

21世纪高等学校规划教材 | 计算机科学与技术

大型数据库管理系统Oracle 11g教程

大型数据库概论

朱辉生 陈琳 李金海 主编



清华大学出版社

21世纪高等学校规划教材 | 计算机科学与技术

A decorative floral graphic in a light gray tone, featuring a large, stylized flower with multiple petals and a long, curved stem with several leaves, positioned behind the main title.

大型数据库概论

朱辉生 陈琳 李金海 主编

清华大学出版社
北京

内 容 简 介

本书以 Oracle 11g 为蓝本,深入浅出地介绍大型数据库系统的相关知识。全书共 10 章,分别介绍数据库的基本理论、Oracle 11g 的安装配置与基本操作、PL/SQL 编程、Oracle 11g 数据库的体系结构、Oracle 11g 数据库的管理、Oracle 11g 数据库对象的管理、Oracle 11g 数据库的安全性、Oracle 11g 数据库的恢复、Oracle 11g 数据库的完整性与并发控制、Oracle 11g 数据库应用系统的开发。

本书内容翔实,示例丰富,结构合理,语言简洁。每章均有针对性很强的示例和适量的习题或实验题,配合读者对相关知识的掌握。

本书既可作为计算机科学与技术、信息管理与信息系统等本科专业“大型数据库概论”课程的教材,也可作为数据库应用系统开发人员的参考书。

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。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

大型数据库概论/朱辉生等主编. —北京:清华大学出版社,2018

(21 世纪高等学校规划教材·计算机科学与技术)

ISBN 978-7-302-49249-8

I. ①大… II. ①朱… III. ①数据库系统—概论 IV. ①TP311.13

中国版本图书馆 CIP 数据核字(2018)第 002556 号

责任编辑:闫红梅 赵晓宁

封面设计:傅瑞学

责任校对:梁毅

责任印制:李红英

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者:三河市君旺印务有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:12.75

字 数:309 千字

版 次:2018 年 2 月第 1 版

印 次:2018 年 2 月第 1 次印刷

印 数:1~2000

定 价:39.00 元

出版说明

随着我国改革开放的进一步深化,高等教育也得到了快速发展,各地高校紧密结合地方经济建设发展需要,科学运用市场调节机制,加大了使用信息科学等现代科学技术提升、改造传统学科专业的投入力度,通过教育改革合理调整和配置了教育资源,优化了传统学科专业,积极为地方经济建设输送人才,为我国经济社会的快速、健康和可持续发展以及高等教育自身的改革发展做出了巨大贡献。但是,高等教育质量还需要进一步提高以适应经济社会发展的需要,不少高校的专业设置和结构不尽合理,教师队伍整体素质亟待提高,人才培养模式、教学内容和方法需要进一步转变,学生的实践能力和创新精神亟待加强。

教育部一直十分重视高等教育质量工作。2007年1月,教育部下发了《关于实施高等学校本科教学质量与教学改革工程的意见》,计划实施“高等学校本科教学质量与教学改革工程”(简称“质量工程”),通过专业结构调整、课程教材建设、实践教学改革、教学团队建设等多项内容,进一步深化高等学校教学改革,提高人才培养的能力和水平,更好地满足经济社会发展对高素质人才的需要。在贯彻和落实教育部“质量工程”的过程中,各地高校发挥师资力量强、办学经验丰富、教学资源充裕等优势,对其特色专业及特色课程(群)加以规划、整理和总结,更新教学内容、改革课程体系,建设了一大批内容新、体系新、方法新、手段新的特色课程。在此基础上,经教育部相关教学指导委员会专家的指导和建议,清华大学出版社在多个领域精选各高校的特色课程,分别规划出版系列教材,以配合“质量工程”的实施,满足各高校教学质量和教学改革的需要。

