

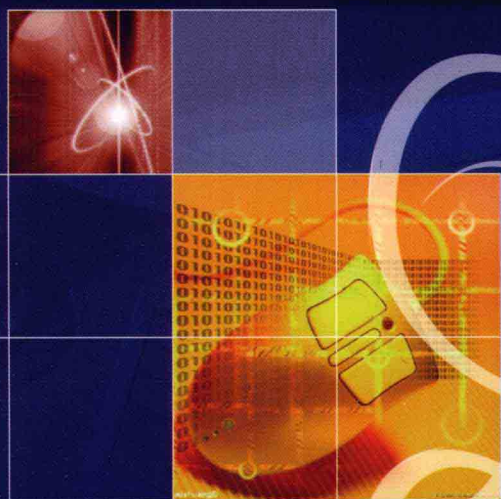


21st CENTURY
实用规划教材

21世纪全国应用型本科计算机案例型规划教材



数据库原理与应用 (SQL Server版)



Database

主 编 毛一梅 郭 红



北京大学出版社
PEKING UNIVERSITY PRESS

21 世纪全国应用型本科计算机案例型规划教材

数据库原理与应用(SQL Server 版)

主 编 毛一梅 郭 红

责任编辑：李 芳
封面设计：李 芳
文字编辑：李 芳
校对：李 芳
出版：清华大学出版社
地址：北京清华大学学研大厦A座
网址：<http://www.tup.tsinghua.edu.cn>
电话：(010) 62770175
电子邮箱：zhangq@tup.tsinghua.edu.cn
印刷：北京清华印刷厂
发行：新华书店
总发行：清华大学出版社

787毫米×1095毫米 16开本 5.22印张 22.7千字

2010年3月第1版 2010年3月第1次印刷

定价：36.00元



北京大学出版社
PEKING UNIVERSITY PRESS

内 容 简 介

本书从数据库原理入手,结合 Microsoft SQL Server 2005 的具体应用详细介绍了数据库技术的相关知识,从整个编写体系来说,本书共分两大部分 12 章,第一部分为基础篇,包括第 1 章~第 7 章,主要介绍数据库的基本概念、关系数据库理论和典型的关系数据库 Microsoft SQL Server 2005 的基本操作方法。第二部分为应用篇,包括第 8 章~第 12 章,主要介绍 Microsoft SQL Server 2005 的具体应用和实用技术。本书的特点在于理论与实际的紧密结合,各章中有大量的应用实例供读者学习和提高。

本书适合作为本科数据库相关课程教材,也可供数据库技术初、中级水平的读者自学用。

图书在版编目(CIP)数据

数据库原理与应用(SQL Server 版)/毛一梅,郭红主编. —北京:北京大学出版社,2010.2

(21 世纪全国应用型本科计算机案例型规划教材)

ISBN 978-7-301-16842-4

I. 数… II. ①毛…②郭… III. 关系数据库—数据库管理系统, SQL Server—高等学校—教材
IV. TP311.138

中国版本图书馆 CIP 数据核字(2010)第 015261 号

书 名:数据库原理与应用 (SQL Server 版)

著作责任者:毛一梅 郭 红 主编

策 划 编 辑:孙哲伟 李 虎

责 任 编 辑:孙哲伟

标 准 书 号:ISBN 978-7-301-16842-4/TP·1079

出 版 者:北京大学出版社

地 址:北京市海淀区成府路 205 号 邮编:100871

网 址:<http://www.pup.cn> <http://www.pup6.com>

电 话:邮购部 62752015 发行部 62750672 编辑部 62750667 出版部 62754962

电 子 邮 箱:pup_6@163.com

印 刷 者:北京大学印刷厂

发 行 者:北京大学出版社

经 销 者:新华书店

787 毫米×1092 毫米 16 开本 23.25 印张 537 千字

2010 年 2 月第 1 版 2010 年 2 月第 1 次印刷

定 价:36.00 元

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究

举报电话:010-62752024

电子邮箱:fd@pup.pku.edu.cn

21 世纪全国应用型本科计算机案例型规划教材

专家编审委员会

(按姓名拼音顺序)

主 任	刘瑞挺		
副主任	陈 钟	蒋宗礼	
委 员	陈代武	胡巧多	黄贤英
	江 红	李 建	娄国焕
	马秀峰	祁亨年	王联合国
	汪新民	谢安俊	解 凯
	徐 苏	徐亚平	宣兆成
	姚喜妍	于永彦	张荣梅

信息技术的案例型教材建设

(代丛书序)

刘瑞挺

北京大学出版社第六事业部在 2005 年组织编写了《21 世纪全国应用型本科计算机系列实用规划教材》，至今已出版了 50 多种。这些教材出版后，在全国高校引起热烈反响，可谓初战告捷。这使北京大学出版社的计算机教材市场规模迅速扩大，编辑队伍茁壮成长，经济效益明显增强，与各类高校师生的关系更加密切。

