

SQL Server 2000 数据库技术 实用教程

主 编：岳国英
副主编：田玉珍 王樱 孟宪明

- ◆ 数名一线教师多年教学经验集萃
- ◆ 图文并茂、条理清晰、易教易学
- ◆ 精心设计上机实验与习题
- ◆ 免费提供PPT格式电子教案



21 世纪高职高专规划教材 计算机系列

内容提要

SQL Server 2000 数据库技术 实用教程

主 编：岳国英
副主编：田玉珍 王樱 孟宪明



中国电力出版社
www.infopower.com.cn

定价：24.00 元
ISBN 7-309-05963-2

内容提要

本书是 21 世纪高职高专规划教材·计算机系列中的一本。

全书共分 11 章,主要内容分两部分:一部分讲述了数据库概述、关系数据库以及其标准语言 SQL 等;另一部分讲述了 SQL Server 2000 的基本知识以及基本操作与应用、SQL Server 2000 的数据查询、SQL Server 2000 Transact-SQL 编程和应用、数据完整性及 SQL Server 的完整性控制、数据库的安全性及 SQL Server 的安全管理、数据库的并发控制及 SQL Server 的并发控制机制、数据库恢复技术及 SQL Server 的数据恢复机制、数据库设计与应用等。本书将数据库技术与 SQL Server 2000 具体实现紧密结合起来,通过贯穿全书的示例数据库,使读者能够轻松地掌握数据库基本原理并应用于实际。通过活期储蓄案例,给出了数据库设计与应用的完整过程,使读者对数据库技术及其应用有一个较全面的认识和理解。为方便学习,每一章都精心设计了习题,并在相应章节安排了实验和课程设计内容,做到了学用结合,使读者能够迅速掌握相应知识。同时为了方便教学,本书配有 PPT 格式电子教案,免费为任课教师提供。

本书可作为高职高专院校相关专业学生学习数据库与 SQL Server 2000 课程的教材,也可作为从事相关工作的人员和广大电脑爱好者学习数据库与 SQL Server 2000 知识的自学教材或参考书。

图书在版编目(CIP)数据

SQL Server 2000 数据库技术实用教程 / 岳国英等编著. —北京:中国电力出版社, 2005

21 世纪高职高专规划教材·计算机系列

ISBN 7-5083-2963-5

I.S... II.岳... III.关系数据库—数据库管理系统, SQL Server 2000—高等学校:技术学校—教材

IV.TP311.138

中国版本图书馆 CIP 数据核字(2004)第 127866 号

丛 书 名: 21 世纪高职高专规划教材·计算机系列

书 名: SQL Server 2000 数据库技术实用教程

出版发行: 中国电力出版社

地 址: 北京市三里河路 6 号 邮政编码: 100044

电 话: (010) 68358031 (总机) 传 真: (010) 68316497, 88383619

本书如有印装质量问题, 我社负责退换

服务电话: (010) 88515918 (总机) 传 真: (010) 88518169

E-mail: infopower@cepp.com.cn

印 刷: 北京丰源印刷厂

开本尺寸: 185×233 **印 张:** 16.5 **字 数:** 368 千字

书 号: ISBN 7-5083-2963-5

版 次: 2005 年 2 月北京第 1 版

印 次: 2005 年 2 月第 1 次印刷

印 数: 0001—4000 册

定 价: 24.00 元

版权所有, 翻印必究

前 言

数据库技术的发展使计算机应用领域又拓展出一片崭新的、广阔的空间。数据库技术已成为计算机信息系统和计算机应用系统的核心与基础。因此,数据库技术是计算机及其相关专业的重要专业课程之一。

本书打破原理与实践相脱节的状况,将数据库基本原理与 SQL Server 2000 两部分内容进行合理的整合,做到了理论与实践紧密结合,并通过贯彻全书的示例和案例,使读者对抽象的理论能够轻松理解和融会贯通,并对利用数据库技术解决实际问题有一个完整的认识和理解。从内容安排上,体现了关系数据库管理系统的理论是共同的、实现只是命令级的差异的理念。学习数据库基本原理的目的是为了掌握数据库技术,学习 SQL Server 2000 不仅因为它是目前最流行、应用最广泛的关系数据库管理系统,更重要的是将 SQL Server 2000 作为实现数据库技术的工具,起到触类旁通的作用。

