数据库系统原理与应用教程

闪四清　编著

清 华 大 学 出 版 社

需要全本请访问 www.ertongbook.com

(京) 新登字 158 号

内 容 简 介

数据库技术是信息技术的一个重要支柱。本书是一本全面介绍数据库技术基本原理、技术实现和应用等知识的教程。全书包括 4 部分、16 章。第一部分对数据库的建模方法进行了研究，第二部分全面介绍了 SQL 语言及其理论基础，第三部分涉及数据库技术深层次的实现，最后一部分详细阐述了数据库应用开发的方法和步骤。

本书内容全面、概念清晰、语言流畅、图文并茂，可作为高等院校计算机相关专业的数据库课程教材，也可供从事计算机软件工作的科技人员和工程技术人员如管理信息系统、电子商务网站开发人员等参阅。

版权所有，翻印必究。
本书封面贴有清华大学出版社激光防伪标签，无标签者不得销售。

书　　名：数据库系统原理与应用教程
作　　者：闪四清　编著
出 版 者：清华大学出版社(北京清华大学学研大厦，邮政编码: 100084)
http://www.tup.tsinghua.edu.cn
责任编辑: **宋方**
印 刷 者:
发 行 者: **新华书店总店北京发行所**
开　　本: 787×1092　1/16　**印张**: 23　**字数**: 545 **千字**
版　　次: 2001 **年** 4 **月第** 1 **版**　2001 **年** 4 **月第** 1 **次印刷**
书　　号: ISBN 7–302–04340–X／TP・2548
印　　数: 0001～6000
定　　价: 35.00 **元**

前　言

当今是一个信息爆炸的时代，信息已经成为社会和经济发展的重要支柱之一。大量信息的产生、处理、存储、传播和使用推动了社会的进步和经济的发展。信息系统是一种以加工处理信息为主的计算机系统。数据库技术作为一种存储和使用信息的信息系统核心技术正在发挥着越来越重要的作用。例如，现在的银行无法再依赖手工记账来处理复杂的储蓄、结算、信贷、信用卡等业务，现在的邮政电信也无法离开信息系统开展各种电信业务，现在的航空运输也无法离开信息系统来管理庞大的分布式售票网点，现在的加工企业无法离开信息系统来管理庞大的零部件仓库、计划、生产、销售、财务等，现在的图书出版业无法依赖手工的方式管理大量图书的出版、库存、销售等业务。

本书对数据库技术进行了全面的阐述和研究。在结合了大量的实例和作者的教学体会基础上，对数据库技术的各个领域进行了深入浅出的剖析，对数据库技术的重点和难点进行了详细的描述，力求做到概念清晰、生动活泼。

本书包括了 4 部分、16 章内容。第一部分是建模篇，重点介绍了 ER 图、ODL 方法、IDEF1x 图、关系模型等数据库建模技术。第二部分是编程篇，对数据库技术中的关系代数、Datalog 语言等理论以及标准 SQL 语言进行了讲述。第三部分是实现篇，讨论了有关索引、查询、事务、并发控制等数据库技术内核的实现问题。最后一部分对典型的桌面型数据库产品、分布式数据库产品以及信息系统的开发方法进行了介绍和分析。

在每一章的开始，对该章将要涉及的内容和作用进行了简单的介绍，然后指出了学完本章应该掌握的重要内容。每一章的正文内容结束之后，有一个对本章所讲述的内容进行总结和评价的本章小结。另外，每一章后面都附有一个思考和练习。思考题侧重于加深读者对本章涉及概念的理解，练习题侧重于培养学生应用本章学到的知识来解决实际问题的能力。

本教程还有一个显著的特点，就是以一个贯穿全书始末的图书管理示例为主线，把各种数据库技术的知识要点，串联成一个逻辑严密的整体。这个示例不断深化、丰富和完善的过程，实际上就是读者不断学习、理解和掌握数据库技术的过程。

