



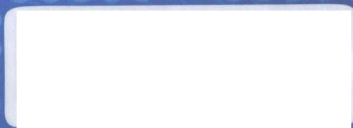
21 世纪高等学校计算机公共课程规划教材
“十一·五”浙江省重点教材建设项目

数据库原理及应用

(第二版)



杨爱民 张文祥 王涛伟 王利霞 编著



中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE



21 世纪高等学校计算机公共课程规划教材
“十一五”浙江省重点教材建设项目

数据库原理及应用

（第二版）

杨爱民 张文祥 王涛伟 王利霞 编著



中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

内 容 简 介

本书是根据教育部制定的关于计算机科学与技术及相关专业学生的培养目标而编写的。本书把数据库理论与当前流行的大型数据库 Oracle 相结合,通过具体应用案例来剖析数据库的理论与实践知识。

本书系统地介绍了关系数据库的基本概念、基本原理、基本方法以及应用实例,内容是按照由浅入深的方法,先介绍数据库技术的基础知识,然后介绍关系数据库理论,以此来指导学生依据什么样的理论来设计数据库,接着介绍了数据库的设计步骤与方法,如何进行数据库的访问,数据库的安全措施以及并发访问的特点,最后以实际应用案例来说明数据库技术的综合应用。

本书既可作为普通高校、成人院校的计算机类、信息类、管理类本科专业的教材,也可作为相关领域技术人员的参考资料及培训教材。

本教材编写组为配合本书的内容自主开发了一套网上实验系统,主要用于 SQL 语言的测评,详见书后附录,有兴趣的学校可以与本教材编写组联系,个人用户可以直接注册使用,网址: <http://datajx.computer.zwu.edu.cn>。

图书在版编目(CIP)数据

数据库原理及应用 / 杨爱民等编著. —2 版. —北京 :
中国铁道出版社, 2012. 12
21 世纪高等学校计算机公共课程规划教材
“十一五”浙江省重点教材建设项目
ISBN 978-7-113-15650-3

I. ①数… II. ①杨… III. ①数据库系统—高等学校—教材 IV. ①TP311.13

中国版本图书馆 CIP 数据核字 (2012) 第 267686 号

书 名: 数据库原理及应用 (第二版)
作 者: 杨爱民 张文祥 王涛伟 王利霞 编著

策 划: 吴宏伟 刘丽丽 读者热线: 400-668-0820
责任编辑: 吴宏伟
编辑助理: 赵 迎
封面设计: 付 巍
封面制作: 白 雪
责任印制: 李 佳

出版发行: 中国铁道出版社 (100054, 北京市西城区右安门西街 8 号)
网 址: <http://www.51eds.com>
印 刷: 北京新魏印刷厂
版 次: 2006 年 8 月第 1 版 2012 年 12 月第 2 版 2012 年 12 月第 1 次印刷
开 本: 787mm×1 092mm 1/16 印张: 15 字数: 359 千
印 数: 1~3 000 册
书 号: ISBN 978-7-113-15650-3
定 价: 29.00 元

版权所有 侵权必究

凡购买铁道版图书, 如有印制质量问题, 请与本社教材图书营销部联系调换。电话: (010) 63550836

打击盗版举报电话: (010) 63549504

随着计算机技术与网络通信技术的发展,数据库技术已成为信息社会中不可缺少的技术之一,这是因为数据库技术应用的范围极其广阔,如金融、保险、超市、企业以及各类办公系统都离不开数据库的支持,它已应用于社会各个领域,而且随着硬件技术与软件技术的发展而不断更新和完善,数据库技术已经成为信息系统的基础和核心。为了适应当今信息社会的需求,各高等院校计算机类、信息类等相关专业都已将数据库技术及应用纳入自己的课程体系之中。

目前市场已有许多关于数据库技术及应用的教材,但一般都缺乏完整的应用案例,并且选用的大都是 Access, FoxPro, SQL Server 等中小型数据库作为样例数据库来介绍,这样学生就无法接触到大型数据库的使用方法,也无从了解数据库管理员 DBA 的权限以及数据库后台管理的知识,对数据库体系无法形成完整概念。

