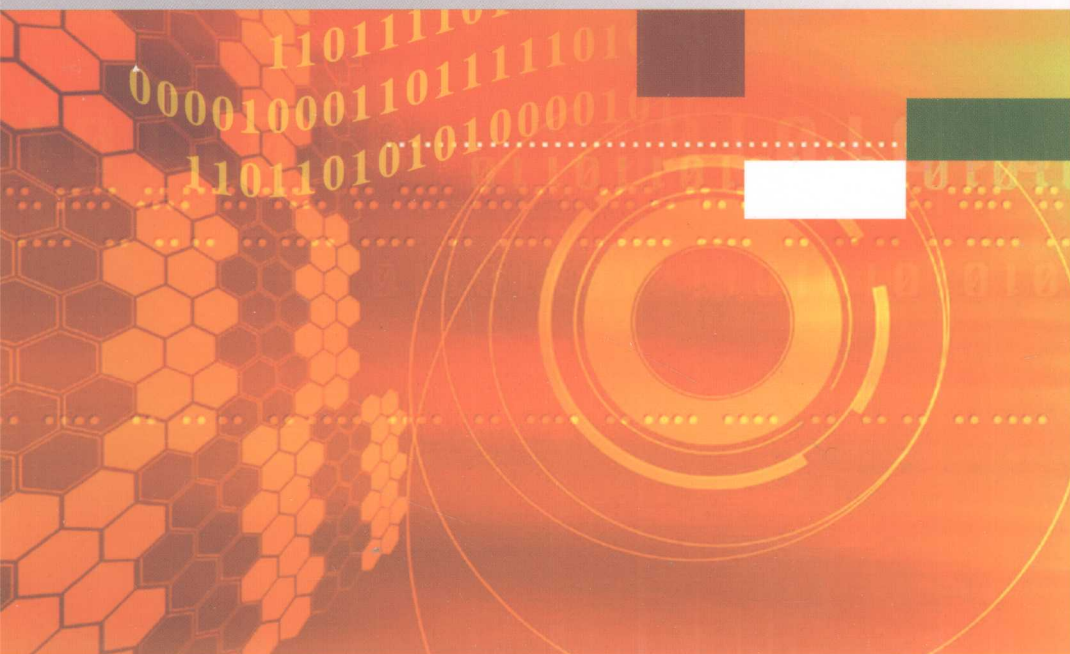




普通高等教育“十一五”规划教材



数据库原理及应用

(Access实现)

余建坤 李春宏/主编



科学出版社

www.sciencep.com

普通高等教育“十一五”规划教材

数据库原理及应用

(Access 实现)

余建坤 李春宏 主编

科学出版社

北京

内 容 简 介

本书是根据教育部《关于进一步加强高等学校计算机基础教学的意见暨计算机基础课程教学基本要求》而编写的教材,以 Microsoft Access 2003 (中文版)数据库系统作为教学数据库,结合非计算机专业学生的特点,以培养学生实践能力为目的,突出应用性和实用性。本书主要内容包括 Access 数据库基础、数据库和表、查询和 SQL 语言、窗体设计、报表设计、宏设计、VBA 应用、数据库安全及管理、数据交换等知识。

本书实例丰富,从易学性和实用性出发,把抽象的数据库原理有机地融入 Access 的具体操作之中,将应用实例和 Access 的相关功能紧密结合,通过任务驱动的方式,循序渐进地讲述了 Access 关系型数据库系统的特点和数据库应用开发技术。

本书可作为普通高等院校非计算机专业学生学习数据库理论和应用的教材,同时也可作为 Access 数据库应用技术培训及全国计算机等级考试(二级 Access)的参考用书。

图书在版编目(CIP)数据

数据库原理及应用: Access 实现/余建坤,李春宏主编. —北京:科学出版社,2008

(普通高等教育“十一五”规划教材)

