



高等学校应用型特色规划教材

Access 2003

数据库教程

解圣庆 主 编
刘永华 徐兴敏 副主编



清华大学出版社

高等学校应用型特色规划教材

Access 2003 数据库教程

解圣庆 主 编

刘永华 徐兴敏 副主编

清华大学出版社

北 京

内 容 简 介

本书全面、系统地介绍了 Access 2003 数据库系统的基础知识与应用开发技术。全书以 Access 2003 关系数据库为背景,介绍数据库的基本概念,数据库的设计与建立及开发与应用,使学生能够掌握数据库设计的思想及 Access 2003 数据库对象的操作技能。本书还结合 VBA,通过实例讲述了程序设计的基本思想与方法,并重点突出面向对象程序设计技能的培养,对读者自主开发数据库应用系统大有帮助。

本书以应用为目的,通过大量的实例,使学生可以尽快掌握 Access 2003 的功能与操作,每章均附有习题,可满足普通高校计算机及相关专业本专科数据库应用教学的需求,亦可作为科技工作者及计算机爱好者的自学用书。

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

Access 2003 数据库教程/解圣庆主编;刘永华,徐兴敏副主编. —北京:清华大学出版社,2006.10
(高等学校应用型特色规划教材)

ISBN 7-302-13761-7

I. A… II. ①解… ②刘… ③徐… III. 关系数据库—数据库管理系统, Access 2003—高等学校—教材
IV. TP311.138 •

中国版本图书馆 CIP 数据核字(2006)第 105271 号

出 版 者:清华大学出版社 地 址:北京清华大学学研大厦

<http://www.tup.com.cn> 邮 编:100084

社 总 机:010-62770175 客户服务:010-62776969

组稿编辑:刘天飞

文稿编辑:闫光龙

印 刷 者:北京密云胶印厂

装 订 者:北京市密云县京文制本装订厂

发 行 者:新华书店总店北京发行所

开 本:185×260 印张:22 字数:525 千字

版 次:2006 年 10 月第 1 版 2006 年 10 月第 1 次印刷

书 号:ISBN 7-302-13761-7/TP·8279

印 数:1~4000

定 价:30.00 元

前 言

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系统核心。随着计算机与网络技术的飞速发展，作为计算机应用的一个重要领域，数据库技术得到了广泛的应用与发展。

在数据库技术教学领域，数据库原理是数据库技术专业的基础理论课程，大型数据库技术（例如 SQL Server 等）作为一门专业课，专注于专业数据库技术的使用与管理。如何使理论与实践相结合，使学生掌握数据库技术的基础理论，掌握数据库的设计与管理、数据的应用与程序设计方法，使学生通过学习能设计一个简单的数据库应用系统，是数据库技术教学的基本目的。

Access 2003 关系型数据库管理系统是 Microsoft Office 系列应用软件的一个重要组成部分。它界面友好，功能全面且操作简单，不仅可以有效地组织与管理、共享与开发应用数据库信息，而且可以把数据库信息与 Web 结合在一起，为在局域网络和互联网共享数据库信息奠定了基础。Access 2003 提供的数据库对象，可以使学生不用编程就能设计出一个桌面数据库应用系统，而且可以使学生从中掌握数据库应用的思想。通过学习 VBA 程序设计，创建 VBA 模块，学生可以学习如何通过面向对象的程序设计，使用对象的属性、事件与方法，建立事件驱动过程，完成面向用户的数据库应用系统的开发设计。

本书共 12 章，内容包括：数据库系统的基础知识、关系数据库设计理论、Access 数据库、数据表设计与应用、查询操作与高级应用、窗体设计及高级应用、报表设计与应用、数据访问页、宏设计、VBA 程序设计基础、在 Access 中创建 VBA 模块和数据库系统应用开发实例。

第 1、2 章，系统介绍了数据库的基础理论，特别是关系数据库的技术思想。

第 3~10 章，详细介绍了 Access 2003 对象的基本功能与操作设计方法。

第 11~12 章，介绍了 VBA 程序设计基础与 VBA 模块的创建方法，运用 DAO 和 ADO 进行网络数据库编程。

本书由解圣庆担任主编并负责统稿整理，刘永华、徐兴敏担任副主编。其中第 6、9、10、11、12 章由解圣庆编写，第 5、7、8 章由刘永华编写，第 1、2、3、4 章由徐兴敏编写。山东大学的石冰教授对全书进行了审阅，并提出了宝贵的修改意见。