本书编写的主要目的是为了适应应用型人才培养的需要,同时也是为生产实践服务。本书的显著特点是以案例为基础,在书中各章节都加入了应用型案例,使学生在学数据库理论的同时,能依据书中提供的案例,动手参与项目实践,充分将所学的书本知识融会贯通。此外,本书的数据库样例除介绍一般的数据库如 Access、SQL Server 的应用外,还融入了大型数据库 Oracle。Oracle 是几种大型数据库的代表之一,它以安全性好、容量大、并发性强以及快速的数据处理能力而著称。目前社会上很多大型企业、大型网站、政府等部门(如银行、税务)的数据管理,采用 Oracle 数据库的居多。学生通过本教材的学习,既可以掌握相关的理论知识,又可以学到很多实际的应用技术,同时还可以掌握一种大型数据库的管理经验。

为了更好地实现应用型人才的培养效果,本教材引入完整的实用型教学案例,如学生信息管理系统、企业网站系统等,使学生在学完本教材后,也完成了一个应用软件系统的制作,这样既让学生学习到了数据库技术的知识,也掌握了数据库应用软件的编程方法,为将来的就业打下基础。

本书共分 10 章,第 1 章主要介绍数据库系统的基本概念、数据库系统的体系结构及数据模型;第 2 章主要介绍关系数据库理论,包括关系数据结构、关系演算理论、函数依赖及范式定理;第 3 章主要介绍数据库的设计步骤与方法,共分为六大步骤;第 4 章主要介绍大型数据库 Oracle 的安装、配置与管理使用;第 5 章主要介绍 SQL 语言对数据库的定义、查询、更新和删除的方法;第 6 章以 Oracle 为例,主要介绍数据库的安全管理技术和方法;第 7 章主要介绍数据库的并发技术以及解决办法;第 8 章主要介绍新一代数据库的发展以及未来数据库所涉及的相关技术;第 9 章主要介绍数据库接口技术,包括 ADO、ODBC 以及 Web 数据库的接口方法;第 10 章主要介绍应用类软件系统(如超市管理、医院管理等)的数据库设计分析,为应用类软件的开发提供了指导。本书为帮助学生理解数据库理论的应用,每章后都配有适量的习题供学生练习。此外,随教材附带了编者自主开发的教学课件及 SQL 测试软件,用于学生学习使用。

本书是由杨爱民、张文祥、王涛伟、王利霞几位老师编写的。目前已列为浙江省宁波市重点建设项目教材。

限于编者水平,书中难免有疏漏之处,敬请广大读者与专家批评指教。

编者

2012 年 8 月

第 1 章 数据库概论	1
1.1 数据库系统	1
1.1.1 数据库系统的组成	1
1.1.2 数据库系统的效益	2
1.2 数据库管理系统 (DBMS)	3
1.2.1 DBMS 的主要功能	3
1.2.2 DBMS 的组成	5
1.2.3 DBMS 的工作过程	6
1.3 数据模型	7
1.3.1 信息和数据	7
1.3.2 数据模型的三个层次	7
1.3.3 信息世界中的基本概念	8
1.3.4 概念模型的 E-R 模型表示方法	9
1.3.5 数据库层次的数据模型	12
1.3.6 物理模型	15
1.4 数据库系统的模式结构	16
1.4.1 数据库系统的三级模式结构	16
1.4.2 数据库的二级映像功能与数据独立性	17
1.5 数据库技术的产生与发展	18
1.5.1 人工管理阶段	18
1.5.2 文件系统管理阶段	18
1.5.3 数据库管理阶段	19
1.6 数据库系统的体系结构	22
1.6.1 单用户数据库系统	22
1.6.2 主/从式结构的数据库系统	22
1.6.3 分布式结构的数据库系统	23
1.6.4 客户机/服务器结构的数据库系统	23
小结	24
思考与练习	24
第 2 章 关系数据库	26
2.1 关系数据库概述	26
2.1.1 数据结构	26
2.1.2 关系操作	26
2.1.3 完整性	27