ISBN 978-7-03-022700-3

I. 数… II. ①余…②李… III. 关系数据库-数据库管理系统, Access-高等学校-教材 IV. TP311.138

中国版本图书馆 CIP 数据核字(2008)第 119135 号

责任编辑:陈晓萍/责任校对:柏连海

责任印制:吕春珉/封面设计:耕者设计工作室

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

铭浩彩色印装有限公司印刷

科学出版社发行 各地新华书店经销

*

2008 年 8 月第 一 版 开本:787×1092 1/16

2008 年 8 月第一次印刷 印张:14 1/2

印数:1—4 000 字数:343 000

定价:32.00 元(共二册)

(如有印装质量问题,我社负责调换<环伟>)

销售部电话 010-62134988 编辑部电话 010-62138978-8003

版权所有,侵权必究

举报电话:010-64030229; 010-64034315; 13501151303

前 言

随着信息技术和社会信息化的发展,以数据库系统为核心的办公自动化系统、信息管理系统、决策支持系统等得到了广泛应用,数据库技术已成为计算机应用的一个重要方面。数据库原理及应用已是高等学校中非计算机专业,尤其是经济类、管理类专业的一门重要公共课程。为配合教育部《关于进一步加强高等学校计算机基础教学的意见暨计算机基础课程教学基本要求》(白皮书(06版))中所提出的要求,本书以任务为驱动,以事例为主线,通过案例式教学,全面讲述关系数据库系统的特点及应用开发技术,旨在提高学生的数据库操作能力和应用能力。

Access 2003 是 Microsoft Office 2003 办公软件的一个重要组成部分,是目前较为流行的、功能强大的桌面数据库管理系统,它是完全面向对象、采用事件驱动机制的关系型数据库系统,使数据库的应用和开发变得更加快捷和灵活。Access 数据库中包含表、查询、窗体、报表、页、宏和模块等数据库对象。Access 的用途非常广泛,既可以创建个人账簿、通信录等个人使用的简单数据库,又可以创建商贸、办公管理等复杂数据库。与 SQL Server 结合还可以创建客户机/服务器数据库应用程序。Access 2003 不但与其他 Office 组件在许多特性上保持一致,可方便地在 Office 组件之间交换数据,使用户更容易操作,同时也具有和 Office 其他组件相同的操作界面和环境,使 Access 易学易用,反映了数据库技术的发展动向和特点。

本书以 Access 2003 作为应用环境,介绍了数据库原理及应用的基本理论和基本方法。全书共有 9 章,各章内容如下:

第 1 章重点介绍了数据库技术的发展、数据库的基本概念、关系数据模型、数据库体系结构、数据库设计基础等内容。

第 2 章对 Access 功能、界面、数据库构成及教材所用案例——罗斯文商贸数据库等内容进行了介绍,使读者对 Access 有一个大体的了解。

第 3~7 章对创建 Access 数据库和表,创建查询、窗体、报表、宏和 VBA 等内容进行了重点介绍,上述章节也是 Access 的基本功能所在。

第 8 章重点介绍了数据交换(数据的导入导出)、数据访问页及数据库安全管理等方面的内容。

第 9 章以罗斯文商贸系统为例,详细介绍了数据库应用系统的开发过程和实现。

结合普通高等学校非计算机专业学生的特点,本书以应用为目的,以案例为引导,章节内容安排循序渐进,阐述由浅入深,层次分明,始终围绕罗斯文商贸系统这个典型的事例进行详细介绍,操作步骤翔实、丰富和具体,最后形成一个完整的商贸数据库管理系统。

为了便于实验教学和学生的学习,同时还编写了与本教材配套的实践教程。

本书由余建坤和李春宏任主编，第1章由余建坤编写，第2章由陈振兴编写，第3章由曾志勇编写，第4章由陶冶编写，第5章由周荣华编写，第6章由尹传娟编写，第7章由谭瑛编写，第8章由李春宏、王冬编写，第9章由冯涛编写。全书由余建坤和李春宏统稿和定稿。