限于作者水平，书中难免有不足之处，敬请广大读者批评指正。

编 者

2006 年 3 月

目 录

第1章 数据库系统的基础知识.....1	2.5.5 BC 范式(BCNF)..... 26
1.1 数据库系统概述.....1	2.5.6 关于规范化..... 26
1.1.1 数据库系统的组成.....1	习题..... 27
1.1.2 数据库系统的特点.....3	第3章 Access 2003 数据库..... 28
1.1.3 高级数据库阶段.....5	3.1 Access 2003 数据库开发环境..... 28
1.2 数据模型.....6	3.1.1 Access 2003 数据库的安装、 启动与关闭..... 28
1.2.1 数据模型的概念.....6	3.1.2 Access 2003 数据库构成..... 29
1.2.2 数据模型的组成要素.....7	3.1.3 Access 2003 数据库 操作环境..... 31
1.2.3 概念模型.....7	3.2 Access 2003 数据库操作与管理..... 32
1.2.4 常用的数据模型.....9	3.2.1 创建数据库..... 33
1.2.5 数据库的体系结构.....11	3.2.2 数据库管理..... 34
1.3 数据库设计.....12	3.3 Access 2003 数据库安全性..... 35
1.4 Access 2003 数据库简介.....13	3.3.1 设置数据库密码..... 35
1.4.1 Access 的发展过程.....13	3.3.2 用户级安全机制..... 37
1.4.2 Access 2003 新增功能.....13	3.3.3 使用权限..... 38
习题.....15	3.4 安全机制管理..... 39
第2章 关系数据库设计理论.....16	3.4.1 增加账户..... 39
2.1 关系模型.....16	3.4.2 删除账户..... 39
2.1.1 关于关系的一些术语.....16	3.4.3 更改账户权限..... 39
2.1.2 关系模型.....17	3.4.4 打印账户和组账户列表..... 40
2.2 关系数据结构及形式化表示.....17	3.5 加密数据库..... 40
2.2.1 关系数据结构.....17	习题..... 41
2.2.2 关系模式.....19	第4章 数据表设计与应用..... 42
2.3 关系代数.....20	4.1 创建表..... 42
2.3.1 集合运算.....20	4.1.1 数据表结构设计..... 42
2.3.2 专门的关系运算.....21	4.1.2 主码..... 49
2.4 关系完整性.....22	4.1.3 索引..... 50
2.5 数据库设计与设计规范.....23	4.1.4 使用向导创建表..... 51
2.5.1 关系数据库设计规范.....23	4.1.5 通过输入数据创建表..... 52
2.5.2 第一范式(1NF).....24	4.2 管理与维护表..... 53
2.5.3 第二范式(2NF).....24	
2.5.4 第三范式(3NF).....25	

4.2.1 表间关系的建立.....	53	5.5.3 SQL 视图.....	97
4.2.2 修改表结构.....	55	习题.....	98
4.3 操作表.....	56	第 6 章 窗体设计及高级应用.....	99
4.3.1 数据输入.....	56	6.1 窗体基础知识.....	99
4.3.2 排序记录.....	59	6.1.1 窗体的概念与作用.....	99
4.3.3 筛选记录.....	59	6.1.2 窗体构成.....	100
4.3.4 设置数据表格式.....	61	6.1.3 窗体类型.....	101
4.4 数据的导入与导出.....	62	6.1.4 窗体视图.....	102
4.4.1 导入、导出数据.....	62	6.2 创建窗体.....	103
4.4.2 链接数据.....	63	6.2.1 使用自动创建窗体向导.....	103
4.5 数据库的转换.....	63	6.2.2 使用“窗体向导”.....	104
4.5.1 在 Access 2003 中使用		6.2.3 使用“图表向导”.....	107
旧版本的数据库.....	63	6.2.4 使用“数据透视表向导”.....	109
4.5.2 将 Access 2003 数据库转换为		6.3 自定义窗体.....	110
Access 97 或 Access 2000		6.3.1 窗体设计视图的组成	
格式.....	64	与主要功能.....	111
4.5.3 将旧版本的 Access 数据库		6.3.2 属性、事件与方法.....	114
转换为 Access 2003 数据库.....	64	6.3.3 窗体与对象的属性	
习题.....	65	及设置方法.....	116
第 5 章 查询操作与高级应用.....	66	6.3.4 窗体与对象的事件.....	122
5.1 查询的种类与应用.....	66	6.3.5 常用方法.....	125
5.1.1 查询的种类.....	66	6.3.6 常用控件的创建方法.....	131
5.1.2 查询的应用.....	67	6.4 窗体与控件的其他应用设计.....	142
5.2 查询的建立方法.....	68	6.4.1 创建计算控件.....	142
5.2.1 使用查询向导.....	68	6.4.2 查找记录.....	144
5.2.2 使用查询设计器.....	71	6.4.3 显示提示信息.....	145
5.3 查询条件.....	73	6.4.4 创建与使用主/子窗体.....	145
5.3.1 表达式.....	74	6.4.5 打印与预览窗体.....	146
5.3.2 标准函数.....	77	6.4.6 设计菜单.....	147
5.4 查询设计.....	78	6.5 窗体外观格式设计.....	150
5.4.1 条件选择查询.....	79	6.5.1 加线条.....	150
5.4.2 交叉表查询.....	81	6.5.2 加矩形.....	151
5.4.3 参数查询.....	84	6.5.3 设置控件格式属性.....	151
5.4.4 操作查询.....	85	6.5.4 使用 TAB 键设置控件次序.....	153
5.4.5 重复项、不匹配项查询.....	89	习题.....	154
5.5 SQL 查询.....	90	第 7 章 报表设计与应用.....	156
5.5.1 SQL 的数据定义.....	90	7.1 报表的作用与组成.....	156
5.5.2 SQL 的数据操纵.....	91		

