

高等院校计算机教育系列教材



钱慎一 编著

数据库基础与应用教程

- ◎ 面向应用，兼顾理论，概念准确，示例丰富
- ◎ 讲解明晰，提供大量上机练习
- ◎ 精心设计项目实例，易于理解掌握

清华大学出版社

高等院校计算机教育系列教材

Oracle 11g 数据库基础与应用教程

钱慎一 编著

清华大学出版社

北 京

内 容 简 介

Oracle 是数据库领域最优秀的数据库系统之一, 本书以 Oracle 11g 为蓝本, 系统地讲述了数据库的原理、Oracle 11g 的功能和应用。

全书共分 11 章, 在讲述数据库原理的基础知识和数据库结构化查询语言 SQL 的同时, 详细地介绍了 Oracle 11g 数据库的安装和卸载、PL/SQL 编程、基本操作、安全管理、存储管理、备份与恢复、闪回技术等, 最后通过实例阐述了基于 Java 开发包和 Oracle 数据库的开发过程, 此外, 还配有大量的图片和翔实的代码, 便于读者自行上机练习。

本书内容翔实、结构合理、示例丰富、语言简洁流畅, 主要用于培养数据库管理人员和数据库开发人员, 适合作为高等院校本、专科计算机软件、信息系统、电子商务等相关专业的数据库课程教材, 同时也适合作为各种数据库技术培训班的教材以及数据库开发人员的参考资料。

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 13701121933

图书在版编目(CIP)数据

Oracle 11g 数据库基础与应用教程/钱慎一编著. —北京: 清华大学出版社, 2011.6

(高等院校计算机教育系列教材)

ISBN 978-7-302-25628-1

I. ①O… II. ①钱… III. ①关系数据库—数据库管理系统, Oracle 11g—高等学校—教材 IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2011)第 087954 号

责任编辑: 张 瑜 宋延清

装帧设计: 杨玉兰

责任校对: 周剑云

责任印制: 何 芊

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

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

印 装 者: 北京国马印刷厂

经 销: 全国新华书店

开 本: 185×260 印 张: 22.5 字 数: 542 千字

版 次: 2011 年 6 月第 1 版 印 次: 2011 年 6 月第 1 次印刷

印 数: 1~4000

定 价: 38.00 元

产品编号: 040131-01

前 言

信息技术的飞速发展大大推动了社会的进步，也逐渐改变了人们的生活、工作和学习方式。数据库技术现已成为计算机科学技术中发展最快的领域之一，也是应用最广的技术之一。数据库技术和网络技术是信息技术中的重要支柱。当今各种热门的信息系统，例如管理信息系统、企业资源计划、供应链管理系统、客户关系管理系统、电子商务系统、决策支持系统、智能信息系统等，都离不开数据库技术强有力的支持。

Oracle 数据库系统是数据库领域最优秀的数据库之一，随着版本的不断升级，功能越来越强大。最新版本的 Oracle 11g 可以为各类用户提供完整的数据库解决方案，可以帮助用户建立自己的电子商务体系，从而增强了用户对外界变化的敏捷反应能力，提高了用户的市场竞争力。

目前，我国应用型和工程型人才短缺，原有的教材已经不适合现在的要求。紧缺型人才的培养注重的是实践能力，应该在学校就开始培养学生的动手能力，使学生在毕业后能直接上岗工作，所以教材的改革就显得尤为重要。本书从实际应用角度出发，系统地介绍了数据库和 Oracle 的相关概念和原理、Oracle 的数据库管理与操作和 Oracle 的应用开发基础，并通过案例来介绍基于 Java 开发包和 Oracle 数据库进行案例开发的详细过程。全书共分 11 章，各章内容介绍如下。

第 1 章 讲述数据库的基本概念、原理和数据库设计的方法与步骤。

第 2 章 讲述 Oracle 的发展历史、产品版本、Oracle 11g 的体系结构和新特性。

第 3 章 讲述 Oracle 在 Windows 上的安装、卸载和配置。

第 4~5 章 讲述 SQL 语言基础与 Oracle PL/SQL 语言及编程技术。

第 6~10 章 讲述 Oracle 的基本操作及其数据库的管理应用操作、安全管理、存储管理、备份与恢复、闪回技术等。

第 11 章 讲述基于 Oracle 数据库和 Java 开发包的综合应用实例。

本书相关源代码可以从 www.wenyuan.com.cn 下载。

本书由资深 Oracle 专家钱慎一老师编著。另外，闫红岩、张保威、金松河、张旭、贺蕾、张阳、王国胜、伏银恋、徐明华、尼春雨、张丽等也参与了编写工作。在本书的编写和出版过程中得到了郑州轻工业学院教务处的大力支持和帮助，在此由衷地向他们表示感谢！本书除了可用作高等院校本专科学生的教材外，也兼顾了普通读者，可供从事计算机应用开发的人员在学习数据库技术时参考。