2008 年 1 月北京大学出版社第六事业部在北京召开了“21 世纪全国应用型本科计算机案例型教材建设和教学研讨会”。这次会议为编写案例型教材做了深入的探讨和具体的部署，制定了详细的编写目的、丛书特色、内容要求和风格规范。在内容上强调面向应用、能力驱动、精选案例、严把质量；在风格上力求文字精练、脉络清晰、图表明快、版式新颖。这次会议吹响了提高教材质量第二战役的进军号。

案例型教材真能提高教学的质量吗？

是的。著名法国哲学家、数学家勒内·笛卡儿(Rene Descartes, 1596—1650)说得好：“由一个例子的考察，我们可以抽出一条规律。(From the consideration of an example we can form a rule.)”事实上，他发明的直角坐标系，正是通过生活实例而得到的灵感。据说是在 1619 年夏天，笛卡儿因病住进医院。中午他躺在病床上，苦苦思索一个数学问题时，忽然看到天花板上有一只苍蝇飞来飞去。当时天花板是用木条做成正方形的格子。笛卡儿发现，要说出这只苍蝇在天花板上的位置，只需说出苍蝇在天花板上的第几行和第几列。当苍蝇落在第四行、第五列的那个正方形时，可以用(4, 5)来表示这个位置……由此他联想到可用类似的办法来描述一个点在平面上的位置。他高兴地跳下床，喊着“我找到了，找到了”，然而不小心把国际象棋撒了一地。当他的目光落到棋盘上时，又兴奋地一拍大腿：“对，对，就是这个图”。笛卡儿锲而不舍的毅力，苦思冥想的钻研，使他开创了解析几何的新纪元。千百年来，代数与几何，井水不犯河水。17 世纪后，数学突飞猛进的发展，在很大程度上归功于笛卡儿坐标系和解析几何学的创立。

这个故事，听起来与阿基米德在浴池洗澡而发现浮力原理，牛顿在苹果树下遇到苹果落到头上而发现万有引力定律，确有异曲同工之妙。这就证明，一个好的例子往往能激发灵感，由特殊到一般，联想出普遍的规律，即所谓的“一叶知秋”、“见微知著”的意思。

回顾计算机发明的历史，每一台机器、每一颗芯片、每一种操作系统、每一类编程语言、每一个算法、每一套软件、每一款外部设备，无不像闪光的珍珠串在一起。每个案例都闪烁着智慧的火花，是创新思想不竭的源泉。在计算机科学技术领域，这样的案例就像大海岸边的贝壳，俯拾皆是。

事实上，案例研究(Case Study)是现代科学广泛使用的一种方法。Case 包含的意义很广：包括 Example 例子，Instance 事例、示例，Actual State 实际状况，Circumstance 情况、事件、境遇，甚至 Project 项目、工程等。

我们知道在计算机的科学术语中，很多是直接来自日常生活的。例如 Computer 一词早在 1646 年就出现于古代英文字典中，但当时它的意义不是“计算机”而是“计算工人”，

即专门从事简单计算的工人。同理, Printer 当时也是“印刷工人”而不是“打印机”。正是由于这些“计算工人”和“印刷工人”常出现计算错误和印刷错误,才激发查尔斯·巴贝奇(Charles Babbage, 1791—1871)设计了差分机和分析机,这是最早的专用计算机和通用计算机。这位英国剑桥大学数学教授、机械设计专家、经济学家和哲学家是国际公认的“计算机之父”。

20 世纪 40 年代,人们还用 Calculator 表示计算机。到电子计算机出现后,才用 Computer 表示计算机。此外,硬件(Hardware)和软件(Software)来自销售人员。总线(Bus)就是公共汽车或大巴,故障和排除故障源自格瑞斯·霍普(Grace Hopper, 1906—1992)发现的“飞蛾子”(Bug)和“抓蛾子”或“抓虫子”(Debug)。其他如鼠标、菜单……不胜枚举。至于哲学家进餐问题,理发师睡觉问题更是操作系统文化中脍炙人口的经典。

以计算机为核心的信息技术,从一开始就与应用紧密结合。例如,ENIAC 用于弹道曲线的计算,ARPANET 用于资源共享以及核战争时的可靠通信。即使是非常抽象的图灵机模型,也受到二战时图灵博士破译纳粹密码工作的影响。

在信息技术中,既有许多成功的案例,也有不少失败的案例;既有先成功而后失败的案例,也有先失败而后成功的案例。好好研究它们的成功经验和失败教训,对于编写案例型教材有重要的意义。