本书由康博金典科技有限公司策划，闪四清博士主编。限于作者水平有限，书中难免有不妥之处，欢迎广大读者批评指正。

作　者

2001 年 1 月

目　　录

第一部分　设计篇

第 1 章　步入数据库系统世界 …… 3

1.1　数据库技术概述 …… 3

1.1.1　数据库系统应用示例 …… 4

1.1.2　数据库系统的作用 …… 7

1.1.3　常见的数据库系统 …… 8

1.1.4　数据库应用程序的开发过程 …… 11

1.2　数据库技术的演变 …… 12

1.2.1　数据库技术的环境 …… 12

1.2.2　传统的文件管理阶段 …… 13

1.2.3　现代的数据库管理系统阶段 …… 14

1.3　数据库系统架构 …… 16

1.3.1　数据存储架构 …… 16

1.3.2　数据视图管理架构 …… 18

1.3.3　数据库应用架构 …… 19

1.4　数据库管理系统组件 …… 22

1.4.1　数据和元数据 …… 23

1.4.2　输入 …… 24

1.4.3　查询处理器 …… 25

1.4.4　存储管理器 …… 25

1.4.5　事务管理器 …… 26

1.5　数据库技术的发展趋势 …… 26

1.5.1　数据库技术面临的挑战 …… 26

1.5.2　数据库技术的未来 …… 27

1.6　本章小结 …… 30

1.7　思考和练习 …… 31

第 2 章　数据库建模——ER 图 …… 32

2.1　ER 图的基本概念 …… 32

2.1.1 实体集和实体集中的属性 32
2.1.2 联系集和联系集中的属性 36
2.1.3 绘制 ER 图 38
2.2 ER 图设计原则 40
2.2.1 真实性原则 41
2.2.2 避免冗余原则 41
2.2.3 简单性原则 42
2.2.4 合适性原则：用实体集还是用属性 42
2.2.5 合适性原则：用实体集还是用联系集 42
2.3 对约束建模 43
2.3.1 映射约束 43
2.3.2 键码约束 46
2.3.3 参照完整性约束 48
2.3.4 其他类型的约束 49
2.4 弱实体集 49
2.4.1 弱实体集产生的原因 50
2.4.2 对弱实体集的要求 51
2.4.3 弱实体集的表示方法 51
2.5 ER 图的高级技术 52
2.5.1 特殊化技术 52
2.5.2 概括技术 53
2.5.3 属性继承技术 54
2.5.4 概括约束设计 54
2.5.5 聚集技术 55
2.6 本章小结 57
2.7 思考和练习 57

第 3 章 数据库建模——IDEF1x 图 59

3.1 IDEF 方法族基本知识 59
3.1.1 IDEF 方法族简介 59
3.1.2 IDEF1x 图的概念和作用 61
3.2 IDEF1x 图的语法和语义 62
3.2.1 实体 62
3.2.2 连接联系 63
3.2.3 分类联系 67
3.2.4 非确定联系 69

需要全本请访问 www.ertongbook.com

3.2.5 属性 …… 70
3.2.6 主键码和次键码 …… 72
3.2.7 外键码 …… 74
3.3 使用 IDEF1x 图建立数据库模型的过程 …… 76
3.3.1 准备阶段 …… 76
3.3.2 定义实体阶段 …… 79
3.3.3 定义联系阶段 …… 80
3.3.4 定义键码阶段 …… 82
3.3.5 定义属性阶段 …… 86
3.3.6 文件编制阶段 …… 88
3.4 本章小结 …… 89
3.5 思考和练习 …… 89

第 4 章　数据库建模——ODL 方法 …… 91

4.1 面向对象的基本概念 …… 91
4.1.1 面向对象模型和关系模型比较 …… 91
4.1.2 面向对象基本概念 …… 92
4.2 ODL 的基本结构 …… 94
4.2.1 ODL 概念 …… 94
4.2.2 ODL 接口说明 …… 95
4.2.3 ODL 中的属性 …… 96
4.2.4 ODL 中的联系 …… 97
4.2.5 反向联系的概念 …… 98
4.2.6 联系的多重性 …… 100
4.2.7 ODL 中的类型 …… 102
4.2.8 ODL 中的方法署名的说明 …… 105
4.3 子类和约束 …… 107
4.3.1 ODL 中的子类 …… 107
4.3.2 ODL 中的约束 …… 111
4.4 本章小结 …… 116
4.5 思考和练习 …… 116

