级模式与二级映射结构构成了数据库的结构体系,如图1.1所示。

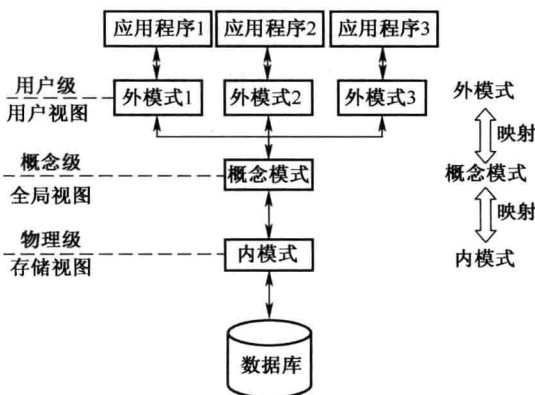


图 1.1 数据库的三级模式与二级映射

1. 数据库的三级模式

美国国家标准协会(American National Standards Institute, ANSI)的数据库管理系统研究小组于1978年提出了标准化的建议,将数据库结构体系分为三级:面向用户或应用程序员的用户级、面向建立和维护数据库人员的概念级、面向系统程序员的物理级。用户级对应外模式,概念级对应概念模式,物理级对应内模式,使不同级别的用户对数据库形成不同的视图。视图是指观察、认识和理解数据的范围、角度和方法,是数据库在用户眼中的反映。很显然,不同层次(级别)用户所看到的数据库是不相同的。

(1) 概念模式。概念模式又称逻辑模式,或简称模式,对应于概念级。它是由数据库设计者综合所有用户的数据,按照统一的观点构造的全局逻辑结构,是对数据库中全部数据的逻辑结构和特征的总体描述,是所有用户的公共数据视图(全局视图)。它由数据库系统提供的数据库定义语言(Data Definition Language, DDL)来描述、定义,体现并反映了数据库系统的整体观。

(2) 外模式。外模式又称子模式或用户模式,对应于用户级。它是某个或某几个用户所看到的数据库的数据视图,是与某一应用有关的数据的逻辑表示。外模式是从概念模式导出的一个子集,包含概念模式中允许特定用户使用的那部分数据。用户可以通过外模式定义语言(外模式DDL)来描述、定义对应于用户的数据记录(用户视图),也可以利用数据库操纵语言(Data Manipulation Language, DML)对这些数据记录进行操作。外模式反映了数据库的用户观。

(3) 内模式。内模式又称存储模式或物理模式, 对应于物理级。它是数据库中全体数据的内部表示或底层描述, 是数据库最低一级的逻辑描述, 它描述了数据在存储介质上的存储方式和物理结构, 对应着实际存储在外存储介质上的数据库。内模式由内模式定义语言 (内模式 DDL) 来描述、定义, 它是数据库的存储观。

在一个数据库系统中只有唯一的数据库, 因而作为定义、描述数据库存储结构的内模式和定义、描述数据库逻辑结构的模式也是唯一的, 但建立在数据库系统之上的应用则是非常广泛、多样的, 所以对应的外模式不是唯一的, 也不可能唯一。

2. 三级模式间的二级映射

数据库的三级模式是数据在 3 个级别 (层次) 上的抽象, 使用户能够逻辑地、抽象地处理数据, 而不必关心数据在计算机中的物理表示和存储方式, 把数据的具体组织交给 DBMS 去完成。为了实现这 3 个抽象级别的联系和转换, DBMS 在三级模式之间提供了二级映射, 正是这二级映射保证了数据库中的数据具有较高的物理独立性和逻辑独立性。

(1) 概念模式/内模式的映射。数据库中的概念模式和内模式都只有一个, 所以概念模式/内模式的映射是唯一的。它确定了数据的全局逻辑结构与存储结构之间的对应关系。存储结构变化时, 概念模式/内模式的映射也应有相应的变化, 使其概念模式仍保持不变, 即把存储结构变化的影响限制在概念模式之下, 这使数据的存储结构和存储方法独立于应用程序, 通过映射功能保证数据存储结构的变化不影响数据的全局逻辑结构的改变, 从而不必修改应用程序, 即确保了数据的物理独立性。

(2) 外模式/概念模式的映射。数据库中的同一概念模式可以有多个外模式, 对于每一个外模式, 都存在一个外模式/概念模式的映射, 用于定义该外模式和概念模式之间的对应关系。当概念模式发生改变时, 如增加新的属性或改变属性的数据类型等, 只需对外模式/概念模式的映射做相应的修改, 而外模式 (即数据的局部逻辑结构) 保持不变。由于应用程序是依据数据的局部逻辑结构编写的, 所以应用程序不必修改, 从而保证了数据与程序间的逻辑独立性。