我国正在实现中华民族的伟大复兴,教育是民族振兴的基石。改革开放 30 年来,我国高等教育在数量上、规模上已有相当的发展。当前的重要任务是提高培养人才的质量,必须从学科知识的灌输转变为素质与能力的培养。应当指出,大学课堂在高新技术的武装下,利用 PPT 进行的“高速灌输”、“翻页宣科”有愈演愈烈的趋势,我们不能容忍用“技术”绑架教学,而是让教学工作乘信息技术的东风自由地飞翔。

本系列教材的编写,以学生就业所需的专业知识和操作技能为着眼点,在适度的基础知识与理论体系覆盖下,突出应用型、技能型教学的实用性和可操作性,强化案例教学。本套教材将会有机融入大量最新的示例、实例以及操作性较强的案例,力求提高教材的趣味性和实用性,打破传统教材自身知识框架的封闭性,强化实际操作训练,使本系列教材做到“教师易教,学生乐学,技能实用”。有了广阔的应用背景,再造计算机案例型教材就有了基础。

我相信北京大学出版社在全国各地高校教师的积极支持下,精心设计,严格把关,一定能够建设出一批符合计算机应用型人才培养模式的、以案例型为创新点和兴奋点的精品教材,并且通过一体化设计、实现多种媒体有机结合的立体化教材,为各门计算机课程配齐电子教案、学习指导、习题解答、课程设计等辅导资料。让我们用锲而不舍的毅力,勤奋好学的钻研,向着共同的目标努力吧!

刘瑞挺教授 本系列教材编写指导委员会主任、全国高等院校计算机基础教育研究会副会长、中国计算机学会普及工作委员会顾问、教育部考试中心全国计算机应用技术证书考试委员会副主任、全国计算机等级考试顾问。曾任教育部理科计算机科学教学指导委员会委员、中国计算机学会教育培训委员会副主任。PC Magazine《个人电脑》总编辑、CHIP《新电脑》总顾问、清华大学《计算机教育》总策划。

前 言

数据库技术是当今世界高新技术潮流中的重要技术之一，也是计算机科学和信息管理应用领域的主要研究对象之一。从它的纵深知识内容来看，它蕴含了较深的数学概念和较多的模型定义及现代技术，从它的横向发展来看，它涉及了各个领域，诸如军事、商业、通信、娱乐等各个领域。

作为应用型本科类的学校，其人才培养目标不是为了造就研究型的知识精英，而是要打造有一定文化素养的、有实用价值的人才。因此，我们在本书的编写过程中，充分考虑到应用型本科学生自身的特点和发展方向，把数据库技术中的原理与具体的应用紧密结合，以数据库原理知识为背景、采用主流数据库 Microsoft SQL Server 2005 为实施工具，把数据库技术的运用方法和技巧融入到具体的数据库应用实例中去，深入浅出、循序渐进地讲解了数据库系统的基本概念和基本理论。

本书共分两大部分 12 章，第一部分为基础篇，包括第 1 章～第 7 章，第二部分为应用篇，包括第 8 章～第 12 章。

第 1 章对数据库系统进行概述，介绍数据库系统相关的概念、数据库系统的产生和发展及数据模型的基本概念。第 2 章讲解关系数据理论，其中包括关系的定义，关系代数、关系演算的相关概念和关系操作的基本方法，关系的完整性定义和关系的规范化法则。第 3 章简略介绍了目前流行的数据库 Microsoft SQL Server 2005 的特点、安装与配置的要求和方法、基本的管理工具和 Microsoft SQL Server 2005 服务器的启动和停用的操作等。第 4 章讲解了 Microsoft SQL Server 2005 的数据库结构、系统数据库及其用途，用户数据库的定义和删除、分离与附加、导入与导出的方法。第 5 章讲解了数据的类型及其创建、数据表结构的创建与修改、数据表中的数据管理、索引操作的基本方法与技巧等。第 6 章讲解了数据查询的方法和技巧，包括基本查询、条件查询、排序查询、分组查询、筛选查询、联结查询、嵌套查询等。第 7 章讲解了视图的基本概念及其使用方法，包括视图的创建、查询、修改等。

第 8 章讲解数据库设计的主要任务和实施过程，结合具体的案例介绍了数据的需求分析、概念结构、逻辑结构和物理结构的设计思想和设计方案的形成。第 9 章介绍了 Microsoft SQL Server 2005 数据库的安全管理机制及其实施策略的具体落实，包括登录账号、权限、角色的管理思想和方法。第 10 章讲解在 Microsoft SQL Server 2005 中数据完整性的具体实施方法和技巧，包括约束、规则的创建和管理。第 11 章讲解了 Microsoft SQL Server 2005 数据库编程的相关内容，包括 Transact-SQL 的基础知识（如标识符、变量、运算符和函数等），Transact-SQL 的编程基础（如批处理、流程控制、错误控制和注释等），事务编程的方法，存储过程的管理和触发器的使用方法等。第 12 章讲解了数据库日常维护和管理的基本知识，包括数据库的备份和还原、代理服务和维护计划的创建及管理。