7.1.1 报表的作用	156	8.2.3 Web 页到数据访问页的 转换.....	183
7.1.2 报表的组成	156	8.3 使用设计视图创建数据访问页.....	183
7.1.3 报表的视图	157	8.3.1 数据访问页设计视图.....	183
7.2 报表的分类	158	8.3.2 数据访问页中控件的使用	184
7.2.1 纵栏式报表	158	8.3.3 添加 Office 电子表格.....	187
7.2.2 表格式报表	158	8.3.4 数据访问页绑定表或查询	188
7.2.3 图表报表	159	8.4 数据访问页的格式设计与应用.....	189
7.2.4 标签报表	159	8.4.1 数据访问页的主题 与背景设置.....	189
7.3 创建报表	160	8.4.2 数据访问页中页对象 的属性.....	191
7.3.1 报表向导	160	8.4.3 在 IE 中访问数据访问页	192
7.3.2 自动创建报表.....	161	习题.....	193
7.3.3 创建子报表	161	第 9 章 宏设计	194
7.3.4 创建多列报表.....	162	9.1 宏的基础知识	194
7.3.5 将窗体转换为报表.....	162	9.1.1 宏的定义与特点.....	194
7.3.6 自定义报表	162	9.1.2 宏组.....	196
7.4 报表编辑	163	9.1.3 条件操作宏.....	196
7.4.1 报表控件及格式设计.....	163	9.1.4 宏与 VBA	197
7.4.2 排序与分组	166	9.1.5 宏向 VBA 代码转换	198
7.4.3 在报表中应用计算.....	169	9.2 创建宏	200
7.5 复杂报表设计	171	9.2.1 宏的创建.....	200
7.5.1 报表属性	171	9.2.2 条件宏的创建.....	202
7.5.2 节属性	172	9.2.3 宏组的创建.....	204
7.5.3 报表加页分割.....	172	9.2.4 宏对象的编辑与修改.....	205
7.5.4 使用报表的预定义格式.....	173	9.2.5 宏调试.....	206
7.5.5 制作图表	173	9.3 宏运行	207
7.6 打印报表	174	9.3.1 在菜单或工具栏中执行宏	207
7.6.1 报表预览	174	9.3.2 在窗体中加入宏	208
7.6.2 打印与保存报表.....	175	9.3.3 宏组中宏的运行.....	208
习题.....	176	9.3.4 自动运行的宏 AutoExec	209
第 8 章 数据访问页	177	9.3.5 在宏组 AutoKeys 中 定义响应组合键.....	209
8.1 数据访问页的基本知识	177	9.3.6 用宏设计系统菜单.....	210
8.1.1 数据访问页的功能.....	177	9.4 宏操作	212
8.1.2 数据访问页的类型.....	177	习题.....	225
8.1.3 数据访问页的数据源.....	178		
8.1.4 数据访问页视图.....	178		
8.2 创建数据访问页	178		
8.2.1 自动创建数据访问页.....	178		
8.2.2 使用向导创建数据访问页	180		

