



普通高等教育“十一五”国家级规划教材

数据库 技术及应用

刘红岩 编著



清华大学出版社



普通高等教育“十一五”国家级规划教材

TP311.13

277

2007

数据库技术及应用

刘红岩 编著

清华大学出版社
北京

内 容 简 介

本书集数据库技术及应用为一体,深入浅出地介绍了数据库技术的基本原理以及数据库应用的开发方法。在数据库原理方面,本书以关系数据库为重点,讲述了数据库的基本概念、概念模型(包括实体联系模型和面向对象模型)的构建方法、关系模型、数据库标准语言 SQL、关系模式的规范化理论及其关系模式的设计方法,以及事务管理技术,包括并发控制和数据库恢复技术。在数据库应用方面,本书以一个实用案例贯穿全书,介绍了从数据库设计到应用实现的详细过程和方法。本书不仅介绍了数据库的理论知识,还介绍了以 IBM UDB DB2 为代表的大型关系数据库系统中 SQL 的使用方法,以及以 Microsoft Access 为代表的桌面数据库系统的使用和应用开发方法。

本书取材新颖、实用,内容丰富,容易理解。本书适合作为信息管理专业或计算机专业的数据库课程的教材,同时也可作为计算机软件相关人员的技术参考书。

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 13501256678 13801310933

图书在版编目(CIP)数据

数据库技术及应用/刘红岩编著. 北京:清华大学出版社,2007.2

ISBN 978-7-302-14440-3

I. 数… II. 刘… III. 数据库系统 IV. TP311.13

中国版本图书馆 CIP 数据核字(2006)第 163967 号

责任编辑:张瑞庆

责任校对:时翠兰

责任印制:王秀菊

出版发行:清华大学出版社

地 址:北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编:100084

c-service@tup.tsinghua.edu.cn

社 总 机:010-62770175

邮购热线:010-62786544

投稿咨询:010-62772015

客户服务:010-62776969

印 装 者:北京鑫海金澳胶印有限公司

经 销:全国新华书店

开 本:185×260 印 张:23.25

字 数:563千字

版 次:2007年2月第1版

印 次:2007年2月第1次印刷

印 数:1~5000

定 价:29.00元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。
联系电话:010-62770177 转 3103 产品编号:021255-01



前 言

数据库技术自 20 世纪 60 年代问世以来,经历了快速的发展,已经成为当今计算机领域中应用最广泛的技术之一,是计算机信息系统的核心技术和重要基础,其应用遍及各行各业。

数据库课程是信息管理和计算机等专业中的一门重要课程。本书作为数据库课程的教材,详细介绍了数据库系统的基本概念、原理和应用方法。本书的特点是,从开发一个数据库应用系统以及从数据库管理的角度讲解数据库系统的基本概念和技术。全书内容可分为 3 个层次,第 1~8 章作为第一层次,讲述了数据库的基本原理以及关系模式的设计方法;第 9~15 章作为第二层次,完整地讲述了开发数据库应用系统的过程和方法;第 16~17 章作为第三层次,讲述了数据库管理的核心内容,适合于课时较多的信息管理专业本科生和研究生使用。本书的具体安排是:第 1、2 章介绍了数据库系统的基本概念以及数据库应用系统的基本过程和方法。第 3、4 章介绍了两种常用的概念模型:实体联系模型和面向对象模型。第 5 章介绍了关系模型的基本概念和数据操纵语言(关系代数和关系演算)。第 6 章介绍了将实体联系模型转换为关系模型的方法。第 7、8 章介绍了关系模型的规范化理论以及关系模型的设计方法。第 9 章以 IBM UDB DB2 数据库系统为例介绍了关系数据库的标准语言 SQL。第 10~15 章讲述了 Microsoft Office 套件中的桌面数据库 Access 的使用方法,包括数据库的创建方法、基于实例的查询(query by example, QBE)方法以及应用模块(包括窗体、报表、图表、宏等)的开发方法。第 16 章介绍了数据库安全管理技术,包括一般策略和方法,以及在 Access 中进行安全管理的方法。第 17 章讲述了事务管理技术,包括基于封锁的并发控制技术和基于日志的数据库恢复技术。附录 A 介绍了 ER 模型的传统图形表示方法。

