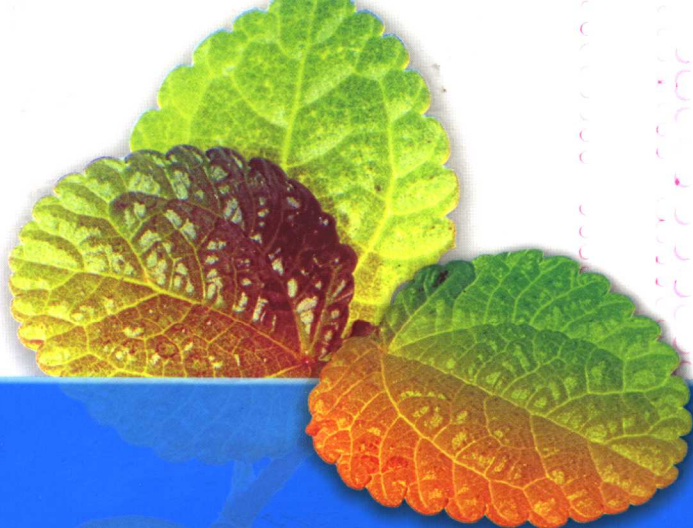


清华 电脑学堂



超值多媒体光盘
多媒体语音视频教程
数据库实例和代码



Oracle 10g

数据库管理 应用与开发 标准教程

马晓玉 孙岩 孙江玮 李洪海 等编著

- 总结了作者多年 Oracle 数据库开发和教学心得
- 系统讲解了 Oracle 10g 数据库的要点和难点
- 提供丰富的实验指导和习题
- 范例典型实用，图文并茂
- 配套光盘提供语音视频教程和数据库实例

清华大学出版社



清华 电脑学堂

Oracle 10g

数据库管理 应用与开发

标准教程

■ 马晓玉 孙岩 孙江玮 李洪海 等编著

清华大学出版社
北 京

内 容 简 介

Oracle 数据库作为世界范围内性能最优异的数据库系统之一, 其在数据库市场的占有率始终处于数据库领域的领先地位。本书以 Oracle 10g for Windows XP 为平台, 由浅入深地介绍了 Oracle 10g 系统的使用方法和基本管理。主要内容包括: Oracle 关系数据库, Oracle 数据库体系结构, SQL 基本查询, 修改 SQL 数据与 SQL*Plus 命令, PL/SQL 编程基础, 用户、模式和表, 高级查询, 过程、函数和程序包, 表类型, 索引, 视图、序列和同义词, 触发器, 事务与并发控制, 安全, 管理存储结构和基本的备份与恢复等知识。

本书全面介绍使用 Oracle 数据库管理应用与开发知识, 适合作为普通高校计算机专业 Oracle 教材, 也可以作为 Oracle 数据库开发和应用人员的参考资料。

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

Oracle 10g 数据库管理、应用与开发标准教程 / 马晓玉等编著. —北京: 清华大学出版社, 2007.11

ISBN 978-7-302-16184-4

I. O… II. 马… III. 关系数据库—数据库管理系统, Oracle 10g—教材 IV. TP311.138

中国版本图书馆 CIP 数据核字 (2007) 第 148526 号

责任编辑: 冯志强

责任校对: 张 剑

责任印制: 李红英

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 刷 者: 北京市清华园胶印厂

装 订 者: 三河市漂源装订厂

经 销: 全国新华书店

开 本: 185×260 印 张: 27.25 字 数: 642 千字

附光盘 1 张

版 次: 2007 年 11 月第 1 版

印 次: 2007 年 11 月第 1 次印刷

印 数: 1~4000

定 价: 39.80 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: 013999-01

Oracle 数据库是世界范围内性能最优异的数据库系统之一，它在数据库市场的占有率远远超过其对手，始终处于数据库领域的领先地位。2003 年，Oracle 公司又推出了代表数据库领域最新技术的网格数据库系统——Oracle 10g，其中的 g 代表网格。这标志着 Oracle 数据库完成了从互联网“i”到网格“g”的演进。

本书以 Oracle 10g for Windows XP 为平台，由浅入深地介绍了 Oracle 10g 系统的使用方法和基本管理技术。主要包括：Oracle 关系数据库、Oracle 数据库体系结构、SQL 语言基础、PL/SQL 程序设计、模式对象管理、存储管理和基本的备份与恢复等知识。

本书内容

第 1 章 Oracle 关系数据库。介绍关系数据库的系统设计，数据库系统的应用系统结构，Oracle 10g 数据库安装注意事项，使用 DBCA 创建数据库，以及 Oracle 数据库服务的启动和可用的数据库用户等知识。

第 2 章 Oracle 数据库体系结构。介绍 Oracle 10g 强大的体系结构，包括数据库的物理存储结构、逻辑存储结构、内存结构和实例的进程结构，以及数据字典。

第 3 章 SQL 基本查询。介绍使用 SQL SELECT 语句对数据库进行各种查询，包括 SELECT 语句及其各种子句的使用，以及 SQL*Plus 中函数的使用。

第 4 章 修改 SQL 数据与 SQL*Plus 命令。介绍对存储的数据进行修改和 SQL*Plus 命令的应用知识，包括添加数据、修改数据、删除数据，以及 SQL*Plus 命令的使用。

