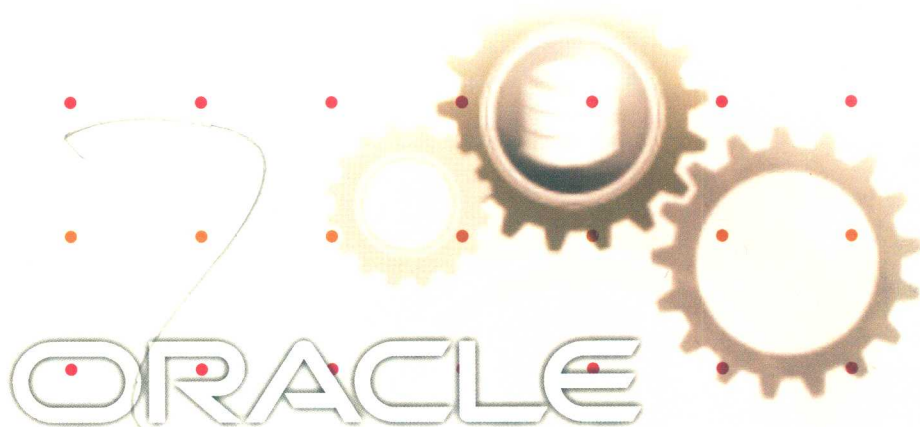




应用型本科信息大类专业“十三五”规划教材



Oracle 11g 数据库教程

◎ 主编 高翠芬 王立平



华中科技大学出版社
<http://www.hustp.com>



应用型本科信息大类专业“十三五”规划教材

Oracle 11g 数据库教程



主 编 高翠芬 王立平
副主编 赵永霞 黄 薇 范桂龄
刘小洋 王维虎 姜 尚



RFID



华中科技大学出版社

<http://www.hustp.com>

中国·武汉

内 容 简 介

Oracle 数据库技术是当今社会主流的一门数据库应用技术,数据库的主要任务是研究如何存储、使用和管理数据。

本书以教务管理系统中的成绩管理子系统的开发与管理为主线,深入浅出地介绍 Oracle 11g 数据库系统开发与管理的知识。本书共 13 章。其中,第 1~5 章为 Oracle 数据库基础知识篇,内容包括数据库概述、Oracle 11g 介绍、Oracle 的常用管理工具、Oracle 数据库体系结构、Oracle 网络管理。第 6~10 章为应用开发篇,以项目开发中数据库编程部分的所有内容为依托,完整地介绍了 Oracle 数据库的应用与开发中需要的各个环节的内容,包括 Oracle 数据库创建、表空间管理、Oracle 模式对象、基本的 SQL 语言查询和高级查询、DML 操作、PL/SQL 编程基础、游标与异常、存储过程与函数的创建、触发器和包的创建与应用等。第 11~13 章为 Oracle 数据库管理与维护篇,内容包括 Oracle 系统安全管理、数据库闪回技术、数据库备份和恢复。

本书内容翔实、结构合理、示例丰富、语言简洁流畅。读者可以从 Oracle 知识零起点开始逐渐全面地了解 Oracle 数据库的基本原理和相关基础管理与应用开发,为将来深入学习 Oracle 数据库奠定基础。书中所有的案例既使用 SQL 语句或各种命令来完成,同时又提供了 SQL Developer 或 OEM 等工具来实现案例的全过程,这样不同程度的读者都可以从本书获取相关的知识。所以本书面向 Oracle 数据库初学者,尤其适合作为高等院校计算机相关专业、信息系统等相关专业的 Oracle 数据库课程教材和教学参考书,同时也适合作为 Oracle 数据库开发人员的参考资料。

为了方便教学,本书还配有电子课件等教学资源包,任课教师和学生可以登录“我们爱读书”网(www.ibook4us.com)免费注册并浏览,或者发邮件至 hustpeiit@163.com 索取。

图书在版编目(CIP)数据

Oracle 11g 数据库教程/高翠芬,王立平主编. —武汉:华中科技大学出版社,2018.12

应用型本科信息类专业“十三五”规划教材

ISBN 978-7-5680-3107-3

I. ①O… II. ①高… ②王… III. ①关系数据库系统-高等学校-教材 IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2017)第 169426 号

Oracle 11g 数据库教程

Oracle 11g Shujuku Jiaocheng

高翠芬 王立平 主编

策划编辑:康 序

责任编辑:段雅婷 柯丁梦

责任监印:朱 玢

出版发行:华中科技大学出版社(中国·武汉)