2.2	关系数据结构	27
2.2.1	关系的定义及性质	27
2.2.2	关系模式与关系数据库	28
2.2.3	关系的完整性规则	29
2.3	关系代数	30
2.3.1	传统的集合运算	30
2.3.2	专门的关系运算	33
2.3.3	关系代数表达式及其应用实例	36
2.4	关系演算	37
2.4.1	元组关系演算	38
2.4.2	域关系演算	39
2.4.3	关系运算的安全性和等价性	39
2.5	查询优化	40
2.5.1	关系代数表达式的优化问题	40
2.5.2	关系代数表达式的等价变换规则	41
2.5.3	优化的一般策略	42
2.5.4	优化算法	43
2.6	函数依赖	46
2.6.1	问题的提出	46
2.6.2	函数依赖定义	47
2.6.3	码	48
2.7	关系的规范化	48
2.7.1	第一范式	49
2.7.2	第二范式	49
2.7.3	第三范式	50
2.7.4	BC 范式	50
2.7.5	模式分解	51
	小结	52
	思考与练习	52
第 3 章	数据库的设计	54
3.1	数据库设计的基本步骤	54
3.2	需求分析	56
3.2.1	需求描述与分析	56
3.2.2	需求分析分类	57
3.2.3	需求分析的内容与方法	57
3.2.4	需求分析的步骤	58
3.2.5	数据字典	59
3.3	概念设计	61
3.3.1	概念设计的必要性及要求	61

3.3.2	概念设计的方法与步骤	62
3.3.3	E-R 模型的操作	63
3.3.4	采用 E-R 方法的数据库概念设计	65
3.4	逻辑结构设计	70
3.4.1	逻辑设计环境	70
3.4.2	逻辑设计步骤	70
3.4.3	从 E-R 图向关系模型转换	71
3.4.4	设计用户子模式	73
3.4.5	对数据模型进行优化	73
3.5	数据库的物理设计	74
3.5.1	数据库设计人员需要掌握的物理设计知识	74
3.5.2	数据库物理设计的主要内容	75
3.5.3	物理设计的性能评价	77
3.5.4	系统数据库表的物理设计	78
3.5.5	数据库表结构	78
	小结	80
	思考与练习	80
第 4 章	Oracle 系统概述及安装配置	81
4.1	Oracle 系统概况	81
4.1.1	Oracle 系统的特点	81
4.1.2	Oracle 产品结构	82
4.2	Oracle 数据库系统的安装	84
4.2.1	Oracle 数据库系统服务器端的安装	84
4.2.2	Oracle 数据库系统客户端的安装	88
4.3	客户机和服务器的连接	92
4.3.1	服务器端的设置	92
4.3.2	客户端的设置	94
4.4	Oracle 数据库服务器的体系结构	97
4.4.1	Oracle RDBMS 文件结构	97
4.4.2	Oracle 数据库的逻辑结构	99
4.4.3	Oracle 数据库的存储结构	102
4.5	管理服务器的配置	102
4.5.1	数据库服务器的管理模式	102
4.5.2	配置管理服务器	103
4.5.3	管理服务器的启动与关闭	105
4.6	数据库服务器的管理用户	106
4.6.1	system 用户	106
4.6.2	sys 用户	106
4.6.3	scott 用户	107

4.6.4	sys 和 system 用户的比较	107
4.7	管理工具的使用	107
4.7.1	SQL Plus 的使用	108
4.7.2	SQL Plus Worksheet 的使用	109
4.7.3	Enterprise Manager Console 的使用	110
4.8	数据库服务器的启动与关闭	112
4.8.1	服务器的关闭	112
4.8.2	服务器的启动	114
	小结	116
	思考与练习	116
第 5 章	SQL 基础	117
5.1	SQL 概述	117
5.2	SQL 数据定义	118
5.2.1	基本表的创建、修改、删除及重命名	118
5.2.2	索引的定义和删除	122
5.2.3	视图的定义和删除	122
5.3	SQL 数据查询	124
5.3.1	查询命令 (SELECT)	124
5.3.2	简单查询	125
5.3.3	表连接操作	127
5.3.4	集合运算	129
5.3.5	聚合和分组查询	130
5.3.6	子查询	130
5.4	SQL 数据操纵	133
5.4.1	向表中插入新行 (记录)	133
5.4.2	表中记录更新 (UPDATE)	134
5.4.3	删除表记录 (DELETE)	134
5.5	SQL 数据控制	135
	小结	136
	思考与练习	136
第 6 章	数据库的安全管理	139
6.1	数据库的安全性	139
6.1.1	数据库的存取控制和用户的建立	140
6.1.2	特权和角色	143
6.1.3	审计	148
6.2	数据完整性	148
6.2.1	完整性约束	149
6.2.2	数据库触发器	149
6.3	Oracle 数据库的安全管理	151