第 5 章 用户、模式和表。介绍用户与模式的概念和模式对象——表的创建，主要包括模式、模式对象、用户的基本创建、基本表的创建、基本表结构的修改、基于表的约束等。

第 6 章 高级查询。介绍多表查询和子查询，主要包括使用逗号连接的多表查询、JOIN 连接的多表查询、集合操作和子查询的使用。

第 7 章 PL/SQL 编程基础。介绍 PL/SQL 程序的编程知识，包括 PL/SQL 程序结构、变量和常量的使用、特殊类型变量、复合变量、条件判断语句、循环语句、游标和异常处理。

第 8 章 过程、函数和程序包。介绍 PL/SQL 的存储过程和函数，以及程序包，包括存储过程和函数的创建与调用，程序包主体和规范，以及过程、函数和程序包的依赖性。

第 9 章 表类型。介绍 Oracle 支持的表类型，包括外部表的使用、索引组织表、临时表、对象和对象表、分区表和簇表的创建与使用。

第 10 章 索引。介绍 Oracle 的另一个重要模式对象——索引，主要包括索引的作用、类型和工作原理，各种索引的创建，以及对索引的管理。

第 11 章 视图、序列和同义词。介绍 Oracle 其他模式对象——视图、序列和同义词，包括视图的创建、视图的使用、可更新视图，以及序列和同义词的创建与使用。

第 12 章 触发器。介绍 Oracle 中触发器的使用，包括触发器的组成、Oracle 支持的触发器类型，以及各种触发器的修改。

第 13 章 事务与并发控制。介绍 Oracle 中事务机制和多事务访问控制，包括事务的概念、对事务 ACID 特性的支持、事务控制语句，以及通过各种锁对多个事务并发访问的控制机制。

第 14 章 安全。介绍 Oracle 的用户、权限和角色管理，包括用户配置文件的创建和应用、对用户账号的管理、系统权限、对象权限的授予与撤销，以及角色的概念和授予、撤销。

第 15 章 管理存储结构。介绍表空间的创建，以及对段、盘区、数据块的管理，主要包括创建本地化管理表空间、调整表空间属性、调整与表空间相关的数据文件、临时表空间、大文件表空间、非标准表空间，以及自动撤销表空间管理。

第 16 章 备份与恢复。介绍为避免数据库出现故障而丢失数据，而进行的备份与恢复，包括数据库的物理备份与恢复、逻辑备份与恢复和 RMAN 管理的备份与恢复。

本书特色

本书是一本完整介绍 Oracle 10g 数据库开发与应用知识的教程，在编写过程中，我们精心设计了丰富的示例，帮助读者顺利学习本书内容。

- 理论结合实践：全书提供了丰富的分析示例，通过示例分析、设计过程，讲解 Oracle 的应用知识。
- 网站互动：我们在网站上提供了本书示例和扩展内容的资料链接，便于学生继续学习相关知识；授课教师也可以下载本书教学课件和其他教学资源。
- 思考与练习：选择题测试读者对本章所介绍内容的掌握程度；简答题理论结合实际，引导学生深入掌握 Oracle 理论知识。

读者对象

本书全面介绍使用 Oracle 数据库应用与开发知识，适合作为普通高校计算机专业 Oracle 教材，也可以作为 Oracle 数据库开发和应用人员的参考资料。

本书作者长期从事应用 Oracle 数据库开发实践和教学培训工作。参与本书编写人员除了封面署名人员之外，还有吴俊海、张瑞萍、董志鹏、祝红涛、郝相林、刘万军、杨宁宁、郭晓俊、康显丽、辛爱军、牛小平、贾栓稳、王立新、苏静、赵元庆、王蕾、亢凤林、韦潜、郝安林等人。

由于时间仓促，书中疏漏之处在所难免，敬请读者批评指正。读者可以通过清华大学出版社网站 www.tup.tsinghua.edu.cn 与我们联系。

编 者

2007 年 8 月

第 1 章 Oracle 关系数据库..... 1	
1.1 关系数据模型..... 1	
1.1.1 数据结构..... 1	
1.1.2 关系操作..... 3	
1.1.3 关系的完整性约束..... 3	
1.2 关系与关系模式..... 4	
1.3 数据库的设计..... 4	
1.3.1 实体-关系模型..... 5	
1.3.2 E-R 图的绘制..... 6	
1.3.3 将 E-R 模型转化为 关系模式..... 7	
1.4 关系数据库规范化理论..... 7	
1.4.1 函数依赖..... 8	
1.4.2 范式理论..... 9	
1.5 Oracle 数据库的应用系统结构..... 12	
1.5.1 Oracle 分布式数据库 系统结构..... 12	
1.5.2 Oracle 客户/服务器 系统结构..... 13	
1.5.3 Oracle 浏览器/服务器 系统结构..... 14	
1.6 Oracle 10g 安装..... 14	
1.6.1 安装 Oracle 10g 系统需求..... 14	
1.6.2 在 Windows 环境下的 安装过程..... 15	
1.6.3 查看 Oracle 系统..... 20	
1.7 创建数据库..... 20	
1.8 数据库的启动与关闭..... 25	
1.9 Oracle 默认用户..... 26	
2.1.2 日志文件..... 29	
2.1.3 控制文件..... 31	
2.1.4 参数文件..... 31	
2.2 逻辑存储结构..... 32	
2.2.1 数据块..... 32	
2.2.2 盘区..... 33	
2.2.3 段..... 33	
2.2.4 表空间..... 34	
2.3 内存结构..... 34	
2.3.1 系统全局区..... 35	
2.3.2 程序全局区..... 37	
2.3.3 排序区..... 37	
2.3.4 大池..... 37	
2.3.5 Java 池..... 37	
2.4 实例的进程结构..... 37	
2.4.1 DBWR..... 39	
2.4.2 LGWR..... 40	
2.4.3 ARCH..... 40	
2.4.4 CKPT..... 41	
2.4.5 SMON..... 42	
2.4.6 PMON..... 42	
2.4.7 RECO..... 42	
2.4.8 Dnnn..... 42	
2.5 数据字典..... 44	
2.6 思考与练习..... 45	
第 2 章 Oracle 数据库体系结构..... 27	
2.1 物理存储结构..... 27	
2.1.1 数据文件..... 27	
第 3 章 SQL 基本查询..... 47	
3.1 SQL 语句概述..... 47	
3.2 SQL*Plus..... 48	
3.3 本书所用的示例模式..... 49	
3.3.1 VENTION 模式..... 49	
3.3.2 SCHOOL 模式..... 51	
3.4 SQL SELECT 语句..... 51	