为了深入贯彻落实教育部《关于加强高等学校本科教学工作,提高教学质量的若干意见》精神,紧密配合教育部已经启动的“高等学校教学质量与教学改革工程精品课程建设工作”,在有关专家、教授的倡议和有关部门的大力支持下,我们组织并成立了“清华大学出版社教材编审委员会”(以下简称“编委会”),旨在配合教育部制定精品课程教材的出版规划,讨论并实施精品课程教材的编写与出版工作。“编委会”成员皆来自全国各类高等学校教学与科研第一线的骨干教师,其中许多教师为各校相关院、系主管教学的院长或系主任。

按照教育部的要求,“编委会”一致认为,精品课程的建设工作从开始就要坚持高标准、严要求,处于一个比较高的起点上。精品课程教材应该能够反映各高校教学改革与课程建设的需要,要有特色风格、有创新性(新体系、新内容、新手段、新思路,教材的内容体系有较高的科学创新、技术创新和理念创新的含量)、先进性(对原有的学科体系有实质性的改革和发展,顺应并符合21世纪教学发展的规律,代表并引领课程发展的趋势和方向)、示范性(教材所体现的课程体系具有较广泛的辐射性和示范性)和一定的前瞻性。教材由个人申报或各校推荐(通过所在高校的“编委会”成员推荐),经“编委会”认真评审,最后由清华大学出版

社审定出版。

目前,针对计算机类和电子信息类相关专业成立了两个“编委会”,即“清华大学出版社计算机教材编审委员会”和“清华大学出版社电子信息教材编审委员会”。推出的特色精品教材包括:

- (1) 21 世纪高等学校规划教材·计算机应用——高等学校各类专业,特别是非计算机专业的计算机应用类教材。
- (2) 21 世纪高等学校规划教材·计算机科学与技术——高等学校计算机相关专业的教材。
- (3) 21 世纪高等学校规划教材·电子信息——高等学校电子信息相关专业的教材。
- (4) 21 世纪高等学校规划教材·软件工程——高等学校软件工程相关专业的教材。
- (5) 21 世纪高等学校规划教材·信息管理与信息系统。
- (6) 21 世纪高等学校规划教材·财经管理与应用。
- (7) 21 世纪高等学校规划教材·电子商务。
- (8) 21 世纪高等学校规划教材·物联网。

清华大学出版社经过三十多年的努力,在教材尤其是计算机和电子信息类专业教材出版方面树立了权威品牌,为我国的高等教育事业做出了重要贡献。清华版教材形成了技术准确、内容严谨的独特风格,这种风格将延续并反映在特色精品教材的建设中。

清华大学出版社教材编审委员会

联系人:魏江江

E-mail: weijj@tup.tsinghua.edu.cn

前言

Oracle 是世界上第一个以 SQL 语言为基础、以分布式数据库为核心的大型数据库管理系统。自 1979 年 Oracle 问世以来,Oracle 公司一直致力于信息管理现代化技术及产品的研究与开发,使 Oracle 在全球数据库市场中稳居龙头位置并成为大型数据库管理系统的工业标准。对于计算机科学与技术、信息管理与信息系统等本科专业学生而言,掌握 Oracle 主流版本 Oracle 11g 的理论及应用,既是对“数据库原理及应用”课程的巩固提高,也是对大型数据库应用系统开发的基础训练。

本书力求通过言简意赅的语言和丰富的示例来介绍 Oracle 11g,所有示例均在编者多年“大型数据库概论”课程的教学过程中实践过。全书共分 10 章。第 1~第 3 章为 Oracle 11g 的基础部分,主要介绍数据库基础、Oracle 11g 的安装配置与基本操作以及 PL/SQL 编程。第 4~第 6 章为 Oracle 11g 的核心部分,介绍 Oracle 11g 数据库的体系结构、数据库的管理和数据库对象的管理。第 7~第 9 章为 Oracle 11g 的管理部分,介绍 Oracle 11g 数据库的安全性、恢复、完整性和并发控制等数据保护技术。第 10 章为 Oracle 11g 的应用部分,主要介绍基于 Visual C++ 开发 Oracle 11g 数据库应用系统的方法。附录为手工创建数据库和初始化参数文件。