数据库技术既具有坚实的理论基础,同时其实用性又很强。本书在介绍理论知识的同时,介绍了实用的数据库系统 IBM UDB DB2 和 Microsoft

Access,便于学生将理论知识在真实的数据库环境中进行实践和体验。通过本书的学习,即使不擅长编写代码的读者也可以开发出一个完整的、界面友好的数据库应用系统。

本书内容可根据各高校数据库课程课时的设置不同,适当地跳过第4章、第7章、第8章、第16章及第17章内容,并不会影响其他章节内容的学习。

本书内容是作者在清华大学十几年讲授数据库课程的经验的基础上总结而成,内容丰富、通俗易懂、实用性强。具有以下特点:

- 适于案例教学模式。为了使数据库技术的讲解更易为读者所理解,全书以一个体育用品销售信息管理系统作为案例贯穿全书,用以说明相关概念和操作。各章内容均由浅入深、逐步深入,便于循序渐进地学习并学以致用。
- 完备的教学资源。本书的每章都配有习题,并且适当设置大作业习题(注:习题序号右上角标有*为大作业),便于学生动手实践。书中在一些重要章节后还提供了丰富的参考文献,便于学生进一步阅读和学习。

通过本书的学习,可以帮助读者建立起一个完整的数据库技术与应用的知识体系,熟练掌握数据库系统的实用开发技术。

本书适合作为大学本科、专科以及高等职业院校、软件职业技术学院的数据库以及数据库技术及应用课程的教材,也可作为初学者学习数据库的入门教材,并可作为数据库应用系统开发人员的技术参考书。本书也适合作为数据库技术的自学教材和参考书。

原香港科技大学计算机系陆宏钧教授生前曾对本书的内容提出了许多宝贵意见,在此表示诚挚的谢意。本书的书写得到了国家自然科学基金(70471006、70621061)的支持,在此表示感谢。

鉴于作者水平有限,书中难免存在一些缺点和不足之处,恳请广大读者批评指正。

刘红岩
2007年元旦于清华园



目 录

| | |
|-------------------------------|----|
| 第 1 章 数据库系统简介 | 1 |
| 1.1 数据库及数据库系统 | 1 |
| 1.2 数据库管理系统 | 2 |
| 1.3 数据库管理系统的发展 | 4 |
| 1.4 数据模型 | 6 |
| 1.5 数据库系统的三级体系结构 | 7 |
| 1.6 本章小结 | 8 |
| 习题 | 9 |
| 参考文献 | 9 |
| 第 2 章 数据库应用开发 | 11 |
| 2.1 文件系统和数据库管理系统 | 11 |
| 2.2 数据库开发过程 | 15 |
| 2.3 产品销售信息管理系统案例 | 17 |
| 2.3.1 切换面板界面的产品销售信息管理系统 | 17 |
| 2.3.2 菜单界面的产品销售信息管理系统 | 23 |
| 2.4 本章小结 | 26 |
| 习题 | 27 |
| 参考文献 | 27 |
| 第 3 章 实体联系模型 | 28 |
| 3.1 实体联系模型 | 28 |
| 3.1.1 实体 | 29 |
| 3.1.2 联系 | 30 |
| 3.1.3 属性 | 31 |