6.3.1	系统安全性与授权	152
6.3.2	审计	156
6.3.3	利用视图实施安全性控制	158
6.4	Oracle 数据库的备份和恢复	159
6.4.1	数据库的卸载实用程序	159
6.4.2	数据库的装载实用程序	161
	小结	163
	思考与练习	163
第 7 章	并发控制	164
7.1	事务的基本概念	164
7.1.1	事务定义	164
7.1.2	事务的性质	165
7.2	并发控制过程	166
7.2.1	丢失修改	166
7.2.2	不可重复读	166
7.2.3	读“脏”数据	167
7.3	封锁及封锁协议	167
7.3.1	封锁	167
7.3.2	封锁协议	168
7.3.3	活锁和死锁	169
7.4	并发调度的可串行性	171
7.4.1	可串行性	172
7.4.2	两段锁协议	173
7.5	封锁的粒度	174
7.5.1	多粒度锁协议	174
7.5.2	意向锁	175
7.6	Oracle 的并发控制机制	176
	小结	177
	思考与练习	177
第 8 章	数据库技术的发展	178
8.1	概述	178
8.2	新一代数据库系统	178
8.2.1	面向对象数据库系统宣言	178
8.2.2	第三代数据库系统宣言	179
8.2.3	数据库技术与相关技术结合形成新型数据库系统	180
	思考与练习	181
第 9 章	数据库接口技术	182
9.1	C++ Builder 编程工具介绍	182

9.1.1	C++ Builder 环境简介	182
9.1.2	C++ Builder 基本功能简介	183
9.2	ADO 数据库接口技术	187
9.2.1	ADO 数据库接口技术简介	187
9.2.2	ADO 数据库接口技术实现方法	187
9.3	ODBC 数据库接口技术	191
9.3.1	ODBC 数据库接口技术简介	191
9.3.2	ODBC 数据库接口技术实现方法	192
9.4	基于 Web 程序的数据库接口技术	195
9.4.1	基于 ASP 的数据库接口技术	195
9.4.2	基于 ASP.NET 的数据库接口技术	201
	小结	215
	思考与练习	215
第 10 章	数据库应用案例分析	216
10.1	学生信息管理系统的数据库设计	216
10.1.1	学生信息管理系统 E-R 模型分析	216
10.1.2	学生信息管理系统数据库的设计	217
10.2	图书信息管理系统的数据库设计	218
10.2.1	图书信息管理系统 E-R 模型分析	218
10.2.2	图书信息管理系统数据库的设计	219
10.3	医院门诊管理系统的数据库设计	220
10.3.1	医院门诊管理系统 E-R 模型分析	220
10.3.2	医院门诊管理系统数据库的设计	222
10.4	小超市管理系统的数据库设计	223
10.4.1	小超市管理系统 E-R 模型分析	223
10.4.2	小超市管理系统数据库的设计	224
	小结	226
	思考与练习	226
附录 A	SQL 进阶训练系统	227
	参考文献	230

第 1 章 数据库概论

本章以数据库技术的基础为主要内容，介绍数据库系统的组成、数据模型、数据库模式结构、数据库技术的产生与发展及数据库的体系结构等知识。通过本章的学习，可以对数据库技术基础知识有一个基本的了解，认识数据库的结构及组成等，为进一步的学习打下基础。

1.1 数据库系统

1.1.1 数据库系统的组成

数据库系统 (DataBase System, DBS) 是指在计算机系统中引入数据库后构成的系统，不仅是一组对数据进行管理的软件 (即通常所说的数据库管理系统)，也不仅是一个数据库。数据库系统是一个可实际运行的，按照数据库方式存储、维护和向应用系统提供数据或信息支持的系统。它是存储介质、处理对象和管理系统的集合体，一般由数据库、硬件、软件、数据库管理员四部分构成。

1. 数据库

数据库 (DataBase, DB) 是与一个特定组织的各项应用相关的所有数据的汇集。通常由两大部分组成：一部分是有关应用所需工作数据的集合，称为物理数据库，是数据库的主体；另一部分是关于各级数据结构的描述，称为描述数据库，通常是由一个数据字典系统管理。