电话:(027)81321913

武汉市东湖新技术开发区华工科技园

邮编:430223

录 排:武汉楚海文化传播有限公司

印 刷:武汉科源印刷设计有限公司

开 本:787mm×1092mm 1/16

印 张:26.75

字 数:718 千字

版 次:2018 年 12 月第 1 版第 1 次印刷

定 价:58.00 元



本书若有印装质量问题,请向出版社营销中心调换
全国免费服务热线:400-6679-118 竭诚为您服务
版权所有 侵权必究

前言

PREFACE

在计算机技术高速发展的今天,数据库管理系统(DBMS)已经成为大型信息系统的重要支撑。在相关的数据库产品中,Oracle 数据库产品以其卓越的性能获得了广泛的应用,已经成为当今世界上最流行的大型关系型数据库管理系统,世界上很多重要的工业或商业领域里都可以看到 Oracle 数据库的应用。当前各种公司对于 Oracle 的相关人才的需求量相当大,很多高校都开设了 Oracle 的相关课程。而 Oracle 11g 作为市场上应用率相当高的产品,目前关于其的图书虽然比较多,但是偏重于技术的深度,很多读者希望有一本难易适度、快速入门又能够比较全面地了解 Oracle 11g 相关知识的书。正是基于以上情况,作者根据多年的开发和教学经验,根据教学和自学的相关规律,编写了《Oracle 11g 数据库教程》。本书内容丰富,注重实践和应用,实用性强。从实用的角度出发,采用理论与实践相结合的方式全面介绍了 Oracle 数据库应用开发与管理技术。该书内容中既包括企业 Oracle 11g 的应用和管理中常用的各种知识点,又包括 Oracle 认证中的相关内容。全书以案例驱动法进行组织,书中通过大量的示例代码和案例分析,由浅入深地介绍 Oracle 数据库系统的使用方法、技术原理、SQL 操作、PL/SQL 应用、数据备份等内容,并配以习题和上机练习,强化基本概念,着重训练学生的动手能力,力求让不同层次的 Oracle 学习人员尽快掌握 Oracle 数据库的相关知识。通过阅读本书,读者能够快速掌握 Oracle 应用与开发和管理的各种知识。

本书的主要特点如下。

(1)每个知识点都使用案例驱动法进行编写,通过丰富的案例分析,提高读者学习的效率。

为了便于读者理解本书内容、提高学习的效率,每章的相关知识点都有详细的案例介绍,大部分的案例都结合实际应用。本书以一个完整的教务管理系统中的成绩管理子系统为依托案例,结合课本理论知识点,进行案例驱动讲解,将每章的理论知识结合实践开发加以灵活运用,力求让应用开发人员将这些知识点尽快应用到实际的开发过程中。

(2)本书内容全面,涵盖 Oracle 11g 的各方面的技术点,并涉及 Oracle 认证考试的相关内容。

本书内容涵盖 Oracle 11g 的体系结构、系统安全、常用工具、SQL 应用及

PL/SQL 语言在实际项目中需要重点掌握的相关技术,同时还提供了 Oracle 认证考试相关内容,课后习题中也有与 Oracle 认证知识点相关的习题。这些可指导并帮助读者在掌握 Oracle 理论与应用的基础上,获得高含金量的 Oracle 认证证书。

(3)提供 Oracle 使用过程中的常用技巧。

本书对 Oracle 使用过程中经常出现的问题和一些常用技巧进行了介绍,另外还对初学者经常出现的一些问题进行了总结归纳,让读者能尽快上手。

(4)提供完善的技术支持和售后服务。

为了方便教学,本书配有电子课件等教学资源包,任课教师和学生可以登录“我们爱读书”网(www.ibook4us.com)免费注册并浏览,或者发邮件至 hustpei-it@163.com 索取。

本书中的案例数据库为学生管理数据库(XSGL),其中的管理员用户 SYS 的密码为 XG501Oracle,XSGL 数据库中的 XSGLADMIN 用户的密码为 manager,样例数据库中的用户 SCOTT 的密码为 tiger,用户 HR 的密码为 hr。

在实际教学中,建议本书按 64 学时安排教学,也可安排 48 学时进行重要内容的讲授,同时要安排至少一半的实验学时进行上机实验。实验环境建议每台计算机都要安装 Oracle 11G R2 的企业版,以便于学生能够全面、深入地了解 Oracle 11g 的全部内容,并建议在分布式网络环境中进行实验练习。

本书由武昌理工学院高翠芬、萍乡学院王立平担任主编,由武汉东湖学院赵永霞、武昌理工学院黄薇、西京学院范桂龄、文华学院刘小洋、湖北工程学院王维虎、大连科技学院姜尚担任副主编。其中,第 1 章由赵永霞编写,第 2 章由黄薇编写,第 3 章至第 8 章由高翠芬编写,第 9 章由王立平编写,第 10 章由刘小洋编写,第 11 章由王维虎编写,第 12 章由姜尚编写,第 13 章由范桂龄编写,最后由高翠芬审稿并统稿。

本书在编写过程中得到了武昌理工学院的各级领导、多位同事和许多同行的大力协助与支持,使编者获益良多,在此表示衷心的感谢。同时感谢 Oracle 教育工程和 OracleWDP 俱乐部给编者多次提供的 OCP、OCM 的培训,本书中的部分案例的思路也源自培训过程中的案例和练习。最后感谢肖鸣老师对本书的校对和指导。