由于时间紧迫，编者水平有限，书中难免有错误和疏漏之处，诚请专家、教师 and 广大读者批评指正。

目 录

第 1 章 概论	1
1.1 数据管理技术的发展	1
1.1.1 人工管理阶段	2
1.1.2 文件系统阶段	2
1.1.3 数据库系统阶段	3
1.1.4 分布式数据库系统阶段	4
1.1.5 数据仓库与数据挖掘阶段	5
1.2 数据库系统	6
1.2.1 数据库与数据库管理系统	6
1.2.2 数据库系统	7
1.3 数据模型	9
1.3.1 数据和数据联系的描述	9
1.3.2 数据模型	12
1.4 关系数据模型	13
1.4.1 关系数据模型的数据结构	13
1.4.2 关系数据模型的运算	15
1.4.3 关系数据模型的完整性规则	16
1.5 数据库系统结构	17
1.5.1 数据库系统内部结构	18
1.5.2 数据库系统外部结构	19
1.6 建立关系数据库	20
1.6.1 关系规范化	20
1.6.2 建立关系数据库	21
1.6.3 一个实例	22
本章小结	24
习题	24
第 2 章 Access 功能浏览	26
2.1 Access 开发环境	26
2.1.1 Access 简介	26
2.1.2 Access 的主要用途	27
2.1.3 Access 2003 的操作窗口	27
2.1.4 Access 的数据库构成	28

2.2	应用案例——罗斯文商贸数据库	31
2.2.1	罗斯文商贸数据库需求分析	31
2.2.2	罗斯文商贸系统设计	32
	本章小结	34
	习题	34
第3章	数据库和表	35
3.1	数据库的打开与关闭	35
3.1.1	打开数据库	35
3.1.2	关闭数据库	39
3.2	创建数据库	39
3.2.1	使用“数据库向导”创建数据库	39
3.2.2	不使用“数据库向导”创建空数据库	41
3.3	创建表	44
3.3.1	使用设计器创建表	45
3.3.2	通过直接输入数据的方式来创建表	50
3.3.3	使用向导创建表	54
3.4	数据的导入和导出	54
3.4.1	导出数据	54
3.4.2	导入数据	57
3.5	常用表数据操作	60
3.5.1	记录排序	60
3.5.2	记录筛选	62
3.6	表的维护	63
3.6.1	表结构的维护	63
3.6.2	创建表之间的关联关系	66
3.6.3	使用子表	68
	本章小结	69
	习题	70
第4章	查询设计和 SQL 语言	71
4.1	查询概念	71
4.1.1	查询	71
4.1.2	记录集	72
4.1.3	查询种类	72
4.1.4	查询视图	72
4.2	用查询向导创建查询	72
4.2.1	简单查询向导	72
4.2.2	交叉表查询向导	75

4.2.3	查找重复项查询向导.....	78
4.2.4	查找不匹配项查询向导.....	80
4.3	用设计视图创建和修改查询.....	83
4.3.1	用设计视图创建查询.....	83
4.3.2	对查询结果排序.....	85
4.3.3	使用准则筛选记录.....	87
4.3.4	查询属性.....	90
4.4	使用查询进行统计计算.....	91
4.5	操作查询.....	92
4.5.1	生成表查询.....	92
4.5.2	更新查询.....	93
4.5.3	删除查询.....	94
4.5.4	追加查询.....	95
4.6	参数查询.....	96
4.7	SQL 查询.....	97
4.7.1	创建 SQL 查询.....	97
4.7.2	联合查询.....	98
	本章小结.....	98
	习题.....	99
第 5 章	窗体设计.....	100
5.1	认识窗体.....	100
5.1.1	窗体概念和功能.....	100
5.1.2	窗体类型.....	101
5.1.3	窗体结构.....	104
5.1.4	窗体使用.....	105
5.2	用向导创建窗体.....	107
5.2.1	创建自动窗体.....	107
5.2.2	通过窗体向导创建窗体.....	108
5.2.3	创建图表窗体.....	110
5.3	用“设计视图”创建窗体.....	112
5.3.1	窗体设计视图.....	112
5.3.2	窗体设计工具.....	113
5.3.3	窗体控件的使用.....	117
5.3.4	修饰窗体.....	123
5.3.5	使用控件设计窗体示例.....	124
5.4	子窗体.....	126
5.5	定制用户界面.....	128
5.5.1	创建切换面板.....	128