本书由朱辉生、陈琳和李金海主编,朱辉生编写了第 1 和第 2 章,陈琳编写了第 3~第 6 和第 8 章,李金海编写了第 7、第 9 和第 10 章及附录。

由于编者水平有限,书中难免有不足之处,敬请广大读者批评指正。对本书的意见请通过 zhs@fudan.edu.cn 反馈给我们,谢谢!

朱辉生

2017 年 8 月

目 录

| | |
|-----------------------------|----|
| 第 1 章 绪论 | 1 |
| 1.1 数据库的基本概念 | 1 |
| 1.1.1 数据 | 1 |
| 1.1.2 数据库 | 1 |
| 1.1.3 数据库管理系统 | 1 |
| 1.1.4 数据库系统 | 2 |
| 1.2 数据模型 | 2 |
| 1.2.1 数据模型的概念 | 2 |
| 1.2.2 数据模型的要素 | 2 |
| 1.2.3 概念模型 | 3 |
| 1.2.4 逻辑模型 | 3 |
| 1.3 数据库系统的体系结构 | 5 |
| 1.3.1 数据库系统的三级模式结构 | 5 |
| 1.3.2 数据库系统的二级映像与数据独立性 | 6 |
| 1.4 数据库管理系统 | 7 |
| 1.4.1 DBMS 的工作模式 | 7 |
| 1.4.2 DBMS 的功能 | 7 |
| 1.4.3 DBMS 的组成 | 8 |
| 1.5 Oracle 11g 的特点 | 9 |
| 1.5.1 Oracle 11g 的特点 | 9 |
| 1.5.2 Oracle 11g 的三个版本 | 9 |
| 1.6 小结 | 10 |
| 习题 1 | 10 |
| 第 2 章 Oracle 11g 的安装配置与基本操作 | 11 |
| 2.1 Oracle 11g 的安装配置 | 11 |
| 2.2 Oracle 11g 的基本操作 | 17 |
| 2.2.1 登录 Oracle 11g 数据库 | 17 |
| 2.2.2 启动 Oracle 11g 数据库 | 19 |
| 2.2.3 关闭 Oracle 11g 数据库 | 20 |
| 2.3 小结 | 23 |
| 习题 2 | 24 |

| | |
|--|-----------|
| 实验 1 Oracle 11g 的安装配置与基本操作 | 24 |
| 第 3 章 PL/SQL 编程 | 25 |
| 3.1 PL/SQL 基础 | 25 |
| 3.1.1 PL/SQL 语句块 | 25 |
| 3.1.2 PL/SQL 变量 | 26 |
| 3.1.3 PL/SQL 常用数据类型 | 27 |
| 3.1.4 PL/SQL 运算符 | 30 |
| 3.1.5 PL/SQL 表达式 | 31 |
| 3.2 PL/SQL 控制结构 | 31 |
| 3.2.1 顺序结构 | 31 |
| 3.2.2 选择结构 | 31 |
| 3.2.3 NULL 结构 | 32 |
| 3.2.4 循环结构 | 32 |
| 3.2.5 GOTO 语句 | 34 |
| 3.3 游标 | 35 |
| 3.3.1 显式游标 | 35 |
| 3.3.2 隐式游标 | 36 |
| 3.3.3 游标的属性 | 36 |
| 3.3.4 带参数的游标 | 37 |
| 3.4 SQL * Plus/Worksheet 的使用 | 38 |
| 3.4.1 SQL * Plus | 38 |
| 3.4.2 SQL Developer | 41 |
| 3.4.3 SQL Developer 与 SQL * Plus 的比较 | 44 |
| 3.5 小结 | 44 |
| 习题 3 | 45 |
| 实验 2 PL/SQL 编程 | 45 |
| 第 4 章 Oracle 11g 数据库的体系结构 | 46 |
| 4.1 Oracle 11g 数据库的逻辑结构 | 46 |
| 4.1.1 表空间 | 46 |
| 4.1.2 段 | 48 |
| 4.1.3 区 | 49 |
| 4.1.4 数据块 | 49 |
| 4.2 Oracle 11g 数据库的物理结构 | 49 |
| 4.2.1 初始化参数文件 | 49 |
| 4.2.2 数据文件 | 50 |
| 4.2.3 控制文件 | 51 |
| 4.2.4 重做日志文件 | 51 |