3.4.1 SELECT 子句和 FROM 子句	52	5.1.3 用户	97
3.4.2 WHERE 子句	54	5.2 表	99
3.4.3 ORDER BY 子句	58	5.2.1 数据类型	99
3.4.4 使用统计函数	60	5.2.2 创建表	103
3.4.5 GROUP BY 子句	62	5.2.3 表特性	106
3.4.6 HAVING 子句	64	5.2.4 默认值和 NULL 值	111
3.5 在 SQL*Plus 中使用函数	65	5.3 修改表	112
3.5.1 字符串函数	65	5.3.1 增加和删除字段	112
3.5.2 数字函数	67	5.3.2 调整表特性	113
3.5.3 日期时间函数	68	5.3.3 手动给表分配存储空间	114
3.5.4 转换函数	69	5.4 删除表定义	115
3.6 实验指导	70	5.5 数据完整性	116
3.7 思考与练习	71	5.5.1 定义 PRIMARY KEY 约束	117
第 4 章 修改 SQL 数据与 SQL*Plus 命令	73	5.5.2 定义 NOT NULL 约束	118
4.1 添加数据	73	5.5.3 定义 UNIQUE 约束	119
4.1.1 简单的 INSERT 语句	74	5.5.4 定义 FOREIGN KEY 约束	120
4.1.2 省略列表清单	75	5.5.5 定义 CHECK 约束	123
4.1.3 INSERT 语句与 SELECT 语句	76	5.5.6 约束的状态	124
4.2 更新数据	77	5.6 查看表信息	125
4.3 删除数据	79	5.6.1 查看表的基本信息	125
4.3.1 DELETE 语句	79	5.6.2 分析表的存储空间	125
4.3.2 TRUNCATE 语句	80	5.7 实验指导	128
4.4 SQL*Plus 命令	81	5.8 思考与练习	131
4.4.1 设置环境变量命令	81	第 6 章 高级查询	133
4.4.2 格式化查询结果命令	84	6.1 简单连接	133
4.4.3 文件操作命令	87	6.1.1 基本连接	133
4.4.4 交互命令	90	6.1.2 表别名	136
4.4.5 帮助命令	92	6.1.3 多个表之间的连接	136
4.4.6 其他的 SQL*Plus 命令	92	6.2 使用 JOIN 连接查询	137
4.5 实验指导	93	6.2.1 内连接	138
4.6 思考与练习	94	6.2.2 外连接	139
第 5 章 用户、模式和表	96	6.2.3 交叉连接	141
5.1 用户和模式	96	6.3 集合操作	141
5.1.1 模式	96	6.3.1 UNION	142
5.1.2 模式对象与非模式对象	97	6.3.2 INTERSECT	143
		6.3.3 MINUS	143