由于编写工作繁忙，书中难免会有疏漏之处，恳请广大读者给予批评指正。

编 者

目 录

第 1 章 数据库技术基础.....1	2.3 Oracle 11g 体系结构概述.....34
1.1 数据库的基本概念.....1	2.3.1 存储结构.....35
1.1.1 数据管理的发展.....1	2.3.2 内存结构.....40
1.1.2 数据库与数据库管理系统.....6	2.3.3 进程结构.....42
1.1.3 数据库系统.....7	2.3.4 数据字典.....44
1.2 数据模型.....7	2.4 Oracle 11g 新特性.....45
1.2.1 E-R 模型.....8	2.4.1 新特性的作用.....45
1.2.2 关系模型.....11	2.4.2 数据库管理方面新特性.....46
1.3 数据库系统结构.....16	2.4.3 PL/SQL 方面新特性.....47
1.3.1 数据库的三级模式结构.....16	2.4.4 其他主要特性.....48
1.3.2 数据库的体系结构.....18	本章小结.....49
1.3.3 数据库的连接.....21	习题.....49
1.4 数据库的规范化.....22	第 3 章 Oracle 11g 的安装、卸载 与配置.....51
1.4.1 数据依赖.....22	3.1 安装前的准备工作.....51
1.4.2 相关概念.....22	3.2 安装 Oracle 11g 数据库.....52
1.4.3 范式.....23	3.2.1 Oracle 通用安装器 OUI.....52
1.5 数据库设计.....25	3.2.2 Oracle 数据库的安装过程.....53
1.5.1 需求分析.....26	3.3 卸载 Oracle 数据库.....59
1.5.2 概念结构设计.....27	3.3.1 停止所有的 Oracle 服务.....59
1.5.3 逻辑结构设计.....27	3.3.2 用 OUI 卸载所有的 Oracle 组件.....60
1.5.4 数据库物理设计.....28	3.3.3 手动删除 Oracle 遗留的成分.....62
1.5.5 数据库的实施.....28	3.4 Oracle 网络与防火墙.....65
1.5.6 数据库的运行和维护.....29	3.4.1 Oracle 网络服务.....65
本章小结.....30	3.4.2 配置 Oracle 网络环境.....65
习题.....30	3.4.3 设置 Oracle 防火墙.....70
第 2 章 初识 Oracle 数据库.....31	3.5 本章小结.....70
2.1 Oracle 的发展历史.....31	习题.....70
2.2 Oracle 11g 版本介绍.....32	第 4 章 SQL 语言基础.....72
2.2.1 企业版.....32	4.1 Oracle 示例数据库.....72
2.2.2 标准版.....33	4.1.1 示例数据库概述.....72
2.2.3 标准版 1.....33	
2.2.4 个人版.....34	



4.1.2	HR 示例方案.....	74	5.2.1	游标的概念.....	130
4.1.3	HR 示例方案中的表结构.....	75	5.2.2	显式游标.....	130
4.2	SQL 语言简介.....	78	5.2.3	隐式游标.....	134
4.2.1	发展历史.....	78	5.2.4	游标的属性.....	134
4.2.2	语言特点.....	79	5.2.5	游标变量.....	136
4.2.3	功能.....	81	5.3	过程.....	138
4.3	数据定义.....	81	5.3.1	过程的创建.....	138
4.3.1	创建操作.....	82	5.3.2	过程的调用.....	139
4.3.2	删除操作.....	84	5.3.3	过程的删除.....	139
4.3.3	修改操作.....	85	5.3.4	参数类型及传递.....	140
4.4	数据查询.....	86	5.4	函数.....	141
4.4.1	简单查询.....	87	5.4.1	函数的创建.....	141
4.4.2	WHERE 子句.....	90	5.4.2	函数的调用.....	142
4.4.3	ORDER BY 子句.....	93	5.4.3	函数的删除.....	143
4.4.4	GROUP BY 子句.....	94	5.5	包.....	143
4.4.5	HAVING 子句.....	96	5.5.1	基本原理.....	143
4.4.6	多表连接查询.....	97	5.5.2	包的创建.....	143
4.4.7	集合操作.....	103	5.5.3	包的调用.....	145
4.4.8	子查询.....	106	5.5.4	删除包.....	145
4.5	数据操纵.....	108	5.6	触发器.....	145
4.5.1	数据插入.....	108	5.6.1	基本原理.....	145
4.5.2	数据修改.....	110	5.6.2	触发器的创建.....	147
4.5.3	数据删除.....	111	5.6.3	触发器的执行.....	147
4.6	数据控制.....	113	5.6.4	触发器的删除.....	148
4.6.1	授权语句.....	113	5.7	同义词.....	148
4.6.2	授权收回语句.....	115	5.7.1	同义词的创建.....	148
本章小结.....		116	5.7.2	同义词的使用.....	149
习题.....		116	5.7.3	同义词的删除.....	149
5.7.4	同义词的查看.....	149	5.8	序列.....	149
第 5 章 Oracle PL/SQL 语言及编程.....	118		5.8.1	序列的创建.....	150
5.1	简介.....	118	5.8.2	序列的使用.....	150
5.1.1	程序结构.....	118	5.8.3	序列的修改.....	151
5.1.2	注释.....	119	5.8.4	序列的删除.....	151
5.1.3	字符集与分隔符.....	120	5.8.5	序列的查看.....	151
5.1.4	数据类型.....	121	本章小结.....		151
5.1.5	变量和常量.....	123	习题.....		152
5.1.6	结构控制语句.....	124	第 6 章 Oracle 操作基础.....	153	
5.1.7	表达式.....	128	6.1	启动和关闭 Oracle.....	153
5.2	游标.....	129			