第 10 章 VBA 程序设计基础	226	11.2.1 数据库引擎与接口.....	294
10.1 VBA 编程环境.....	226	11.2.2 数据库对象.....	295
10.1.1 认识 VBA.....	226	11.2.3 DAO(数据访问对象).....	296
10.1.2 VBA 编程环境: VBE 窗口.....	227	11.2.4 ADO(活动数据对象).....	310
10.1.3 VBA 代码窗口的使用.....	230	习题.....	317
10.2 模块的基础知识.....	232	第 12 章 数据库应用系统开发实例 ..	318
10.2.1 类模块.....	233	12.1 教职工人事管理系统功能.....	318
10.2.2 标准模块.....	233	12.1.1 主界面设计.....	318
10.2.3 创建模块.....	233	12.1.2 系统主要功能.....	319
10.3 面向对象程序设计基础.....	235	12.2 数据库设计.....	319
10.3.1 面向对象的程序设计概念.....	235	12.2.1 数据表结构设计与表创建... ..	319
10.3.2 对象和类的概念.....	236	12.2.2 建立表间关系.....	321
10.4 VBA 编程基础.....	236	12.3 数据处理功能设计.....	321
10.4.1 数据库对象.....	236	12.4 数据查询功能设计.....	326
10.4.2 VBA 中的基本数据类型.....	237	12.4.1 创建单项查询.....	326
10.4.3 变量、常量与数组.....	241	12.4.2 创建综合信息查询.....	328
10.4.4 数据库对象变量.....	247	12.5 数据统计功能设计.....	330
10.4.5 运算符与表达式.....	248	12.5.1 统计操作窗体设计 (jzg_sjtj).....	330
10.4.6 函数.....	253	12.5.2 统计查询设计.....	332
10.4.7 程序语句.....	268	12.6 系统服务功能设计.....	334
10.5 过程调用与参数传递.....	276	12.6.1 用户名与密码管理.....	334
10.6 VBA 程序的运行错误 处理与调试.....	279	12.6.2 数据备份与恢复.....	337
10.6.1 程序的运行错误处理.....	279	12.7 宏的设计.....	337
10.6.2 程序的调试.....	281	12.7.1 用宏设计下拉子菜单.....	337
习题.....	283	12.7.2 用宏设计水平菜单 (主菜单).....	338
第 11 章 创建 VBA 模块	284	12.7.3 设计主窗体激活菜单系统... ..	339
11.1 模块创建.....	284	12.7.4 用 Autoexec 宏启动教职工 人事管理系统.....	339
11.1.1 过程与变量的作用域.....	284	习题.....	341
11.1.2 创建事件过程.....	288	参考文献	342
11.2 数据库对象的使用.....	294		

第 1 章 数据库系统的基础知识

本章学习目标

本章从数据库系统的基础知识入手,对数据库的基本概念、数据模型、数据库体系结构、数据库设计及 Access 2003 数据库进行介绍。通过本章的学习,读者应该掌握以下内容:

- 数据库系统的概念与组成
- 数据模型
- 数据库体系结构
- 数据库设计方法与步骤
- Access 2003 数据库的特点

1.1 数据库系统概述

数据库技术产生于 20 世纪 60 年代末,是数据管理的最新技术,计算机科学的重要分支。在当今信息社会中,信息已成为各个行业、部门的重要财富和资源,信息系统也越来越显示出它的重要性。数据库技术是信息系统的核心和基础,它的出现极大地促进了计算机应用向各行各业的渗透。从一般的小型事务处理到大型的信息系统,越来越多的新应用领域开始采用数据库技术存储与处理其信息资源。数据库的建设规模、数据库信息量的大小和使用频度已成为衡量一个国家信息化程度的重要标志。

1.1.1 数据库系统的组成

数据(Data)是数据库中存储的基本对象。数据的种类很多,例如文字、图形、图像和声音等都是数据。

数据可定义为描述事物的符号记录。数据有多种形式,它们均可以经过数字化后储存在计算机中。在描述事物的过程中,数据与其解释是不可分的。

数据库是指长期储存在计算机内的、有组织的、可共享的数据集合。数据库中的数据是按一定的数据模型组织、描述和储存的,具有较小的冗余度、较高的数据独立性和易扩展性,并且可以被多个用户、多个应用程序共享。

数据库管理系统(Database Management System, DBMS)是位于用户与操作系统(OS)之间的一层数据管理软件,是数据库系统的中心枢纽。数据库管理系统能科学地组织和存储数据、高效地获取和维护数据。用户对数据库进行的各种操作,如数据库的建立、使用和维护,都是在 DBMS 的统一管理和控制下进行的。

数据库管理系统的主要功能有以下几个方面:

(1) 数据定义功能

提供数据定义语言(DDL), 用于定义数据库中的数据对象。

(2) 数据操纵功能

提供数据操纵语言(DML), 用于操纵数据实现对数据库的基本操作, 例如查询、插入、删除和修改等。

(3) 数据库的运行管理

保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后的系统恢复。

(4) 数据库的建立和维护功能

提供数据库数据输入、批量装载、数据库转储、介质故障恢复、数据库的重组织及性能监视等功能。

数据库系统(Database System, DBS)是指在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(硬件和基本软件)、数据库、数据库管理系统(及其开发工具)、应用系统和有关人员(数据库管理员、应用设计人员、最终用户)组成的具有高度组织性的总体。