| | | |
|--------------|----------------------------------|-----------|
| 4.3 | Oracle 11g 实例 | 51 |
| 4.3.1 | 实例的概念 | 51 |
| 4.3.2 | 实例的内存结构 | 52 |
| 4.3.3 | 实例的进程 | 53 |
| 4.4 | 小结 | 55 |
| | 习题 4 | 55 |
| 第 5 章 | Oracle 11g 数据库的管理 | 56 |
| 5.1 | 管理数据库 | 56 |
| 5.1.1 | 使用数据库配置助手创建数据库 | 56 |
| 5.1.2 | 查看数据库 | 63 |
| 5.2 | 管理表空间 | 66 |
| 5.2.1 | 创建表空间 | 66 |
| 5.2.2 | 查看、修改表空间 | 71 |
| 5.2.3 | 删除表空间 | 72 |
| 5.3 | 管理段 | 72 |
| 5.3.1 | 创建段 | 72 |
| 5.3.2 | 查看和修改段 | 73 |
| 5.3.3 | 删除段 | 74 |
| 5.4 | 小结 | 74 |
| | 习题 5 | 74 |
| | 实验 3 Oracle 11g 数据库的管理 | 74 |
| 第 6 章 | Oracle 11g 数据库对象的管理 | 75 |
| 6.1 | 管理表 | 75 |
| 6.1.1 | 表的概念 | 75 |
| 6.1.2 | 创建表 | 75 |
| 6.1.3 | 查看、编辑表 | 85 |
| 6.1.4 | 使用 PL/SQL 语句对表中数据行进行更新 | 87 |
| 6.1.5 | 使用 PL/SQL 语句对表中数据进行查询 | 88 |
| 6.1.6 | 删除表 | 88 |
| 6.1.7 | 更改表名 | 89 |
| 6.2 | 管理索引 | 89 |
| 6.2.1 | 索引的概念 | 89 |
| 6.2.2 | 创建索引 | 90 |
| 6.2.3 | 查看、编辑索引 | 92 |
| 6.2.4 | 删除索引 | 93 |
| 6.3 | 管理视图 | 94 |
| 6.3.1 | 视图的概念 | 94 |