本书由浙江水利水电专科学校岳国英教授担任主编,郑州电力高等专科学校田玉珍、南昌工程学院王樱、山东水利职业学院孟宪明等老师担任副主编,参加编写的还有郑州电力高等专科学校刘忠箐老师。本书第 1 章、第 2 章由孟宪明和岳国英编写,第 3 章、第 6 章和实训由王樱编写,第 4 章、第 5 章由刘忠箐编写,第 7 章、第 8 章和第 11 章由田玉珍编写,第 9 章、第 10 章和案例由岳国英编写,全书由岳国英修改并统稿。

浙江水利水电专科学校王英华教授、李大庆副教授对本书的部分内容提出了宝贵意见,在此,一并表示诚挚的谢意。

本书适合作为高等职业技术学校、普通高等院校计算机及其相关专业培养应用型人才的教材,也可作为计算机应用系统开发与设计者学习数据库知识的参考用书。

由于作者水平有限,加之时间仓促,书中错误和不足之处在所难免,恳请广大读者批评指正。也可通过电子邮件与我们联系。

E-mail: yuegy@zjwhc.com。

作 者

2004 年 12 月

21 世纪高职高专规划教材·计算机系列

编 委 会

主任委员:

宗 健 岳国英

副主任委员:(以姓氏笔画为序)

丁亚明 马敬卫 王树勇 王晓光 冯玉东 刘广峰

朱世同 刘克兴 刘治安 齐现伟 孙奕学 孙春临

孙 辉 陈 东 李亚生 陈希球 陈 炜 寿建平

罗 众 林逢春 崔凤磊 黄华国 彭同明

委 员:(以姓氏笔画为序)

马冬生 万朝阳 王卫东 王建华 王展运 石文华

付晓波 朱卫红 安丰彩 吕 来 刘 阳 李大庆

何万敏 陈忠文 张国锋 李 娜 张海波 陈 智

罗亚东 胡文红 姚发洲 侯仰东 胡顺增 秦昌平

康玉忠 黄泽均 黄逵中 梁 曦 廖立军

21 世纪高职高专规划教材参编院校

(排名不分先后)

保定电力职业技术学院
山东电力高等专科学校
黄河水利职业技术学院
湖北水利水电职业技术学院
长江工程职业技术学院
郑州电力高等专科学校
武汉电力职业技术学院
江西电力职业技术学院
浙江水利水电高等专科学校
福建水利电力职业技术学院
广东水利电力职业技术学院
四川水利职业技术学院
兰州电力技术学院
兰州电力学校
南昌水利水电高等专科学校
贵州电力职业技术学院
福建电力职业技术学院
广西电力职业技术学院
内蒙古电力学校
浙江电力职业技术学院
四川电力职业技术学院
石家庄职业技术学院
秦皇岛职业技术学院
唐山工业职业技术学院
唐山科技职业技术学院
天津职业大学
天津大学职教学院

天津理工大学职业技术学院
北京科技大学(管庄校区)
天津渤海职业技术学院
天津轻工职业技术学院
天津中德职业技术学院
天津石油职业技术学院
北京联合大学
太原理工大学
长治职业技术学院
湖南工业职业技术学院
广西工学院职业技术学院
苏州职业大学
南通职业大学
常熟理工学院
常州工学院
徐州工程学院
常州纺织服装职业技术学院
常州轻工职业技术学院
常州信息职业技术学院
连云港职业技术学院
南京工程学院
武汉公交职业技术学院
湖北轻工职业技术学院
武汉职业技术学院
四川工程职业技术学院
四川托普信息技术职业学院
泸州职业技术学院