1.2.3 数据库系统的特点

数据库系统的出现是计算机数据管理技术的重大进步, 它克服了文件系统的缺陷, 提供了对数据更高级、更有效的管理。

1. 数据结构化

在文件系统中, 文件的记录内部是有结构的。例如, 学生数据文件的每个记录是由学号、姓名、性别、出生年月、籍贯、简历等数据项组成的。但这种结构只适用于特定的应用, 对其他应用并不适用。

在数据库系统中, 每一个数据库都是为某一应用领域服务的。例如, 学校信息管理涉及多个方面的应用, 包括对学生的学籍管理、课程管理、学生成绩管理等, 还包括教工的人事管理、教学管理、科研管理、住房管理和工资管理等, 这些应用彼此之间都有着密切的联系。因此, 在数据库系统中不仅要考虑某个应用的数据结构, 还要考虑整个组织 (即多个应用) 的数据结构。这种数据组织方式使数据结构化了, 这就要求在描述数据时不仅要描述数据本身, 还要描述数据之间的联系。而在文件系统中, 尽管其记录内部已有了某些结构, 但记录之间没有联系。数据库系统实现整体数据的结构化, 这是数据库的主要特点之一, 也是数据库系统与文件系统的本质区别。

2. 数据共享性高、冗余度低

数据共享是指多个用户或应用程序可以访问同一个数据库中的数据,而且 DBMS 提供并发和协调机制,保证在多个应用程序同时访问、存取和操作数据库数据时不产生任何冲突,从而保证数据不遭到破坏。

数据冗余既浪费存储空间,又容易产生数据的不一致。在文件系统中,由于每个应用程序都有自己的数据文件,所以数据存在着大量的重复。

数据库从全局观念来组织和存储数据,数据已经根据特定的数据模型结构化,在数据库中用户的逻辑数据文件和具体的物理数据文件不必一一对应,从而有效地节省了存储资源,减少了数据冗余,保证了数据的一致性。

3. 具有较高的数据独立性

数据独立性是指应用程序与数据库的数据结构之间相互独立。在数据库系统中,因为采用了数据库的三级模式结构,保证了数据库中数据的独立性。在数据存储结构改变时,不影响数据的全局逻辑结构,这样保证了数据的物理独立性。在全局逻辑结构改变时,不影响用户的局部逻辑结构和应用程序,这样就保证了数据的逻辑独立性。

4. 有统一的数据控制功能

在数据库系统中,数据由 DBMS 进行统一控制和管理。DBMS 提供了一套有效的数据控制手段,包括数据安全性控制、数据完整性控制、数据库的并发控制和数据库的恢复等,增强了多用户环境下数据的安全性和一致性保护。

1.3 数据模型

数据库是现实世界中某种应用环境(一个单位或部门)所涉及的数据集合,它不仅要反映数据本身的内容,而且要反映数据之间的联系。由于计算机不能直接处理现实世界中的具体事物,所以必须将这些具体事物转换成计算机能够处理的数据。在数据库技术中,用数据模型(Data Model)来对现实世界中的数据进行抽象和表示。

1.3.1 数据模型的组成要素

一般而言,数据模型是一种形式化描述数据、数据之间的联系以及有关语义约束规则的方法,这些规则分为3个方面:描述实体静态特征的数据结构、描述实体动态特征的数据操作规则和描述实体语义要求的数据完整性约束规则。因此,数据结构、数据操作及数据的完整性约束也被称为数据模型的3个组成要素。

1. 数据结构

数据结构研究数据之间的组织形式(数据的逻辑结构)、数据的存储形式(数据的物理结构)、数据对象的类型等。存储在数据库中的对象类型的集合是数据库的组成部分。例如,在教学管理系统中,要管理的数据对象有学生、课程、选课成绩等,在课程对象集合中,每门课程包括课程号、课程名、学分等信息,这些基本信息描述了每门课程的特性,构成在数据库中存储的框架,即对象类型。

数据结构用于描述系统的静态特性,是刻画一个数据模型性质最重要的方面。因此,在数据库系统中,通常按照数据结构的类型来命名数据模型,如层次结构、网状结构和关系结构的数据模型分别命名为层次模型、网状模型和关系模型。

2. 数据操作

数据操作用于描述系统的动态特性,是指对数据库中的各种数据所允许执行的操作的集合,包括操作及有关的操作规则。数据库主要有查询和更新(包括插入、删除和修改等)两大类操作。数据模型必须定义这些操作的确切含义、操作符号、操作规则(如优先级)、实现操作的语言等。