6.4 子查询.....	144	7.9.4 异常传播.....	190
6.4.1 使用返回多行的子查询.....	144	7.9.5 在 PL/SQL 中使用 SQLCODE 和 SQLERRM.....	193
6.4.2 使用单值子查询.....	147	7.10 实验指导.....	194
6.4.3 使用嵌套子查询.....	148	7.11 思考与练习.....	197
6.5 实验指导.....	149		
6.6 思考与练习.....	151		
第 7 章 PL/SQL 编程基础.....	154	第 8 章 过程、函数和程序包.....	199
7.1 PL/SQL 程序结构.....	154	8.1 存储过程.....	199
7.2 变量与常量.....	156	8.1.1 创建存储过程.....	199
7.2.1 PL/SQL 标识符.....	156	8.1.2 调用存储过程.....	200
7.2.2 标量变量.....	157	8.1.3 修改存储过程.....	201
7.2.3 为变量和常量赋值.....	158	8.1.4 参数.....	201
7.2.4 作用域.....	158	8.1.5 局部变量和子过程.....	205
7.3 PL/SQL 块中的 SQL 语句.....	159	8.2 函数.....	206
7.4 使用%TYPE 和%ROWTYPE 类型的变量.....	160	8.2.1 创建函数.....	206
7.4.1 %TYPE 变量.....	161	8.2.2 调用函数.....	207
7.4.2 %ROWTYPE 变量.....	161	8.3 程序包.....	208
7.5 复合变量.....	162	8.3.1 规范.....	209
7.5.1 记录类型.....	163	8.3.2 主体.....	210
7.5.2 记录表类型.....	165	8.3.3 私有成员.....	212
7.6 条件判断语句.....	168	8.3.4 实例化.....	214
7.6.1 IF 语句.....	169	8.3.5 重载.....	215
7.6.2 CASE 语句.....	171	8.3.6 管理程序包和执行权限.....	217
7.7 循环语句.....	172	8.4 依赖性.....	218
7.7.1 无条件循环.....	173	8.5 实验指导.....	219
7.7.2 WHILE 循环.....	174	8.6 思考与练习.....	222
7.7.3 FOR 循环.....	175		
7.8 游标.....	176	第 9 章 表类型.....	225
7.8.1 隐式游标.....	176	9.1 外部表.....	225
7.8.2 显式游标.....	178	9.1.1 建立外部表.....	226
7.8.3 游标 FOR 循环.....	181	9.1.2 处理外部表错误.....	228
7.8.4 游标变量.....	181	9.1.3 修改外部表.....	230
7.8.5 使用游标更新数据库.....	184	9.1.4 外部表的局限性.....	231
7.9 异常处理.....	185	9.2 索引组织表.....	231
7.9.1 预定义的 Oracle 异常.....	185	9.2.1 什么是索引组织表.....	231
7.9.2 非预定义的异常.....	188	9.2.2 创建索引组织表.....	232
7.9.3 用户自定义的异常处理.....	189	9.2.3 维护索引组织表.....	235
		9.3 临时表.....	235
		9.4 对象和对象表.....	237

9.4.1 创建对象类型	237	第 11 章 视图、序列和同义词	285
9.4.2 构造函数	238	11.1 视图	285
9.4.3 引用对象类型	239	11.1.1 创建视图	285
9.4.4 方法	240	11.1.2 检索视图定义	286
9.4.5 继承	244	11.1.3 连接视图	287
9.4.6 重写	246	11.1.4 视图的相关性	287
9.4.7 对象表	247	11.1.5 带错误创建视图	288
9.5 分区表	250	11.1.6 删除视图	289
9.5.1 创建分区表	250	11.2 可更新的视图	289
9.5.2 分区表的操作	253	11.3 管理序列	291
9.5.3 分区表的维护	254	11.3.1 创建序列	291
9.6 簇表	255	11.3.2 修改序列	293
9.6.1 创建簇和簇表	255	11.3.3 查询序列	293
9.6.2 管理簇	256	11.4 管理同义词	294
9.7 散列聚簇表	257	11.5 实验指导	295
9.8 实验指导	258	11.6 思考与练习	297
9.9 思考与练习	263	第 12 章 触发器	298
第 10 章 索引	266	12.1 触发器的组成	298
10.1 了解索引	266	12.2 Oracle 触发器的类型	299
10.1.1 B 树索引	267	12.2.1 DML 触发器	300
10.1.2 位图索引	267	12.2.2 替代触发器	305
10.1.3 反向键索引	269	12.2.3 系统事件触发器	307
10.1.4 基于函数的索引	269	12.2.4 用户事件触发器	307
10.1.5 全局索引和局部索引	270	12.3 ALTER TRIGGER 语句	308
10.2 创建索引	271	12.4 与触发器相关的数据字典	309
10.2.1 创建 B 树索引	272	12.5 实验指导	310
10.2.2 创建位图索引	274	12.6 思考与练习	313
10.2.3 创建反向键索引	274	第 13 章 事务与并发控制	315
10.2.4 创建基于函数的索引	274	13.1 了解事务	315
10.2.5 创建全局和局部索引	275	13.2 事务的 ACID 特性	316
10.3 索引与约束	277	13.2.1 原子性	316
10.4 管理索引	278	13.2.2 一致性	316
10.4.1 合并索引和重建索引	278	13.2.3 隔离性	317
10.4.2 监视索引	279	13.2.4 持久性	318
10.4.3 删除索引	281	13.3 事务控制语句	318
10.5 何时使用索引	281	13.3.1 设置事务属性	319
10.6 实验指导	282		
10.7 思考与练习	283		