6.1.1 Oracle 数据库的启动.....	153	7.5.1 概要文件中的参数.....	217
6.1.2 Oracle 数据库的关闭.....	164	7.5.2 概要文件中的管理.....	217
6.2 表.....	168	7.6 数据库审计.....	219
6.2.1 设计表.....	168	7.7 使用 OEM 进行安全管理.....	220
6.2.2 创建表.....	171	本章小结.....	228
6.2.3 修改表.....	176	习题.....	228
6.3 视图.....	178	第 8 章 数据库存储管理.....	229
6.3.1 创建视图.....	178	8.1 数据文件.....	229
6.3.2 修改视图.....	182	8.1.1 数据文件概述.....	229
6.3.3 删除视图.....	184	8.1.2 数据文件的管理.....	229
6.4 索引.....	185	8.2 控制文件.....	232
6.4.1 创建索引.....	185	8.2.1 控制文件概述.....	233
6.4.2 删除索引.....	187	8.2.2 控制文件的管理.....	233
6.5 数据查询及操纵.....	189	8.3 重做日志文件.....	235
6.5.1 数据查询.....	189	8.3.1 重做日志文件概述.....	235
6.5.2 批量插入记录.....	191	8.3.2 重做日志文件的管理.....	236
6.5.3 通过视图操纵数据.....	192	8.4 归档重做日志文件.....	237
本章小结.....	194	8.4.1 归档重做日志文件概述.....	237
习题.....	194	8.4.2 归档重做日志文件的管理.....	237
第 7 章 数据库安全管理.....	196	8.5 表空间.....	240
7.1 数据库安全性概述.....	196	8.5.1 表空间概述.....	240
7.2 用户管理.....	196	8.5.2 创建表空间.....	242
7.2.1 创建用户.....	198	8.5.3 修改表空间.....	245
7.2.2 修改用户.....	199	8.5.4 删除表空间.....	247
7.2.3 删除用户.....	200	8.5.5 表空间信息的查询.....	248
7.2.4 查询用户信息.....	200	本章小结.....	249
7.3 权限管理.....	201	习题.....	249
7.3.1 授予权限.....	201	第 9 章 数据库的备份与恢复.....	251
7.3.2 回收权限.....	208	9.1 备份与恢复概述.....	251
7.4 角色管理.....	210	9.2 逻辑备份与恢复.....	252
7.4.1 创建角色.....	211	9.2.1 使用 expdp 导出数据.....	253
7.4.2 角色权限的授予与回收.....	212	9.2.2 使用 impdp 导入数据.....	254
7.4.3 修改角色.....	213	9.2.3 使用 OEM 进行逻辑备份 与恢复.....	256
7.4.4 角色的生效与失效.....	213	9.3 脱机备份与恢复.....	272
7.4.5 删除角色.....	214	9.3.1 脱机备份.....	272
7.4.6 使用角色进行权限管理.....	214	9.3.2 脱机恢复.....	274
7.4.7 查询角色信息.....	216		
7.5 概要文件管理.....	217		

9.4 联机备份与恢复	274	11.1.4 需求规定	293
9.4.1 归档日志模式的设置	274	11.1.5 数据流图	293
9.4.2 创建恢复目录所用的表 空间	275	11.2 系统结构设计	295
9.4.3 创建 rman 用户并授权	276	11.2.1 系统结构	296
9.4.4 创建恢复目录	276	11.2.2 系统角色和业务流程 分析	296
9.4.5 注册目标数据库	277	11.3 数据库设计	296
9.4.6 使用 RMAN 程序进行 备份	277	11.3.1 数据库设计概述	297
9.4.7 使用 RMAN 程序进行 恢复	279	11.3.2 数据库概念和结构设计	297
9.5 各种备份与恢复方法的比较	280	11.3.3 数据库逻辑结构设计	301
本章小结	281	11.3.4 数据库实施	304
习题	281	11.4 系统功能实现	306
第 10 章 闪回技术	282	11.4.1 图书信息维护模块的主要 界面	306
10.1 闪回技术概述	282	11.4.2 读者信息维护模块的主要 界面	308
10.2 闪回查询技术	282	11.4.3 管理员信息维护模块的主要 界面	310
10.2.1 闪回查询	283	11.4.4 系统维护模块的主要界面	311
10.2.2 闪回版本查询	283	11.4.5 馆藏检索模块的主要界面	312
10.2.3 闪回事务查询	285	11.4.6 读者借阅信息查询的主要 界面	313
10.3 闪回错误操作技术	286	11.4.7 读者借/还书模块的主要 界面	315
10.3.1 闪回数据库	286	11.4.8 系统外观设定	316
10.3.2 闪回表	287	11.4.9 部分 Java 类的设计 与实现	316
10.3.3 闪回回收站	288	11.5 系统开发运行环境	348
本章小结	289	11.5.1 运行环境	348
习题	289	11.5.2 系统的运行	349
第 11 章 图书管理系统	291	11.6 本章小结	350
11.1 系统概述	291	习题	350
11.1.1 开发背景	291		
11.1.2 功能介绍	291		
11.1.3 需求描述	292		