| | | |
|--------------|-------------------------------------|------------|
| 6.3.2 | 创建视图 | 94 |
| 6.3.3 | 查看、编辑视图 | 97 |
| 6.3.4 | 删除视图 | 98 |
| 6.4 | 管理同义词和序列 | 99 |
| 6.4.1 | 同义词的概念 | 99 |
| 6.4.2 | 管理同义词 | 99 |
| 6.4.3 | 序列的概念 | 102 |
| 6.4.4 | 管理序列 | 102 |
| 6.5 | 管理簇 | 106 |
| 6.5.1 | 簇的概念 | 106 |
| 6.5.2 | 创建簇 | 107 |
| 6.5.3 | 查看、编辑簇 | 108 |
| 6.5.4 | 删除簇 | 109 |
| 6.6 | 管理过程、函数和包 | 110 |
| 6.6.1 | 过程 | 110 |
| 6.6.2 | 函数 | 112 |
| 6.6.3 | 包 | 114 |
| 6.7 | 小结 | 118 |
| | 习题 6 | 119 |
| | 实验 4 Oracle 11g 数据库对象的管理(综合一) | 119 |
| 第 7 章 | Oracle 11g 数据库的安全性 | 120 |
| 7.1 | 用户管理 | 120 |
| 7.1.1 | 用户认证 | 120 |
| 7.1.2 | 创建用户 | 121 |
| 7.1.3 | 查看用户 | 124 |
| 7.1.4 | 修改用户 | 125 |
| 7.1.5 | 删除用户 | 126 |
| 7.2 | 权限管理 | 127 |
| 7.2.1 | 权限的分类 | 127 |
| 7.2.2 | 管理系统权限 | 128 |
| 7.2.3 | 管理对象权限 | 129 |
| 7.3 | 角色管理 | 130 |
| 7.3.1 | 预定义角色 | 130 |
| 7.3.2 | 创建角色 | 131 |
| 7.3.3 | 查看角色 | 132 |
| 7.3.4 | 为角色授予或撤销权限 | 133 |
| 7.3.5 | 设置默认角色及启用、禁用角色 | 134 |
| 7.3.6 | 删除角色 | 134 |

| | |
|--|------------|
| 7.4 小结 | 135 |
| 习题 7 | 135 |
| 实验 5 Oracle 11g 数据库的安全性 | 135 |
| 第 8 章 Oracle 11g 数据库的恢复 | 136 |
| 8.1 数据库备份概述 | 136 |
| 8.1.1 数据库备份的概念 | 136 |
| 8.1.2 数据库备份的模式 | 137 |
| 8.1.3 数据库备份的策略 | 137 |
| 8.2 Oracle 11g 数据库的备份 | 137 |
| 8.2.1 脱机备份 | 138 |
| 8.2.2 联机备份 | 138 |
| 8.2.3 使用企业管理器进行联机备份 | 139 |
| 8.3 数据库恢复概述 | 142 |
| 8.3.1 数据库恢复的概念 | 142 |
| 8.3.2 实例恢复与介质恢复 | 142 |
| 8.3.3 完全恢复和不完全恢复 | 142 |
| 8.4 Oracle 11g 数据库的恢复 | 143 |
| 8.4.1 使用企业管理器进行数据库恢复 | 143 |
| 8.4.2 使用 PL/SQL 命令进行数据库恢复 | 145 |
| 8.5 小结 | 145 |
| 习题 8 | 146 |
| 实验 6 Oracle 11g 数据库的恢复 | 146 |
| 第 9 章 Oracle 11g 数据库的完整性与并发控制 | 147 |
| 9.1 完整性控制 | 147 |
| 9.1.1 完整性约束 | 147 |
| 9.1.2 触发器 | 148 |
| 9.2 并发控制 | 151 |
| 9.2.1 事务 | 151 |
| 9.2.2 并发操作可能引起的数据不一致 | 154 |
| 9.2.3 Oracle 11g 的封锁机制 | 154 |
| 9.3 小结 | 156 |
| 习题 9 | 156 |
| 实验 7 Oracle 11g 数据库的完整性与并发控制 | 156 |
| 第 10 章 Oracle 11g 数据库应用系统的开发 | 158 |
| 10.1 Visual C++ 开发数据库应用系统概述 | 158 |
| 10.1.1 Visual C++ 简介 | 158 |