第 5 章　关系模型 …… 118

5.1 关系模型的基本概念 …… 118
5.2 从 ER 图到关系模型设计的转换 …… 123
5.2.1 ER 图和关系模式的比较 …… 124

5.2.2 实体集到关系模式的转换……124
5.2.3 联系集到关系模式的转换……125
5.2.4 弱实体集到关系模式的转换……127
5.2.5 ER 图中其他元素的转换……128
5.2.6 键码的转换……130
5.3 从 IDEF1x 图到关系模型设计的转换……131
5.4 从 ODL 设计到关系模型设计的转换……132
5.4.1 ODL 设计和关系模式设计的比较……132
5.4.2 从 ODL 属性到关系模式的属性……133
5.4.3 ODL 设计中非原子属性的转换……133
5.4.4 ODL 设计中联系的转换……137
5.4.5 ODL 设计中联系和反向联系问题……138
5.4.6 ODL 设计中子类结构的转换……138
5.5 本章小结……140
5.6 思考和练习……140
第 6 章 关系模式的规范化设计……142
6.1 关系模式的设计问题……142
6.1.1 异常问题……142
6.1.2 泛关系模式和数据库模式……143
6.2 函数依赖……144
6.2.1 函数依赖的定义……144
6.2.2 函数依赖和键码……145
6.2.3 逻辑蕴含……146
6.2.4 函数依赖的推理规则……147
6.2.5 属性集的闭包……148
6.2.6 正则覆盖……150
6.3 关系模式的分解……151
6.4 关系模式的范式……152
6.4.1 第一范式……152
6.4.2 BC 范式……153
6.4.3 分解成 BCNF 模式的算法……153
6.4.4 函数依赖的投影……155
6.4.5 第三范式……156
6.4.6 第二范式……158
6.5 多值依赖……158

6.5.1 多值依赖的概念……159
6.5.2 多值依赖的推理规则……160
6.5.3 第四范式……160
6.6 范式之间的关系……161
6.7 数据库模式的实例……162
6.7.1 图书管理数据库模式……162
6.7.2 计算机产品信息管理数据库模式……164
6.8 本章小结……165
6.9 思考和练习……166

第二部分　编程篇

第 7 章　关系代数基本理论……171

7.1 关系代数概述……171
7.2 关系代数的基本运算……172
7.2.1 集合运算……172
7.2.2 选择运算……174
7.2.3 投影运算……176
7.2.4 笛卡尔积运算……177
7.2.5 改名运算……178
7.3 关系代数的附加运算……179
7.3.1 自然连接运算……179
7.3.2 θ 连接运算……180
7.3.3 除法运算……181
7.3.4 赋值运算……182
7.4 扩充的关系代数运算……183
7.4.1 广义投影运算……183
7.4.2 外连接运算……183
7.4.3 聚集运算……185
7.5 关系演算……185
7.5.1 元组关系演算……186
7.5.2 域关系演算……186
7.6 关系代数的修改运算……187
7.6.1 插入运算……187
7.6.2 删除运算……188

7.6.3 修改运算……188
7.7 关系代数表达式的优化策略……189
7.8 本章小结……190
7.9 思考和练习……190

第 8 章 Datalog 语言……192

8.1 基本概念……192
8.1.1 基本结构……192
8.1.2 一般规则……194
8.1.3 安全规则……195
8.1.4 外延谓词和内涵谓词……197
8.2 关系代数向 Datalog 规则的转换……198
8.2.1 从集合运算到 Datalog 规则……198
8.2.2 从投影运算到 Datalog 规则……199
8.2.3 从笛卡尔乘积到 Datalog 规则……199
8.2.4 从选择运算到 Datalog 规则……200
8.2.5 从连接运算到 Datalog 规则……202
8.2.6 从多重运算到 Datalog 规则……202
8.3 递归原理……203
8.3.1 关系代数存在的问题……203
8.3.2 计算最小固定点……205
8.3.3 使用 Datalog 规则表示固定点公式……206
8.4 包的运算……207
8.4.1 包的意义……207
8.4.2 包的关系运算……208
8.4.3 包的逻辑运算……211
8.5 本章小结……211
8.6 思考和练习……212