本书是作者多年从事数据库教学的经验 and 感受的总结，本书的特点是涵盖知识比较全面，既包括了数据库的基础理论知识，又包括了数据库的应用技术，并提供了大量实例进行讲解、分析，为读者理解相关知识点、提高实际应用能力提供方便。

本教材由毛一梅、郭红任主编，秦福建、罗代忠任副主编，第1章、第12章由毛一梅编写，第2章至第7章由毛一梅、秦福建、马新强、罗代忠编写，第8章至第11章由郭红、毛一梅编写，最后由毛一梅统稿。

在整个编写过程中，得到了沈群力、龙青云等热情支持并提出了宝贵的意见，在此深表谢意。鉴于作者水平有限，书中不当之处望广大读者不吝赐教。

编者

2009年12月

目 录

第1章 绪论	1
1.1 数据库系统概述	2
1.1.1 数据	2
1.1.2 数据库	2
1.1.3 数据库管理系统	3
1.1.4 数据库系统	3
1.2 数据管理技术的产生和发展	4
1.2.1 人工管理阶段	4
1.2.2 文件系统阶段	4
1.2.3 数据库系统阶段	4
1.3 数据库模型	5
1.3.1 基本概念	5
1.3.2 数据库结构模型	5
1.4 本章小结	9
习题1	10
第2章 关系数据库理论	11
2.1 关系的数学定义	12
2.1.1 基本概念	12
2.1.2 关系模式与关系数据库	14
2.2 关系代数	15
2.2.1 集合运算	16
2.2.2 专门的关系运算	17
2.2.3 应用实例	20
2.3 关系演算	22
2.3.1 元组关系演算	22
2.3.2 域关系演算	26
2.3.3 应用实例	26
2.4 关系操作	30
2.5 关系的完整性	30
2.5.1 实体完整性	30
2.5.2 参照完整性	31
2.5.3 用户定义的完整性	32
2.6 关系的规范化	33
2.6.1 关系模式的设计问题	33
2.6.2 函数依赖	36
2.6.3 码的形式定义	37
2.6.4 关系范式	38
2.7 案例分析	41

2.8 本章小结	45
习题2	45
第3章 关系数据库——SQL Server 2005	
基础	48
3.1 SQL Server 2005 的特点	49
3.1.1 SQL Server 2005 的版本	49
3.1.2 SQL Server 2005 新特性	51
3.2 SQL Server 2005 的安装与配置	52
3.2.1 硬件要求	52
3.2.2 软件要求	53
3.2.3 SQL Server 2005 的安装	54
3.3 SQL Server 2005 的工具和实用程序	63
3.4 SQL Server 2005 的服务器管理	64
3.5 疑难分析	65
3.5.1 身份验证模式的选择	65
3.5.2 SQL Server 2005 的版本升级	67
3.6 本章小结	68
习题3	69
第4章 数据库的管理	70
4.1 SQL Server 2005 数据库的结构	71
4.1.1 SQL Server 2005 数据库的目录结构	71
4.1.2 SQL Server 2005 数据库的存储结构	71
4.2 SQL Server 2005 系统数据库	72
4.2.1 master 系统数据库	72
4.2.2 model 模板数据库	72
4.2.3 msdb 系统数据库	72
4.2.4 tempdb 临时数据库	73
4.2.5 resource 系统数据库	73
4.3 用户数据库的创建与修改	73
4.3.1 预备知识	73
4.3.2 在对象资源管理器中创建与修改数据库	76
4.3.3 用命令语句创建与修改数据库	80
4.3.4 应用实例	84