| | | |
|--------------|--------------------------|-----------|
| 3.1.4 | 结构约束 | 33 |
| 3.1.5 | 设计 ER 模型的常见问题 | 37 |
| 3.2 | 增强实体联系模型 | 41 |
| 3.2.1 | 特殊化 | 41 |
| 3.2.2 | 概括 | 43 |
| 3.2.3 | 聚集 | 44 |
| 3.3 | 体育用品公司的 ER 模型 | 44 |
| 3.4 | 本章小结 | 46 |
| | 习题 | 46 |
| | 参考文献 | 48 |
| 第 4 章 | 面向对象数据模型 | 49 |
| 4.1 | 对象和类 | 49 |
| 4.2 | 关联 | 51 |
| 4.3 | 关联类 | 51 |
| 4.4 | 概括 | 53 |
| 4.5 | 多态和重载 | 54 |
| 4.6 | 聚合和合成 | 55 |
| 4.7 | 体育用品公司的面向对象模型 | 56 |
| 4.8 | 本章小结 | 57 |
| | 习题 | 57 |
| | 参考文献 | 58 |
| 第 5 章 | 关系模型 | 59 |
| 5.1 | 关系模型的数据结构 | 59 |
| 5.1.1 | 关系 | 59 |
| 5.1.2 | 数据库模式 | 61 |
| 5.1.3 | 键 | 62 |
| 5.2 | 关系模型的约束 | 63 |
| 5.2.1 | 域完整性约束 | 63 |
| 5.2.2 | 实体完整性约束 | 63 |
| 5.2.3 | 参照完整性约束 | 63 |
| 5.2.4 | 一般性约束 | 64 |
| 5.3 | 关系模型的数据操纵语言 | 64 |
| 5.3.1 | 关系代数 | 64 |
| 5.3.2 | 元组关系演算 | 72 |
| 5.3.3 | 域关系演算 | 74 |
| 5.3.4 | 关系代数和关系演算的表达能 力 | 75 |
| 5.4 | 本章小结 | 76 |

| | |
|----------------------------------|------------|
| 习题 | 76 |
| 参考文献 | 77 |
| 第 6 章 实体联系模型转换为关系模型 | 79 |
| 6.1 实体及其属性的转换 | 79 |
| 6.2 联系及其属性的转换 | 80 |
| 6.2.1 二元联系的转换 | 80 |
| 6.2.2 一元联系的转换 | 82 |
| 6.2.3 多元联系的转换 | 83 |
| 6.3 概括/特殊化的转换 | 83 |
| 6.4 本章小结 | 84 |
| 习题 | 84 |
| 第 7 章 关系数据库的规范化理论 | 85 |
| 7.1 更新异常问题 | 85 |
| 7.2 函数依赖和 Armstrong 公理 | 87 |
| 7.2.1 函数依赖 | 88 |
| 7.2.2 Armstrong 公理 | 89 |
| 7.2.3 属性组的闭包和函数依赖集的最小覆盖 | 90 |
| 7.3 范式 | 92 |
| 7.3.1 1NF | 93 |
| 7.3.2 2NF | 95 |
| 7.3.3 3NF | 95 |
| 7.3.4 BCNF | 97 |
| 7.4 其他依赖及范式 | 98 |
| 7.4.1 多值依赖和 4NF | 99 |
| 7.4.2 连接依赖和 5NF | 101 |
| 7.4.3 包含依赖 | 102 |
| 7.5 本章小结 | 103 |
| 习题 | 103 |
| 参考文献 | 105 |
| 第 8 章 关系模式的设计方法 | 107 |
| 8.1 规范化关系模式 | 107 |
| 8.1.1 分解的原则 | 107 |
| 8.1.2 规范化为 BCNF | 112 |
| 8.1.3 规范化为 3NF | 113 |
| 8.2 关系模式的合并和调整 | 115 |
| 8.3 本章小结 | 116 |