第 9 章 SQL 语言初步……214

9.1 SQL 概述……214
9.2 简单查询语句……215
9.2.1 最简单的查询语句……216
9.2.2 投影……217
9.2.3 选择……219
9.2.4 字符串比较……222

9.2.5 日期和时间的比较……223
9.2.6 排序结果……223
9.2.7 副本……225
9.3 高级查询语句……226
9.3.1 连接查询……226
9.3.2 集合查询……232
9.3.3 聚集查询……233
9.3.4 子查询……236
9.4 数据库更新语句……238
9.4.1 插入数据……238
9.4.2 删除数据……239
9.4.3 修改数据……240
9.5 数据定义语句……241
9.5.1 数据类型和域……241
9.5.2 管理表……242
9.5.3 默认值……244
9.5.4 索引定义……245
9.5.5 管理视图……245
9.6 递归查询……248
9.7 本章小结……249
9.8 思考和练习……249

第 10 章 SQL 语言高级功能……251

10.1 约束……251
10.1.1 键码约束……251
10.1.2 外键码约束……253
10.1.3 非空约束……254
10.1.4 CHECK 约束……255
10.1.5 断言约束……256
10.2 触发器……258
10.2.1 触发器的结构……258
10.2.2 触发器示例……260
10.3 游标……261
10.3.1 定义游标……261
10.3.2 游标的选项……262
10.4 嵌入式 SQL 语言……263

10.4.1 接口和声明……263
10.4.2 动态的 SQL 语句……264
10.5 安全控制和权限管理……265
10.5.1 权限的类型……265
10.5.2 授权和收权……266
10.6 本章小结……267
10.7 思考和练习……267

第三部分　实现篇

第 11 章　索引和散列技术……271

11.1 文件组织……271
11.1.1 文件组织方式……271
11.1.2 顺序文件组织……272
11.1.3 聚集文件组织……273
11.2 索引技术……274
11.2.1 基本概念……274
11.2.2 顺序索引……274
11.2.3 B 树索引文件……278
11.3 散列技术……278
11.3.1 基本概念……279
11.3.2 散列索引……279
11.4 Microsoft SQL Server 系统中的索引……280
11.4.1 索引类型……280
11.4.2 创建索引的方法……282
11.5 本章小结……283
11.6 思考和练习……283

第 12 章　查询处理技术……284

12.1 基本步骤……284
12.2 查询处理的代价模型……286
12.3 单个关系运算的代价估计……287
12.3.1 选择运算的代价估计……287
12.3.2 连接运算的代价估计……288
12.3.3 集合运算的代价估计……289

需要全本请访问 www.ertongbook.com

12.4　表达式运算的代价估计……289
12.4.1　实体化方法……289
12.4.2　流水线方法……290
12.5　Microsoft SQL Server 系统的查询处理器……291
12.5.1　特性概述……291
12.5.2　交互式操作……292
12.5.3　执行计划……293
12.5.4　优化索引……294
12.5.5　优化隐藏……295
12.6　本章小结……296
12.7　思考和练习……297

第 13 章　事务和并发控制……298

13.1　事务的概念……298
13.1.1　事务的基本概念和特性……298
13.1.2　事务的状态变迁图……299
13.1.3　事务的并发执行……300
13.1.4　事务的一致性级别……302
13.2　并发控制……303
13.2.1　并发操作带来的问题……303
13.2.2　锁……303
13.2.3　死锁……304
13.3　Microsoft SQL Server 系统的事务和锁……305
13.3.1　事务的类型……305
13.3.2　事务的工作原理……306
13.3.3　锁的粒度……308
13.3.4　锁的类型……311
13.3.5　死锁和死锁的解决方案……313
13.4　本章小结……314
13.5　思考和练习……314