13.3.2 结束事务.....	322	15.1.3 调整表空间的 数据文件.....	366
13.3.3 存储点.....	323	15.1.4 临时表空间.....	371
13.3.4 设置约束延期性.....	325	15.1.5 大文件表空间.....	373
13.4 并发控制.....	327	15.1.6 非标准块表空间.....	374
13.4.1 锁.....	327	15.1.7 删除表空间.....	375
13.4.2 锁模式.....	327	15.2 撤销表空间.....	375
13.5 锁粒度.....	329	15.2.1 什么是自动撤销管理.....	375
13.5.1 TX 锁.....	330	15.2.2 撤销表空间管理参数.....	376
13.5.2 TM 锁.....	330	15.2.3 管理撤销表空间.....	377
13.5.3 数据库级锁.....	331	15.2.4 监视撤销表空间.....	379
13.6 查询锁.....	331	15.3 段.....	379
13.7 死锁.....	334	15.4 盘区.....	381
13.8 实验指导.....	335	15.5 数据块.....	382
13.9 思考与练习.....	336	15.6 实验指导.....	383
第 14 章 安全.....	338	15.7 思考与练习.....	384
14.1 用户账号.....	338	第 16 章 备份与恢复.....	385
14.1.1 用户配置文件.....	339	16.1 数据库备份概述.....	385
14.1.2 修改用户账号.....	342	16.2 数据库备份模式.....	386
14.1.3 删除用户账号.....	344	16.3 备份数据库.....	387
14.1.4 监视用户.....	345	16.3.1 物理备份.....	387
14.2 权限管理.....	345	16.3.2 逻辑备份.....	389
14.2.1 系统权限.....	346	16.4 数据库手动恢复.....	395
14.2.2 对象权限.....	349	16.4.1 实例恢复.....	395
14.3 角色管理.....	352	16.4.2 介质恢复.....	396
14.3.1 系统预定义角色.....	352	16.5 恢复管理器 (RMAN).....	401
14.3.2 自定义角色.....	353	16.5.1 设置 RMAN.....	401
14.3.3 启用和禁用角色.....	355	16.5.2 连接到目标数据库.....	402
14.3.4 修改用户时设置角色.....	357	16.5.3 RMAN 操作目标 数据库.....	404
14.3.5 删除角色.....	357	16.5.4 RMAN 通道.....	404
14.4 实验指导.....	358	16.5.5 RMAN 备份数据库.....	408
14.5 思考与练习.....	359	16.5.6 RMAN 恢复数据库.....	413
第 15 章 管理存储结构.....	360	16.5.7 RMAN 管理恢复目录.....	416
15.1 表空间.....	360	16.6 实验指导.....	421
15.1.1 创建表空间.....	360	16.7 思考与练习.....	422
15.1.2 表空间属性.....	364		

第 1 章 Oracle 关系数据库

本章首先介绍关系数据模型，包括如何根据实际的应用创建数据库，并规范化数据库；然后介绍最优秀的关系型数据库产品——Oracle 数据库的应用模式，这也是数据库领域最常用的应用模式。最后介绍 Oracle 10g 数据库系统的安装和数据库的创建，为读者学习本书以后各章奠定基础。

本章学习要点：

- 关系模型的组成
- 理解关系模型中的一些常用术语
- 掌握关系数据库的设计
- 掌握数据库的规范化
- 正确安装 Oracle 10g
- 了解 Oracle 数据库的应用模式
- 能够使用 DBCA 创建数据库
- 掌握如何启动和关闭 Oracle 服务
- 能够通过 SQL*Plus 连接到 Oracle 数据库

1.1 关系数据模型

数据库技术始于 20 世纪 60 年代。当时计算机的主要应用领域正由科学计算逐步转向事务处理，这一转变促进了数据库技术的发展。到 20 世纪 70 年代，数据库技术有了飞速的发展。E.F.Codd 教授发表了“大型共享数据库数据的关系模型”论文，他提出了关系数据库理论和模型，为关系数据库模型奠定了理论基础，从而使数据库技术成为计算机科学的一个重要分支。

从 1968 年美国 IBM 公司推出层次模型的数据管理系统开始，几十年来，数据库技术的发展经历了层次模型数据库、网状模型数据库，再到关系模型数据库，现在又出现了第三代新型的数据库技术——面向对象的数据数据库技术。

关系数据库系统是目前应用最为广泛的数据库系统。关系数据库系统采用关系数据模型作为数据的组织方式。关系模型建立在严格的数学概念的基础上，它由关系的数据结构、关系的操作集合和关系的完整性约束三部分组成。

1.1.1 数据结构

在关系数据库系统中，关系是数据模型的核心。关系数据库系统中最主要的数据结构就是关系。在实际的表示中，关系可以用一个由“行”和“列”组成的二维表来直观

地表示。表中的每一列表示关系的一个属性，每列的名字即为一个属性名，每一行表示一个记录，代表一个物理实体。关系数据库中，所有的数据都是通过表来进行存储的，没有表数据就无法进行存储和表示。

如图 1-1 所示为一个学生信息表，表中的每一行表示一个物理实体——学生。例如，第一行记录的是学生“黄小明”的信息。学生信息表中的每一列代表一个学生的一项数据，记录学生的一部分信息。例如，“SID”列用来存储每个学生的学号，“Sname”列用来存储每名学生的姓名。第一行的“SID”列存储了学生“黄小明”的学号 063721，“Sbirth”列存储了学生“黄小明”的出生日期 1985 年 10 月 03 日。而第二行记录的是学生“李文斌”的信息。

SID	Sname	Ssex	Sbirth	Sdept
063721	黄小明	男	1985/10/03	法律系
063245	李文斌	男	1985/04/15	历史系
053654	李 娜	女	1985/01/25	外语系
053546	王晓明	男	1984/11/09	机电系

图 1-1 学生信息表

在关系模型中，列称为属性或字段，行称为元组或记录，元组的集合称为关系或实例。有时也称关系为表格，而属性的取值范围为域。

关系数据模型中数据结构的核心规则是：所有数据都以二维表的形式呈现给用户。它提供了数据表达和存储之间的一级抽象概念，这样使得数据具有独立性，即无论采用哪种关系数据库产品，也不论其采用的是什么物理存储方式，查询和处理数据的方式都是相同的。