目 录

前 言	
第 1 章 数据库技术基础知识	1
1.1 数据库技术概述	1
1.2 数据模型	6
1.3 数据库设计	9
1.4 概念模型	10
习题	14
第 2 章 关系数据库	16
2.1 关系数据模型	16
2.2 E-R 模型到关系模型的转换	24
2.3 关系模型的三类完整性约束	26
2.4 案例: 活期储蓄管理系统数据库设计	28
习题	31
第 3 章 关系数据库标准语言 SQL	33
3.1 SQL 概述	33
3.2 数据定义	35
3.3 数据查询	40
3.4 数据更新	51
3.5 视图	54
3.6 案例: 活期储蓄管理系统数据库上的查询	56
习题	59
第 4 章 SQL Server 2000 基本知识	63
4.1 SQL Server 2000 简介	63
4.2 SQL Server 2000 的安装	63
4.3 SQL Server 2000 配置	70
4.4 SQL Server 2000 常用的管理器	75
习题	80
实训 SQL Server 2000 安装与配置	80
第 5 章 SQL Server 2000 基本操作与应用	82
5.1 SQL Server 2000 数据库基本操作	82
5.2 表和视图的基本操作	92
5.3 案例: 创建活期储蓄管理系统数据库——数据库的物理设计与实现	114
习题	117
实训 1 SQL Server 2000 系统目录和数据库的基本操作	117
实训 2 表和索引的基本操作	118
第 6 章 SQL Server 2000 的数据查询	120
6.1 简单查询	120
6.2 连接查询	125

6.3 子查询	133
6.4 联合查询	136
6.5 案例: 活期储蓄管理系统数据库上的一般操作	141
习题	143
实训 1 简单数据查询	143
实训 2 高级数据查询	144
第 7 章 SQL Server 2000 Transact-SQL 编程和应用	145
7.1 SQL Server 2000 Transact-SQL 编程	145
7.2 存储过程	163
7.3 触发器	173
7.4 案例: 建立活期储蓄管理系统数据库的触发器和存储过程	180
习题	184
实训 1 SQL Server 2000 Transact-SQL 编程	184
实训 2 存储过程和触发器的基本操作	185
第 8 章 数据库完整性及 SQL Server 的完整性控制	187
8.1 完整性约束条件及完整性控制	187
8.2 SQL Server 的数据完整性及其实现	188
8.3 案例: 活期储蓄管理系统数据库的完整性控制	194
习题	198
第 9 章 数据库的安全性及 SQL Server 安全管理	199
9.1 数据库安全性控制的一般方法	199
9.2 SQL Server 的安全体系结构	203
9.3 SQL Server 数据库安全性管理	205
习题	217
实训 1 视图的基本操作	217
实训 2 SQL Server 的安全管理	218
第 10 章 数据库并发控制及 SQL Server 的并发控制机制	220
10.1 事务及并发控制的基本概念	220
10.2 封锁机制	223
10.3 SQL Server 的并发控制机制	226
习题	228
第 11 章 数据库恢复技术与 SQL Server 的数据恢复机制	229
11.1 故障的种类	229
11.2 数据恢复的实现技术	230
11.3 数据库恢复策略	232
11.4 SQL Server 的数据备份和数据恢复机制	233
习题	243
附录 A SQL Server 2000 的常用的内置函数	244
附录 B SQL Server 2000 的常用系统表	248
附录 C SQL Server 2000 的常用系统存储过程	250
参考文献	254

第1章 数据库技术基础知识

数据库技术作为数据管理的实现技术,已成为计算机应用技术的核心。随着计算机技术、通信技术、网络技术的迅速发展,人类社会进入了信息时代。建立一个行之有效的管理信息系统已成为每个企业或组织生存和发展的重要条件。从某种意义上讲,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 数据库技术概述

1.1.1 数据库基本概念

1. 数据

数据(Data)是描述现实世界事物的符号记录,是用物理符号记录的可以鉴别的信息。物理符号有多种表现形式,包括:数字、文字、图形、图像、声音及其他特殊符号。数据的各种表现形式,都可以经过数字化后存入计算机。

2. 数据库

数据库(DataBase, DB)是长期存储在计算机内、有组织的、可共享的数据集合。这种集合具有如下特点:

- (1) 最小的冗余度。以一定的数据模型来组织数据,数据尽可能不重复。
- (2) 应用程序对数据资源共享。以最优方式为某个特定组织或企业的多种应用服务。
- (3) 数据独立性强。数据结构较强地独立于使用它的应用程序。
- (4) 统一管理和控制。对数据的定义、操纵和控制,由数据库管理系统统一进行管理和控制。

3. 数据库管理系统

数据库管理系统(DataBase Management System, DBMS)是位于用户与操作系统之间的一个数据管理软件,它的基本功能包括以下几个方面:

