

普通高等学校计算机专业系列教材



关系数据库理论及应用

周定康 许 婕 李云洪 马明磊



华中科技大学出版社
<http://press.hust.edu.cn>

普通高等学校计算机专业系列教材

关系数据库理论及应用

周定康 许婕 李云洪 马明磊

华中科技大学出版社

图书在版编目(CIP)数据

关系数据库理论及应用/周定康 许婕 李云洪 马明磊

武汉:华中科技大学出版社,2005年9月

ISBN 7-5609-3538-9

I. 关…

II. ①周… ②许… ③李… ④马…

III. 数据库系统-高等学校-教材

IV. TP311.13

关系数据库理论及应用

周定康 许 婕 李云洪 马明磊

策划编辑:谢燕群

封面设计:潘 群

责任编辑:刘 勤

责任校对:刘 竣

责任监印:张正林

出版发行:华中科技大学出版社

武昌喻家山 邮编:430074 电话:(027)87557437

录 排:华中科技大学惠友文印中心

印 刷:湖北新华印务有限公司

开本:787×960 1/16

印张:19

字数:337 000

版次:2005年9月第1版

印次:2005年9月第1次印刷

定价:24.80元

ISBN 7-5609-3538-9/TP·590

(本书若有印装质量问题,请向出版社发行部调换)

内 容 提 要

本书系统、完整地介绍了当前数据库技术的基本概念、基本原理及应用实践。主要内容包括:体系结构、数据模型、关系数据库设计理论、数据库设计、数据库保护、SQLserver 2000 简介、SQL 语言 SQL Server 2000 安全及其数据库备份与恢复、Powerbuilder 9.0 编程基础、Power Script 语言与数据库连接以及开发应用实例。每章后都附有习题。

本书可作为高等院校计算机专业及相关专业的教科书,也可供从事计算机开发与应用的科研人员、工程技术人员以及其他有关人员参考。

前 言

数据库技术自 20 世纪 60 年代后期出现以来,一直流行于数据处理领域,并得到了迅速的发展,已经成为计算机科学的一个重要分支。由于数据库技术已经成为开发各种信息系统的重要工具,它的推广使用,使得计算机应用深入到工农业生产、商业管理、办公自动化、科学研究和工程技术的各个领域。随着社会各个领域的迫切需要和广泛应用,人们对数据库技术重要性的认识越来越明确,对它的应用和学习要求也越来越迫切。

本书是作者在多年从事数据库系统教学和科研的基础上编著的。在编写过程中,本着由易到难,循序渐进的原则,力求做到概念清楚,深入浅出,理论联系实际,面向应用,重点突出应用最广泛的关系数据库系统,对它的概念、原理作了较深入的讨论。为了强调实际应用,本书还较详细地介绍了目前国内外较为流行的 SQL Server 2000 关系数据库,Power Builder 9.0 编程技术,以便初学者在学习过程中能更好地把理论与实际结合起来,从而达到更好的学习效果。另外,在内容选取上,我们还注意尽量反映这一领域的新方法、新技术,以使学能能对数据库领域的前沿动态有初步的了解。

全书共分 12 章。第 1 章介绍数据库系统的基本概念。第 2 章介绍关系模型。第 3 章介绍关系代数及其优化技术。第 4 章介绍关系数据库规范化理论。第 5 章介绍数据库设计的基本技术和方法,以指导应用。第 6 章介绍数据库保护技术。第 7 章介绍 SQL Server 2000 数据库系统管理。第 8 章介绍 Transact-SQL 语言。第 9 章介绍 SQL Server 2000 安全性及数据库备份与恢复。第 10 章介绍 Power Builder 9.0 的编程基础。第 11 章介绍 PowerScript 语言与数据库连接。第 12 章介绍开发应用实例《同学通信录管理系统》。