第四部分　应用篇

第 14 章　Microsoft Access……317

14.1　Microsoft Access 2000 的基本特性……317

14.2 Microsoft Access 2000 的基本操作……319
14.2.1 创建数据库……319
14.2.2 在数据库中创建 Book 表……321
14.2.3 在数据库中创建查询……322
14.2.4 创建窗体……324
14.3 本章小结……326
14.4 思考和练习……326
第 15 章 Microsoft SQL Server……327
15.1 Microsoft SQL Server 2000 的特点……327
15.2 Microsoft SQL Server 2000 系统的基本工具……329
15.2.1 SQL Server Enterprise Manager……329
15.2.2 SQL Query Analyzer……332
15.2.3 其他重要的工具……334
15.3 本章小结……337
15.4 思考和练习……337
第 16 章 信息系统的开发过程……338
16.1 信息系统概述……338
16.2 开发信息系统的方法论……340
16.2.1 方法类型……340
16.2.2 开发策略……341
16.3 信息系统开发阶段……342
16.3.1 调查研究阶段……342
16.3.2 系统分析阶段……344
16.3.3 系统设计阶段……346
16.3.4 系统实施阶段……349
16.3.5 系统评价阶段……349
16.4 影响信息系统开发成功和失败的因素……350
16.5 本章小结……351
16.6 思考和练习……352
参考书目……353

第一部分　设计篇

- 步入数据库系统世界
- 数据库建模——ER 图
- 数据库建模——IDEF1x 图
- 数据库建模——ODL 方法
- 关系模型
- 关系模式的规范化设计

第1章　步入数据库系统世界

本书将要全面讲述数据库技术的原理和应用。通过本书的学习，读者不仅可以掌握数据库技术的基本原理，而且可以掌握使用具体的数据库产品技术和从事数据库应用开发技术。对于一个数据库技术专业人员来说，仅掌握“怎样做”的技术应用是不够的，更重要的是需要理解和掌握“为什么这样做”的技术理论。

本章作为开场白，将要把读者带入到数据库技术的世界，为他们提供一个有关数据库技术的清晰的轮廓。首先，通过介绍一些典型的数据库应用，使读者了解数据库技术到底可以起到什么作用。为了加深读者对数据库的理解，本章快速浏览了当前市场上的主流数据库产品，讨论了数据库应用开发的过程。回顾历史是为了更好地珍惜现在和展望未来。第二，谈了数据库的发展简史，这些内容有助于读者了解数据库技术的当前状况、存在问题和发展趋势。第三，对数据库技术的结构进行深入的研究，这些内容包括数据存储架构、数据管理架构和应用开发架构。第四，对数据库管理系统的架构进行了剖析，把数据库技术的内幕展示出来。最后，总结数据库技术面临的挑战和机遇，展望未来数据库技术的发展趋势。

本章的学习目标：

- 了解数据库技术的应用现状和作用；
- 了解当前数据库市场上的主流产品；
- 了解数据库应用开发的基本过程；
- 掌握数据库技术的发展历史和各阶段的特征；
- 理解数据库中数据的存储、管理和应用架构；
- 理解数据库管理系统的架构和各个组成部分的作用；
- 了解当前数据库面临的挑战；
- 掌握数据库技术的未来发展趋势。

1.1　数据库技术概述

信息技术的发展极大地推动了社会的进步和发展。作为信息技术主要支柱之一的数据库技术在社会各个领域中有着广泛的应用。数据库技术可以为各种用户提供及时的、准确的、相关的信息，满足这些用户的各种不同的需要。

1.1.1 数据库系统应用示例

现在，通过研究几个数据库系统应用的典型示例，使读者对这门信息技术的支撑技术有一个感性的认识。

1. 航空售票系统

该系统可能是最早使用数据库技术的应用领域。在这种系统中，包括这些数据项：

(1) 座位预定信息：座位分配、座位确认、餐饮选择等；

(2) 航班信息：航班号、飞机型号、机组号、起飞地、目的地、起飞时间、到达时间、飞行状态等；

(3) 机票信息：票价、折扣、有无等。

对这种系统的查询主要是在某一段时间内从某个指定的城市到另一个指定的城市的航班，是否还有可以选择的座位、是否有其他飞机型号、飞机票价、是否折扣等信息。对该系统的主要更新操作包括为乘客登记航班、分配座位、选择餐饮等。任何时候，都会有许多航空售票代理商访问这些数据，并且要避免出现多个代理商同时卖出同一个座位的情况。这些数据还可以自动统计出经常乘坐某一个航班的乘客的信息，为这些常旅客提供特殊的优惠服务。