| | |
|--------------------------------|------------|
| 习题 | 117 |
| 参考文献 | 118 |
| 第 9 章 关系数据库实用语言 | 119 |
| 9.1 数据库的查询 | 120 |
| 9.1.1 简单查询 | 122 |
| 9.1.2 连接查询 | 125 |
| 9.1.3 聚集和分组查询 | 128 |
| 9.1.4 嵌套查询 | 129 |
| 9.1.5 集合操作查询 | 133 |
| 9.2 数据的更新 | 135 |
| 9.2.1 数据的插入 | 135 |
| 9.2.2 数据的删除 | 136 |
| 9.2.3 数据的修改 | 137 |
| 9.3 数据库对象的定义 | 137 |
| 9.3.1 表的定义及修改 | 137 |
| 9.3.2 表的完整性约束 | 141 |
| 9.3.3 索引的创建和修改 | 145 |
| 9.3.4 视图 | 146 |
| 9.4 触发器 | 149 |
| 9.5 嵌入式 SQL | 151 |
| 9.5.1 声明变量 | 153 |
| 9.5.2 简单的嵌入式 SQL 语句 | 153 |
| 9.5.3 游标 | 154 |
| 9.5.4 动态 SQL | 156 |
| 9.6 SQL:1999 的新特性 | 157 |
| 9.6.1 关系方面的新特性 | 157 |
| 9.6.2 面向对象方面的新特性 | 158 |
| 9.7 本章小结 | 159 |
| 习题 | 159 |
| 参考文献 | 162 |
| 第 10 章 在 MS Access 中创建表 | 163 |
| 10.1 MS Access 简介 | 163 |
| 10.2 启动 MS Access | 164 |
| 10.3 创建数据库 | 165 |
| 10.3.1 创建空数据库 | 166 |
| 10.3.2 利用向导创建数据库 | 166 |
| 10.3.3 打开数据库 | 168 |

| | | |
|---------------|----------------------------|------------|
| 10.4 | 数据库中表的创建 | 169 |
| 10.4.1 | 使用设计器创建表 | 169 |
| 10.4.2 | 定义字段属性 | 174 |
| 10.4.3 | 查看和筛选表中数据 | 180 |
| 10.4.4 | 导入、导出和链接表 | 184 |
| 10.4.5 | 定义表的其他方法 | 189 |
| 10.4.6 | 定义表之间的关系 | 191 |
| 10.5 | 如何获得帮助 | 192 |
| 10.5.1 | Office 助手 | 192 |
| 10.5.2 | 利用?按钮获取帮助 | 192 |
| 10.5.3 | “Access 帮助”任务窗格 | 194 |
| 10.6 | 本章小结 | 195 |
| | 习题 | 195 |
| 第 11 章 | 在 MS Access 中创建查询对象 | 197 |
| 11.1 | Microsoft Access 的查询对象 | 197 |
| 11.2 | 选择查询 | 198 |
| 11.2.1 | 简单查询 | 198 |
| 11.2.2 | 连接查询 | 200 |
| 11.2.3 | 分组查询 | 202 |
| 11.3 | 参数查询 | 204 |
| 11.4 | 交叉表查询 | 205 |
| 11.5 | 动作查询 | 208 |
| 11.5.1 | 生成表查询 | 208 |
| 11.5.2 | 追加查询 | 209 |
| 11.5.3 | 更新查询 | 209 |
| 11.5.4 | 删除查询 | 210 |
| 11.6 | 利用向导创建其他查询 | 211 |
| 11.6.1 | 简单查询向导 | 211 |
| 11.6.2 | 查找重复项查询向导 | 214 |
| 11.6.3 | 查找不匹配项查询向导 | 216 |
| 11.7 | 本章小结 | 218 |
| | 习题 | 219 |
| 第 12 章 | 在 MS Access 中创建窗体 | 220 |
| 12.1 | 窗体简介 | 220 |
| 12.2 | 自动创建窗体 | 220 |
| 12.2.1 | 自动创建纵栏式窗体 | 221 |
| 12.2.2 | 利用窗体查看和编辑数据 | 223 |