通常情况下, 把数据库系统简称为数据库。数据库系统组件之间的关系如图 1.1 所示。

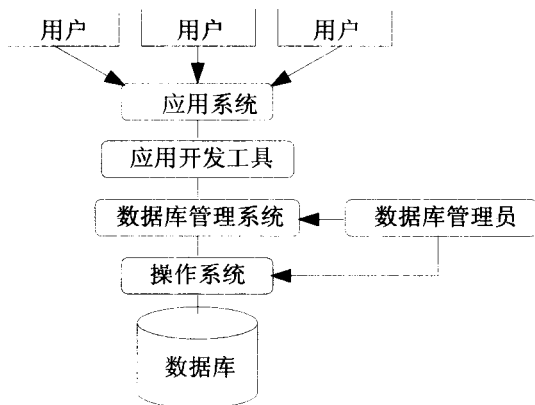


图 1.1 数据库系统组成

数据库技术的核心任务是数据处理。数据处理是指对各种数据进行收集、存储、加工和传播等一系列活动的总和。数据管理则是指对数据进行分类、组织、编码、存储、检索和维护, 它是数据处理的中心问题。

数据管理技术的发展, 与计算机硬件(主要是外部存储器)、系统软件及计算机应用的范围有着密切的联系。数据管理技术的发展经历了以下几个阶段: 人工管理阶段、文件系统阶段、数据库系统阶段和分布式数据库系统阶段。

20 世纪 50 年代中期以前, 计算机主要用于科学计算。那时的计算机硬件方面, 外存只有卡片、纸带及磁带, 没有磁盘等可直接存取的存储设备; 软件方面, 只有汇编语言, 没有操作系统和高级语言, 更没有管理数据的软件; 数据处理的方式是批处理。这些决定了当时的数据管理只能依赖人工来进行。

人工管理阶段的特点是:

(1) 数据不进行保存。当时的计算机主要用于科学计算, 一个程序对应一个数据。在

计算某一问题时,把程序和对应的数据装入,计算完就退出,没有将数据长期保存的必要。

(2) 没有专门的数据管理软件。数据需要由应用程序自己管理,因此应用程序的设计者不仅要考虑数据的逻辑结构,还考虑数据的物理结构,比如存储结构、存取方法、输入输出方式等。存储结构发生变化,应用程序也要做相应的修改,程序员的负担非常重,数据的独立性也很差。

(3) 数据面向应用。一组数据对应一组程序。倘若多个程序使用相同的数据,必须各自定义,不能共享。所以程序之间存在大量的数据冗余。

(4) 只有程序的概念,基本上没有文件的概念。

20 世纪 60 年代,计算机技术有了很大提高,计算机的应用范围不断扩大,不仅用于科学计算,还大量用于管理。这时的计算机硬件已经有了磁盘、磁鼓等直接存取的外存设备;软件则有了操作系统、高级语言,操作系统中的文件系统专门用于数据管理;处理方式不仅有批处理,还增加了联机实时处理。

文件系统阶段的特点如下:

(1) 数据可以长期保存在磁盘上。用户可以反复对文件进行查询、修改、插入和删除等操作。

(2) 文件系统提供了数据与程序之间的存取方法。应用程序和数据有了一定的独立性,数据存储结构的改变也不一定反映在程序上,大大减轻了程序员的负担。

(3) 数据冗余量大。文件系统中,文件仍然是面向应用的,一个文件基本上对应于一个应用程序,不能对数据项进行共享,因此数据冗余大,存储空间浪费。由于数据可能有多个副本,对其中之一进行修改时还容易造成数据的不一致性。

(4) 文件之间缺乏联系,相对孤立,仍然不能反映客观世界各个事物之间错综复杂的联系。

数据库系统阶段的特点如下:

这个阶段基本实现了数据共享,减少了数据冗余,数据库采用特定的数据模型,具有较高的数据独立性,有统一的数据控制和管理功能。

分布式数据库系统阶段的特点如下:

分布式数据库在逻辑上是一个整体,是分布在不同地理位置上的数据集合,它受分布式数据库管理系统的控制和管理。

- (1) 分布透明性
- (2) 局部与集中控制相结合
- (3) 高可靠性和可用性
- (4) 高效率和灵活性

1.1.2 数据库系统的特点

20 世纪 60 年代末以来,计算机的应用更为广泛,用于数据管理的应用系统规模也更为庞大,由此带来数据量的急剧膨胀;计算机磁盘技术有了很大发展,出现了大容量的磁盘;在处理方式上,联机实时处理的要求更多。这些变化促使了数据管理手段的进步,数据库技术应运而生。与人工管理和文件系统相比,数据库系统的特点主要有以下几个方面:

(1) 数据的结构化

在文件系统中,只考虑了同一文件记录内部数据项之间的联系,而不同文件的记录之间是没有联系的,从整体上看数据是无结构的,不能反映客观世界各种事物之间的错综复杂的联系。在数据库系统中,实现了整体数据的结构化,把文件系统中简单的记录结构变成了记录和记录之间的联系所构成的结构化数据。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。

(2) 数据共享性好

数据库系统从整体角度看待与描述数据,使数据不再面向某个应用而是面向整个系统,这些数据可以供多个部门使用,实现了数据的共享。各个部门的数据基本上没有重复的存储,数据的冗余量小。