作者在从事 Oracle 11g 基础教程的教研实践中得到了 2015 年度江西省高校人文社会科学研究项目(TQ1516)、2017 年度高校人文社会科学重点研究基地项目(JD17127)、2017 年度江西省教育科学规划课题(17YB276、17YB273)、2017 年度江西省重点研发计划项目(20071BBE50049)的资助。在本书的撰写过程中,我们也得到了萍乡学院的支持。华中科技大学出版社负责本书编辑出版工作的同志为本书付出了大量的劳动,他们严格细致、一丝不苟的工作精神,使得本书的质量得以进一步提升。在此谨向每一位关心本书编写工作的各界人士表示由衷的感谢。尽管我们认真撰写、力求完美,但书中难免存在不足,恳请广大读者和专家批评指正。

由于作者水平有限,加之时间匆促,书中错误在所难免,敬请广大读者和专家批评指正。

编 者

2018 年 2 月

目
录

CONTENTS

第 1 章 数据库基础	(1)
1.1 数据库管理系统概述	(1)
1.2 关系数据库基础	(2)
1.3 数据库设计	(6)
第 2 章 Oracle 11g 数据库概述	(15)
2.1 Oracle 11g 概述	(15)
2.2 Oracle 11g 的安装	(18)
2.3 Oracle 11g 的卸载	(23)
第 3 章 Oracle 11g 数据库常用工具	(26)
3.1 SQL * Plus	(26)
3.2 Oracle 企业管理器	(38)
3.3 SQL Developer	(42)
第 4 章 Oracle 11g 体系结构	(47)
4.1 Oracle 11g 体系结构概述	(47)
4.2 Oracle 的内存结构	(48)
4.3 Oracle 服务器的进程结构	(51)
4.4 Oracle 数据库的存储结构	(55)
4.5 数据字典	(63)
第 5 章 Oracle 11g 网络管理	(67)
5.1 Oracle Net 概述	(67)
5.2 本地 Net 服务名的配置与管理	(70)
5.3 监听程序的配置与管理	(78)
第 6 章 Oracle 数据库创建	(86)
6.1 创建 Oracle 数据库	(86)
6.2 Oracle 数据库的启动与关闭	(98)

第 7 章	Oracle 11g 表空间的管理	(106)
7.1	Oracle 表空间概述	(106)
7.2	创建表空间	(108)
7.3	管理表空间	(114)
7.4	删除表空间	(122)
7.5	数据文件的管理	(122)
第 8 章	数据库常用对象创建与管理	(130)
8.1	Oracle 11g 数据库常用对象概述	(130)
8.2	表	(130)
8.3	约束	(162)
8.4	索引	(180)
8.5	视图	(191)
8.6	序列	(197)
8.7	数据库链接	(200)
8.8	同义词	(202)
第 9 章	Oracle 数据库 DML 操作和数据查询	(210)
9.1	Oracle 数据 DML 操作	(210)
9.2	Oracle 中的常用查询	(221)
9.3	连接查询	(231)
9.4	子查询	(240)
9.5	集合运算	(246)
9.6	Oracle 支持的 SQL 函数	(247)
9.7	Oracle 复杂查询	(259)
第 10 章	PL/SQL 程序设计	(270)
10.1	PL/SQL 基础	(270)
10.2	数据类型	(273)
10.3	PL/SQL 中的 SELECT 语句	(281)
10.4	PL/SQL 基本程序结构和语句	(286)
10.5	异常	(290)
10.6	游标	(294)
10.7	存储过程和函数	(300)
10.8	包	(310)
10.9	触发器	(319)
第 11 章	Oracle 11g 系统安全管理	(336)
11.1	Oracle 11g 数据库安全性	(336)
11.2	权限管理	(340)
11.3	用户管理	(348)

11.4	角色	(354)
11.5	概要文件	(358)
11.6	审计	(363)
第 12 章	闪回技术	(369)
12.1	闪回技术概述	(369)
12.2	闪回查询	(370)
12.3	闪回版本查询	(373)
12.4	闪回事务查询	(375)
12.5	闪回表	(376)
12.6	闪回删除	(378)
12.7	闪回数据库	(381)
12.8	闪回数据归档	(385)
第 13 章	Oracle 11g 数据库备份和恢复	(389)
13.1	备份和恢复概述	(389)
13.2	物理备份	(391)
13.3	物理恢复	(394)
13.4	RMAN	(398)
13.5	逻辑备份和恢复	(403)
13.6	数据泵	(410)
参考文献		(420)

第1章

数据库基础



1.1 数据库管理系统概述

数据库技术的产生使计算机应用进入了一个新的时期,社会的各个领域都与计算机发生了联系。数据库技术聚集了数据处理最精华的思想,是管理信息最先进的工具。

数据库在我们的生活中已经无处不在,对事物的描述,充满了信息与数据的概念。因此在介绍数据库技术之前,先简单介绍数据、信息与数据处理三个重要的概念。

1.1.1 数据、信息与数据处理

数据就是对客观事物的一种反映或描述,它是用一定方式记录下来的客观事物的特征。数据是数据库中存储的基本对象。例如,某学生的学号、姓名、性别、出生日期、地址、成绩等,就是反映该学生基本状况的数据,它们是学生信息数据库的基本对象。数据的形式可以是文字、数值、图形、声音、视频等。

信息是人围绕某个目的从相关数据中提取的有价值的意义。例如,从成绩这个数据,可以得到该学生是否可以获得奖学金等信息。数据是承载信息的物理符号或称之为载体,而信息是数据的内涵。二者的区别是:数据可以表示信息,但不是任何数据都能表示信息,同一数据也可以有不同的解释。信息是抽象的,同一信息可以有不同的数据表示方式。