- (1) 数据定义功能。DBMS 提供数据定义语言(Data Definition Language, DDL),用户通过它可以方便地对数据库中的数据对象进行定义。
- (2) 数据操纵功能。DBMS 还提供数据操纵语言(Data Manipulation Language, DML),用户可以使用 DML 操纵数据,实现对数据的基本操作。如查询、插入、删除和修改等。
- (3) 数据库的运行管理功能。数据库在建立、运行和维护时由数据库管理系统统一管理和控制,以保证数据的安全性、完整性,以及对并发操作的控制和发生故障后的系统恢复等。

(4) 数据库的建立和维护功能。它包括数据库初始数据的输入、转换功能, 数据库的转储、恢复功能, 数据库的重组功能和性能监视、分析功能等。

数据库管理系统软件有多种, 比较著名的有 Oracle、Informix、Sybase、SQL Server、DB2 等。

4. 数据库系统

数据库系统 (DataBase System, DBS) 是指在计算机系统中引入数据库后构成的系统。一般由数据库、操作系统、数据库管理系统 (及其开发工具)、应用系统、数据库管理员和用户构成。应当指出的是, 数据库的建立、使用和维护等工作只有 DBMS 远远不够, 还要有专门的人员来完成, 这些人被称为数据库管理员 (DataBase Administrator, DBA)。

1.1.2 数据库技术的产生与发展

数据库技术就是数据管理技术, 是对数据的分类、组织、编码、存储、检索和维护的技术。数据库技术的发展是和计算机技术及其应用的发展联系在一起的。数据管理技术经历了如下三个阶段: 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

这一阶段是指 20 世纪 50 年代中期以前, 计算机主要用于科学计算, 当时的计算机硬件状况是: 外存只有磁带、卡片、纸带, 没有磁盘等直接存取的存储设备; 软件状况是: 没有操作系统, 没有管理数据的软件, 数据处理方式是批处理。

人工管理阶段的特点是: 数据不保存, 数据无专门软件进行管理, 数据不共享 (冗余度大), 数据不具有独立性 (完全依赖于程序), 数据无结构。

2. 文件系统阶段

这一阶段从 20 世纪 50 年代后期到 60 年代中期, 计算机硬件和软件都有了一定的发展。计算机不仅用于科学计算, 还大量用于管理。这时硬件方面已经有了磁盘、磁鼓等直接存取的存储设备。在软件方面, 操作系统中已经有了数据管理软件, 一般称为文件系统。处理方式上不仅有了文件批处理, 而且能够联机实时处理。

文件阶段的数据管理特点是: 数据可以长期保存, 由文件系统管理数据, 程序与数据有一定的独立性, 数据共享性差 (冗余度大), 数据独立性差, 记录内部有结构 (但整体无结构)。

3. 数据库系统阶段

从 20 世纪 60 年代后期以来, 计算机硬件和软件技术得到了飞速发展, 计算机用于管理的规模更为庞大, 应用越来越广泛, 数据量急剧增长, 数据的共享要求越来越强。出现了内存大、运算速度快的主机和大容量的磁盘, 计算机硬件价格下降, 软件价格上升, 为编制和维护系统软件及应用程序所需的成本相对增加。处理方式上, 联机实时处理要求更多, 并开始提出和考虑分布处理。

为了解决多用户、多应用共享数据, 使数据为尽可能多的应用服务, 文件系统已不能满足应用需求, 一种新的数据管理技术——数据库技术应运而生。

与人工管理和文件系统阶段相比较, 数据库系统阶段具有以下的特点:

(1) 数据结构化。在文件系统中,各文件相互独立,文件记录内部结构的最简单形式是等长同格式记录的集合。这种思想就是数据库方法的雏形。它把文件系统中记录内部有结构的思想扩大到了两个记录之间。但这种方法还存在着局限性。因为这种灵活性只是对某一个应用而言的,而一个组织或企业包括许多应用。从整体来看,不仅要考虑一个应用(程序)的数据结构,而且要考虑整个组织的数据结构问题。整个组织的数据结构化,要求在描述数据时不仅描述数据本身,还要描述数据之间的联系。文件系统中记录内部已有了某些结构,但记录之间是没有联系的。因此,数据的结构化是数据库的主要特征之一,也是数据库与文件系统的根本区别。

(2) 数据共享性高、冗余度小、易扩充。数据的冗余度是指数据重复的程度。数据库系统从整体角度描述数据,使数据不再是面向某一应用,而是面向整个系统。因此,数据可以被多个应用共享。这不仅大大减小了数据的冗余度、节约存储空间、减少存取时间,而且可以避免数据之间的不相容性和不一致性。