3. 数据的完整性约束

数据的完整性约束是一组完整性规则的集合。完整性规则是给定的数据模型中数据及其联系所具有的约束和依存规则,用以限定符合数据模型的数据库状态和状态的变化,以保证数据的正确、有效和相容。

数据模型应该反映和规定数据必须遵守的、基本的、通用的完整性约束。此外,数据模型还应该提供定义完整性约束条件的机制,以反映具体涉及的数据必须遵守的、特定的语义约束条件,如学生信息中的“性别”只能为“男”或“女”、学生选课信息中的“课程号”的值必须取自学校已开设课程的课程号等。

1.3.2 数据抽象的过程

从现实世界中的客观事物到数据库中存储的数据是一个逐步抽象的过程,这个过程经历了现实世界、观念世界和机器世界3个阶段,对应于数据抽象的不同阶段采用不同的数据模型。首先将现实世界的事物及其联系抽象成观念世界的概念模型,然后再转换成机器世界的数据库模型。概念模型并不依赖于具体的计算机系统,它不是DBMS所支持的数据模型,它是现实世界中客观事物的抽象表示。概念模型经过转换成为计算机上某一DBMS支持的数据模型。所以说,数据模型是对现实世界进行抽象和转换的结果,这一过程如图1.2所示。

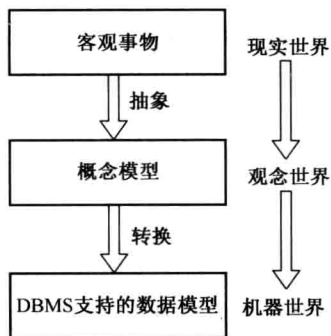


图 1.2 数据抽象的过程

1. 对现实世界的抽象

现实世界就是客观存在的世界,其中存在着各种客观事物及其相互之间的联系,而且每个事物都有自己的特征或性质。计算机处理的对象是现实世界中的客观事物,在对其实施处理的过程中,首先应了解和熟悉现实世界,从对现实世界的调查和观察中抽象出大量描述客观事物的事实,再对这些事实进行整理、分类和规范,进而将规范化的事实数据化,最终实现数据库系统的存储和处理。

2. 观念世界中的概念模型

观念世界是对现实世界的一种抽象,通过对客观事物及其联系的抽象描述构造出概念模

型 (Conceptual Model)。概念模型的特征是按用户需求观点对数据进行建模,表达了数据的全局逻辑结构,是系统用户对整个应用项目涉及数据的全面描述。概念模型主要用于数据库设计,它独立于现实世界的 DBMS,也就是说选择何种 DBMS 不会影响概念模型的设计。

概念模型的表示方法很多,目前较常用的是实体联系模型 (Entity Relationship Model),简称 E-R 模型。E-R 模型主要用 E-R 图来表示。

3. 机器世界中的逻辑模型和物理模型

机器世界是指现实世界在计算机中的体现与反映。现实世界中的客观事物及其联系在机器世界中以逻辑模型 (Logical Model) 描述。在选定 DBMS 后,就要将 E-R 图表示的概念模型转换为具体的 DBMS 支持的逻辑模型。逻辑模型的特征是按计算机实现的观点对数据进行建模,表达了数据库的全局逻辑结构,是设计人员对整个应用项目数据库的全面描述,逻辑模型服务于 DBMS 的应用实现。通常也把数据的逻辑模型直接称为数据模型。数据库系统中主要的逻辑模型有层次模型、网状模型和关系模型。

物理模型 (Physical Model) 是对数据最低层的抽象,用以描述数据在物理存储介质上的组织结构,与具体的 DBMS、操作系统和硬件有关。

从概念模型到逻辑模型的转换是由数据库设计人员完成的,从逻辑模型到物理模型的转换是由 DBMS 完成的,一般人员不必考虑物理实现细节,因而逻辑模型是数据库系统的基础,也是应用过程中要考虑的核心问题。

1.3.3 概念模型

当分析某种应用环境所需的数据时,总是首先找出涉及的实体及实体之间的联系,进而得到概念模型,这是数据库设计的先导。

1. 实体与实体集

实体 (Entity) 是现实世界中任何可以相互区分和识别的事物,它既可以是能触及的客观对象 (如一位教师、一名学生、一种商品等),也可以是抽象的事件 (如一场足球比赛、一次借书等)。

性质相同的同类实体的集合称为实体集 (Entity Set),如一个系的所有教师、2014 年南非世界杯足球赛的全部 64 场比赛等。

2. 属性

每个实体都具有一定的特征或性质,这样才能区分一个个实体。例如,教师的编号、姓名、性别、职称等都是教师实体具有的特征,足球赛的比赛时间、地点、参赛队、比分、裁判姓名等都是足球赛实体的特征。实体的特征称为属性 (Attribute),一个实体可用若干属性来刻画。