数据处理是将数据转换成信息的过程。这个过程主要是指对所输入的数据进行加工整理,包括对数据的收集、存储加工、分类、检索、传播等一系列活动。其目的就是大量的、已知的数据出发,根据事物之间的固有联系和运动规律,采用分析、推理、归纳等手段,提取对人们有价值、有意义的信息,作为某种决策的依据。

我们可以用图 1-1 简单地表示出数据和信息的关系。从图 1-1 中我们可以看到,只有经过了数据处理的数据才有可能成为信息。注意数据与信息的关系是相对的,而不是绝对的。

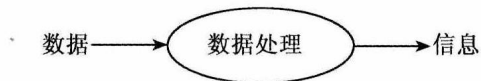


图 1-1 数据和信息的关系

1.1.2 数据库系统的有关概念

数据库系统(database system, DBS)是指一个完整的、能为用户提供信息服务的系统。数据库系统是引进数据库技术后的计算机系统,它实现了有组织地、动态地存储大量相关数据的功能,提供了数据处理和信息资源共享的便利手段。而数据库技术是一门研究数据库

结构、存储、管理和使用的软件学科。下面介绍几个相关概念。

数据库(database,DB)是以一定的数据模型组织和存储的、能为多个用户共享的、独立于应用程序的、相互关联的数据集合,或者可以理解为它是一个存放数据的“仓库”。数据库本身不是独立存在的,它是数据库系统的一部分,在实际应用中,人们常面对的是数据库系统。

数据库管理系统(database management system,DBMS)是处理数据库访问的软件,可以把它看成是操作系统的一个特殊用户,它向操作系统申请所需的软硬件资源,并接受操作系统的控制和调度。DBMS 提供数据库的用户接口。目前常用的数据库管理系统有 FoxPro、SQL Server、Oracle 和 Infomix 等。只有在计算机上配置了 DBMS,才能建立所需要的数据库。DBMS 是数据库系统的核心部分。DBMS 主要用于提供一个可以方便、有效地存取数据库信息的环境。

数据库应用系统是指系统开发人员利用数据库资源开发出来的、面向某一类信息处理问题而建立的软件系统,例如学生信息管理系统和人事管理系统等。



1.2 关系数据库基础

关系数据库应用数学方法来处理数据库中的数据。美国 IBM 公司的 E. F. Codd 系统而严格地提出了关系模型的概念,他从 1970 年起连续发表了多篇论文,奠定了关系数据库的理论基础。在 20 世纪 70 年代末,关系方法的理论研究和软件系统的研制取得了较好的成果,其中,美国 IBM 公司的 System R 和美国加州大学 Berkeley 分校的 Ingres 的关系数据库实验系统在功能和技术上最有代表性。1981 年 IBM 公司在 System R 的基础上先后推出了两个商品化的 RDBMS(关系数据库管理系统):SQL/DS 和 DB2。同时,Berkeley 分校也研制出了商品化的 Ingres 系统,使数据库走向了实用化和商品化。20 世纪 80 年代,关系数据库管理系统成为发展的主流,越来越多地用到微机上。随着计算机技术与网络技术的发展,数据库管理系统向着分布式和面向对象式数据库系统发展,产生了网络数据库、多媒体数据库和对象-关系数据库及其他扩充的关系数据库系统。这些年来,关系数据库管理系统的研究取得了辉煌的成就,涌现出许多性能良好的商品化的 RDBMS,如 DB2、Oracle、SQL Server、Informix 等,进入了关系数据库的鼎盛时代,并在此基础上向新一代数据库系统发展。

关系数据库是以关系模型为基础的数据库。在数据库技术发展初期,人们普遍使用的是层次数据库管理系统和网状数据库管理系统,它们分别以层次模型和网状模型为基础。现在用得比较成熟的数据库管理系统则是关系数据库管理系统,它有很好的用户界面,并具有简单灵活的数据模型、较高的数据独立性、良好的语言接口和坚实的理论基础。

关系模型是关系数据库的数据模型,它有严格的数学基础,抽象级别比较高,简单清晰,便于理解和使用,也是目前用的较为成熟的数据模型。所谓关系模型就是用二维表结构来表示实体及其联系的模型。关系模型是建立在集合代数的基础上的。

1.2.1 关系的基本概念

1. 域

在关系数据模型中,每个属性都有一个取值范围。域是用来描述这个属性取值范围的,或者说域是一组具有相同数据类型的值集合。如, $\{\text{若干整数集合}\},\{\text{男,女}\},\{0,1\},\{A,B,C,D,E\}$ 等,都可以看作域。域有如下特点:

(1)域必须命名。命名后的域如下所示:

$D_1 = \{\text{张三, 李四, 王五, 赵六}\}$,表示某些姓名的集合,域名为 D_1 。

$D_2 = \{\text{男, 女}\}$,表示性别的集合,域名为 D_2 。

$D_3 = \{18, 19, 20\}$,表示年龄的集合,域名为 D_3 。

(2)域中数据的个数称为域的基数。

例如上述命名后的域, D_1 的基数为 4, D_2 的基数为 2, D_3 的基数为 3。

2. 笛卡儿积

设 $D_1, D_2, \dots, D_n$ 为给定的域,则 $D_1, D_2, \dots, D_n$ 的笛卡儿积为:

$$D_1 \times D_2 \times \dots \times D_n = \{(d_1, d_2, \dots, d_n) \mid d_i \in D_i, i=1, 2, \dots, n\}$$

其中,每一个元素 $(d_1, d_2, \dots, d_n)$ 称为一个 n 元组,简称元组,元组中的每一个 d_i 称为元组的一个分量, d_i 必须是 D_i 中的一个值。

这里注意元组与集合的区别。元组不是 d_i 的集合,因为元组中的分量是按序排列的;集合是一组相关数据的组合,集合中的数据是无序的。二者的表示方法不同。

例如,元组: $(a, b, c) \neq (b, a, c) \neq (c, b, a)$;

$(a, a, a) \neq (a, a) \neq (a)$ 。

集合: $\{a, b, c\} = \{b, a, c\} = \{c, b, a\}$ 。

例 1-1 求笛卡儿积 $D_1 \times D_2$ 。

设 $D_1 = \{0, 1\}, D_2 = \{a, b, c\}$, 则:

$$D_1 \times D_2 = \{(0, a), (0, b), (0, c), (1, a), (1, b), (1, c)\}。$$

可以把笛卡儿积看成是一张二维表。上述 $D_1 \times D_2$ 的结果还可以画成表格形式,如表 1-1 所示。

表 1-1 中的第一个分量来自 D_1 ,第二个分量来自 D_2 。笛卡儿积就是由所有这样的元组组成的集合。

3. 关系

笛卡儿积 $D_1 \times D_2 \times \dots \times D_n$ 的任意一个子集称为集合 $\{D_1, D_2, \dots, D_n\}$ 上的一个 n 元关系。关系中属性个数称为元数,元组个数称为基数。在实际应用中,关系往往是从笛卡儿积中选取的有意义的子集。在计算机里,一个关系可以存储为一个文件。

例 1-2 设 $D_1 = \{\text{张强, 李林, 王孝文}\}$ 是一个学生集合, $D_2 = \{\text{高等数学, 英语, C 语言, 电子技术}\}$ 是一个课程集合, 则 $R = D_1 \times D_2$ 如表 1-2 所示。

关系 R_1 是从 R 中提取的一个有意义的子集。

$R_1 \subseteq R, R_1$ 如表 1-3 所示。

表 1-1 $D_1 \times D_2$ 的结果	
D_1	D_2
0	a
0	b
0	c
1	a
1	b
1	c

表 1-3 关系 R_1	
D_1	D_2
张强	高等数学
张强	电子技术
李林	英语
李林	C 语言
王孝文	高等数学
王孝文	英语
王孝文	C 语言

表 1-2 关系 R	
D_1	D_2
张强	高等数学
张强	英语
张强	C 语言
张强	电子技术
李林	高等数学
李林	英语
李林	C 语言
李林	电子技术
王孝文	高等数学
王孝文	英语
王孝文	C 语言
王孝文	电子技术

笛卡儿积 $D_1 \times D_2$ 的结果集 R 本身是无意义的,而关系 R_1 是从 $D_1 \times D_2$ 中选取出来的有意义的子集,它表示了学生与课程之间存在的一种选修关系。所以说,关系是从笛卡儿积中选出的有意义的子集。

4. 属性

属性对应于关系中的列,也称为字段。关系中的属性名称必须是互不相同的。

5. 关键字

如果一个属性(集)的值能唯一标识一个关系的元组,则称该属性(集)为候选关键字(关键字),简称为键。

关键字有以下特点:

- (1)一个关系中可以有多个候选关键字。
- (2)在对关系进行插入、删除或检索的时候,可以选取其中一个候选关键字作为主关键字(简称关键字或主键)。每个关系都有一个并且只有一个主关键字。
- (3)凡可作为候选关键字的属性称为主属性,否则,称为非主属性。
- (4)当关系中的某个属性集并非主键,但却是另一个关系的主键时,则称该属性集为外部键,简称为外键。

6. 元组

元组对应于关系中的行,也称为记录。一个元组对应一个实体,一个关系可由一个或多个元组构成,一个关系中的元组必须互不相同。

7. 关系模式

一个关系的属性名表称为关系模式,一个关系模式描述了一个实体,是对关系的描述。它包括关系名、组成关系的属性名、属性间的数据依赖关系等。关系模式实际上就是关系框架,

即二维表的表结构。如,设关系名为 REL,其属性为 $A_1, A_2, \dots, A_n$,则关系模式可表示为:

$REL(A_1, A_2, \dots, A_n)$

8. 关系模型

关系模型是所有的关系模式、属性名和关键字的汇集,是模式描述的对象。一个关系模型描述了若干个实体及其相互联系,反映了客观世界的逻辑抽象。

9. 关系数据库

关系数据库是对应于一个关系模型的所有关系的集合。关系数据库可以用型和值去描述。关系数据库的型是指数据库的结构描述,它包括关系数据库名、若干属性的定义以及这些属性上的若干关系模式。关系数据库的值是指符合这些关系模式的多个关系在某一时刻各自所取的值。