本书第 1~3 章由周定康编写,第 4~6 章由许婕编写,第 7~9 章由李云洪编写,第 10~12 章由马明磊编写,全书由周定康负责结构的设计和内容的组织。

在本书编写过程中,得到江西师范大学计算机信息工程学院谢旭升、朱明华、张练、李宏伟和周勇等老师的无私帮助和支持,在此谨向他们表示诚挚的谢意。

本书可作为高等院校计算机有关专业的数据库课程教材,也可供从事信息领域工作的科技人员和工程技术人员及其他有关人员参阅。

由于水平有限,加之时间匆忙,书中难免有错误和不妥之处,敬请读者批评指正。

编 者

2005 年 3 月

目 录

第 1 章 数据库系统概论	(1)
1.1 数据库的基本概念	(1)
1.1.1 数据处理技术的发展历程	(1)
1.1.2 信息描述与数据模型	(5)
1.1.3 数据库技术的研究领域	(6)
1.2 数据库体系结构	(7)
1.2.1 数据库系统模式概论	(7)
1.2.2 数据库系统的三级模式结构	(8)
1.2.3 数据库的二级映像功能	(9)
1.3 数据库系统	(10)
1.3.1 数据库系统的基本特点	(10)
1.3.2 数据库系统的组成	(12)
1.3.3 数据库管理系统的功能	(14)
1.3.4 数据库系统的视图	(14)
1.3.5 数据库系统的访问	(15)
1.3.6 数据库系统发展阶段的划分	(16)
习题一	(16)
第 2 章 数据库系统的数据模型	(17)
2.1 概念模型	(18)
2.1.1 概念模型的数据结构	(18)
2.1.2 概念模型的表示方法	(20)
2.2 结构数据模型(Structure Data Model)	(22)
2.2.1 结构数据模型的三要素	(23)
2.2.2 层次数据模型	(23)
2.2.3 网状数据模型	(28)
2.2.4 关系数据模型	(30)
2.3 关系数据库的基本概念	(32)
2.3.1 笛卡儿乘积	(33)
2.3.2 关系	(33)
2.3.3 关键字	(34)

2.3.4	关系模式	(34)
2.3.5	关系数据模型	(34)
2.3.6	关系数据库	(35)
2.4	关系数据库模型及其描述	(35)
2.4.1	关系数据库模型(Relational Model)	(35)
2.4.2	关系数据库模式建立	(36)
2.4.3	关系数据库局部模式建立	(37)
2.4.4	关系模型的完整性规则	(38)
	习题二	(39)
第3章	关系代数及其优化技术	(41)
3.1	关系代数	(42)
3.1.1	传统的集合运算	(43)
3.1.2	专门的关系运算	(44)
3.1.3	检索操作实例	(49)
3.2	查询优化技术	(50)
3.2.1	查询优化的一般策略	(51)
3.2.2	关系代数等价变换规则	(52)
3.2.3	关系代数表达式的优化算法	(53)
	习题三	(56)
第4章	关系数据库设计理论	(58)
4.1	概述	(58)
4.2	函数依赖	(59)
4.2.1	函数依赖	(60)
4.2.2	候选码	(61)
4.3	范式	(61)
4.3.1	第一范式(1NF)	(62)
4.3.2	第二范式(2NF)	(63)
4.3.3	第三范式(3NF)	(65)
4.3.4	BC范式(BCNF)	(66)
4.4	多值依赖与第四范式(4NF)	(67)
4.4.1	多值依赖	(69)
4.4.2	第四范式(4NF)	(71)
4.5	关系模式的规范化	(72)
4.6	数据依赖的公理系统	(73)
4.7	关系模式的分解	(80)