具体来说，数据的独立性体现在两个重要的方面：物理独立性和逻辑独立性。物理独立性是指用户所看到的数据与数据的物理存储是完全独立、不同的，即可以改变或重新安排数据的物理存储而不会影响到数据的使用和逻辑数据库的设计。逻辑独立性指的是，可以改变表、行和列之间的相互关系，即对数据结构进行修改而不会削弱应用程序，特别是查询的功能。

若表中的某一个属性或属性组的取值能够唯一标识一行记录，则称该属性或属性组为候选键。一个表中可能具有多个候选键，这时可以选择一个作为该表的主键，即主键是从候选键中选出的，剩下的候选键称为备用键。候选键中的属性称为键属性，其他属性则称为非键属性。主键中的属性称为主属性，其他属性则称为非主属性。

判断属性是否可以成为候选键有两条严格的规则：

- 表中任意两行记录的候选键的属性值不相同，即唯一性规则（Uniqueness Property）。
- 作为候选键的属性组合中，其属性组合的任意子集都不具有唯一性，即不可简化性规则（Irreducibility Property）。

例如,在图 1-1 中,属性组合{SID, Sname, Sbirth}和属性{SID},二者均可独立地标识表中的每一行,都表现出唯一性;但是属性{SID}显然是属性组合{SID, Sname, Sbirth}的子集,即属性组合中的子集{SID}表现出唯一性,所以属性组合{SID, Sname, Sbirth}不满足不可简化性规则,因此属性组合{SID, Sname, Sbirth}不是主键。在“学生信息表”中,由于 Sbirth 和 Sname 都可能重复,所以{SID}是唯一的一个主键,“Sname”是这个主键的主属性,表中其他的属性称为非主属性。

现在介绍一下什么是外键。假设 X 是表 A 中的一个属性组合,若 X 同时是另一个表 B 的主属性,则称 X 是表 A 关于表 B 的外键。

例如,与表“学生信息表”相对应的有一个表“院系信息”,在表“院系信息”中包含了一个系的系名、办公地点、联系电话、系主任信息,对于一个学校而言,院系的名称是唯一的,因此可以将“系名”作为一个主键。该属性同时也是表“学生信息表”中的一个属性,所以“Sdept”是表“学生信息表”关于表“院系信息”的外键。

1.1.2 关系操作

关系模型中常用的关系操作包括选择(Select)、投影(Project)、连接(Join)、除(Divide)、并(Union)、交(Intersection)、差(Difference)等查询操作,以及增加(Insert)、删除(Delete)、修改(Update)操作两大部分。其中,查询操作是其最重要的部分。

关系操作的特点是集合操作,即操作的对象和结果都是集合。早期的关系操作能力通常用代数方式和逻辑方式表示,它们分别被称为关系代数和关系演算。关系代数是使用对关系的运算来表示查询要求的方式。关系演算是用谓词来表达查询要求的方式。实际的关系数据库系统的查询语言除了提供关系代数或关系演算功能外,还提供了许多附加功能,例如,集函数、算术运算等。

在实际的关系数据库系统中,为用户提供了一种介于关系代数和关系演算之间的语言 SQL (Structured Query Language, 结构化查询语言)。SQL 不仅具有丰富的查询功能,而且具有数据定义和数据控制功能。它充分体现了关系数据库语言的特点和优点,是关系数据库的标准语言,例如,Oracle 数据库系统中的 SQL*Plus, SQL Server 中的 Transact_SQL。

1.1.3 关系的完整性约束

关系模型的完整性规则是对关系的某种约束条件。关系模型允许定义 3 类完整性约束:实体完整性、参照完整性和用户定义的完整性。其中实体完整性和参照完整性是关系模型必须满足的完整性约束条件,被称做关系的两个不变性,应该由关系数据库系统自动支持。用户定义的完整性是应用领域需要遵循的约束条件,体现了具体领域中的语义约束。

- **实体完整性(Entity Integrity)** 实体完整性规则规定关系的所有元组的主键属性不能取空值;当主键由属性组组成时,属性组中所有的属性均不能取空值。例如,当选定如图 1-1 所示的“学生信息表”中的“SID”为主键时,则“SID”属性不能取空值。实体完整性是针对基本

关系的, 一个基本表通常对应于现实世界中的一个实体集。例如, “学生信息表”对应于该学校中的所有学生的集合。

- **参照完整性 (Referential Integrity)** 现实世界中的实体之间往往存在某种联系, 这样就会存在关系与关系之间的引用。例如, 表“学生信息表”和表“院系信息”之间存在着属性的引用, 即“学生信息表”引用了“院系信息”表中的主键“Sdept”。显然, “学生信息表”表中的“Sdept”属性的取值必须存在于表“院系信息”中。
- **用户定义的完整性 (User-defined Integrity)** 实体完整性和参照完整性是任何关系数据库系统都必须支持的。除此之外, 不同的关系数据库系统根据其应用环境的不同, 往往还需要一些特殊的约束条件, 用户定义的完整性就是针对某一具体关系数据库的约束条件。它反映某一具体应用所涉及的数据必须满足的语义要求。例如, 在“学生信息表”中, 可以根据具体的情况规定学生的年龄不得大于 60 岁。