5.5.2 设置启动式窗体.....	129
5.6 数据透视图表窗体.....	130
5.6.1 数据透视表窗体.....	130
5.6.2 数据透视图窗体.....	131
本章小结.....	131
习题.....	132
第 6 章 报表设计	133
6.1 报表概述.....	133
6.1.1 Access 中报表的分类.....	133
6.1.2 了解报表节.....	135
6.1.3 报表的视图方式.....	137
6.2 报表的创建.....	137
6.2.1 使用“自动创建报表”的方法建立报表.....	137
6.2.2 使用向导创建报表.....	139
6.2.3 在设计视图中手动创建报表.....	151
6.3 报表的修改和美化.....	153
6.3.1 在报表设计器中修改已有报表.....	153
6.3.2 在报表设计器中美化已有报表.....	157
6.3.3 应用举例——修改美化报表.....	158
6.4 报表的打印.....	162
6.4.1 页面设置.....	162
6.4.2 预览及打印.....	163
本章小结.....	164
习题.....	164
第 7 章 宏与 VBA	165
7.1 宏的基本概念.....	165
7.2 宏的基本操作.....	165
7.2.1 创建宏.....	165
7.2.2 宏组.....	166
7.2.3 条件操作宏.....	168
7.2.4 运行宏.....	170
7.2.5 宏的调试.....	171
7.2.6 宏的编辑与修改.....	171
7.2.7 常用宏操作.....	172
7.3 VBA 简介.....	173
7.4 宏与 VBA.....	173
7.5 VBA 语法基础.....	174

7.5.1	常量和变量.....	174
7.5.2	程序结构.....	176
7.5.3	VBA 过程和函数.....	179
	本章小结.....	182
	习题.....	182
第 8 章	Access 高级应用.....	183
8.1	数据的导入.....	183
8.1.1	从 Access 数据库导入对象.....	183
8.1.2	从文本文件导入数据.....	184
8.1.3	从 Word 导入数据.....	187
8.1.4	从 XML 导入数据.....	187
8.2	数据的导出.....	188
8.2.1	导出数据到 Access 数据库.....	189
8.2.2	导出到 XML.....	189
8.3	数据链接.....	191
8.3.1	数据链接与导入的区别.....	191
8.3.2	数据导入或链接的选用规则.....	191
8.3.3	数据链接的操作过程.....	192
8.3.4	链接表的管理.....	193
8.4	数据访问页.....	194
8.4.1	数据访问页概述.....	194
8.4.2	窗体、报表与数据访问页的比较.....	194
8.4.3	创建数据访问页.....	194
8.4.4	编辑数据访问页.....	198
8.5	数据库安全与管理.....	203
8.5.1	数据库安全措施.....	203
8.5.2	设置数据库密码.....	203
8.5.3	用户级安全.....	204
8.5.4	数据库编码/解码.....	204
8.5.5	生成 MDE 文件.....	205
8.5.6	设置“启动”选项.....	205
8.5.7	数据存储安全.....	206
8.5.8	数据库拆分.....	206
8.5.9	优化数据库性能.....	206
8.5.10	复制与同步数据库.....	207
8.5.11	数据库转换.....	207
	本章小结.....	207