| | | |
|---------------|--------------------------|------------|
| 12.2.3 | 改变窗体中控件的类型 | 223 |
| 12.3 | 修改窗体的设计 | 227 |
| 12.3.1 | 替换已有控件 | 227 |
| 12.3.2 | 调整控件布局 | 229 |
| 12.3.3 | 创建命令按钮 | 231 |
| 12.4 | 利用向导创建窗体 | 233 |
| 12.5 | 创建主子窗体 | 236 |
| 12.5.1 | 利用“自动窗体”创建主子窗体 | 236 |
| 12.5.2 | 利用“窗体向导”创建主子窗体 | 236 |
| 12.5.3 | 利用子窗体控件创建主子窗体 | 239 |
| 12.6 | 创建图表 | 244 |
| 12.7 | 利用设计视图创建窗体 | 246 |
| 12.8 | 本章小结 | 251 |
| | 习题 | 251 |
| 第 13 章 | 在 MS Access 中创建报表 | 253 |
| 13.1 | 报表简介 | 253 |
| 13.2 | 自动创建报表 | 253 |
| 13.2.1 | 自动创建表格式报表 | 253 |
| 13.2.2 | 调整报表的外观 | 257 |
| 13.3 | 利用向导创建报表 | 259 |
| 13.4 | 创建带参数报表 | 264 |
| 13.4.1 | 创建带参数查询 | 265 |
| 13.4.2 | 创建报表 | 265 |
| 13.4.3 | 创建窗体 | 265 |
| 13.5 | 创建图表 | 269 |
| 13.6 | 利用设计视图创建报表 | 272 |
| 13.6.1 | 创建分组表格式报表 | 272 |
| 13.6.2 | 创建组合报表 | 277 |
| 13.6.3 | 创建排名报表 | 280 |
| 13.7 | 本章小结 | 283 |
| | 习题 | 283 |
| 第 14 章 | 宏和应用查询系统 | 285 |
| 14.1 | 宏 | 285 |
| 14.1.1 | 创建宏 | 285 |
| 14.1.2 | 创建有条件宏 | 287 |
| 14.1.3 | 创建宏组 | 291 |
| 14.2 | 面向用户的查询系统的实现 | 292 |

| | | |
|---------------|-------------------------------|------------|
| 14.2.1 | 利用窗体创建查询界面 | 292 |
| 14.2.2 | 创建查询对象 | 294 |
| 14.2.3 | 创建包含查询结果的窗体 | 295 |
| 14.2.4 | 创建展示查询结果的报表 | 295 |
| 14.2.5 | 进一步设计查询界面窗体 | 296 |
| 14.2.6 | 创建宏 | 298 |
| 14.2.7 | 进一步完善查询界面窗体 | 299 |
| 14.3 | 本章小结 | 300 |
| | 习题 | 300 |
| 第 15 章 | 在 MS Access 中创建切换面板和菜单 | 301 |
| 15.1 | 简介 | 301 |
| 15.2 | 切换面板 | 302 |
| 15.2.1 | 创建切换面板 | 302 |
| 15.2.2 | 修改切换面板 | 309 |
| 15.3 | 菜单 | 311 |
| 15.3.1 | 创建菜单 | 311 |
| 15.3.2 | 调用菜单 | 317 |
| 15.4 | 本章小结 | 318 |
| | 习题 | 318 |
| 第 16 章 | 数据库安全性 | 319 |
| 16.1 | 简介 | 319 |
| 16.2 | 编码/解码数据库 | 320 |
| 16.3 | 数据库密码 | 321 |
| 16.4 | 用户级安全机制 | 323 |
| 16.4.1 | 创建工作组信息文件和管理员帐户 | 323 |
| 16.4.2 | 创建组和用户帐户 | 325 |
| 16.4.3 | 权限管理 | 326 |
| 16.4.4 | 利用向导进行用户级安全管理 | 328 |
| 16.5 | 本章小结 | 334 |
| | 习题 | 334 |
| 第 17 章 | 事务管理 | 335 |
| 17.1 | 事务 | 335 |
| 17.2 | 并发控制 | 337 |
| 17.2.1 | 并发调度异常 | 338 |
| 17.2.2 | 基于封锁的并发控制技术 | 339 |
| 17.3 | 数据库恢复技术 | 343 |

| | |
|----------------------------|------------|
| 17.3.1 故障类型 | 345 |
| 17.3.2 日志 | 345 |
| 17.3.3 系统故障恢复技术 | 347 |
| 17.3.4 介质故障恢复技术 | 348 |
| 17.4 本章小结 | 351 |
| 习题 | 353 |
| 参考文献 | 354 |
| 附录 A ER 模型的传统图形表示方法 | 356 |
| A.1 实体 | 356 |
| A.2 属性 | 356 |
| A.3 联系 | 356 |
| A.4 特殊化 | 357 |
| A.5 聚集 | 357 |