4.7.1	模式分解的三个定义	(80)
4.7.2	分解的无损联接性和依赖保持	(81)
4.7.3	转换成 3NF 的分解	(84)
4.7.4	转换成 BCNF 的分解	(86)
4.7.5	模式设计方法的原则	(88)
习题四		(89)
第 5 章	数据库设计	(91)
5.1	数据库设计概述	(91)
5.1.1	数据库设计的定义	(91)
5.1.2	数据库设计的特点	(92)
5.1.3	数据库设计方法简述	(93)
5.1.4	数据库设计的基本步骤	(94)
5.2	需求分析	(97)
5.2.1	需求分析的任务	(97)
5.2.2	需求分析的方法	(98)
5.2.3	数据字典	(100)
5.3	概念结构设计	(102)
5.3.1	概念结构设计的方法与步骤	(102)
5.3.2	局部视图设计	(103)
5.3.3	视图的集成	(105)
5.4	逻辑结构设计	(107)
5.4.1	E-R 图向关系模型的转换	(108)
5.4.2	数据模型的优化	(109)
5.4.3	设计用户子模式	(110)
5.5	数据库的物理设计	(110)
5.5.1	数据库的物理设计的内容	(111)
5.5.2	关系模式存取方法选择	(111)
5.5.3	确定数据库的存储结构	(113)
5.5.4	评价物理结构	(114)
5.6	数据库的实施	(114)
5.6.1	数据的载入和应用程序的调试	(114)
5.6.2	数据库的试运行	(115)
5.7	数据库的运行与维护	(116)
习题五		(117)

第 6 章 数据库的保护	(118)
6.1 事务的基本概念	(118)
6.1.1 事务(transaction)	(118)
6.1.2 实例	(118)
6.1.3 事务的特性	(119)
6.2 数据库恢复	(120)
6.2.1 故障的种类	(120)
6.2.2 恢复的实现技术	(122)
6.2.3 恢复策略	(124)
6.3 并发控制	(126)
6.3.1 封锁并发控制方法	(128)
6.3.2 活锁和死锁	(131)
6.3.3 并发调度的可串行性	(133)
6.3.4 两段锁协议	(134)
6.4 数据库的安全性	(136)
6.4.1 数据库安全性控制	(136)
6.4.2 安全系统的基本要求	(139)
6.4.3 自主存取控制(DAC)和强制存取控制(MAC)	(140)
6.5 数据库完整性	(141)
6.5.1 完整性约束条件	(141)
6.5.2 完整性控制	(143)
习题六	(146)
第 7 章 SQL Server 2000 数据库管理系统	(148)
7.1 SQL Server 2000 简介	(148)
7.1.1 SQL Server 2000 的结构	(148)
7.1.2 SQL Server 2000 的特点	(149)
7.1.3 SQL Server 2000 的安装	(151)
7.1.4 SQL Server 2000 的系统数据库	(154)
7.2 SQL Server 2000 服务器的注册和启动	(155)
7.3 SQL Server 2000 数据库管理	(158)
7.3.1 数据库文件组和数据库文件	(158)
7.3.2 建立、管理和删除数据库	(159)
7.3.3 查看数据库信息	(160)
习题七	(161)

第 8 章 Transact-SQL 语言	(162)
8.1 Transact-SQL 语言基础	(162)
8.1.1 SQL 简介	(162)
8.1.2 标识符、数据类型	(162)
8.1.3 运算符、变量与批处理	(167)
8.1.4 流控制语句	(170)
8.2 函数	(173)
8.3 表、视图和索引	(175)
8.3.1 表	(175)
8.3.2 约束	(177)
8.3.3 索引	(180)
8.3.4 视图	(181)
8.4 存储过程与触发器	(185)
8.4.1 存储过程概述	(185)
8.4.2 存储过程的使用和管理	(185)
8.4.3 触发器概述	(187)
8.4.4 触发器的使用与管理	(188)
8.4.5 触发器的特殊功能	(190)
8.5 Transact-SQL 查询	(190)
8.5.1 简单查询	(190)
8.5.2 统计查询	(201)
8.5.3 联合查询	(202)
8.5.4 子查询	(203)
8.5.5 连接查询	(205)
习题八	(206)
第 9 章 管理 SQL Server2000 的安全性及数据库的备份与恢复	(208)
9.1 管理 SQL Server2000 安全性	(208)
9.1.1 安全性概述	(208)
9.1.2 SQL Server2000 登录验证模式	(209)
9.1.3 角色	(209)
9.1.4 许可	(213)
9.2 数据库的备份	(217)
9.2.1 数据库备份概念	(217)
9.2.2 数据库备份类型	(217)
9.2.3 执行数据库备份	(217)