(3) 数据独立性好

数据库系统有三层结构:用户(局部)数据的逻辑结构、整体数据的逻辑结构和数据的物理结构。在这三层结构之间,数据库系统提供了两层映像功能。首先是用户数据逻辑结构和整体数据逻辑结构之间的映像,这一映像保证了数据的逻辑独立性:当数据库的整体逻辑结构发生变化时,通过修改这层映像可使局部的逻辑结构不受影响,因此不必修改应用程序。另外一级映像是整体数据逻辑结构和数据物理结构之间的映像,它保证了数据的物理独立性:当数据的存储结构发生变化时,通过修改这层映像可使数据的逻辑结构不受影响,因此应用程序同样不必修改。

(4) 数据存储粒度小

在文件系统中,数据存储的最小单位是记录;而在数据库系统中,数据存储的粒度可以小到记录中的一个数据项。因此数据库中数据存取的方式非常灵活,便于对数据的管理。

(5) 数据管理系统对数据进行统一的管理和控制

DBMS 不仅具有基本的数据管理功能,还具有如下的控制功能:

① 保证数据的完整性

数据的完整性指数据的正确性、有效性和相容性,要求数据在一定的取值范围内或相互之间满足一定的关系。例如,规定考试的成绩在 0 分到 100 分之间,血型只能是 A 型、B 型、AB 型、O 型中的一种等。

② 保证数据的安全性

让每个用户只能按指定的权限访问数据,防止不合法地使用数据,造成数据的破坏和丢失。比如学生对于课程的成绩只能进行查询,不能修改。

③ 并发控制

对多用户的并发操作加以协调和控制,防止多个进程同时存取、修改数据库中的数据时发生冲突、造成错误。比如在学生选课系统中,某门课只剩下最后一个名额,但有两个学生在两台选课终端上同时发出了选这门课的请求,必须采取某种措施,确保两名学生不能同时拥有这最后的一个名额。

④ 数据库的恢复

当数据库系统出现硬件或软件故障时,DBMS 具有把数据库恢复到最近某个时刻的正确状态上来的能力。

(6) 为用户提供了友好的接口

用户可以使用交互式的命令语言,如 SQL(Structured Query Language,结构化查询语言),对数据库进行操作;也可以把普通的高级语言(如 C++语言等)和 SQL 结合起来,从而把对数据库的访问和对数据的处理有机地结合在一起。

1.1.3 高级数据库阶段

1. 分布式数据库系统

20 世纪 70 年代后期之前,数据库系统多数是集中式的。分布式数据库系统是数据库技术和计算机网络技术相结合的产物,在 80 年代中期已有商品化产品问世。分布式数据库是一个逻辑上统一、地域上分布的数据集合,是计算机网络环境中各个结点局部数据库的逻辑集合,同时受分布式数据库管理系统的控制和管理。

分布式数据库在逻辑上像一个集中式数据库系统,实际上数据存储处于不同地点的计算机网络的各个结点上。每个结点都有自己的局部数据库管理系统,使每个节点有很高的独立性。用户可以由分布式数据库管理系统(网络数据库管理系统),通过网络通信相互传输数据。分布式数据库系统有高度透明性,每台计算机上的用户并不需要了解他所访问的数据究竟在什么地方,就像在使用集中式数据库一样。其主要优点有:

(1) 局部自主

网络上每个结点的数据库系统都具有独立处理本地事物的能力(大量的),而且各局部结点之间也能够互相访问、有效地配合处理更复杂的事物。因此,分布式数据库系统特别适合各个部门的地理位置分散的组织机构。例如,银行业务、飞机订票和企业管理等。

(2) 可靠性和可用性

分布式系统比集中式系统有更高的可靠性,在个别结点或个别通信链路发生故障的情况下可以继续工作。一个局部系统发生故障不至于导致整个系统停顿或破坏,只要有一个结点上的数据备份可用,数据就是可用的。

(3) 效率和灵活性

分布式系统分散了工作负荷,缓解了单机容量的压力。数据可以存储在临近的常用结点,如果本结点的数据子集包含了要查询的全部内容,显然会比集中式数据库在全集上查找节省时间。

系统易于实现扩展。例如,一个单位要增加新的机构,分布式数据库系统能够在对现有系统影响较小的情况下实现扩充。因此,扩大系统规模比集中式系统更加方便、经济和灵活。

2. 数据库的发展方向

目前,数据库技术的发展方向主要有两个:一是改造和扩充关系数据库,以适应新的应用要求;二是改用新的数据库模型。目前,这两个方面都取得了很大发展,主要出现了下列新的数据库技术。