1.2 关系与关系模式

在数据库中要区分型和值, 型是指对某一类数据的结构和属性的说明, 值是型的一个具体赋值。在关系数据库中, 关系模式是型, 关系是值, 关系模式是对关系的描述。因为关系实质上是一个二维表, 表中的每一行在关系模式中被称为元组, 一个元组由许多属性组成, 也就是表中的列。关系是元组的集合, 因此关系模式必须指出这个元组集合的结构, 即它由哪些属性组成, 以及这些属性的取值等。关系模式可以形式化地表示为:

$$R(U, D, \text{dom}, F)$$

其中, R 表示关系名; U 是组成该关系的属性名集合; D 是属性的域; dom 是属性向域的映像集合; F 为属性之间数据的依赖关系集合。

关系模式通常可以简记为:

$$R(U) \text{ 或 } R(A_1, A_2, \dots, A_n)$$

其中, R 为关系名, $A_1, A_2, \dots, A_n$ 为属性名。而域以及属性向域的映像常常直接声明为属性的类型、长度。例如, 描述“学生信息表”的关系模式如下:

$$S(\text{SID}, \text{Sname}, \text{Ssex}, \text{Sbirth}, \text{Sdept})$$

关系是关系模式在某一时刻的状态或内容。关系模式是静态的、稳定的, 而关系是动态的、随时间不断变化的, 因为关系操作在不断地更新数据库中的数据。

1.3 数据库的设计

在数据库的设计过程中, 有许多需要考虑的因素, 如数据库的背景、应用环境等方面, 只有对这些因素都进行了详细的分析, 设计的数据库才能易于使用和维护, 并且具有高效和一致的特征。

对数据库的设计包括如下几个过程:

① 信息的收集。研究和考虑所要建立的数据库的信息环境,对数据库应用领域中各种信息要求和操作要求进行详细的分析;了解应用领域中数据项、数据项之间的关系和所有的数据操作的详细要求,了解哪些因素对响应时间、可用性和可靠性有较大的影响等。收集信息需要数据库设计人员与相关人员进行各种交流,了解数据库需要完成的功能、数据项的定义等。

② 确定数据。这是数据库设计最基本、最重要的工作,在收集到所需的全部相关信息后,将所有收集到的数据类型,以及这些类型的属性进行明确的定义,同时需要仔细地考虑属性与数据类型的关系。

③ 建立实体-关系模型。按照定义的数据类型和属性创建实体和实体属性列表。实体形成表,属性则为表中的列。

④ 进行规范化。按照规范化准则对数据进行整体分析,以找到逻辑上的错误,更正任何违背了规范化形式的做法,使数据符合规范化的要求。

⑤ 编写 SQL 代码以创建数据库。

1.3.1 实体-关系模型

实体-关系模型也称为 E-R (Entity-Relationship) 数据模型。设计 E-R 数据模型的出发点是为了更有效和更好地模拟现实世界,而不需要考虑在计算机中如何实现。下面介绍 E-R 数据模型的 3 个抽象概念。

1. 实体 (Entity)

实体是 E-R 模型的基本对象,是现实世界中各种事物的抽象。凡是可以相互区别,并可以被识别的事、物、概念等均可认为是实体。在一个单位中,具有共性的一类实体可以划分为一个实体集。例如,学生李小明、王杰等都是实体。为了便于描述,可以定义“学生”这样一个实体集,所有学生都是这个集合的成员。

2. 属性 (Attribute)

实体一般具有若干特征,称之为实体的属性。例如,学生具有学号、姓名、性别、所属院系等属性。实体的属性值是数据库中存储的主要数据,一个属性实际上相当于关系数据库中表的一个列。

能唯一标识实体的属性或属性组称为实体集的实体键。如果一个实体集有多个实体键,则可以选择一个作为实体主键。

3. 联系 (Relationship)

实体之间会存在各种关系,例如,学生实体与课程实体之间有选课关系,人与人之间可能有领导关系等。这种实体与实体之间的关系被抽象为联系。E-R 数据模型将实体之间的联系区分为一对一 (1:1)、一对多 (1:N) 和多对多 (M:N) 3 种,并在模型中明确地给出这些联系的语义。

□ 一对一联系 (1:1) 对于实体集 A 和实体集 B 来说,如果对于 A 中的每一个实体 a, B 中

至多有一个实体 b 与之联系；而且，对于实体集 B 和实体集 A 来说，也是如此，则称实体集 A 与实体集 B 具有一对一联系，表示为 $1:1$ 。

- **一对多联系 ($1:N$)** 对于实体集 A 中的每一个实体，在实体集 B 中有 N 个实体与之联系，而且，对于实体集 B 中的每一个实体，实体集 A 中至多有一个实体与之联系，则称实体集 A 和实体集 B 具有 1 对多联系，表示为 $1:N$ 。
- **多对多联系 ($M:N$)** 如果对于实体集 A 中的每一个实体，实体集 B 中有 N 个实体与之联系；同时，对于实体集 B 中的每一个实体，实体集 A 中有 M 个实体与之联系，则称它们具有多对多联系，记为 $M:N$ 。