| | | |
|--------|--------------------------------------|-----|
| 10.1.2 | Visual C++开发数据库应用系统的特点 | 159 |
| 10.1.3 | Visual C++开发数据库应用系统的相关技术 | 160 |
| 10.2 | Visual C++开发数据库应用系统前的准备工作 | 162 |
| 10.2.1 | 数据库和数据库表的创建 | 162 |
| 10.2.2 | 数据源的配置 | 162 |
| 10.3 | 基于MFC ODBC类开发Oracle 11g数据库应用系统 | 165 |
| 10.3.1 | 创建应用系统框架 | 165 |
| 10.3.2 | 应用系统框架的资源分析 | 167 |
| 10.3.3 | 应用系统框架的文件分析 | 168 |
| 10.3.4 | 制作应用系统的主窗体 | 168 |
| 10.3.5 | 实现应用系统的增加、删除、修改与查询记录功能 | 171 |
| 10.4 | 基于ADO技术开发Oracle 11g数据库应用系统 | 176 |
| 10.4.1 | 创建应用系统框架 | 176 |
| 10.4.2 | 制作应用系统的启动界面 | 177 |
| 10.4.3 | 制作应用系统的主界面 | 178 |
| 10.4.4 | 制作“学生信息管理”对话框 | 179 |
| 10.4.5 | 制作“课程信息管理”对话框 | 183 |
| 10.4.6 | 制作“成绩信息管理”对话框 | 183 |
| 10.4.7 | 制作“关于”对话框 | 183 |
| 10.5 | 小结 | 184 |
| 习题 10 | | 184 |
| 实验 8 | Oracle 11g数据库应用系统的开发(综合二) | 184 |
| 附录 A | 手工创建数据库和初始化参数文件 | 186 |
| A.1 | 手工创建数据库 | 186 |
| A.2 | 手工初始化参数文件 | 188 |

数据库技术产生于 20 世纪 60 年代中期,是数据管理的最新技术,是计算机科学的重要分支,它的出现极大地促进了计算机应用向各行各业的渗透。数据库的基本概念、数据模型、数据库系统的体系结构、数据库管理系统等相关知识是掌握大型数据库管理系统 Oracle 11g 的基础。

本章学习目标:

- (1) 理解数据库的 4 个基本概念。
- (2) 掌握数据模型的概念、要素、分类和常见数据模型的特点。
- (3) 掌握数据库系统的三级模式结构和二级映像。
- (4) 掌握数据库管理系统的工作模式、功能和组成。
- (5) 了解 Oracle 11g 的特点。

1.1 数据库的基本概念

1.1.1 数据

数据(Data)是能够被计算机识别、存储和处理的信息。在计算机中,为了存储和处理现实世界中的具体事物,就要抽取这些事物的特性组成一个记录来描述。例如,学校管理部门对学生感兴趣的是学生的学号、姓名、性别、年龄和班级等,数据“20170801,张军,男,20,计算机科学与技术 1 班”表示张军同学的学号为 20170801、性别为男、年龄为 20 岁、就读于计算机科学与技术 1 班。

1.1.2 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。数据库中的数据按一定的数据模型组织、描述和存储,具有较小的冗余度、较高的数据独立性和易扩展性,并可为多个用户共享。

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。数据库管理系统集中管理和控制着数据库的建立、运行和维护,它使

得用户可以方便地定义和操纵数据,并能够保证数据的安全性、完整性、多用户并发访问以及故障发生后的数据库恢复。

1.1.4 数据库系统

数据库系统(Database System, DBS)是指引入数据库后的计算机系统,它一般由操作系统、数据库管理系统、数据库、应用程序、数据库管理员(Database Administrator, DBA)和用户构成。

在不引起混淆的情况下,可以把数据库系统简称为数据库。

1.2 数据模型

数据库是某家企业、组织或部门所涉及的数据集合,它不仅要反映数据本身的内容,而且要反映数据之间的联系。由于计算机不可能直接处理现实世界中的具体事物,所以人们必须事先把事物转换成能够处理的数据。

1.2.1 数据模型的概念

数据模型是对现实世界的模拟,是能够描述实体及实体之间联系的一种模型。

数据模型应满足三方面的要求:一是能比较真实地模拟现实世界;二是容易为人们所理解;三是便于在计算机上实现。但目前一种数据模型很难同时很好地满足这三方面的要求。

根据应用目的,可以将数据模型划分为两类,它们分属于两个不同的层次。一类是概念模型,它是按用户的观点对数据建模;另一类是逻辑模型(也称为结构模型),主要有层次模型、网状模型、关系模型、面向对象模型等,它是按计算机系统的观点对数据建模。