由于数据库中的数据面向整个应用系统,所以容易增加新的应用,适应各种应用需求。当应用需求改变或增加时,只要重新选取整体数据的不同子集,便可以满足新的要求,这就使得数据库系统具有弹性大,易扩充的特点。

(3) 数据独立性高。数据独立性包括物理独立性和逻辑独立性。

数据的物理独立性是指当数据的物理存储改变时,应用程序不用改变。换言之,用户的应用程序与数据库中的数据是相互独立的。数据在数据库中的存储形式是由 DBMS 管理的,用户程序不需要了解,应用程序要处理的只是数据的逻辑结构。

数据的逻辑独立性是指当数据的逻辑结构改变时,用户应用程序不用改变。换言之,用户的应用程序与数据库的逻辑结构是相互独立的。

数据和程序的独立性,可以将数据的定义和描述从应用程序中分离出来。数据的存取由 DBMS 管理,用户不必考虑存取路径等细节,从而简化了应用程序的编制,大大减少了应用程序的维护和修改工作量。

(4) 统一的数据管理和控制。数据库对系统中的用户是一种共享资源。计算机的共享一般是并发的,即多个用户可以同时存取数据库中的数据,甚至可以同时存取数据库中同一个数据。因此,数据库管理系统必须提供以下几个方面的数据控制保护功能。

① 数据的安全性(security)保护。数据的安全性是指保护数据以防止不合法的使用所造成的数据泄密和破坏,使每个用户只能按规定,对某种数据以某些方式进行使用和处理。例如,用身份鉴别、检查口令或其他手段来检查用户的合法性,合法用户才能进入数据库系统。

② 数据的完整性(integrity)控制。数据的完整性指数据的正确性、有效性和相容性。完整性检查提供必要的功能,保证数据库中的数据在输入和修改过程中始终符合原来的定义和规定,即在有效的范围内或保证数据之间满足一定的关系。例如,月份是1~12之间的正整数,性别是“男”或“女”,大学生的年龄是大于15小于45的整数,学生的学号是惟一的,等等。

③ 数据库恢复(recovery)。计算机系统的硬件、软件故障、操作员的失误以及人为

的攻击和破坏,会影响数据库中数据的正确性,甚至会造成数据库部分或全部数据的丢失。因此数据库管理系统必须能够进行应急处理,将数据库从错误状态恢复到某一已知的正确状态。

④ 并发 (concurrency) 控制。当多个用户的并发进程同时存取、修改数据库时,可能会发生由于相互干扰而导致结果错误的情况,并使数据库完整性遭到破坏。因此,必须对多用户的并发操作加以控制和协调。

1.1.3 数据库系统的体系结构

数据库系统的体系结构划分为五类,即集中式系统、个人计算机系统、分布式系统、客户/服务器系统和浏览器/服务器系统。目前,客户/服务器系统和浏览器/服务器系统是数据库系统中最常用的结构。

1. 集中式系统

在集中式系统中,DBMS 和应用程序以及与用户终端进行通信的软件等都运行在一台宿主计算机上,所有的数据处理都是在宿主计算机中进行。宿主计算机一般是大型机、中型机或小型机。应用程序和 DBMS 之间通过操作系统管理的共享内存或应用任务区来进行通信,DBMS 利用操作系统提供的服务来访问数据库。终端通常是非智能的,本身没有处理能力。近年来,微处理器的出现引起了智能化终端的发展,这种终端可以完成某些用户的输入、输出处理。

集中系统的主要优点是:具有集中的安全控制,以及处理大量数据和支持大量并发用户的能力。集中系统的主要缺点是:购买和维持这样的系统一次性投资太大,并且不适合分布处理。

2. 个人计算机系统

当 DBMS 在个人计算机(PC)上运行时,PC 机起到宿主机的作用,同时也起到了终端的作用。与大型系统不同,通常个人计算机(微机)上的 DBMS 功能和数据库应用功能是结合在一个应用程序中的。这类 DBMS(如 FoxPro、Access)的功能灵活,系统结构简洁,运行速度快,但这类 DBMS 的数据共享性、安全性、完整性等控制功能比较薄弱。