10. 关系的性质

关系就是一个二维表,可以用二维表来理解关系的性质。

- (1)关系的每一列属性必须具有不同的名字。
- (2)关系的每一列属性是同一类型的域值,不同属性的域值可以相同。
- (3)关系的任意两行不能完全相同,即不可能出现两个完全相同的元组。
- (4)关系的每一分量都是不可再分的最小数据单位,即所有的属性值都是原子的。
- (5)关系中行的顺序、列的顺序可以任意互换,不会改变关系的意义。
- (6)每个关系都有一个主关键字唯一标识它的各个元组。

1.2.2 关系的完整性规则

1. 实体完整性规则

实体完整性是指关系中元组在组成主键的属性上不允许出现空值(NULL)。空值就是“不知道”或“无意义”。关系中的每一行都代表一个实体,而任何实体都应是可以区分的,主键的值正是区分实体的唯一标识。如果出现空值,那么主键值就起不到唯一标识元组的作用。

例如,关系 STUDENT(sno, sname, sex, birthday)中的主键是 sno,它在任何时候都不能取空值。

例如,学生选课关系 SC(sno, cno, grade),属性组学号 sno 和课程号 cno 构成了选课关系的主键,所以 sno 和 cno 这两个属性在任何时候都不能取空值。

实体完整性的意义在于,如果主键中的属性取空值,就说明存在某个不可标识实体,即存在不可区分的实体,这与关键字的意义相矛盾。

2. 参照完整性规则

所谓参照完整性规则是指表的外键必须是另一个表主键的有效值,或者是空值。如果外键存在一个值,则这个值必须是另一个表中主键的有效值,也就是说,外键可以没有值,即空值,但不允许是一个无效值。

例如,除了上述学生关系 STUDENT(sno, sname, sex, birthday),选课关系 SC(sno, cno, grade),还有课程关系 COURSE(cno, cname, credit),每个关系中有下画线的属性表示主键。可知,选课关系 SC 中的属性学号 sno 是一个外键,课程号 cno 也是一个外键,它们分别是关系 STUDENT 和 COURSE 的主键。所以,选课关系 SC 中的学号值必须是实际存在的学号,即学生关系 STUDENT 中有这个学生的记录;选课关系 SC 中的课程号的值也必须是确实存在的课程的课程号,即课程关系 COURSE 中有该课程的记录。也就是说,选课关

系 SC 中某些属性的取值需要参照其他关系的取值。

3. 用户定义完整性

任何关系数据库系统都应该支持实体完整性和参照完整性。除此之外,不同的关系数据库根据它的应用环境不同,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的,是用户按照实际数据库运行环境的要求对关系中的数据所定义的约束条件,反映的是某一具体应用所涉及的数据必须满足的条件。系统提供定义和检验这类完整性的机制,以使用统一的方法处理它们,而不再由应用程序承担这项工作。例如,将学生的性别定义为字符型数据,范围太大,我们可以写一个规则,把性别限制为男或女。



1.3 数据库设计

数据库的应用非常广泛,它是现代各种计算机信息系统的基础和核心,也是信息资源开发、管理和服务的最有效手段。因此,数据库的设计显得尤为重要,因为数据库中存储的信息能否正确地反映现实世界,以及在实际应用中能否及时、准确地为各应用程序提供所需要的数据,达到各种目标和处理要求,关键的问题在于数据库的设计和构造。

1.3.1 信息系统

信息系统是提供信息、辅助人们对环境进行控制和决策的系统,当今社会有 80% 的软件系统是信息系统,这些信息系统遍布在人们的生产和生活的各个环节,也极大地提高了人们的工作效率,加快了人们的生活节奏。

数据库是信息系统的一部分,它是信息系统的核心和基础。数据库把信息系统中大量的数据按一定的模型组织起来,提供数据存储、数据维护和数据检索等功能,使信息系统可以方便、及时和准确地从数据库中获得所需的信息。数据库是信息系统各个部分能够紧密结合的关键。只有对数据库进行合理的设计才能开发出高效而完美的信息系统。因此,数据库设计是信息系统开发和建设的核心技术。数据库及其相关信息处理功能是作为一个完整的信息系统开发项目的一个重要部分而被开发的。

1.3.2 数据库设计概述

1. 数据库设计的内容

数据库设计是指相对于一个给定的应用环境,提供一个确定最优数据模型与处理模式的逻辑设计,以及一个确定数据库存储结构与存储方法的物理设计,建立起能反映现实世界的信息和联系,满足用户的数据要求和加工要求,又能被某个数据库管理系统所接受,同时能实现系统目标,并有效存取数据的数据库。

数据库设计包括以下几个方面的内容:

1) 静态特性设计

静态特性设计又称结构特性设计,也就是根据给定的功能环境,设计数据库的数据模型或数据库模式,它包含数据库的概念结构设计和逻辑结构设计两个方面。

2) 动态特性设计

动态特性设计又称数据库行为特性设计,主要包括数据库查询、事务处理和报表处理等

应用程序设计。

3) 物理设计

根据动态特性,即应用处理要求,在选定的数据库管理系统环境下,把静态设计得到的数据库模式加以物理实现,即设计数据库的存储模式和存取方法。

显然,从使用方便和改善性能的角度考虑,结构特性必须适应行为特性。目前建立数据模型的方法并没有给行为特性的设计提供有效的工具和技巧,所以结构设计和行为设计不得不分离进行,但是它们又必须相互参照。

2. 数据库设计的特点

(1)数据库设计是一项综合性技术。“三分技术,七分管理,十二分基础数据”是数据库建设的基本规律。

(2)结构(数据)设计和行为(处理)设计相结合,数据库设计应该与应用系统相结合,在整个设计过程中把结构设计和行为设计密切结合起来。也就是说,整个设计过程中要把数据库结构设计和对数据的处理设计密切结合起来,是一种“反复探寻,逐步求精的过程”。首先从数据模型开始设计,以数据模型为核心进行展开,将数据库设计 and 应用设计相结合,建立一个完整、独立、共享、冗余小和安全有效的数据库系统。

设计与应用系统相结合,就要求设计者不仅仅具备计算机专业知识,诸如程序设计、数据库设计技术、软件工程、算法等,还要有相应应用对象的专业知识,如设计图书数据库需要图书管理方面的知识,设计人事档案数据库需要人事管理方面的知识。在这里可以要求用户协助,因为用户是数据库应用系统的提出者,也是最终的使用者,所以用户参与数据库设计的全部过程是满足用户要求的关键,设计者和用户合作的程度直接影响数据库设计的质量和进度。

数据库的静态结构设计与动态行为设计是分离进行的。静态结构设计侧重数据库的模式框架设计,而动态行为设计侧重应用程序设计。这就导致数据库应用系统的设计表现出分离设计、相互参照、反复探寻的特点。

除上述特点外,要求设计人员还要有战略眼光,要求其设计好的系统应该有生命力。因为事物是在不断发展变化的,设计好的系统不仅应能满足用户目前的需求,也应满足近期的需求,对远期需求也应有相应的处理方案,即设计人员应充分考虑到系统可能的扩充与改变,这样系统才有生命力。

3. 数据库设计的方法

数据库设计是一项工程技术,应有科学的理论和方法做指导,否则工程的质量难以保证,常常是数据库运行了一段时间后会不同程度地发生各种问题,增加了系统维护的代价。数据库设计有许多方法,主要分为四类:直观设计法、规范设计法、计算机辅助设计法和自动化设计法。

直观设计法与设计人员的技巧、经验和水平直接相关,但往往缺乏科学理论和工程原则的支持,很难保证设计质量。

为改变这种设计人员直观、手工试凑的状况,人们提出了运用软件工程的思想来设计数据库的方法,提出了各种设计准则和规程,这些都属于规范设计法。目前常用的规范设计法大多起源于新奥尔良方法。1978年10月,来自欧美国家的30多个主要数据库专家在美国的新奥尔良市讨论了数据库的设计问题,提出了相应的工作规范,并取名为新奥尔良方法。它将数据库设计分为四个阶段:需求分析、概念设计、逻辑设计和物理设计。

这些规范设计方法中基于 E-R 模型的数据库设计方法、基于 3NF 的数据库设计方法、基于抽象语法规则的设计方法等,是在数据库设计的不同阶段上支持实现的具体技术和方法。从本质上看,规范设计法仍然是手工设计方法,其基本思想是过程迭代,逐步求精。就目前的技术条件,这种按照一定设计规程,用工程化方法设计数据库是一种很实用的选择。

计算机辅助设计是指数据库设计的某些过程模拟某一规范设计方法,通过人机交互的方式实现设计中的某些部分,这要求设计者有一定的相关知识和经验。

用来帮助设计数据库或数据库应用软件的工具称为自动化设计工具,它可以自动并加速完成设计数据库系统的任务。用自动化设计工具(如 Oracle Designer,PowerDesigner 等)设计数据库的方法称为自动化设计法。。

总之,一个好的数据库设计方法应该能在合理的期限内,以合理的工作量产生一个有合理利用价值的数据库结构。

4. 数据库设计的基本步骤

成功的数据库设计最重要的是基于终端用户的要求。设计者的工作基于与终端用户的交互,但是支配设计结构的技术必须由设计者做出。就像如果你希望设计出成功的汽车,就必须花费大量的时间与该汽车的预期购买者和驾驶者进行交流,但是你不能希望驾驶员来确定活塞的点火顺序或发动机组的最佳铸造方法。

所以在做数据库设计时,要充分地和企业中的用户进行交流。了解企业现有系统的运行情况,了解当前系统不适用的地方(尤其要了解是否有数据方面的问题),了解他们的数据需求,了解他们的操作方式。在具体进行设计之前,要充分地了解企业现状,包括企业现有的工作流程、企业中出现的问题、企业想要利用新的系统实现的目标,了解这些后,才能开始数据库设计工作。