4.4 用户数据库的分离与附加	89	6.2.4 空值判断查询条件	138
4.4.1 用户数据库分离	89	6.2.5 使用逻辑运算符查询	139
4.4.2 用户数据库附加	91	6.2.6 应用实例	139
4.4.3 应用实例	93	6.3 排序查询	143
4.5 用户数据库的删除	94	6.3.1 基本语法	143
4.6 数据之间的导入与导出	95	6.3.2 应用实例	143
4.6.1 数据的导入	95	6.4 分组查询	144
4.6.2 数据的导出	99	6.4.1 基本语法	145
4.6.3 应用实例	102	6.4.2 应用实例	145
4.7 本章小结	103	6.5 筛选查询	146
习题4	104	6.5.1 基本语法	146
第5章 数据表的管理	105	6.5.2 应用实例	147
5.1 SQL Server 2005 的数据类型	106	6.6 联结查询	147
5.1.1 SQL Server 2005 提供的系统 数据类型	106	6.6.1 等值与非等值联结查询	147
5.1.2 用户自定义的数据类型	109	6.6.2 复合条件联结	148
5.2 表结构的创建与修改	111	6.6.3 自身联结	148
5.2.1 表结构的创建	111	6.6.4 外联结	148
5.2.2 表结构的修改	115	6.6.5 应用实例	149
5.2.3 表的删除	117	6.7 嵌套查询	153
5.2.4 应用实例	118	6.7.1 使用 IN 和 NOT IN	154
5.3 表数据的操作	119	6.7.2 使用比较运算符	154
5.3.1 插入数据	119	6.7.3 使用 EXISTS 和 NOT EXISTS	154
5.3.2 修改数据	121	6.7.4 应用实例	154
5.3.3 删除数据	122	6.8 综合实例	155
5.3.4 应用实例	122	6.9 本章小结	155
5.4 索引操作	123	习题6	156
5.4.1 索引的分类	123	第7章 视图	157
5.4.2 索引的创建	125	7.1 基本概念	158
5.4.3 索引的修改	127	7.1.1 视图概述	158
5.4.4 索引的删除	128	7.1.2 视图的优点和注意事项	158
5.4.5 应用实例	128	7.2 视图的创建	159
5.5 本章小结	131	7.2.1 使用对象资源管理器创建 视图	159
习题5	131	7.2.2 用命令语句创建视图	161
第6章 数据查询	135	7.2.3 应用实例	161
6.1 基本的 SELECT 查询	136	7.3 视图的查询	164
6.1.1 基本语法	136	7.3.1 视图属性的查询	164
6.1.2 应用实例	136	7.3.2 视图记录的查询	165
6.2 条件查询	137	7.3.3 应用实例	165
6.2.1 比较查询	137	7.4 视图的修改	166
6.2.2 范围查询	138	7.4.1 修改视图定义	166
6.2.3 模糊查询	138	7.4.2 重命名视图	167

7.4.3 编辑视图记录	168	9.2.2 SQL Server 登录账号的 管理	207
7.4.4 删除视图	169	9.2.3 应用实例	213
7.4.5 应用实例	169	9.3 权限管理	214
7.5 疑难分析	170	9.3.1 SQL Server 权限分类	214
7.5.1 视图数据更新的限制条件 ..	170	9.3.2 利用“对象资源管理器” 管理用户权限	215
7.5.2 保证列名唯一性的限制 条件	171	9.3.3 利用 T-SQL 语句管理用户 权限	217
7.5.3 视图权限的限制条件	171	9.3.4 应用实例	218
7.6 本章小结	171	9.4 角色管理	220
习题7	172	9.4.1 系统预定义角色	220
第8章 数据库设计	173	9.4.2 用户自定义角色	222
8.1 数据库设计概述	174	9.4.3 应用实例	224
8.1.1 数据库设计的特点	174	9.5 疑难分析	225
8.1.2 数据库设计的基本步骤 ..	175	9.6 本章小结	225
8.2 需求分析	177	习题9	226
8.2.1 需求分析的任务	177	第10章 数据的完整性	227
8.2.2 需求分析的方法	178	10.1 数据完整性概述	228
8.2.3 数据流图和数据字典	179	10.1.1 数据操作中存在的 问题	228
8.2.4 应用实例	184	10.1.2 数据完整性的分类	228
8.3 概念结构设计	186	10.2 数据完整性的实现	230
8.3.1 概念结构设计的方法和 步骤	186	10.3 约束	231
8.3.2 数据抽象与局部视图设计 ..	187	10.3.1 PRIMARY KEY 约束 ..	231
8.3.3 全局概念结构的集成	189	10.3.2 UNIQUE 约束	232
8.3.4 应用实例	191	10.3.3 CHECK 约束	232
8.4 逻辑结构设计	191	10.3.4 FOREIGN KEY 约束 ..	233
8.4.1 E-R 图向关系模型的转换 ..	192	10.3.5 NOT NULL 约束	234
8.4.2 逻辑结构的优化	194	10.3.6 DEFAULT 约束	235
8.4.3 外模式的设计	195	10.3.7 应用实例	235
8.4.4 应用实例	196	10.4 规则的创建和管理	242
8.5 物理结构设计	196	10.4.1 创建规则	242
8.5.1 确定数据库的物理结构 ..	196	10.4.2 绑定规则	242
8.5.2 评价物理结构	200	10.4.3 解除规则的绑定	243
8.6 数据库的实施	200	10.4.4 删除规则	243
8.7 数据库的运行和维护	202	10.4.5 查看规则	244
8.8 本章小结	203	10.4.6 应用实例	244
习题8	203	10.5 疑难分析	246
第9章 数据的安全性	204	10.5.1 完整性的检查次序	246
9.1 数据库的安全性机制	205	10.5.2 FOREIGN KEY 约束中关 联表的设置	246
9.2 SQL Server 的安全策略	205	10.6 本章小结	247
9.2.1 SQL Server 的安全认证 模式	206		