在 E-R 模型中，不但实体可以有属性，联系也可以有属性。例如，“选课”这种联系可以具有选修时间、成绩等属性。

1.3.2 E-R 图的绘制

实体-关系图是表现实体-关系模型的图形工具，简称 E-R 图。E-R 图提供了用图形表示实体、属性和联系的方法。在 E-R 图中，约定实体用方框表示，属性用椭圆表示，联系用菱形表示，并在其内部填上实体名、属性名、联系名，如图 1-2 所示。

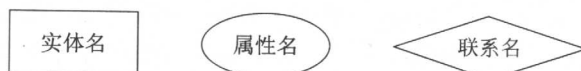


图 1-2 E-R 图的基本元素

如图 1-3 所示分别表示了实体间的 3 种不同的联系方式。

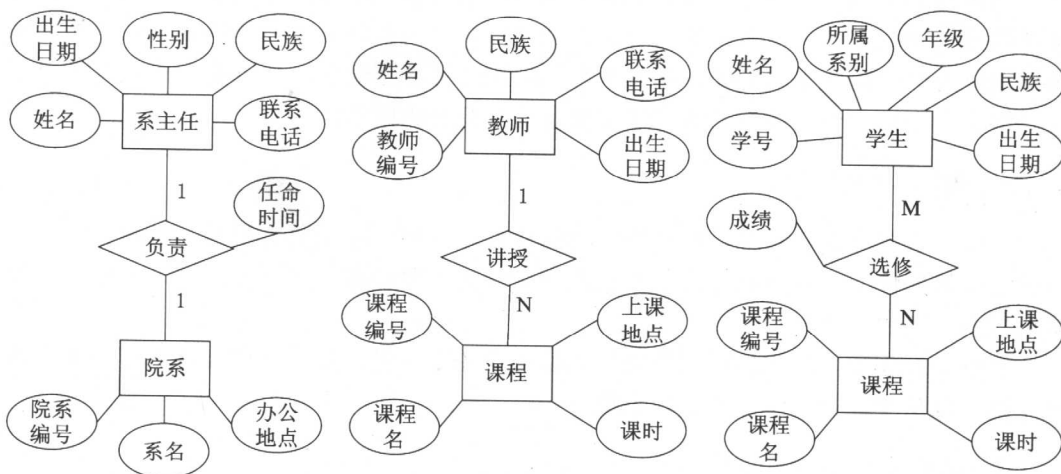


图 1-3 E-R 图示例

注意，不仅实体具有属性，而且联系也可能有属性。例如，学生与课程联系上的“成绩”，它既不是实体“学生”的属性，也不是实体“课程”的属性，而是联系“选修”的属性。有时为了使 E-R 图简洁明了，常将图中的属性省略，而着重反映实体的联系，而

将属性以表格的形式单独列出来。注意，由于现实世界的复杂性，这将导致增加 E-R 模型的复杂性。例如上例中的系主任也是教师，也可以讲授课程。

由于 E-R 图直观易懂，在概念上表示了一个数据库的信息组织情况，所以如果能够画出数据库系统的 E-R 图，也就意味着弄清楚了应用领域中的问题，此后就可根据 E-R 图，并结合具体的数据库管理系统 (DBMS)，将 E-R 模型演变为 DBMS 支持的数据模型。

1.3.3 将 E-R 模型转化为关系模式

在 E-R 模型建立后，就可以将实体、实体间的联系等模型转变为关系模式，即生成数据库中的表，并确定表的列。下面讨论由 E-R 模型生成表的方法。

1. 实体转化为表

对 E-R 模型中的每个实体，在创建数据库时相应地为其建立一个表，表中的列对应实体所具有的属性，主属性就作为表的主键。如上面建立的 E-R 模型中，可以将院系、学生和课程实体转化为表，如图 1-4 所示。

表 Teacher	表 Curricula	表 Student
教师编号 姓名 民族 联系电话 出生日期	课程编号 课程名 课时 上课地点	学号 姓名 所属系别 年级 民族 出生日期

图 1-4 实体转化为表

2. 实体间联系的处理

对于实体间的一对一关系，为了加快查询的速度，可以将一个表中的列添加到两个表中。一对一关系的转换比较简单，一般情况下不需要再建立一个表，而是直接将一个表的主键作为外键添加到另一个表中，如果联系也有属性，则还需要将联系的属性添加到该表中。例如在上例中，可以直接在“院系”实体中添加一个“负责人”属性，而将“系主任”实体的“姓名”属性值添加到该表中，并将“任命时间”属性也添加到表中。

实体间的一对多关系的转换也不需要再为其创建一个表。设表 A 与表 B 之间是 1:N 关系，则转换时可以将表 A 的主键作为外键添加到表 B 中。

多对多关系的转换要比一对多关系复杂得多，通常情况是创建一个称为连接表的特殊表，以表达两个实体之间的关系。连接表的列包含其连接的两个表的主键列，同时包含一些可能在关系中存在的特定的列。

为了保证设计的数据库能够有效、正确地运行，往往还需要对表进行规范化，以消除数据库中的各种异常现象。

1.4 关系数据库规范化理论

数据库理论与设计中有一个重要的问题，就是在一个数据库中如何构造合适的关系模式，它涉及一系列的理论与技术，从而形成了关系数据库设计理论。由于合适的关系模式要符合一定的规范化要求，所以又称为关系数据库的规范化理论。