在数据库的设计过程中把数据库的设计和对数据库中数据处理的设计紧密结合起来,将这两个方面的需求分析、抽象、设计和实现在各个阶段同时进行,相互参照,相互补充,以完善两个方面的设计。事实上,如果不了解应用环境对数据的处理要求,或没有考虑如何去实现这些处理要求,是不可能设计出一个良好的数据库结构的。数据库设计的过程就是将现实世界的信息经过人为的选择、加工进入计算机存储处理又回到现实世界中去的过程。可以把它分为三大部分、六个阶段,如图 1-2 所示。

1) 数据库结构设计

(1)需求分析。在设计数据库时,必须首先准确了解与分析用户需求(包括数据与处理)。需求分析做得是否充分和准确,会影响整个数据库的设计,所以需求分析是整个设计过程的基础,其设计结果将直接影响后面各个阶段的设计,并影响设计结果的合理性。若需求分析做得不好,可能会导致整个数据库设计返工重做。所以,它是最困难、最耗费时间的一步。

(2)概念设计。概念设计是整个数据库设计的关键,它是对用户需求进行综合、归纳与抽象,形成一个独立于具体 DBMS 的概念模型。

E-R 模型是用来描述现实世界的概念模型。描述概念模型的方法有很多种,E-R 数据模型是最为著名也最为常用的数据模型,即由美籍华人陈平山(P. P. S. Chen)于 1976 年提出的实体-联系模型(entity-relationship model,E-R 模型)。用来表示 E-R 模型的图称为实体-联系图,简称 E-R 图,它主要用于描述概念世界。E-R 图提供了表示实体、属性和联系的方法。

E-R 图的三个要素如下。

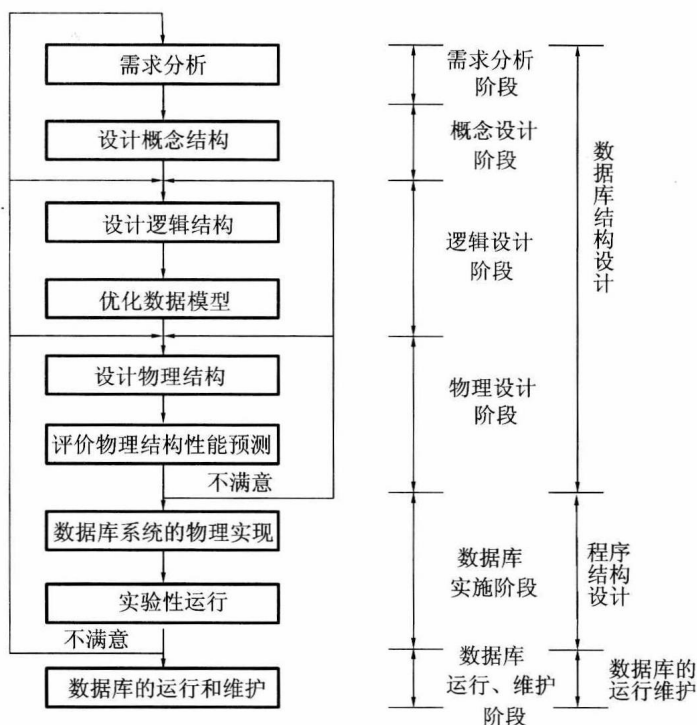


图 1-2 数据库设计步骤

- 实体:用矩形表示实体,矩形内标注实体名称。
- 属性:用椭圆表示属性,椭圆内标注属性名称,并用连线与实体连接起来。
- 实体之间的联系:用菱形表示,菱形内注明联系名称,并用连线将菱形分别与相关实体相连,同时在连线上注明联系类型($1:1$, $1:n$ 或 $m:n$)。

在画 E-R 图时,在实体与属性的连线上画线段,用此表示该属性是关键属性。学生的 E-R 图如图 1-3 所示。

若一个联系具有属性,则这个属性也要用连线与该联系连接起来。如图 1-4 所示,销售量是销售的属性。在后续图中,如果不加特殊说明,都表示此含义。

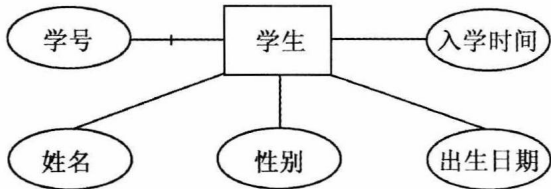


图 1-3 学生的 E-R 图

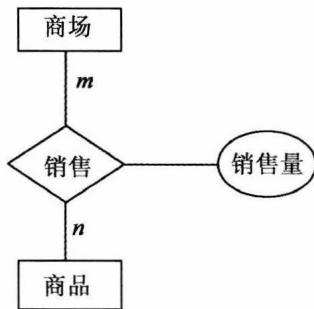


图 1-4 联系的属性

实体之间的联系如下。

① 一对一联系:如果对于实体集 A 中的每个实体,实体集 B 中至多有一个(可以没有)与之相对应,反之亦然,则称实体集 A 与实体集 B 具有一对一联系,记作: $1:1$ 。如班长和班级之间的联系就是一对一联系,如图 1-5(a)所示。

② 一对多联系:如果对于实体集 A 中的每个实体,实体集 B 中有 n ($n \geq 0$) 个实体与之