第 1 章

数据库系统简介

通过本章的学习,读者应能了解及掌握:

- ◆ 有关数据库系统的基本概念;
- ◆ 数据库技术的发展历史和应用前景;
- ◆ 数据模型的概念和分类;
- ◆ 数据库系统的体系结构。

数据库技术在过去几十年中被广泛地应用于各行各业中,为实现数据管理、完成业务处理的自动化、提高业务运营的速度起到了重要的作用。另一方面,数据库应用的广泛普及也进一步推动了数据库技术的研究和发展,使得数据库应用的分析、设计、实现等各方面都有相应的理论和方法指导,也使得当今的数据库应用已不仅仅局限于数据的存储、更新和检索等数据管理功能,而且已扩展到对数据的分析处理,从而辅助管理人员进行科学的决策等高级应用。

本章将对数据库系统的基本概念、组成、发展历史以及系统结构作简要介绍。

1.1 数据库及数据库系统

数据库(database)是逻辑上相关的数据的集合,它通常是为了实现某一目的、具有特定用户的、反映真实世界的**数据聚集**。**数据**是企事业单位生产或业务运营过程中涉及的信息的符号表示。**信息**是现实世界事物存在方式或运动状态的反映。例如,某公司的每个客户的信息包括客户的编号、名称、联系电话以及所在地区、省、市和通信地址等,这些信息可以通过文本格式的符号记录下来,转换成数据存放。公司生产和销售的每个产品的信息包括产品的编号、类型、名称、成本、价格、外观、使用方法等,其中,除了使用文本符号记录大部分信息之外,还可以使用图像(如产品的外观的图片),以及视频(如使用方法的演示录像)等各种格式来记录信息。信息通过数据的形式存储起来,不仅便于信息的存储,更重要的是便于信息的加工和传递。例如,通过将客户的信息聚集到一起并以数据的形式存放起来,销售员

就可以利用该数据与客户联系销售业务;财务人员可以利用该数据为客户提供发票,将客户数据分地区进行人数汇总统计;根据统计结果,主管销售的管理人员可以适当地调整各个地区的销售方案等。当然这些功能,如数据的存储、数据的检索以及数据的加工不能仅仅依靠将数据收集起来,聚集到一起就能完成,还需要利用一组程序对数据进行有效的管理,这组程序就是数据库管理系统(database management system,DBMS)。

利用数据库管理系统和其他开发工具创建数据库,实现对某一特定应用背景进行数据管理的功能的应用软件系统称为**数据库应用系统**。一个**数据库系统**指的是一个典型的数据库环境相关的各种成分组成的一个整体,它通常包括操作系统、数据库管理系统、数据库、设计和开发工具、应用程序以及数据库用户等组成成分。

操作系统是 DBMS 运行的基础,DBMS 经常借助计算机的操作系统完成某些功能。数据库是由与应用背景相关的需要管理的数据,以及便于对这些数据进行管理而产生的元数据(meta data)组成。**元数据**,又称为关于数据的数据,是用于描述其他数据的结构、特性和语义的数据。数据库用户又可细分为最终用户(end user)、数据库管理员(database administrator,DBA)和系统开发人员(system developer)等。最终用户指的是使用数据库应用系统完成各种日常业务操作的人员。**数据库管理员**则是维护数据库正常和高效运作的必不可少的重要人员,在数据库应用系统的生命周期参见 2.2 节的每个阶段都需要数据库管理员的参与。系统开发人员通常是指利用设计和开发工具,对系统进行分析和编程实现的系统分析员和程序员。