3. 客户/服务器系统

在客户/服务器(Client/Server, C/S)结构的数据库系统中,数据处理任务被划分为两部分:一部分运行在客户端;另一部分运行在服务器端。划分的方案可以有多种,常用的方案是:客户端负责应用处理,数据库服务器完成 DBMS 的核心功能。

在 C/S 结构中,客户端软件和服务器端软件可以运行在一台计算机上,但大多是分别运行在网络中不同的计算机上。客户端软件一般运行在 PC 上,服务器端软件可以运行在从 PC 机到大型机等各类计算机上。数据库服务器把数据处理任务分开在客户端和服务器上运行,因而充分利用了服务器的高性能数据库处理能力以及客户端灵活的数据表示能力。通常从客户端发往数据库服务器的只是查询请求,从数据库服务器传回给客户端的只是查询结果,不需要传送整个文件,从而大大减少了网络上的数据传输量。

C/S 结构是一个简单的两层模型，一端是客户机，另一端是服务器。这种模型中，客户机上都必须安装应用程序和工具，使客户端过于庞大、负担太重，而且系统安装、维护、升级和发布困难，从而影响效率。

4. 分布式系统

一个分布式数据系统由一个逻辑数据库组成，整个逻辑数据库的数据，存储在分布于网络中的多个节点上的物理数据库中。

在分布式数据库中，由于数据分布于网络中的多个节点上，因此与集中式数据库相比，存在一些特殊的问题，例如，应用程序的透明性、节点自治性、分布式查询和分布式更新处理等，这就增加了系统实现的复杂性。

较早的分布式数据库是由多个宿主系统构成的，数据在各个宿主系统之间共享。在当今的客户/服务器结构的数据库系统中，服务器的数目可以是一个或多个。当系统中存在多个数据库服务器时，就形成了分布系统。

5. 浏览器/服务器系统

随着 Internet 的迅速普及，出现了三层客户机/服务器模型：客户机→应用服务器→数据库服务器。这种结构的客户端只需安装浏览器就可以访问应用程序，这种系统称为浏览器/服务器（Browser/Server, B/S）系统。B/S 结构克服了 C/S 结构的缺点，是 C/S 的继承和发展。

1.1.4 数据库系统三级模式结构

从 DBMS 方面考虑，数据库系统通常采用三级模式结构，这是 DBMS 内部的系统结构。

在数据库中，数据模型可以分为三个层次，分别称为外模式、模式和内模式。

外模式反映的是一种局部的逻辑结构，它与应用程序相对应，由用户自己定义。一个数据库可以有多个外模式。模式反映的是总体的逻辑结构，一个数据库只有一个模式，它是由数据库管理员（DBA）定义的。内模式是反映物理数据存储的模型，它也是由数据库管理员（DBA）定义的。

在数据模型中有“型”（Type）和“值”（Value）的概念。型是对某一类数据的结构和属性的说明（或定义），值是对型的一个具体赋值。例如，图书记录定义为（编号，书名，作者，出版社，出版日期，定价）这是记录的型，而（G11.11, C 语言程序设计，张大海，蓝天，2004.4, 26.30）是记录型的一个记录值。

模式是数据库中全体数据的逻辑结构和特征的描述，它仅仅涉及到型的描述，不涉及具体的值。模式的一个具体值称为模式的一个实例。同一模式可以有很多实例。模式反映的是数据的结构及其联系，而实例反映的是数据库某一时刻的状态。所以模式是相对稳定的，而实例是相对变动的。

1. 模式

模式（schema）也称为逻辑模式，是数据中全体数据的逻辑结构和特征描述，是所有用户的公共数据视图。它是数据库系统模式结构的中间层，既不涉及数据的物理存储细节和硬件环境，也与具体的应用程序及其所使用的开发工具（如 C、Visual Basic、Power Build、ASP、JSP 等）无关。

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(包括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),而且要定义数据之间的联系,定义与数据有关的安全性、完整性要求。

DBMS 提供描述语言(模式 DDL)来严格定义模式。

2. 外模式

外模式(external schema)也称为子模式(subschema)或用户模式,它是数据库用户(包括应用程序员和最终用户)能够看到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,如果不同的用户的应用需求、看待数据的方式、对数据保密的要求等存在差异,则其外模式描述就是不同的。即使对模式中同一数据记录,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,同一个外模式也可作为某一用户的多个应用系统所用,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,数据库中其余数据是不可见的。