9.3 数据库的恢复	(221)
习题九	(222)
第 10 章 PowerBuilder 9.0 的编程基础	(223)
10.1 PowerBuilder 简介	(223)
10.2 PB 9.0 开发环境	(223)
10.3 PB 9.0 的主要画板	(224)
10.3.1 创建新的工作区	(224)
10.3.2 应用程序画板	(225)
10.3.3 窗口画板	(226)
10.3.4 菜单画板	(230)
10.3.5 工程画板	(232)
10.4 常用控件	(232)
10.4.1 窗口控件简介	(233)
10.4.2 控件属性的修改	(234)
10.5 DataWindow 数据窗口对象	(235)
10.5.1 数据窗口的创建	(235)
10.5.2 数据窗口显示风格	(237)
10.5.3 数据源	(239)
习题十	(239)
第 11 章 PowerScript 语言与数据库连接	(240)
11.1 PowerScript 语言	(240)
11.1.1 PowerScript 语言基础	(240)
11.1.2 变量的类型和作用域	(241)
11.1.3 数组及使用	(242)
11.1.4 结构的使用	(243)
11.1.5 常用 PowerScript 语句	(243)
11.1.6 常用函数和语句	(246)
11.2 SQL 语句的使用	(250)
11.2.1 静态 SQL 语句	(250)
11.2.2 动态 SQL 语句	(254)
11.3 PB 9.0 与常见数据库的连接配置	(258)
11.3.1 PB 9.0 与 Oracle 数据库的连接	(258)
11.3.2 PB 9.0 连接任意 ODBC 数据源	(261)
习题十一	(267)

第 12 章 开发应用实例“通信录管理系统”	(269)
12.1 系统需求分析	(269)
12.1.1 系统目标设计	(269)
12.1.2 系统功能分析	(269)
12.2 后台数据库设计	(270)
12.3 数据库的实现	(271)
12.3.1 数据库的创建	(271)
12.3.2 数据库表单的创建	(272)
12.3.3 数据库连接的建立	(273)
12.4 应用程序对象的创建	(275)
12.5 系统各对象设计及脚本编写	(276)
12.5.1 系统主窗口 w_main 和主菜单 m_main	(276)
12.5.2 数据窗口(DataWindow)的创建	(276)
12.5.3 普通窗口 Window 的创建	(280)
12.6 各对象脚本的编写	(281)
12.6.1 应用程序对象 test 和主菜单 m_main 的脚本	(281)
12.6.2 窗口对象脚本的编写	(283)
12.7 应用程序运行和编译	(287)
12.7.1 应用程序的运行	(287)
12.7.2 应用程序的编译	(288)
习题十二	(290)

第 1 章 数据库系统概论

计算机应用主要有三个方面,科学计算、数据处理和过程控制,而数据处理几乎占计算机应用的 70% 左右。数据库技术是计算机数据处理的核心,是现代信息科学与技术的重要组成部分。随着计算机技术与网络通信技术的发展,数据库技术已成为信息社会中对大量数据进行组织与管理的重要技术手段,是网络信息化管理系统的基础。

从 20 世纪 70 年代后期开始,数据库作为计算机专业的一门重要课程进入国外各个大学的课堂,我国高等院校从 20 世纪 80 年代开始,将数据库作为主要课程对计算机专业的本科生和研究生讲授,是计算机专业的核心课程之一。作为一门学科,数据库技术是计算机科学技术的一个重要分支,没有数据库技术,人们在浩瀚的信息世界中将显得手足无措。