习题 10	248	11.6.5 应用实例	292
第 11 章 数据库编程	249	11.7 疑难分析	293
11.1 Transact-SQL 概述	250	11.7.1 有条件的 INSERT 触发器	293
11.1.1 Transact-SQL 的产生	250	11.7.2 指定 AFTER 触发器的 顺序	294
11.1.2 Transact-SQL 特点与 分类	250	11.8 本章小结	294
11.2 Transact-SQL 基础	252	习题 11	295
11.2.1 标识符	252	第 12 章 数据库的日常维护与管理	296
11.2.2 变量	253	12.1 数据的备份与还原	297
11.2.3 运算符	254	12.1.1 数据库的恢复模式	297
11.2.4 函数	256	12.1.2 数据库的备份	299
11.2.5 应用实例	261	12.1.3 数据库的还原	313
11.3 Transact-SQL 编程	264	12.1.4 应用实例	321
11.3.1 批处理	264	12.2 代理服务	327
11.3.2 流程控制语句	265	12.2.1 启动和停止 SQL Server 代理服务	327
11.3.3 错误捕获语句	269	12.2.2 设置 SQL Server 代理服务 的属性	329
11.3.4 注释	270	12.2.3 自动管理组件	329
11.3.5 应用实例	271	12.2.4 作业管理	331
11.4 事务编程	273	12.2.5 警报管理	333
11.4.1 事务概述	273	12.2.6 操作员管理	336
11.4.2 事务处理语句	275	12.2.7 应用实例	338
11.4.3 应用实例	276	12.3 数据库的维护	346
11.5 存储过程	277	12.3.1 维护任务	346
11.5.1 存储过程概述	277	12.3.2 维护计划的创建	347
11.5.2 存储过程的创建	279	12.3.3 维护计划的修改	347
11.5.3 存储过程的执行	280	12.3.4 维护计划的执行	347
11.5.4 存储过程的管理	280	12.3.5 应用实例	347
11.5.5 应用实例	282	12.4 本章小结	355
11.6 触发器	286	习题 12	355
11.6.1 触发器概述	286	参考文献	357
11.6.2 创建触发器	287		
11.6.3 DML 触发器的工作 原理	290		
11.6.4 管理触发器	290		

绪 论

教学目标

1. 掌握数据库的基本概念。
2. 了解数据管理的发展历史。
3. 熟悉基本的数据模型。

在信息时代的今天,各行各业都有各自不同的信息管理系统,几乎所有的信息管理系统都要用到数据库,相信大家也曾听过许多关于数据库的专业术语,那么,也许大家会问:

- 什么是数据?数据和信息之间有什么关系?
- 什么是数据库?数据库、数据库系统、数据库管理系统之间又有什么联系?
- 什么是数据库结构模型?不同的数据库结构模型各有什么优缺点?

本章将详细介绍数据库及数据库技术的相关概念和知识。

1.1 数据库系统概述

在科技飞速发展的今天,信息无处不在,为了及时获取有效的信息,人们通常要把数据收集起来,然后进行加工处理,从中发现有用的信息。因此,数据处理是当前计算机的主要应用之一,数据库技术是作为一门数据处理技术而发展起来的,所研究的问题就是如何科学地组织和存储数据,如何高效地获取和处理数据。

在当今这信息爆炸的年代,随着数据量的日益膨胀,数据库技术作为信息分析的核心和基础得到了越来越广泛的应用。

1.1.1 数据

众所周知,做任何一件事情决策很重要,正确的决策必须有正确的信息作为依据,这些信息来源于事实,对事实的记录称之为数据。

什么是数据?数据(Data)就是对客观事实的记录,它是可以鉴别的符号,这种符号可以是数字、文字、图形、图像、声音等多种表现方式。

在现代计算机系统中,数据的概念是广义的。早期的计算机系统主要用于解决烦琐的数字计算,处理的数据主要是整数、实数、浮点数等传统数学中的数据。现代计算机能够存储和处理的对象十分广泛,不仅可以是数字、文本、图形,还可以是音频、视频等多媒体数据,因此,数据的形式越来越复杂。

数据和其语义是不可分的。所谓数据的语义就是指对数据的解释,例如,402是一个数据,它可能是一个门牌号,也可能是一个货品的编号或是价格,如果只有一个数据而没有对它的解释,那么这个数据是无意义的。因此,数据不是一个孤立的符号,伴随着数据的出现必须有对该数据含义的说明,也就是说,数据是要有语义的。