习题	207
第 9 章 罗斯文商贸系统的实现	208
9.1 系统简介	208
9.2 系统功能概览	209
9.2.1 “主切换面板”界面	209
9.2.2 “产品类别”界面	209
9.2.3 “供应商”界面	209
9.2.4 “产品信息”界面	209
9.2.5 “订单查询”界面	210
9.2.6 “打印销售额报表”界面	210
9.2.7 “显示数据库窗口”功能	212
9.2.8 “退出 Microsoft Access”功能	212
9.3 系统功能实现	212
9.3.1 创建数据库	212
9.3.2 创建数据库表	213
9.3.3 功能窗体创建	213
本章小结	220
习题	221
参考文献	222

第 1 章 概 论

[本章导读] 数据库技术是计算机应用领域中最重要的技术之一，是软件学科的一个独立分支。本章将介绍数据管理技术的发展过程及数据库技术所涉及的基本概念，包括数据库、数据模型、数据库系统的体系结构、关系数据库的基本理论，最后给出建立关系数据库的方法及实例，通过对本章的学习使读者对数据库技术有一个全面的了解。

数据库技术是信息社会中信息资源管理与利用的基础，是计算机软件学科的一个重要分支，是研究如何存储、使用和管理数据的一门学科。随着计算机应用技术的提高，数据库应用领域已从数据处理、信息管理、事务处理扩大到计算机辅助设计、人工智能、办公信息系统和网络应用等新的应用领域。

经过 30 多年的发展，数据库技术已形成完整的理论体系和一大批实用系统。关系运算理论和模式设计理论不断完善，数据库管理系统软件日益丰富，为数据库的应用与开发奠定了基础。本章将介绍数据管理的发展过程及数据库系统所涉及的基本概念，包括数据库、数据模型、数据库系统的体系结构、关系数据库的基本理论，最后给出建立关系数据库的方法及实例。

1.1 数据管理技术的发展

数据管理技术是指对数据的组织、分类、编码、存储、检索和维护。与任何其他技术的发展一样，计算机数据管理也经历了由低级到高级的发展过程。计算机数据管理技术随着计算机硬件（主要是外存储器）、软件技术和计算机应用范围的发展不断发展，多年来大致经历了以下五个阶段。

- 人工管理阶段。
- 文件系统阶段。
- 数据库系统阶段。

- 分布式数据库系统阶段。
- 数据仓库与数据挖掘阶段。

1.1.1 人工管理阶段

20 世纪 50 年代中期以前, 计算机主要用于科学计算。当时在硬件方面, 外存储器只有卡片、纸带、磁带, 没有像磁盘这样的可以随机访问、直接存取的外部存储设备; 软件方面, 没有专门管理数据的软件, 数据由计算或处理它的程序自行携带, 数据处理方式基本是批处理, 这种方式使得程序依赖于数据, 如果数据的类型、格式或者数据量、存取方法、输入输出方式等改变了, 程序必须作相应的修改, 数据与程序不具有独立性; 由于数据是面向应用程序的, 在一个程序中定义的数据, 程序运行结束后就退出计算机系统, 数据占用的空间随程序空间一起被释放, 一个程序中的数据无法被其他程序利用, 因此程序与程序之间存在大量的重复数据; 另一方面, 由于系统中没有对数据进行管理的软件, 对数据管理任务, 包括存储结构、存取方法、输入输出方式等完全由程序设计人员自负其责, 这就给应用程序设计人员增加了很大的负担。

1.1.2 文件系统阶段

20 世纪 50 年代后期至 60 年代中后期, 计算机开始大量地用于管理中的数据处理工作, 大量的数据存储、检索和维护成为紧迫的需求。在硬件方面, 可直接存取的磁鼓、磁盘成为联机的主要外存; 在软件方面, 出现了高级语言和操作系统, 操作系统中的文件系统 (有的也称为信息处理模块) 是专门管理外存的数据管理软件, 数据处理方式有批处理, 也有联机实时处理。