本章先回顾数据库技术的发展历程,然后介绍数据库的基本概念、体系结构和数据库管理系统。

1.1 数据库的基本概念

了解数据库技术的发展历程,掌握数据库这门学科的术语、思想和体系结构是很有必要的。

1.1.1 数据处理技术的发展历程

从 20 世纪 50 年代中期开始,计算机应用从科学研究部门扩展到企业管理及政府行政部门,数据库技术随着计算机技术,应用技术的发展而发展,它与硬件(主要是外部存储器)、软件和计算机应用的领域有密切联系。数据处理技术的发展主要经历了四个阶段:人工管理阶段、文件系统阶段、数据库系统阶段和高级数据库系统阶段。

1. 人工管理阶段

20 世纪 50 年代中期以前,计算机的外部存储设备只有磁带、卡片和纸带,没有磁盘等可直接存取的设备,没有操作系统和数据管理软件,数据处理方式是批处理,所有的数据完全由人工进行管理,因此这个阶段称为人工管理阶段。主要有以下特点。

① 一组数据对应于一个应用程序,应用程序与其处理的数据结合成一个整

体。在进行计算时,系统将应用程序与数据一并装入内存,用完后就将其撤销,释放被占用的数据空间与程序空间。其数据管理模型如图 1.1 所示。

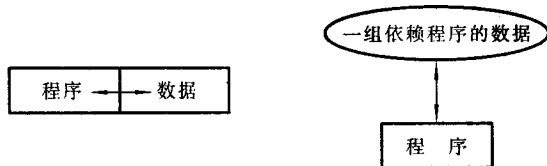


图 1.1 人工管理阶段数据管理模型

② 没有对数据进行管理的管理软件。程序员不仅要规定数据的逻辑结构,还要在程序中设计物理结构,如:存储结构,存取方法,输入输出方式等,因此,数据与程序不具有独立性,如果数据在存储上改变了,程序员就必须修改程序。

③ 没有文件概念,数据的组织方式由应用程序开发人员自行设计和安排。

④ 数据面向应用。即如果两个应用程序使用相同的数据,也必须各自定义自己的数据存储和存取方式,不能共享相同的数据定义,因此造成了程序之间大量的冗余数据。

2. 文件系统阶段

从 20 世纪的 50 年代后期到 60 年代中期,计算机的软、硬件都得到发展,磁盘、磁鼓等直接存取设备也相继投入使用,软件中也有了高级语言和操作系统。在操作系统中已经有了专门的数据管理软件(文件系统),数据处理方式不再是单一的批处理,产生了联机实时处理方式。计算机不仅用于科学计算,而且大量用于计算机数据管理。因此这个阶段称为文件系统阶段,主要有以下特点。

① 数据以文件方式长期保存在磁盘上,应用程序与其处理的数据分离,应用程序(用户)可经常通过文件管理系统实现对保存在外存上的数据进行查询、修改、插入和删除等操作,一定程度上实现了数据共享。其数据管理模型如图 1.2 所示。



图 1.2 文件系统阶段数据管理模型

② 文件的逻辑结构与存储结构之间由文件管理系统进行转换,程序与数据具有一定的独立性。程序员不必关心数据的物理位置和物理结构,用户可以集中精力考虑算法,从而节省了维护程序的工作量。

③ 文件形式的多样化,由于有了直接存取存储设备,也就有了索引文件,链接

文件,直接存取文件等。对文件的记录可以顺序和(或)随机访问等。文件之间相互独立,必须用程序来实现文件与文件之间的联系。

④ 数据存取以记录为单位。

文件系统比第一阶段的数据管理有很大的改进,但仍然存在许多问题,主要表现在以下几个方面。

① 数据和程序之间独立性差。在文件管理系统看来,文件中的数据是数据流,文件中的记录结构体现在应用程序之中。即文件系统中的数据是面向应用的,数据的逻辑结构改变,则必须修改应用程序。而应用程序的修改又会影响到数据结构的改变,因此文件系统仍然是一个不具有弹性的数据集合。

② 数据冗余度大。文件系统中的数据是面向应用的,通常一个文件对应于相应的应用程序。当一个文件被多个应用程序共享时,共享该文件的多个应用程序中应具有相同的文件记录结构,这给文件记录结构的修改带来不便。为了解决这个问题,只好一个文件对应于一个应用程序。因此,造成相同数据在多个应用程序的文件中重复存储。