第 1 章 数据库技术基础

数据库技术已成为计算机科学的一个重要分支，是数据管理的最新技术，是计算机科学技术中发展最快的领域之一。许多信息系统都是以数据库为基础建立的，数据库已成为计算机信息系统的核心技术和重要基础，成为人们储存数据、管理信息、共享资源的最先进、最常用的技术。

本章将介绍数据库系统的基本概念、数据管理技术的发展过程、数据模型、数据库系统设计、数据库应用系统结构和数据库的规范化理论等，最后还将阐述高级数据库技术的相关知识。读者从中可以学习到为什么要使用数据库技术以及明确数据库技术的重要性。本章是学习后面各章节的预备和基础。

1.1 数据库的基本概念

数据库技术是计算机技术中发展最为迅速的领域之一，已成为人们存储数据、管理信息和共享资源的最常用、最先进的技术。数据库技术已经在科学、技术、经济、文化和军事等各个领域发挥着重要的作用。

1.1.1 数据管理的发展

自计算机产生以来，人类社会进入了信息时代，对数据处理速度及规模的需求远远超出了过去人工或机械方式的能力范围，计算机以其快速准确的计算能力和海量的数据存储能力在数据处理领域得到了广泛的应用。随着数据处理的工作量呈几何方式的不断增加，数据管理技术应运而生，其演变过程随着计算机硬件或软件的发展速度以及计算机应用领域的不断拓宽而不断变化。总地来说，数据管理的发展经历了人工管理、文件管理和数据库管理 3 个阶段。

1. 人工管理阶段

在计算机没有应用到数据管理领域之前，数据管理的工作是由人工完成的。这种数据处理经历了很长一段时间。

20 世纪 50 年代中期以前，计算机主要用于科学计算。当时外存的状况是只有纸带、卡片、磁带等设备，并没有磁盘等直接存取的存储设备；而计算机系统软件的状况是没有操作系统，没有管理数据的软件，在这种情况下的数据管理方式为人工管理数据。人工管理数据具有如下特点。

(1) 数据不被保存

由于当时计算机主要用于科学计算，一般不需要将数据长期保存，只是在计算某一课

题时将数据输入，用完就撤走。

(2) 应用程序管理数据

数据需要由应用程序自己管理，没有相应的软件系统负责数据的管理工作。应用程序中不仅要规定数据的逻辑结构，而且要设计物理结构，包括存储结构、存取方法、输入方式等，因此程序员负担很重。

(3) 数据不能共享

数据是面向应用的，一组数据只能对应一个程序。当多个应用程序涉及某些相同的数据时，由于必须各自定义，无法互相利用、互相参照，因此程序与程序之间有大量的冗余数据，如图 1.1 所示。

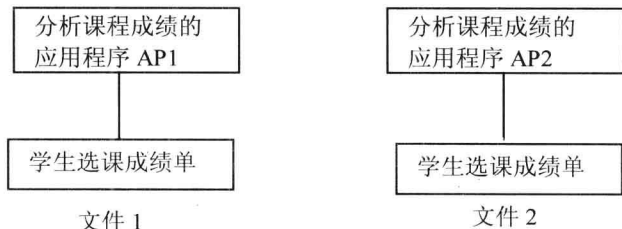


图 1.1 两个应用程序各自使用同一数据

(4) 数据不具有独立性

数据的逻辑结构或物理结构改变后，必须对应用程序做相应的修改，这就进一步加重了程序员的负担。

在人工管理阶段，程序与数据之间的一一对应关系如图 1.2 所示。

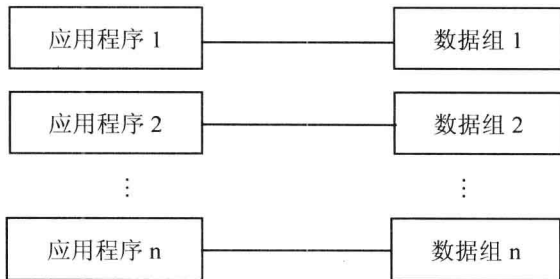


图 1.2 人工管理阶段应用程序与数据之间的对应关系

2. 文件系统阶段

到了 20 世纪 50 年代后期，及 60 年代中期，这时已有了磁盘、磁鼓等直接存储设备；而在计算机系统方面，不同类型的操作系统的出现极大地增强了计算机系统的功能。操作系统中用来进行数据管理的部分是文件系统。这时可以把相关的数据组织成一个文件存放在计算机中，在需要的时候只要提供文件名，计算机就能从文件系统中找出所要的文件，把文件中存储的数据提供给用户进行处理。但是，由于这时数据的组织仍然是面向程序，所以存在大量的数据冗余，无法有效地进行数据共享。