在这一阶段, 程序与数据有了一定的独立性, 程序和数据分开存储, 有了程序文件和数据文件的区别。数据文件可以长期保存在外存储器上多次存取, 如进行查询、修改、插入、删除等操作。数据的存取以记录为基本单位, 并出现了多种文件组织形式, 如顺序文件、索引文件、随机文件等。

在文件系统的支持下, 数据的逻辑结构与物理结构之间可以有一定的差别, 逻辑结构与物理结构之间的转换由文件系统的存取方法来实现。数据与程序之间有设备独立性, 程序只需用文件名访问数据, 不必关心数据的物理位置。这样, 程序员可以集中精力在数据处理的算法上, 而不必考虑数据存储的具体细节, 如图 1.1 所示。

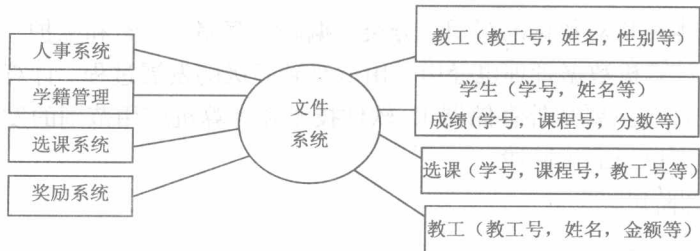


图 1.1 文件管理中数据与程序的关系

文件系统阶段对数据的管理虽然有了长足的进步,但一些根本性问题仍然没有彻底解决,主要表现在以下三个方面。

(1) 数据冗余度大

数据冗余是指不必要的重复存储,同一数据项重复出现在多个文件中。在文件系统下,数据文件基本上与各自应用程序相对应,数据不能以记录和数据项为单位共享。即使有部分数据相同,只要逻辑结构不同,用户必须各自建立自己的文件,这不仅浪费存储空间,增加更新开销,更严重的是,由于不能统一修改,容易造成数据的不一致性。

(2) 数据无集中管理

除了对记录的存取由文件系统承担以外,文件没有统一的管理机制,其安全性与完整性无法保障。数据的维护任务仍然由应用程序来承担。

(3) 文件是无弹性、无结构的数据集合

所谓无弹性,是指由于记录的内部结构是由应用程序自己定义,而不是由系统来统一管理,因此对现有数据文件的应用不易扩充、不易移植,也难以增、删数据项以适应新的应用要求;无结构是指各个数据文件之间是独立的,缺乏联系,不能反映现实世界事物之间的联系。

这些问题阻碍了数据管理技术的发展,不能满足日益增长的信息需求,这正是数据库技术产生的原动力,也是数据库系统产生的背景,应用需求和计算机技术的发展促使人们研究一种新的数据管理技术——数据库技术。

1.1.3 数据库系统阶段

从20世纪60年代后期开始,计算机应用于管理的规模更加庞大,需要计算机管理的数据量急剧增长,并且对数据共享的需求日益增强,大容量磁盘(数百兆字节以上)系统的采用,使计算机联机存取大量数据成为可能;软件价格上升,硬件价格相对下降,使独立开发系统维护软件的成本增加,文件系统的管理方法已无法适应开发应用系统的需要。为解决数据的独立性问题,实现数据的统一管理,达到数据共享的目的,发展了数据库技术。

数据库(Database, DB)是通用化的相关数据集合,它不仅包括数据本身,而且包括相关数据之间的联系。数据库中的数据不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。例如,某个企业、组织或行业所涉及的全部数据的汇集。其数据结构独立于使用数据的程序,对于数据的增加、删除、修改和检索由系统进行统一控制,而且数据模型也有利于将来应用的扩展。

为了让多种应用程序并发地使用数据库中具有最小冗余度地共享数据,必须使数据与程序具有较高的独立性。这就需要有一个软件系统对数据实行专门管理,提供安全性和完整性等统一控制机制,方便用户以交互命令或程序方式对数据库进行操作。