DBMS 提供子模式描述语言(子模式 DDL)来严格定义子模式。

3. 内模式

内模式(internal schema)也称为存储模式(storage schema),一个数据库只有一个内模式。它是数据物理结构和存储方式的描述,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 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,存储记录有何规定等。

DBMS 提供内模式描述语言(内模式 DDL,或者存储模式 DDL)来严格定义内模式。

1.2 数据模型

1.2.1 数据模型及其组成要素

模型是现实世界特征的模拟和抽象。数据模型(Data Model)也是一种模型,它是实现数据特征的抽象。数据库系统的核心是数据库,数据库是根据数据模型建立的,因而数据模型是数据库系统的基础。

一般地讲,任何一种数据模型都是严格定义的概念的集合。这些概念必须能够精确地描述系统的静态特性、动态特性和完整性约束条件。因此,数据模型通常都是由数据结构、数据操作和完整性约束 3 个要素组成。

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)以及数据对象的类型等。存储在数据库中的对象类型的集合是数据库的组成部分。例如在图书馆管理中,要管理的数据对象有图书、读者、借阅等基本情况。图书对象集中,每本图书包括编号、书名、作者、出版社、出版日期、定价等信息,这些基本信息描述了每本图书的特性,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。

数据结构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,通常按照其数据结构的类型来命名数据模型。例如层次结构、网状结构、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。

数据操作是指对数据库中的各种对象(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。数据库主要有查询和更新(包括插入、删除、修改)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)以及实现操作的语言。

3. 数据完整性约束

数据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制约和储存规则,用以符合数据模型的数据库状态以及状态的变化,以保证数据正确、有效和相容。

数据模型应该反映和规定本数据模型必须遵守的、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的机制,以反映具体所涉及的数据必须遵守的特定的语义约束。例如,在图书信息中,图书的“定价”只能取大于零的值;人员信息中的“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经开设的课程的课程号等。

数据模型是数据库技术的关键,它的3个要素完整地描述了一个数据模型。

1.2.2 数据模型的种类

目前,数据库领域中,最常用的数据模型有:层次模型、网状模型和关系模型。其中,层次模型和网状模型统称为非关系模型。非关系模型的数据库系统在20世纪70年代非常流行,到了20世纪80年代,关系模型的数据库系统以其独特的优点逐渐占据了主导地位,成为数据库系统的主流。

1. 层次模型

层次模型(hierarchical model)是数据库中最早出现的数据模型,层次数据库系统采用层次模型作为数据的组织方式。用树型(层次)结构表示实体类型以及实体间的联系是层次模型的主要特征。其示意图如图1-1所示。

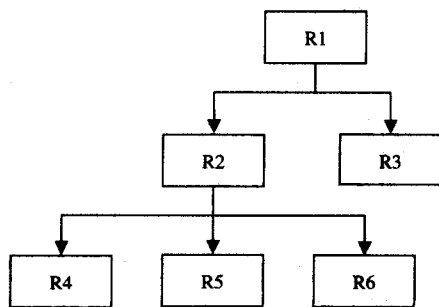


图 1-1 层次模型示意图

层次结构是一棵树。树的节点是记录类型，根节点只有一个，根节点以外的节点只有一个双亲节点，每一个节点可以有多个孩子节点。

层次模型的另一个最基本的特点是，任何一个给定的记录值（也称为实体）只有按照其路径查看时，才能显出它的全部意义。没有一个子记录值能够脱离双亲记录值而独立存在。

层次数据库系统的典型代表是 IBM 公司的 IMS (Information Management Systems) 数据库管理系统，这是 1968 年 IBM 公司推出的第一个大型的商用数据库管理系统。曾经在 20 世纪 70 年代的商业上广泛应用。目前，仍有某些特定用户在使用。

2. 网状模型

在现实世界中事物之间的联系更多的是非层次关系的，用层次模型表示非树形结构是很不直接的，网状模型 (network model) 则可以克服这一弊端。

用网状结构表示实体类型及实体之间联系的数据模型称为网状模型，其示意图如图 1-2 所示。在网状模型中，一个子节点可以有多个父节点，在两个节点之间可以有一种或多种联系。记录之间联系是通过指针实现的，因此，数据的联系十分密切。网状模型的数据结构在物理上易于实现，效率较高，但是编写应用程序较复杂，程序员必须熟悉数据库的逻辑结构。