(1) 文件系统管理数据具有如下优点。

- 数据可以长期保存：数据可以组织成文件长期保存在计算机中反复使用。
- 由文件系统管理数据：文件系统把数据组织成内部有结构的记录，实现“按文件名访问，按记录进行存取”的管理技术。

文件系统使应用程序与数据之间有了初步的独立性，程序员不必过多地考虑数据存储的物理细节。例如，文件系统中可以有顺序结构文件、索引结构文件、Hash 等。数据在存储上的不同不会影响程序的处理逻辑。如果数据的存储结构发生改变，应用程序的改变很小，节省了程序的维护工作量。

(2) 但是，文件系统仍存在以下缺点。

- 数据共享性差，冗余度大：在文件系统中，一个(或一组)文件基本上对应于一个应用(程序)，即文件是面向应用的。当不同的应用(程序)使用部分相同的数据时，也必须建立各自的文件，而不能共享相同的数据。因此数据的冗余度大，浪费存储空间。同时由于相同数据的重复存储、各自管理，容易造成数据的不一致性，给数据的修改和维护带来了困难。
- 数据独立性差：文件系统文件是为某一特定应用服务的，文件的逻辑结构对该应用来说是优化的，因此要相对现有的数据再增加一些新的应用会很困难，系统不容易扩充。一旦数据的逻辑结构发生改变，就必须修改应用程序，修改文件结构的定义。因此数据与程序之间仍缺乏独立性。文件系统阶段程序与数据之间的关系如图 1.3 所示。

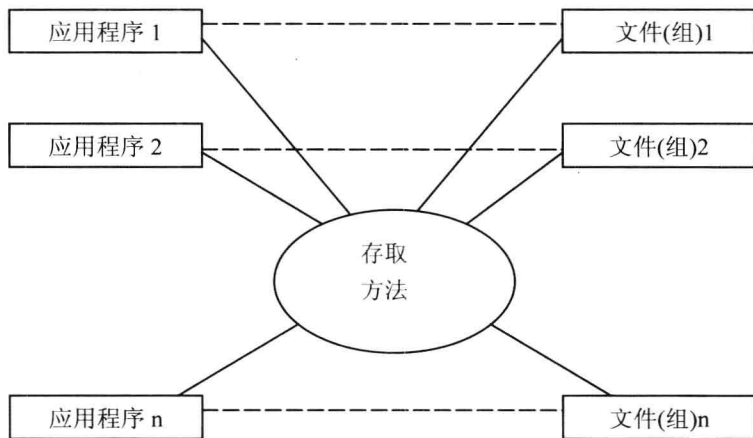


图 1.3 文件系统阶段应用程序与数据之间的对应关系

3. 数据库系统阶段

20 世纪 60 年代后期，计算机用于管理的规模越来越大，应用越来越广泛，数据量急剧增长，同时多种应用、多种语言互相覆盖的共享数据集合的要求也越来越强烈。

这时已有大容量磁盘，硬件价格下降，软件价格则上升，为编制和维护系统软件及应用程序所需的成本相对增加。在这种背景下，以文件系统作为数据管理手段已经不能满足应用的需求，于是为解决多用户、多应用共享数据的需求，使数据为尽可能多的应用服务，

数据库技术便应运而生，出现了统一管理数据的专用软件系统——数据库管理系统。

用数据库系统来管理数据比文件系统具有明显的优点，从文件系统到数据库系统，标志着数据管理技术的飞跃。

在数据库系统阶段，应用程序与数据之间的对应关系如图 1.4 所示。

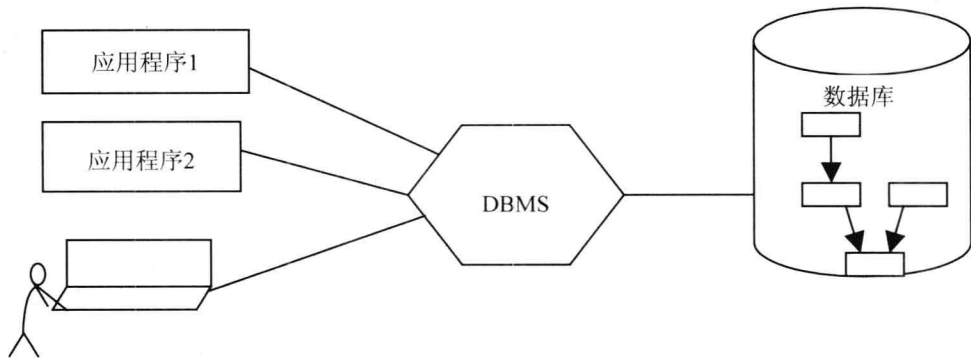


图 1.4 数据库系统阶段应用程序与数据之间的对应关系

由于数据库是以数据为中心组织数据、减少了数据的冗余，可提供更高的数据共享能力，同时要求程序和数据具有较高的独立性，当数据的逻辑结构改变时，不涉及数据的物理结构，也不影响应用程序，这样就降低了应用程序研制与维护的费用。