(1) 并行数据库技术

并行数据库技术包括对数据库的分区管理和并行查询。它通过将一个数据库任务分割

成多个子任务的方法由多个处理机协同完成这个任务,从而极大地提高了事务处理能力,并且通过数据分区可以实现数据的并行 I/O 操作。

(2) 数据仓库和数据采掘技术

所谓数据仓库就是按决策目标将传统的事务型数据库中的数据重新组织划分,由此造成一种面向主题的、集成的、稳定的及随时间发展的数据集合。数据仓库与传统数据库的区别在于存储的数据容量大、存储的数据时间跨度大、存储的数据来源复杂、可用于企业与组织的决策分析处理等。所谓数据采掘(Data Mining, DM)就是从大型数据库或数据仓库的数据中提取人们感兴趣的、隐含的、事先未知的、潜在的知识。数据采掘方法的提出使人们有能力从过去若干年时间里积累的海量的、以不同的形式存储的、十分繁杂的数据资料中认识数据的真正价值。目前,数据采掘的研究已与数据库的研究结合起来。

(3) 多媒体数据库

人们通常把能够管理数值、文本、图形、图像和声音等媒体类型的数据库称为多媒体数据库(Multimedia Database)。与传统数据库管理系统(DBMS)一样,多媒体 DBMS 也要进行数据的处理、查询和事物的管理等,但是多媒体数据库有不同的用户接口和存储构造,在多媒体 DBMS 中特别强调“媒体独立性”(所谓媒体独立性是指不论管理的多媒体数据的媒体如何变化,都不需改变 DBMS)。

(4) 模糊数据库

传统的数据库仅允许对精确的数据进行存储和处理,而客观世界中许多事物是不精确的。模糊数据库技术的研究和实践就是为了解决模糊数据的表达和处理问题,使得数据库描述的模型更自然、更贴切地反映客观世界。

(5) 网络数据库

网络数据库是数据库技术与 Web 技术相互融合的技术。

此外,还有其他一些新的数据库技术,如:模糊演绎数据库、主动数据库、集中式工程数据库和面向对象数据库等。

1.2 数据模型

模型是现实世界特征的模拟和抽象。数据模型也是一种模型,只不过它模拟的对象是数据。根据模型应用的不同层次和目的,可以将模型分为两类,第一类是概念模型,按用户的观点来对数据和信息建模,主要用于数据库设计;另一类数据模型,主要包括网状模型、层次模型和关系模型等,它是按计算机系统的观点对数据建模。

1.2.1 数据模型的概念

数据模型是现实世界数据特征的抽象。数据模型是工具,是用来抽象、表示和处理现实世界中的数据和信息的工具。在数据库中用数据模型这个工具来抽象、表示和处理现实世界的数据和信息,现有数据库系统均是基于某种数据模型的。

数据模型应满足三个方面的要求:

- (1) 能够比较真实地模拟现实世界;
- (2) 容易被人理解;
- (3) 便于在计算机系统中实现。

1.2.2 数据模型的组成要素

数据模型是由数据结构、数据操作和数据的约束条件三部分组成的。

数据结构是所研究对象的集合, 这些对象是数据库的组成成分, 例如表中的字段、名称等。数据结构分为两类, 一类是与数据类型、内容、性质有关的对象; 一类是与数据之间联系有关的对象。

数据操作是指对数据库中各种对象(型)的实例(值)允许执行的操作的集合, 包括操作及有关的操作规则。数据库的操作主要有检索和更新两大类。数据模型必须定义数据操作的确切含义、操作符号、操作规则以及实现操作的语言。

数据的约束条件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制约和依存规则, 用以限定符合数据模型的数据库状态以及状态的变化, 以保证数据的正确、有效和相容。

数据模型给出了在计算机系统上描述和动态模拟现实数据及其变化的一种抽象方法, 数据模型不同, 描述和实现方法亦不相同, 相应的支持软件, 即数据库管理系统也就不同。严格地讲, 一个数据模型应由上述三部分组成, 但数据结构是数据模型的本质标志。

1.2.3 概念模型

概念模型是现实世界到信息世界的第一层抽象, 是现实世界到计算机的一个中间层次。概念模型是数据库设计的有力工具和数据库设计人员与用户之间进行交流的语言。它必须具有较强的语义表达能力, 能够方便、直接地表达应用中的各种语义知识, 且简单、清晰、易于用户理解。

在现实世界中, 事物之间的联系是客观存在的。概念世界是现实世界在人们头脑中的反映, 是对客观事物及其联系的一种抽象描述, 不是现实世界的简单录像, 而是把现实世界中的客观对象抽象为某一种信任结构, 这种信任结构不是某一个 DBMS 支持的数据模型, 而是概念级的模型。

建立概念模型涉及以下几个术语:

(1) 实体(entity)