1.2.2 数据模型的要素

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三个要素组成的。

1. 数据结构

数据结构用于描述系统的静态特性(各种对象类型)。

数据结构是刻画一个数据模型最重要的方面,因此在数据库系统中,常常按照其数据结构的类型来命名数据类型。例如,层次结构、网状结构、关系结构中的数据类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(各种对象类型的实例允许执行的操作的集合)。

数据库中的数据操作主要有检索和更新(插入、删除和修改)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. 完整性约束

完整性约束是指给定数据模型中的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、相容和有效。

数据模型应该规定本数据模型必须遵守的基本完整性约束。例如,在关系模型中,任何关系必须满足实体完整性和参照完整性两个条件。

此外,数据模型还应该提供定义完整性约束的机制,以反映具体应用所涉及的数据必须遵守的特定的语义约束。例如,学生数据库中规定学生的年龄必须取正整数值、性别必须取男或女两个值之一等。

1.2.3 概念模型

为了把现实世界中的具体事物抽象、组织为某一 DBMS 支持的数据类型,人们常常首先将现实世界抽象为信息世界,再将信息世界转换为机器世界。概念模型就是现实世界到信息世界的第一层抽象,是对信息世界建立的不依赖于具体的计算机系统、不为某个 DBMS 所支持的数据模型,它是用户与数据库设计人员之间进行交流的语言。

概念模型用于信息世界的建模,应该能方便、准确地表示信息世界的常用概念(如实体、属性、联系等)。概念模型的表示方法很多,其中最为常用的是 P. P. S. Chen 于 1976 年提出的实体—联系方法(Entity-Relationship Approach, ER 方法),该方法用 ER 图来描述现实世界。

ER 图提供了表示实体、属性和联系的方法。其中,实体用矩形表示,矩形框内写明实体名;属性用椭圆表示,并用无向边将其与对应的实体连接起来;联系用菱形表示,菱形框内写明联系名,并用无向边分别与有关实体连接起来,同时,在无向边旁标上联系的类型(1:1, 1:n 或 m:n);联系本身也可以有属性,这些属性用无向边与该联系连接起来。

1.2.4 逻辑模型

逻辑模型与 DBMS 有关,直接面向数据库的逻辑结构。目前最常用的逻辑模型有层次模型、网状模型、关系模型和面向对象模型。

1. 层次模型

现实世界中许多实体之间的联系本来就呈现为一种很自然的层次关系,如行政机构、家族关系等。层次模型是数据库系统中最早出现的数据模型,典型代表是 IBM 公司于 1968 年推出的信息管理系统(Information Management System, IMS)。

1) 层次模型的数据结构

层次模型用树形结构表示各类实体以及实体之间的联系,因此它有两个限制:

- (1) 只有一个结点没有双亲结点,称之为根结点。
- (2) 根以外的其他结点有且只有一个双亲结点。

2) 层次模型的数据操作

层次模型的数据操作主要有查询、插入、删除和修改。

3) 层次模型的完整性约束

层次模型的完整性约束是指进行插入、删除和修改操作时要满足的约束:

- (1) 进行插入操作时,如果没有相应的双亲结点,就不能插入子女结点。
- (2) 进行删除操作时,如果删除双亲结点,则相应的子女结点也被同时删除。
- (3) 进行修改操作时,应修改所有的相应记录,以保证数据的一致性。