为数据库的建立、使用和维护而配置的软件称为数据库管理系统(DataBase Management System, DBMS),它是在操作系统支持下运行的。目前较流行的数据库管理系统包括Oracle、Informix、Sybase、DBII等大型数据库管理系统和在微机上应用较

广泛的数据库管理系统 Access 2003、Visual FoxPro 6.0、SQL Server 2005 等。

现在, 数据库已成为各类信息系统的核心。在数据库管理系统支持下, 数据与程序的关系如图 1.2 所示。

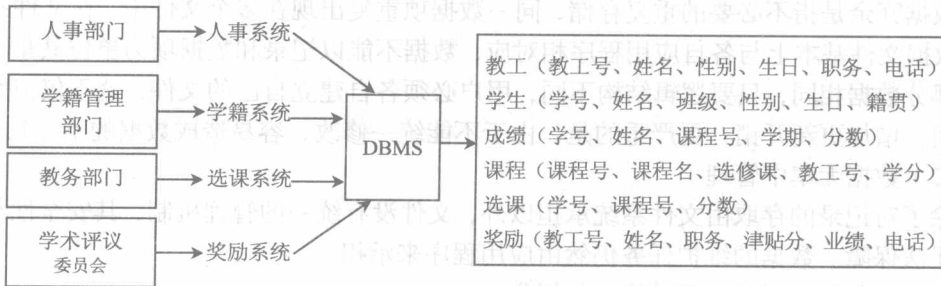


图 1.2 数据库管理系统中数据与程序的关系

数据库系统的主要特点如下。

(1) 实现数据共享, 减少数据冗余

在数据库系统中, 对数据的定义和描述已经从应用程序中分离出来, 通过数据库管理系统来统一管理。数据的最小访问单位是数据项, 既可以按数据项的名称存取数据库中某一个或某一组数据项, 也可以存取一条记录或一组记录。

建立数据库时, 应当以面向全局的观点组织数据库中的数据, 而不能像文件系统那样仅仅考虑某一部门的局部应用, 这样才能发挥数据共享的优势。

(2) 采用特定的数据模型

整个组织的数据不是一盘散沙, 必须表示出数据之间所存在的有机关联, 才能反映现实世界事物之间的联系。也就是说, 数据库中的数据是有结构的, 这种结构由数据模型表示出来 (如关系模型)。

(3) 具有较高的数据独立性

在数据库系统中, 数据库管理系统提供映像的功能, 确保应用程序对数据结构和存取方法有较高的独立性。数据的物理存储结构与用户看到的逻辑结构可以有很大差别。用户只以简单的逻辑结构来操作数据, 无需考虑数据在存储器上的物理位置与结构。

(4) 有统一的数据控制功能

数据库作为多个用户和应用程序的共享资源, 对数据的存取往往是并发使用, 即多个用户同时使用同一个数据库。数据库管理系统必须提供并发控制功能、数据的安全性控制功能和数据的完整性控制功能。

1.1.4 分布式数据库系统阶段

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节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, 如图 1.3 所示。

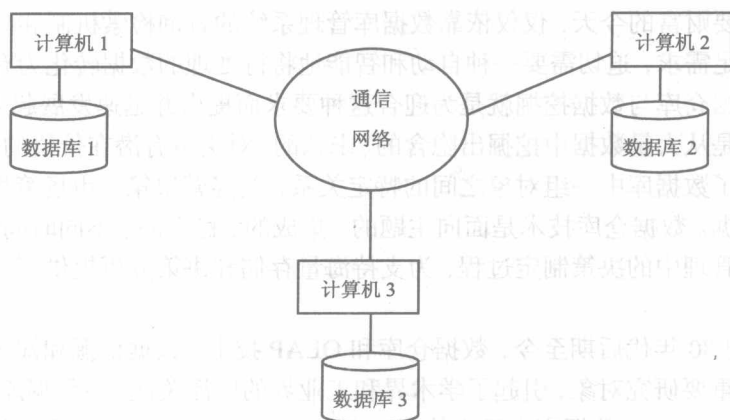


图 1.3 分布式数据库系统

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计算机网络的各个节点上。每个节点有自己的局部数据库管理系统，它有很高的独立性。用户可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。其主要优点如下。