能唯一标识实体的属性或属性集称为实体标识符,如教师的编号可以作为教师实体的标识符。

3. 类型与值

属性和实体都有类型 (Type) 和值 (Value) 之分。属性类型就是属性名及其取值类型,属性值就是属性所取的具体值。例如,教师实体中的“姓名”属性,属性名“姓名”和取字符类型的值是属性类型,而“黎德瑟”、“王德浩”等是属性值。每个属性都有特定的取值范围,即值域 (Domain),超出值域的属性值则认为无实际意义,如“性别”属性的值域为 (男,女)、“职称”属性的值域为 (助教,讲师,副教授,教授) 等,由此可见,属性类型是个变量,属

性值是变量所取的值, 而值域是变量的取值范围。

实体类型 (Entity Type) 就是实体的结构描述, 通常是实体名和属性名的集合; 具有相同属性的实体有相同的实体类型。实体值是一个具体的实体, 是属性值的集合。例如, 教师实体类型是:

教师 (编号, 姓名, 性别, 出生日期, 职称, 基本工资, 研究方向)

教师“王德浩”的实体值是:

(T6, 王德浩, 男, 09/21/65, 教授, 2750, 数据库技术)

由此可见, 属性值所组成的集合表征一个实体, 相应的这些属性名的集合表征一个实体类型, 相同类型实体的集合称为实体集。

在 SQL Server 中, 用“表”来表示同一类实体, 即实体集, 用“记录”来表示一个具体的实体, 用“字段”来表示实体的属性。显然, 字段的集合组成一个记录, 记录的集合组成一个表。实体类型则代表了表的结构。

4. 实体间的联系

实体之间的对应关系称为联系 (Relationship), 它反映了现实世界事物之间的相互关联。例如, 图书和出版社之间的关联关系为: 一个出版社可以出版多种书, 同一种书只能在一个出版社出版。

实体间的联系是指一个实体集中可能出现的每一个实体与另一实体集中多少个具体实体存在联系。实体之间有各种各样的联系, 归纳起来有 3 种类型:

(1) 一对一联系。如果对于实体集 A 中的每一个实体, 实体集 B 中至多只有一个实体与之联系, 反之亦然, 则称实体集 A 与实体集 B 具有一对一联系, 记为 1:1。例如, 一个工厂只有一个厂长, 一个厂长只在一个工厂任职, 厂长与工厂之间的联系是一对一的联系。

(2) 一对多联系。如果对于实体集 A 中的每一个实体, 实体集 B 中可以有多个实体与之联系, 反之, 对于实体集 B 中的每一个实体, 实体集 A 中至多只有一个实体与之联系, 则称实体集 A 与实体集 B 具有一对多联系, 记为 1:n。例如, 一个公司有许多职员, 但一个职员只能在一个公司就职, 所以公司和职员之间的联系是一对多的联系。

(3) 多对多联系。如果对于实体集 A 中的每一个实体, 实体集 B 中可以有多个实体与之联系, 而对于实体集 B 中的每一个实体, 实体集 A 中也可以有多个实体与之联系, 则称实体集 A 与实体集 B 之间有多对多的联系, 记为 m:n。例如, 一个读者可以借阅多种图书, 任何一种图书可以被多个读者借阅, 所以读者和图书之间的联系是多对多的联系。

5. E-R 图

概念模型是反映实体及实体之间联系的模型。在建立概念模型时, 要逐一给实体命名以示区别, 并描述它们之间的各种联系。E-R 图是用一种直观的图形方式建立现实世界中实体及其联系模型的工具, 也是数据库设计的一种基本工具。

E-R 模型用矩形框表示现实世界中的实体, 用菱形框表示实体间的联系, 用椭圆框表示实体和联系的属性, 实体名、属性名和联系名分别写在相应框内。对于作为实体标识符的属性, 在属性名下画一条横线。实体与相应的属性之间、联系与相应的属性之间用线段连接。联系与其涉及的实体之间也用线段连接, 同时在线段旁标注联系的类型 (1:1、1:n 或 m:n)。

图 1.3 所示为学生信息系统中的 E-R 图, 该图建立了学生、课程和学院 3 个不同的实体及其联系的模型。其中“课程号”属性作为课程实体的标识符 (不同课程的课程号不同), “学号”属性作为学生实体的标识符, “编号”属性作为学院实体的标识符。联系也可以有自己的属性,