1.2 数据库管理系统

数据库管理系统是用于定义、创建和维护数据库中大量数据,并且能够控制对数据的访问的软件系统。概括地说,数据库管理系统应该具有以下几方面的功能。

- 提供一定的方法便于用户定义数据库的存储结构和逻辑结构。DBMS 通常提供一种数据定义语言(data definition language,DDL),用户利用它定义数据库的结构。另外,系统也可以通过提供其他方法,如可视化界面,通过可视化操作完成此功能。
- 提供一定的方法便于用户对数据库中的数据进行操纵,如插入、删除、修改和查询等。DBMS 通常提供一种数据操纵语言(data manipulation language,DML),用于描述用户的数据操纵请求。系统同样也可以提供其他方法或机制以完成此功能。
- 系统能够管理大量数据,即使系统发生故障仍能保证对数据的有效访问,并且保证数据的安全性,避免对数据的非法使用。
- 系统能够允许多用户对数据库的并发访问,避免异常结果的产生。

一个实用的、商品化的 DBMS 通常不仅具有以上几方面的功能,还具有其他功能。被称为关系系统之父的 E. F. Codd 在 1981 年获得图灵奖的演讲中,提出 DBMS 应能提供的 8 个方面功能:数据存储、检索和更新;用户可访问的系统目录;支持事务的概念;恢复;并发控制;授权;支持数据通信功能以及保证数据的完整性。

1. 数据存储、检索和更新的功能(data storage, retrieval, and update)

这是一个 DBMS 应该具备的基本功能。它指的是 DBMS 能够有效地将数据存储起来,能够以方便的方式让用户插入、删除、修改以及查询数据。

2. 提供用户可访问的系统目录(a user-accessible system catalog)

DBMS 应该将数据库中数据的描述、用户的有关信息以及应用方面的信息存放于系统目录(system catalog)或数据字典(data dictionary)中,并且系统应该允许用户访问这些数据。存放和管理的这些数据通常又称为元数据(meta-data),它们是有关数据的数据(data about data),尽管各个系统管理的元数据不尽相同,但是应该包含以下内容。

- 外模式、概念模式和内模式的定义(参见 1.5 节)。
- 数据项的名称、数据类型、精度要求和含义等。
- 数据项和数据项之间的各种约束。
- 数据库的合法用户的用户名、口令和存取权限等。
- 用户访问数据库中各种对象的统计数字,如事务的执行次数、频率等(事务的概念请参见第 17 章)。
- 数据库中数据的分布情况等。

利用这些数据,用户易于理解数据库中的数据,也便于控制数据的冗余性。利用这些数据,数据库的安全性、完整性也容易实现,通过检查用户的用户名、口令和存取权限,可以控制数据库的安全性;通过检查定义的各种约束,可以实施数据库的完整性检查(参见 5.2 节)。

3. 支持事务的概念(transaction support)

一个事务是由一组相关的操作组成的,用于完成对数据库的查询或更新任务。DBMS 必须保证事务的原子性(事务的其他特性请参见第 17 章)。原子性指的是组成一个事务的各个操作要么都做要么都不做。例如,每销售一个产品,将销售的细节信息插入数据库中,并且将该产品的库存量进行相应的修改。DBMS 必须保证无论发生什么事件,如在执行完第一个操作之后系统发生故障,该事务中的两个操作的结果要么都反映到数据库中,要么都不反映到数据库中,否则就会使数据库处于一个不一致的状态。

4. 恢复功能(recovery services)

在 DBMS 由于发生各种事件,如系统故障、介质故障以及其他软硬件故障,而中断或停止运行时,有些事务可能还没有提交,或者已经提交的事务尚没有将结果反映到数据库中,或者数据库中的数据无法访问,这时都需要 DBMS 具有将系统恢复到某个最近的一致状态的功能,这就是数据库系统应该具备的可恢复性。

5. 并发控制功能(concurrency control services)