4) 层次模型的优点

(1) 对于实体间联系是固定的,且预先定义好的应用系统,采用层次模型来实现,其性能优于关系模型,不次于网状模型。

(2) 提供了良好的完整性支持。

5) 层次模型的缺点

(1) 现实世界中很多联系是非层次的,如多对多联系,用层次模型表示这些联系须引入冗余数据,容易产生数据的不一致。

(2) 对插入、删除操作的限制较多。

(3) 数据的独立性差。由于实体之间的联系本质上是通过存取路径指示的,因此应用程序在访问数据时要指定存取路径。

2. 网状模型

现实世界中实体之间的联系更多的是非层次关系,用层次模型表示非树形结构是很不直接的,网状模型则可以弥补这一不足。网状模型的典型代表是1969年由数据系统语言研究会(Conference on Data Systems Language, CODASYL)下属的数据库任务组(Database Task Group, DBTG)提出的DBTG报告。

1) 网状模型的数据结构

网状模型用图形结构表示各类实体以及实体之间的联系,它突破了层次模型数据结构的两个限制,允许多个结点没有双亲结点,允许一个结点可以有多个双亲结点。

2) 网状模型的数据操作

网状模型的数据操作主要有查询、插入、删除和修改。

3) 网状模型的完整性约束

(1) 进行插入操作时,允许插入尚未确定双亲结点的子女结点。

(2) 进行删除操作时,只需删除双亲结点,相应的子女结点仍然保留。

4) 网状模型的优点

(1) 能够更为直接地描述现实世界。

(2) 具有良好的性能,存取效率高。

5) 网状模型的缺点

(1) DDL语言极其复杂。

(2) 数据的独立性差。由于实体之间的联系本质上是通过存取路径指示的,因此应用程序在访问数据时要指定存取路径。

3. 关系模型

关系模型是目前最重要的一种数据模型。美国IBM公司的研究员E. F. Codd于1970

年发表了题为“大型共享系统的关系数据库的关系模型”的论文,文中首次提出了数据库系统的关系模型。20世纪80年代以来,计算机厂商新推出的DBMS大都支持关系模型。

1) 关系模型的数据结构

关系模型的数据结构是一张二维表,由行和列组成。但关系模型要求关系必须是规范化的,即要求关系模式必须满足一系列的规范条件,这些规范条件中最基本的一条是关系的每一分量必须是一个不可分割的数据项。

2) 关系模型的数据操作

关系模型的数据操作主要有查询、插入、删除和修改。

3) 关系模型的完整性约束

关系模型的完整性约束包括三大类:实体完整性、参照完整性和用户自定义完整性。

4) 关系模型的优点

(1) 关系模型建立在严格的数学概念基础上。

(2) 概念单一,实体及其联系均用关系表示,数据操作的对象及结果都是一个关系。

(3) 存取路径对用户透明,具有较高的数据独立性和安全性。

5) 关系模型的缺点

由于存取路径对用户透明,查询效率较差,为提高查询性能,一般要进行查询优化,这就增加了额外的开销。

4. 面向对象模型

虽然关系模型比层次模型、网状模型简单灵活,但还不能表达现实世界中存在的许多复杂的数据结构,如CAD数据、图形数据、嵌套递归的数据等,面向对象模型可以很好地解决这一问题。

面向对象概念最早出现在1968年的Smalltalk语言中,随后迅速渗透到计算机领域的每个分支,现已使用在数据库技术中。该模型的基本概念是对象和类。

对象是现实世界中实体的模型化,与记录相仿,但比记录复杂。每个对象有唯一的标识符,把状态和行为封装在一起。其中,对象的状态是该对象属性值的集合,对象的行为是在对象状态上操作的方法集。

类是相同对象组成的集合,可以从其父类中继承所有的属性和方法。

1.3 数据库系统的体系结构

虽然实际的数据库系统种类很多,它们支持不同的数据模型,使用不同的数据库语言,建立在不同的操作系统之上,数据的存储结构也各不相同,但从数据库管理系统的角度看,它们在体系结构上通常都具有相同的特征,即采用三级模式结构,并提供二级映像功能。

1.3.1 数据库系统的三级模式结构

数据库系统的三级模式结构是指数据库系统由外模式、模式和内模式三级构成,如