随着计算机应用的进一步发展和网络的出现，有人提出数据管理的高级数据库阶段，这一段的主要标志是 20 世纪 80 年代的分布式数据库系统、90 年代的对象数据库系统和 21 世纪初的网络数据库系统的出现。

(1) 分布式数据库系统

在这一阶段以前的数据库系统是集中式的。在文件系统阶段，数据分散在各个文件中，文件之间缺乏联系。集中式数据库把数据库集中在一个数据库中进行管理，减少了数据冗余的不一致性，而且数据联系比文件系统强得多。但集中式系统也有弱点：一是随着数据量增加，系统相当庞大，操作复杂，开销大；二是数据集中存储，大量的通信都要通过主机，造成拥挤现象。随着小型计算机和微型计算机的普及，以及计算机网络软件和远程通信的发展，分布式数据库系统崛起了。

分布式数据库系统主要有以下 3 个特点：

- 数据库的数据物理上分布在各个场地，但逻辑上是一个整体。
- 各个场地既可执行局部应用(访问本地 DB)，又可执行全局应用(访问异地 DB)。
- 各地的计算机由数据通信网络相联系。本地计算机单独不能胜任的处理任务，可以通过通信网络取得其他 DB 和计算机的支持。

分布式数据库系统兼顾了集中管理和分布处理两个方面，因而有良好的性能。

(2) 对象数据库系统

在数据处理领域，关系数据库的使用已相当普遍、相当出色。但是现实世界存在着许多具有更复杂数据结构的实际应用领域，已有的层次、网状、关系三种数据模型对这些应用领域都显得力不从心。例如多媒体数据、多维表格数据、CAD 数据等应用问题，需要更高级的数据库技术来表达，以便于管理、构造与维护大容量的持久数据，并使它们能与大

型复杂程序紧密结合。对象数据库正是适应这种形势发展起来的，它是面向对象的程序设计技术与数据技术结合的产物。

对象数据库系统主要有以下两个特点：

- 对象数据库模型能完整地描述现实世界的数据结构，能表达数据间嵌套、递归的联系。
- 具有面向对象技术的封装性(把数据与操作定义在一起)和继承性(继承数据结构和操作)的特点，提高了软件的可重用性。

(3) 网络数据库系统

随着 C/S(客户机/服务器)结构的出现，人们可以最有效地使用计算机资源。但在网络环境中，如果要隐藏各种复杂性，就要使用中间件。中间件是网络环境中保证不同的操作系统、通信协议和 DBMS 之间进行对话、互操作的软件系统。其中涉及数据访问的中间件，就是 20 世纪 90 年代提出的 ODBC 技术和 JDBC 技术。

现在，计算机网络已成为信息化社会中十分重要的一类基础设施。随着广域网(WAN)的发展，信息高速公路已发展成为采用通信手段将地理位置分散的、各自具备自主功能的若干台计算机和数据库系统有机地连接起来，组成因特网(Internet)，用于实现通信交往、资源共享或协调工作等目标。这个目标在 20 世纪末已经实现，正在对社会的发展起着极大的推进作用。

上述三个阶段可以做一比较，如表 1.1 所示。

表 1.1 数据管理三个阶段的比较

		人工管理阶段	文件系统阶段	数据库系统阶段
背景	应用背景	科学计算	科学计算、管理	大规模管理
	硬件背景	无直接存取设备	磁盘、磁鼓	大容量磁盘
	软件背景	没有操作系统	有文件系统	有数据库管理系统
	处理方式	批处理	联机实时处理、批处理	联机实时处理、分布处理、批处理
特点	数据库的管理者	用户(程序员)	文件系统	数据库管理系统
	数据的共享程度	某一应用程序	某一应用	现实世界
	数据面向的对象	无共享，冗余度极大	共享性差，冗余度大	共享性高，冗余度小
	数据的独立性	不独立，完全依赖于程序	独立性差	具有高度的物理独立性和一定的逻辑独立性
	数据的结构化	无结构	记录内有结构、整体无结构	整体结构化，用数据模型描述
	数据控制能力	应用程序自己控制	应用程序自己控制	由数据库管理系统提供数据安全性、完整性、并发控制和恢复能力

1.1.2 数据库与数据库管理系统

数据库与数据库管理系统是密切相关的两个基本概念，我们先可以简单地理解为：数据库是指存放数据的文件，而数据库管理系统是用来管理和控制数据库文件的组织、存储以及如何访问数据库中数据的专门工具。事实上，数据库管理系统是计算机系统中有着非常重要地位的系统软件。

1. 数据库

数据库(Database)，顾名思义，就是存放数据的仓库。只不过这个仓库是在计算机的存储设备上，而且数据是按照一定的数据模型组织并存放在外存上的一组相关数据集合，通常这些数据是面向一个组织、企业或部门的。例如学生成绩管理系统中，学生的基本信息、课程信息、成绩信息等都是从学生成绩管理数据库的。