数据与信息有什么关系?数据与信息不同,数据指的是用符号记录下来的可区别的一种事物的特征或事实,信息是反映现实世界的知识。信息以数据的形式表示,即数据是信息的载体;信息是抽象的,而数据是具体的,信息不随数据设备所决定的数据形式而改变,信息是经过对数据的加工,对客观世界产生影响的数据,信息是对数据的解释。而数据的表示方式可以是不同的,数据是对客观事实的记录,也是对信息的一种描述。

同一个数据经过不同人的处理可以产生不同的信息,同一个数据在不同背景下也会产生不同的信息。例如,同样一个产品的销售数据对一个大型企业来说,可能会觉得销售量太小,需要减少产量,调整产品销售策略,而这个销售数据对一个小型企业来说,可能就觉得销售量很大,需要增加该产品的产量等。

数据处理是指将数据向信息转换的过程,它包括对数据的收集、存储、传播、检索、分类、加工和输出等活动。

1.1.2 数据库

什么是数据库?数据库(Database)是长期存储在计算机内的、有组织的、可共享的大量数据的集合。数据库中的数据不是杂乱无章地堆积在一起的,而是按照一定的数据模型组织、描述和存储的。数据库中的数据相互关联,它可以为多个用户、多个程序所共享,

具有较小冗余度,数据间联系密切,而又有较高的数据独立性。

数据库技术要解决的主要问题就是如何科学地组织和存储数据,如何高效地获取、更新和加工处理数据,并保证数据的安全性、可靠性和共享性。

数据库技术从诞生到现在,在不到半个世纪的时间里,形成了坚实的理论基础,成熟的商业产品和广泛的应用领域,吸引越来越多的研究者加入。数据库的诞生和发展给计算机信息管理带来了一场巨大的革命。近30多年来,国内外已经开发建设了成千上万个数据库,它已成为企业、部门乃至个人日常工作、生产和生活的基础设施。同时,随着应用的扩展与深入,数据库的数量和规模越来越大,数据库的研究领域也已经大大地拓广和深化了。

1.1.3 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层数据管理软件,为用户或应用程序提供访问数据库的方法,是用来管理数据库的计算机应用软件,可以让用户很方便地对数据库进行维护、排序、检索和统计等操作。数据库管理系统的主要目标就是使数据成为方便用户使用的资源,易于为各类用户所共享,它建立在操作系统的基础之上,对数据库进行统一的管理和控制。

数据库管理系统是用户与数据库的接口,应用程序只有通过数据库管理系统才能和数据库打交道。数据库管理系统的基本功能有:定义数据、组织和管理数据、数据库运行管理、数据库创建和维护等。

数据库管理系统是一个大型的、复杂的软件系统,是计算机中的基础软件。目前专门研制数据库管理系统的厂商及其研制的DBMS产品很多,比较著名的有IBM公司的DB2关系数据库管理系统和IMS层次数据库管理系统、Oracle公司的Oracle关系数据库管理系统、Sybase公司的Sybase关系数据库管理系统和微软公司的Access、SQL Server关系数据库管理系统等。

1.1.4 数据库系统

数据库系统(Database System, DBS)是实现有组织地、动态地存储大量关联数据,方便多用户访问的计算机软件、硬件和数据资源组成的系统,即采用了数据库技术的计算机系统。从狭义上来讲,数据库系统主要是指数据库、数据库管理系统和用户。从广义上来讲,它不仅包括数据库、数据库管理系统和用户,还包括计算机硬件、操作系统和维护人员。其中,数据库管理系统是数据库系统的核心和主体,它保证了数据库的独立性和共享性。

数据库、数据库系统、数据库管理系统之间又有什么联系呢?可以用一个图书馆系统来比拟一个数据库系统,把数据库看作图书馆里的书库,数据库中的数据看作图书馆中的图书,把数据库管理系统看作是图书馆管理的操作规程,图书馆中的一切操作如书的存储、查阅、借还等以及所有的日常管理都必须按照图书馆指定的操作规程进行,而数据库中对数据的任何操作包括数据的定义、数据查询、数据维护、数据库运行控制等也都必须在数据库管理系统的管理之下进行。

1.2 数据管理技术的产生和发展

数据管理指的是对数据的分类、组织、编码、存储、检索和维护等。计算机的数据管理主要经历了人工管理、文件系统、数据库系统3个阶段。

1.2.1 人工管理阶段

在20世纪50年代中期以前,计算机主要用于科学计算。由于当时的外存只有纸带、卡片、磁带,没有磁盘等直接存取的存储设备,数据只是在需要时输入,用完后撤走,且没有专门用于管理硬件设备的操作系统,没有管理数据的专门软件,数据开发人员在应用程序中自己设计、定义和管理数据,应用程序中不仅要规定数据的逻辑结构,还要设计物理结构,包括存储结构、存取方法、输入方式等,数据处理方式是批处理。所有的数据完全由人工进行管理,因此这个阶段被称为人工管理阶段。在这个阶段中,数据本身不能独立存储和提供应用,数据只能是附属于计算机程序的一部分,不能在应用程序中共享,它随着应用程序一起运行与消失。