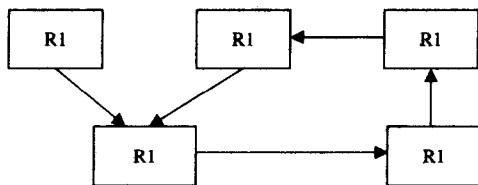


图 1-2 网状模型示意图

3. 关系模型

关系模型 (relational model) 是目前最常用的一种数据模型。关系数据库系统采用关系模型作为数据的组织方式。1970 年美国 IBM 公司 San Jose 研究室的研究员 E.F.Codd 首次提出了数据库系统的关系模型，开创了数据库关系方法和关系数据理论的研究，为关系数据库技术奠定了理论基础。由于 E.F.Codd 的杰出工作，他于 1981 年获得 ACM 图灵奖。

20 世纪 80 年代以来，计算机厂商推出的数据库管理系统几乎都支持关系模型，非关系模型系统的产品也大都加上了接口。数据库领域当前的研究工作也都是以关系方法为基础。

在现实世界中，人们经常用表格形式表示数据信息。但是日常生活中使用的表格往往比较复杂，在关系模型中基本数据结构被限制为二维表格。因此，在关系模型中，数据在用户观点下的逻辑结构就是一张二维表。每一张二维表称为一个关系 (relation)。

关系模型比较简单，容易为初学者接受。关系在用户看来是一个表格，记录是表中的行，属性是表中的列。

关系模型是数学化的模型，可把表格看成一个集合，因此集合论、数理逻辑等知识可引入到关系模型中来。关系模型已得到广泛应用。本书以后各章节的讨论均是基于关系模型的。

1.3 数据库设计

1.3.1 数据库设计概述

数据库设计是数据库技术的主要内容之一。数据库设计是指对于给定的应用环境（包括硬件环境和操作系统、DBMS 等软件环境），构建一个性能良好的、能满足用户要求的、能够被选定的 DBMS 所接受的数据库模式，建立数据库以及应用系统，使之能够有效地、合理地采集、存储、操作和管理数据，满足企业或组织中各类用户的应用需求。

数据库设计的主要内容有数据库的结构特性设计和数据库的行为特性设计。数据库的结构特性设计起着关键作用。

数据库的结构特性是静态的，一般情况下不会轻易变动。因此，数据库的结构特性设计又称为静态结构设计。其设计过程是：先将现实世界中的事物、事物之间的联系用 E-R 图（下一小节介绍）表示，再将各个分 E-R 图汇总，得出数据库的概念结构模型，最后将概念结构模型转换为数据库的逻辑结构模型表示。

数据库的行为结构设计是指确定数据库用户的行为和动作。数据库用户的行为和动作是指数据查询和统计、事务处理及报表处理等。这些都需要通过应用程序表达和执行。因而设计数据库的行为特征要与应用系统的设计结合进行。由于用户的行为是动态的，所以，数据库的行为特性设计也称为数据库的动态设计。其设计过程是：首先将现实世界中的数据及应用情况用数据流图和数据字典表示，并详细描述其中的数据操作要求，进而得出系统的功能结构和数据库的子模式。

本书以讨论数据库的结构特性设计为主。

1.3.2 数据库设计的基本步骤

考虑数据库及其应用系统开发的全过程，可以将数据库设计过程分为以下 6 个阶段。

1. 需求分析阶段

进行数据库应用软件的开发，首先必须准确了解与分析用户需求（包括数据处理）。需求分析是整个开发过程的基础，是最困难、最耗费时间的一步。作为“地基”的需求分析是否做得充分与准确，决定了在其上建造数据库“大厦”的速度与质量。需求分析做得不好，会导致整个数据库应用系统开发返工重做的严重后果。

2. 概念结构设计阶段

概念结构设计是整个数据库设计的关键，它通过对用户需求进行综合、归纳与抽象，形成一个独立于具体 DBMS 的概念模型，一般用 E-R 图表示概念模型。

3. 逻辑结构设计阶段

逻辑结构设计是将概念结构转化为选定的 DBMS 所支持的数据模型，并使其在功能、性能、完整性约束、一致性和可扩充性等方面均满足用户的需求。