如果没有使用数据库技术，那么就会因为数据量庞大和更新缓慢，航空部门无法提供及时、准确、有竞争力的服务。

2. 银行业务系统

银行业务的繁忙状态是社会经济发展的标志。银行业务系统也是最早使用信息技术或数据库技术的系统之一。在银行业务系统中，典型的数据项包括：

(1) 顾客信息：姓名、身份证号码、地址、电话等信息；

(2) 账户信息：账号、存款、余额、取款、日期等信息；

(3) 顾客和账户关系信息：顾客、账户等信息。

对银行业务系统的操作既可以通过各地的银行营业网点，也可以通过安装在各地的ATM(Automated Teller Machine，自动出纳机)进行。该系统的主要查询操作包括询问顾客的账户、账户的余额以及更新账户的数据等。就像航空售票系统一样，允许对同一个账户进行并发访问而不会出现任何错误，这是非常重要的。即使系统发生了故障，例如 ATM 突然断电了，正在处理的账户数据也不会出现任何不一致的记录。当前的数据库技术已经完全可以解决这种表面简单而实质复杂的难题。

使用了信息技术和数据库技术的银行业务系统给人们的生活和工作带来了巨大的便利。例如，人们可以在任何银行网点存款和取款，避免随身携带大量的货币，保证安全。可以快速地异地汇兑和结算，减少资金的在途时间，提高企业的信誉。

3. 超市的销售业务系统

现在，超市购物由于其种类繁多的商品、较低的价格、选物的便利已经成为人们日常生活的一个重要组成部分。数据库技术是超市取得成功的重要技术基础。在超市的销售业务系统中，主要的数据项为：

(1) 销售信息：连锁点、日期、时间、顾客、商品、数量、总价等；

(2) 商品信息：商品名称、单价、进货数量、供应商、商品类型、摆放位置等；

(3) 供应商信息：供应商名称、地点、商品、信誉等。

对超市销售业务系统的主要操作是记录顾客的购买信息，查询超市现有商品的结构，分析当天连锁店的销售情况，确定明天进货的内容和货物的摆放位置。超市的经营决策主要是依赖销售业务系统中存储的大量数据，从这些表面似乎孤立的大量数据中发掘出真正有价值的销售规律，提高经营者的决策水平。

我们经历过了许多艰难的岁月，如商品短缺、服务态度恶劣、短斤少两、假冒伪劣等一度影响和扰乱了我们正常的生活秩序。信息技术，特别是数据库技术，推动了社会的前进，提高了人们的生活水平。

4. 工厂的管理信息系统

工厂的管理信息系统(Management Information System，MIS)是最早依据数据库技术建立的一个比较完整的集成系统。在这种 MIS 系统中，主要包括了下面一些数据项：

(1) 销售记录：产品、服务、客户、销售人员、时间等；

(2) 雇员信息：姓名、地址、工资、津贴、所得税款等；

(3) 财务信息：合同、应收货款、应付货款等。

在这种系统中，典型的查询操作包括打印雇员的工资、打印应收应付货款清单、打印销售人员的业绩、打印工厂的各种统计报表等。每进行一次采购和销售，收到每一个账单、收据，雇员的聘用、解聘、提职、加薪等都将导致对数据库的更新。

一个典型的 MIS 系统的窗口如图 1-1 所示。这是一个典型的工厂 MIS 系统的业务导航窗口，包括了进货、销售、仓库、账务、人事、系统维护等功能。使用这个导航系统，可以执行各种业务操作。在这种 MIS 系统的背后，是存储了大量业务数据的数据库管理系统。

如果没有 MIS 系统，那么许多企业就会陷入这种混乱的状态：货款迟迟没有到位却没有人及时发现，财务报表不能及时提供，领导不知道库存的零件、产品有多少，作业计划的安排不符合实际情况，等等。

5. 学校教学管理系统

学校教学管理系统主要是涉及学生、教师、教室、课程、排课等信息的管理。该系统包括的典型数据项如下：

(1) 学生信息：姓名、学号、性别、班级、年龄、宿舍、电话、E-mail 地址等；

需要全本请访问 www.ertongbook.com