除了用户可以直接使用的数据外，还有另外一种数据，它们是有关数据库的定义信息的，如数据库的名称、表的定义、数据库用户名及密码、权限等。这些数据用户不会经常性地使用，但是对数据库来说非常重要。这些数据通常存放在“数据字典(Data Dictionary)”中。数据字典是数据库管理系统中非常重要的组成部分之一，它是由数据库管理系统自动生成并维护的一组表和视图。数据字典是数据库管理系统工作的依据，数据库管理系统借助于数据字典来理解数据库中数据的组织，并完成对数据库中数据的管理与维护。数据库用户可通过数据字典获取有用的信息，如用户创建了哪些数据库对象，这些对象是如何定义的，这些对象允许哪些用户使用等。但是，数据库用户是不能随便改动数据字典中的内容的。

人们收集并抽取出一个应用所需要的大量数据之后，应将其保存起来供进一步查询和加工处理，以获得更多有用的信息。过去人们把数据存放在文件柜里，当数据越来越多时从大量的文件中查找数据显得十分困难。现在人们借助计算机和数据库科学地保存和管理大量复杂的数据，能方便而充分地利用这些宝贵的信息资源。

严格地讲，数据库是长期存储在计算机内、有组织的、大量的、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，并可为各种用户共享。简单地讲，数据库数据具有永久存储、有组织和可共享这3个基本特点。

2. 数据库管理系统

在建立了数据库之后，下一个问题就是如何科学地组织和存储数据，如何高效地获取和维护数据，完成这个任务的机制是一个系统软件——数据库管理系统 DBMS(Database Management System)。

DBMS 是指数据库系统对数据进行管理的软件系统，它是数据库系统的核心组成部分，数据库系统的一切操作，包括查询、更新及各种控制，都是通过 DBMS 进行的。DBMS 总是基于数据模型，因此可以把它看成是某种数据模型在计算机系统上的具体实现。根据所采用数据模型的不同，DBMS 可以分成网状型、层次型、关系型、面向对象型等。但在

不同的计算机系统中，由于缺乏统一的标准，即使是同种数据模型的 DBMS，它们在用户接口、系统功能等方面也常常是不同的。

如果用户要对数据库进行操作，是由 DBMS 把操作从应用程序带到外部级和概念级，再导向内部级，进而操纵存储器中的数据。一个 DBMS 的主要目标是使数据作为一种可管理的资源来处理。DBMS 应使数据易于为各种不同的用户所共享，应该增进数据的安全性、完整性及可用性，并提供高度的数据独立性。

1.1.3 数据库系统

数据库系统是指在计算机系统中引入数据库后的系统，一般由数据库、数据库管理系统(及其开发工具)、应用系统和数据库管理员构成。应当指出的是，数据库的建立、使用和维护等工作只靠一个 DBMS 是远远不够的，还要有专门的人员来完成，这些人被称为数据库管理员 DBA(Database Administrator)。

在一般不引起混淆的情况下，人们常常把数据库系统简称为数据库。数据库系统组成如图 1.5 所示。数据库系统在计算机系统中的地位如图 1.6 所示。

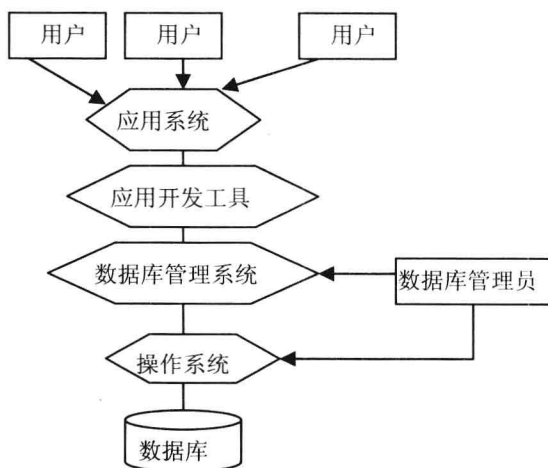


图 1.5 数据库系统

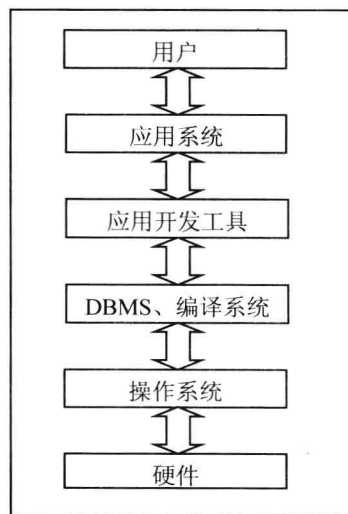


图 1.6 数据库在计算机系统中的地位

1.2 数据模型

模型是现实世界特征的模拟与抽象。比如一组建筑规划沙盘、精致逼真的飞机航模，都是对现实生活中的事物的描述和抽象，见到它就会让人们联想到现实世界中的实物。

数据模型(Data Model)也是一种模型，它是现实世界数据特征的抽象。由于计算机不可能直接处理现实世界中的具体事物，因此人们必须事先把具体事物转换成计算机能够处理的数据，即首先要数字化，要把现实世界中的人、事、物和概念用数据模型这个工具来抽