(1) 局部自主

网络上每个节点的数据库系统都具有独立处理本地大量事务的能力，而且各局部节点之间也能够互相访问、有效地配合处理更复杂的事务。

(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，在个别节点或个别通信链路发生故障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，只要有一个节点上的数据备份可用，则数据是可用的。可见，支持一定程度的数据冗余是充分发挥分布式数据库系统优点的先决条件之一。

(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。数据可以存储在邻近的常用节点上，如果本节点的数据子集包含了要查询的全部内容，显然比集中式数据库在全集上查找节省时间。

1.1.5 数据仓库与数据挖掘阶段

随着数据库应用技术的日趋成熟，大量管理信息系统在企事业单位得到了广泛的应用，人们积累了大量的数据资料，但数据库中隐藏的丰富知识远远没有得到充分地发掘和利用。随着 Internet 的日益普及，每个网络用户都可以借助先进的通信手段，获取信息或进行信息发布，促进了信息的倍增，信息量呈几何放大式增长，在这样大量的信息环境中，如何提取有用信息、抛却冗余信息，已成为信息管理者日益关注的问题。在知

识是全世界主要财富的今天,仅仅依靠数据库管理系统的查询检索机制和统计学方法已经远远不能满足需求,迫切需要一种自动和智能地将待处理的数据转化为有用信息和知识的技术,数据仓库与数据挖掘就是为迎合这种要求而提出并迅速发展起来的。

数据挖掘是从大量数据中挖掘出隐含的、未知的、对决策有潜在价值的知识和规则,这些规则蕴含了数据库中一组对象之间的特定关系,为经营决策、市场策划、政策法规制定等提供依据。数据仓库技术是面向主题的、集成的、稳定的、不同时间的数据集合,用以支持经营管理中的决策制定过程,为支持海量存储和决策分析提供了一种很好的解决方案。

从 20 世纪 80 年代后期至今,数据仓库和 OLAP 技术、数据挖掘和知识发现已成为数据库技术的重要研究对象,引起了学术界和工业界的广泛关注,在数据库产品 Oracle、SQL Server 2005 等大型数据库中已有体现,IBM Almaden 和 GTE 及众多的学术单位都在这个领域开展了各种各样的研究计划,研究的主要目标是发展有关的方法论、理论和工具,以支持从大量数据中提取有用的知识和模式。

1.2 数据库系统

1.2.1 数据库与数据库管理系统

数据库是长期储存在计算机内、有组织的、可共享的大量数据集合。数据库具有下列特征:

- 1) 数据按一定的数据模型组织、描述和存储。
- 2) 可为各种用户共享。
- 3) 冗余度较小。
- 4) 数据独立性较高。
- 5) 易扩展。

数据库管理系统是位于用户与操作系统之间的一层数据管理软件。常见的数据库管理系统软件有桌面型的 Visual FoxPro、Access,大型的 Oracle、SQL Server、MySQL 等,其主要功能如下。

(1) 数据库定义功能

提供数据定义语言(Data Description Language, DDL)或者操作命令以便对各级数据模式进行精确地描述。为此,系统必须包含 DDL 的编译程序或解释程序。

(2) 数据库操纵功能

为了对数据库中的数据进行追加、插入、修改、删除、检索等操作,数据库管理系统提供的语言或者命令,称为数据操纵语言(Data Manipulation Language, DML)。不同的数据库管理系统语言的语法格式可以不相同,从实现方法来讲,DML 可分为两种类型。一类 DML 不依赖于任何程序设计语言,可以独立交互式使用,称为自含型或自