1.2.2 文件系统阶段

20世纪50年代后期到20世纪60年代中期,随着计算机硬件的发展,有了磁盘、磁鼓等直接存取的存储设备,计算机的应用范围不再局限于科学计算,操作系统中已经有了专门的数据管理软件,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。

在文件系统阶段,数据可以以文件的组织方式长期保存在外存上,供应用程序反复进行查询、修改、插入、删除等操作;程序和数据之间有了一定的独立性,操作系统提供了文件管理功能和访问文件的存取方法,程序和数据之间有了数据存取的接口,数据具有一定的共享性,但是它的共享性是有一定局限的,当不同的应用程序使用具有部分相同的数据时,仍必须建立各自的文件,而不能共享相同的数据,造成数据的冗余度大,不能确保数据的一致性。文件系统阶段的数据与程序之间相互依赖性还是比较强的。

1.2.3 数据库系统阶段

文件系统中存在的各种问题使人们把希望寄托在数据库系统中。数据库是由逻辑上关联的数据组成的,它存储在一个数据“储藏室”中,数据库在最终用户数据的存储、访问和管理方面采用了不同的方式,通过数据库系统中的数据库管理系统软件,人们可以在很大程度上消除在文件系统中存在的数据不一致、数据异常、数据依赖和结构依赖等问题。

数据库系统与文件系统的主要区别在于:文件系统是操作系统的重要组成部分,而数据库管理系统是独立于操作系统、在操作系统之上实现的软件,数据库中数据的组织和存储是通过操作系统中的文件系统来实现的;文件系统上的文件是为某一特定应用服务的,文件的逻辑结构对该应用程序来说是优化的,但要想对现有的数据再增加一些新的应用会很困难,系统不容易扩充,而数据库系统面向现实世界,它共享性高、冗余度

小、易扩充,具有较高的物理独立性和一定的逻辑独立性,整体数据结构化,用数据模型来描述,由数据库管理系统提供数据的安全性、完整性、并发控制能力和数据恢复能力。

使用数据库系统可以大大提高应用开发的效率。因为在数据库系统中,应用程序不必考虑数据的定义、存储和数据存取的具体路径,这些工作都由数据库管理系统来完成。

使用数据库系统可以减少开发人员的工作量。当因应用逻辑发生改变而需要改变数据的逻辑结构时,由于数据库系统提供了数据与程序之间的独立性,数据逻辑结构改变时开发人员不必修改应用程序,或者只需要修改很少的应用程序,从而既简化了应用程序的编制,又减少了应用程序的维护和修改。

使用数据库系统可以减轻数据库系统管理人员维护系统的负担。因为数据库管理系统在数据库建立、运行和维护时对数据库进行统一的管理和控制,包括数据的完整性、安全性、多用户并发控制、故障恢复等。

总之,使用数据库系统的优点是很多的,既便于数据的集中管理,控制数据冗余,提高数据的利用率和一致性,又有利于应用程序的开发和维护。

1.3 数据库模型

1.3.1 基本概念

由于计算机不能直接处理现实世界中的具体事物,因此,必须将那些具体的事物转换成计算机能够处理的数据。数据库中用数据模型来抽象、描述和处理现实世界中的数据。数据模型是数据库的核心概念,每个数据库中的数据都是按照某种特定的数据模型来组织的。

数据库模型是数据库中用来表示数据结构和数据联系的逻辑概述的集合,数据库模型可以分为两种类型:概念模型和结构模型。

数据库的概念模型是独立于计算机系统的模型,它强调的是数据库中描述的是什么,而不是如何描述它,概念模型通常用来描述某个特定组织所关心的信息结构。关于概念数据模型本书将在第8章中做详细的介绍。这里重点要讨论的是数据的结构模型,它直接面向数据库的逻辑结构,是现实世界的第二层抽象。

数据结构模型的好坏直接影响数据库的性能,因此,选择数据的结构模型是设计数据库的一项首要任务,现有的各种数据库管理系统软件都是基于某种结构模型的。数据库的结构模型包含数据结构、数据操作、数据完整性约束3个部分。

数据结构是所研究的对象类型的集合,常见的数据结构有层次结构、网状结构和关系结构等;数据操作是指对数据库各种对象的实例允许执行的操作的集合,它包括操作及有关的操作规则;数据完整性约束条件是数据完整性规则的集合,完整性规则是给定的数据模型中数据及其联系所具有的依存规则,用以限定符合数据模型的数据库状态及其状态的变化。

1.3.2 数据库结构模型

数据库结构模型的3个方面内容完整地描述了一个数据模型,其中的数据结构是描述