象、表示和加工处理。数据模型是数据库中用来对现实世界进行抽象的工具，是数据库中用于提供信息表示和操作手段的形式构架，是现实世界的一种抽象模型。

数据模型按不同的应用层次分为 3 种类型，分别是概念数据模型(Conceptual Data Model)、逻辑数据模型(Logic Data Model)和物理数据模型(Physical Data Model)。

(1) 概念数据模型又称概念模型，是一种面向客观世界、面向用户的模型，与具体的数据库管理系统无关，与具体的计算机平台无关。

人们通常先将现实世界中的事物抽象到信息世界，建立所谓的“概念模型”，然后再将信息世界的模型映射到机器世界，将概念模型转换为计算机世界中的模型。因此，概念模型是从现实世界到机器世界的一个中间层次。

(2) 逻辑数据模型又称逻辑模型，是一种面向数据库系统的模型，它是概念模型到计算机之间的中间层次。概念模型只有在转换成逻辑模型之后才能在数据库中得以表示。目前，逻辑模型的种类很多，其中比较成熟的有层次模型、网状模型、关系模型、面向对象模型等。

这几种数据模型的根区别在于数据结构不同，即数据之间联系的表示方式不同：

- 层次模型用“树结构”来表示数据之间的联系。
- 网状模型是用“图结构”来表示数据之间的联系。
- 关系模型是用“二维表”来表示数据之间的联系。
- 面向对象模型是用“对象”来表示数据之间的联系。

(3) 物理数据模型又称物理模型，它是一种面向计算机物理表示的模型，此模型是数据模型在计算机上的物理结构表示。

数据模型通常由三部分组成，分别是数据结构、数据操纵和完整性约束，也称为数据模型的三大要素。

1.2.1 E-R 模型

概念模型中最著名的是实体联系模型(Entity Relationship Model, E-R 模型)。E-R 模型是 P.P.Chen 于 1976 年提出的。这个模型直接从现实世界中抽象出实体类型及实体间联系，然后用实体联系图(E-R 图)表示数据模型。设计 E-R 图的方法称为 E-R 方法。E-R 图是设计概念模型的有力工具。下面先介绍一下有关的名词术语及 E-R 图。

1. 实体

现实世界中客观存在并可相互区分的事物叫做实体。实体可以是一个具体的人或物，如王伟、汽车等；也可以是抽象的事件或概念，如购买一本图书等。

2. 属性

实体的某一特性称为属性。如学生实体有学号、姓名、年龄、性别、系等方面的属性。属性有“型”和“值”之分，“型”即为属性名，如姓名、年龄、性别是属性的型；“值”即为属性的具体内容，如(990001, 张立, 20, 男, 计算机)这些属性值的集合表示了一个学生实体。

3. 实体型

若干个属性型组成的集合可以表示一个实体的类型,简称实体型。如学生(学号,姓名,年龄,性别,系)就是一个实体型。

4. 实体集

同型实体的集合称为实体集。如所有的学生、所有的课程等。

5. 码

能唯一标识一个实体的属性或属性集称为实体的码。如学生的学号,学生的姓名可能有重名,不能作为学生实体的码。

6. 域

属性值的取值范围称为该属性的域。如学号的域为 6 位整数,姓名的域为字符串集合,年龄的域为小于 40 的整数,性别的域为(男,女)。

7. 联系

在现实世界中,事物内部以及事物之间是有联系的,这些联系同样也要抽象和反映到信息世界中来。在信息世界中将被抽象为实体型内部的联系和实体型之间的联系。

实体内部的联系通常是指组成实体的各属性之间的联系;实体之间的联系通常是指不同实体集之间的联系。

两个实体型之间的联系有如下 3 种类型。

(1) 一对一联系(1:1)

实体集 A 中的一个实体至多与实体集 B 中的一个实体相对应,反之亦然,则称实体集 A 与实体集 B 为一对一的联系。记作 1:1。如班级与班长、观众与座位、病人与床位。

(2) 一对多联系(1:n)

实体集 A 中的一个实体与实体集 B 中的多个实体相对应,反之,实体集 B 中的一个实体至多与实体集 A 中的一个实体相对应。记作 1:n。如班级与学生、公司与职员、省与市。

(3) 多对多(m:n)

实体集 A 中的一个实体与实体集 B 中的多个实体相对应,反之,实体集 B 中的一个实体与实体集 A 中的多个实体相对应。记作(m:n)。如教师与学生、学生与课程、工厂与产品。

实际上,一对一联系是一对多联系的特例,而一对多联系又是多对多联系的特例。可以用图形来表示两个实体型之间的这三类联系,如图 1.7 所示。

在 E-R 图中有下面 4 个基本成分:

- 矩形框,表示实体类型(研究问题的对象)。
- 菱形框,表示联系类型(实体间的联系)。
- 椭圆形框,表示实体类型和联系类型的属性。
- 直线,联系类型与其涉及的实体类型之间以直线连接,用来表示它们之间的联系,并在直线端部标注联系的种类(1:1、1:N 或 M:N)。