客观存在并可相互区别的事物称为实体。实体可以是实际事物, 也可以是抽象事件。比如, 一个职工、一个部门属于实际事物; 一次定货、借阅若干本图书、一场演出是比较抽象的事件。

同一类实体的集合称为实体集。例如, 全体学生的集合、全馆图书等。我们用命名的实体型表示抽象的实体集, 实体型“学生”表示全体学生的概念, 并不具体指职工甲或学生乙。

(2) 属性(attribute)

描述实体的特性称为属性。例如, 学生实体用若干个属性(学号、姓名、性别、出生日期、籍贯等)来描述。属性的具体取值称为属性值, 用以刻画一个具体实体。

(3) 关键字

如果某个属性或属性组合能够唯一地标识出实体集中的各个实体, 可以选作关键字, 也称为码。

(4) 联系(relationship)

实体集之间的对应关系称为联系, 它反映现实世界事物之间的相互关联。联系分为两种, 一种是实体内部各属性之间的联系, 另一种是实体之间的联系。

(5) E-R 图

概念模型的表示方法有很多, 常用实体-联系方法(E-R 方法或 E-R 图)来描述现实世界的概念模型, E-R 方法也称为 E-R 模型。

E-R 图有 3 个要素:

实体——用矩形并在框内标注实体名称来表示;

属性——用椭圆形表示, 并用连线将其与相应的实体连接起来;

联系——用菱形表示, 菱形框内写明联系名, 并用连线分别与有关实体连接起来, 同时在连线上标上联系的类型(1:1、1:n 或 m:n)。图 1.2 为 E-R 图的示例。

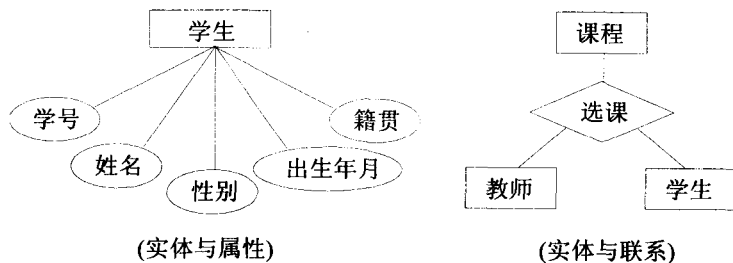


图 1.2 E-R 图示例

联系有三种类型:

(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。

(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 中有 n 个实体($n \geq 0$)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 有一对多联系, 记为 1:n。

(3) 多对多联系(m:n)

如果对于实体集 A 中的每一个实体, 实体集 B 中有 n 个实体($n \geq 0$)与之联系, 反之, 对于实体集 B 中的每一个实体, 实体集 A 中也有 m 个实体($m \geq 0$)与之联系, 则称实体集 A 与实体集 B 具有多对多联系, 记为 m:n。

1.2.4 常用的数据模型

每个数据库管理系统都是基于某种数据模型的。在目前数据库领域中，常用的数据模型有4种：层次模型、网状模型、关系模型和面向对象模型。

1. 层次模型

层次和网状模型是最早用于数据库系统的数据模型。层次模型的基本数据结构是层次结构，也称树型结构，树中每个结点表示一个实体类型，这些结点应满足：

- (1) 有且只有一个结点无双亲结点，这个结点称为根结点；
- (2) 其他结点有且仅有一个双亲结点。

在层次结构中，每个结点表示一个记录类型(实体)，结点之间的连线(有向边)表示实体间的联系。现实世界中许多实体间存在着自然的层次关系，如组织机构、家庭关系和物品分类等。图 1.3 就是一个层次模型的例子。

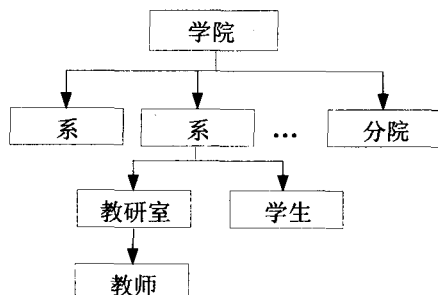


图 1.3 层次模型示例

2. 网状模型

网状模型的数据结构是一个网络结构。在数据库中，把满足以下两个条件的基本层次联系集合称为网状模型：

- (1) 一个结点可以有多个双亲结点；
- (2) 多个结点可以无双亲结点。

在网状模型中每个结点表示一个实体类型，结点间的连线表示实体间的联系。与层次模型不同，网状模型中的任意结点间都可以有联系，适用于表示多对多的联系，因此，与层次模型相比网状模型更具有普遍性。

网状模型虽然可以表示实体间的复杂关系，但它与层次模型没有本质的区别，它们都用连线表示实体间的联系，在物理实现上也有许多相同之处，如都用指针表示实体间的联系。层次模型是网状模型的特例，它们都称为格式化的数据模型。图 1.4 就是一个网状层次模型的例子。