当多个用户同时并发访问数据库中的数据时,DBMS 必须能够保证这些用户对数据库正确地进行读取和修改,即与它们顺序执行的某个结果相同。

假设甲、乙两个事务同时都要完成将同一个产品的库存量减去销售量的操作,设库存量的原值为 100,甲事务要从中减去 20,乙事务要从中减去 30,如果没有适当的并发控制机制,由于甲、乙两个事务同时读取库存量的原值 100,甲(乙)事务将其减去 20(30)后将 80(70)写回数据库,乙(甲)事务将结果 70(80)写回数据库,将 80(70)覆盖。而正确结果应为 50。因此,没有合适的并发控制无法保证事务操作的正确性。

6. 授权功能(authorization services)

DBMS 应该采用一定的机制保证数据库中的数据只被合法的用户访问,系统可以控制某些用户只能访问部分数据,从而保证数据库的安全性。例如,非销售主管不能查看每个销售员的个人信息,如年龄、工资等。

7. 支持数据通信功能(support for data communication)

DBMS 必须具有与通信软件集成的功能。由于一个数据库通常由多个用户共享,每个用户通常都是通过网络与安装 DBMS 系统的计算机相连,DBMS 系统将用户的请求作为通信的消息进行接受和回应,因此,DBMS 需要集成各种进行数据通信管理的软件。

8. 保证数据的完整性(integrity services)

数据完整性指的是数据的正确性和一致性。通常 DBMS 采用约束的方法来保证完整性。首先提供给用户一定的方法定义各种约束,如性别的取值只能为男或女,年龄的取值在 1~100 之间,订单的客户编号与客户列表中的客户编号必须一致等;其次,系统能够在适当的时候对约束进行检查,验证对数据库的各种更新操作是否遵守完整性约束的定义。

除了以上 8 项功能外,有些 DBMS,尤其是关系模型的数据库管理系统,还具有以下的功能。

9. 保证数据的独立性(services to promote data independence)

DBMS 应该能够通过一定的机制在一定程度上保证数据的逻辑独立性和物理独立性。对于 DBMS 来说,完全达到两种独立性要求是很难的。尽管如此,DBMS 通过某些机制(如视图)可以在一定程度上满足独立性的要求。例如,至少对数据库逻辑结构的部分更改操作可以避免对应用程序的更改。

10. 提供实用工具(utility services)

DBMS 应该能够提供一系列实用工具,便于 DBA 对数据库进行日常维护和调优操作。例如,数据的加载工具常用于将存放于各种类型文件中的数据导入数据库,而数据的卸载工具则完成相反的过程。性能监视工具通常用于监视对数据库的各种操作情况。统计分析工具用于分析系统的性能,统计各种数据的分布和使用情况。索引和数据库重构工具可以重构索引,重新组织数据库中的数据,减少由于各种数据库更新操作导致的数据分布的零散状况,从而改善性能。

1.3 数据库管理系统的发展

数据库管理系统于 20 世纪 60 年代提出。最早的数据库管理系统是为克服文件系统的缺陷而诞生的,之后,在应用和研究过程中,数据库技术不断发展,经历了多个发展阶段。每个阶段都不同程度地从以下 3 个方面对数据库技术进行完善和发展。

- 数据的独立性:提高应用程序和数据之间的独立性,以便降低应用程序的维护开销。
- 可处理的数据的复杂性:能够处理复杂的数据类型,以适用于更广泛的应用领域。
- 系统的易用性:使系统更容易使用,没有数据库知识背景的用户也能够轻松地使用数据库,这样可以降低数据库系统的使用成本,使更多的人可以使用数据库。

数据库技术的发展可以划分为以下几个阶段。

1. 20 世纪 60 年代

数据库管理系统提出于 20 世纪 60 年代,最早可追溯于阿波罗登月计划。当时为了响应肯尼迪总统在 20 世纪 60 年代末实现人类登上月球的愿望,北美航空公司(North American Aviation, NAA,现在的 Rockwell International 公司)开发了一个称为 GUAM (general update access method)的软件,用于管理大量有关项目的数据,该软件采用树的数