③ 易产生不一致性。由于数据冗余大,在进行数据修改时,易造成同样的数据在不同的文件中不一致的现象。

④ 数据联系弱。由于文件中的数据是数据流,这就使得文件之间及文件中数据之间联系差。

⑤ 操作功能简单。在文件系统阶段对文件的操作是一些基本操作,仅有文件的打开、关闭、读、写和指针移动等。

3. 数据库系统阶段

从20世纪60年代后期开始,计算机用于管理的规模越来越大,存储技术也取得了很大发展,有了大容量的磁盘存储器,对联机实时处理的要求越来越多,并开始提出和考虑分布式处理。在这样的背景下,为了解决多用户、多应用共享数据的需求,提高对文件操作的效率,人们着手开发和研制更加完善的数据管理模式,出现了数据库这样的数据管理技术,推出了数据库管理系统软件。进入20世纪70年代,数据库技术得到迅猛发展,成为计算机科学的一个重要分支。因此,这个阶段为数据库管理系统阶段,主要有以下几个特点。

① 数据库中的数据是结构化的。数据库系统不仅考虑数据项之间的联系,而且考虑记录的联系。

② 数据库中的数据是面向系统的,多个不同应用通过数据库管理系统(DBMS),实现对数据库的操作,其数据管理模型如图1.3所示。这样就实现了不同应用间的数据共享,减少了数据冗余。

③ 数据具有较高的独立性。数据库系统提供了两方面的映象或转换功能:一方面是数据的总体逻辑结构与存储结构之间的映象或转换功能,另一方面是数据



图 1.3 数据库系统阶段数据管理模型

的总体逻辑结构与某类应用所涉及的局部逻辑结构之间的映象或转换功能。前者使得当数据的存储结构改变时,数据的逻辑结构可以不变,从而应用程序可以不必修改,称为数据对程序的物理独立性,简称数据的物理独立性;后者使得当总体逻辑结构改变时,通过对映象的相应改变而保持局部逻辑结构不变,程序员根据局部逻辑结构编写应用程序,因此应用程序可以不必改变。这就是数据对程序的逻辑独立性,简称数据的逻辑独立性。

④ 数据库系统为用户提供了方便的接口,用户可以用数据库系统提供的查询语言和交互式命令操纵数据库,也可以用高级语言编程序通过数据库系统提供的接口操纵数据库,拓宽了数据库的应用范围。

⑤ 数据的最小存取单位是数据项。在数据库中,用户既可以存取数据库中某一个数据项或一组数据项,也可以存取一个记录或一组记录。

⑥ 数据库系统提供了统一的数据控制功能,主要有数据的安全性(Security)控制,数据的完整性(Integrity)约束,并发(Concurrency)控制,数据库恢复(Recovery)。

4. 高级数据库系统阶段

从 20 世纪 70 年代中期以来,数据库技术得到迅速的发展,出现了分布式数据库,面向对象数据库和智能型知识数据库,通常被称为高级数据库技术。进入 20 世纪 80 年代以后,数据库技术与应用领域中相关技术的结合,形成了很多新技术、新产品,组成了数据库大家族。例如,数据库技术与工程设计技术结合形成工程数据库,数据库技术与客户/服务器技术、分布式技术、并行技术等相关技术相结合,形成了客户/服务器结构的数据库技术、分布式数据库技术、并行数据库技术。当前常见的数据库管理系统有如下几种。

① 客户机/服务器结构的数据库管理系统。应用客户机/服务器网络结构,将某个或某些结点机用来专门执行数据库管理系统的功能,称为“数据库服务器”。其他结点上的计算机运行数据库管理系统的外围应用开发工具,支持用户的应用,称为“客户机”。客户机执行应用程序,并对服务器提出服务请求,服务器完成客户机所委托的服务,并且把服务结果返回给客户机,即形成通常所说的客户机/服务器结构。