数据库构建主要是通过综合各个用户的文件，除去不必要的冗余，使之相互联系形成的数据结构。数据结构的实现取决于数据库的类型。

2. 硬件支持系统

硬件是数据库赖以存在的物理设备，包括 CPU、内存、外存、数据通道等各种存储、处理和传输数据的设备。对数据库系统来说，要求有较大的内存，用来存放系统程序、应用程序以及开辟系统和用户工作区缓冲区；外部存储一般要配备高速、大容量的直接存取设备，如磁盘或光盘等；I/O 存取速度、可支持终端数和性能稳定性等指标也是影响数据库系统性能的因素。在许多应用中还要考虑系统支持联网的能力和配备必要的后备存储设备等因素。此外，还要求系统有较高的通道能力，以提高数据的传输速度。

3. 软件支持系统

软件支持系统主要包括操作系统、数据库管理系统、各种宿主语言和支持开发的实用程序等。数据库管理系统 (DataBase Management Systems, DBMS) 是管理数据库的软件系统，DBMS 是在操作系统 (OS) 支持下工作的，选用 DBMS 时还要考虑选择提供支持的操作系统；为开发

应用系统，需要各种宿主语言（如 COBOL、PL/I、FORTRAN、C 等）及其编译系统，这些语言应与数据库有良好的接口。需要支持开发的实用程序，如报表生成器、表格系统、图形系统、具有数据库存取和表格 I/O 功能的软件、数据字典等，是系统为应用开发人员和最终用户提供的高效率、多功能的交互式程序设计系统，它们为数据库应用系统的开发和应用提供了良好的环境，使用户生产率提高 20~100 倍。

4. 开发和使用数据库系统的人员

管理、开发和使用数据库系统的人员主要有系统分析员、数据库管理员（DBA）、应用程序员和用户，分别有不同的职责。

① 系统分析员：负责应用系统的需求分析和规范说明。系统分析员要与用户及 DBA 配合，确定系统的软硬件配置并参与数据库各级模式的概要设计。

② 数据库管理员：对数据库系统监督、管理的人员，又称 DBA。

③ 应用程序员：负责设计应用系统的程序模块，根据外模式编写应用程序和编写对数据库的操作过程程序。

④ 用户：分为应用程序和终端用户两类。用户通过应用系统的用户接口使用数据库，目前常用的接口方式有菜单驱动、表格操作、图形显示报表生成等，这些接口给用户提供简明直观的数据表示。

大型的系统中由于数据库具有共享性，要想成功地运转数据库，需要配上 DBA，维护和管理数据库，使之处于最佳的状态。DBA 可以是个人或几个人组成的小组，其主要职责是：

① 决定数据库的信息内容和结构，确定某现实问题的实体联系模型，建立与 DBMS 有关的数据模型和概念模式。

② 决定存储结构和存取策略，建立内模式和模式/内模式映像。使数据的存储空间利用率和存取效率两方面都较优。

③ 充当用户和 DBS 的联络员，建立外模式和外模式/模式映像。

④ 定义数据的安全性要求和完整性约束条件，以保证数据库的安全性和完整性。安全性要求是用户对数据库的存取权限，完整性约束条件是对数据进行有效性检验的一系列规则和措施。

⑤ 确定数据库的后援支持手段及制定系统出现故障时数据库的恢复策略。

⑥ 监视并改善系统的“时空”性能，提高系统效率。

⑦ 当系统需要扩充和改造时，负责修改和调整外模式、模式和内模式。

总之，DBA 承担创建、监控和维护整个数据库结构的责任。由于职责重要和任务复杂，要求具有系统程序员和运筹学专家的品质和知识，一般由业务水平较高、资历较深的人员担任。

1.1.2 数据库系统的效益

数据库系统的应用，使计算机应用深入到社会的各个领域。这是因为从数据库系统可获得很大的效益，具体有以下几个方面：

① 灵活性：数据库容易扩充以适应新用户的要求，同时也容易移植以适应新的硬件环境和更大的数据容量。

② 简易性：由于精心设计的数据库能模拟企业的运转情况，并提供该企业数据逼真的描述，使管理部门和使用部门能很方便地使用和理解数据库。

③ 面向用户：由于数据库反映企业的实际运转情况，因此基本上能满足用户的要求，同时数据库又为企业的信息系统奠定了基础。

④ 数据控制：对数据进行集中控制，就能保证所有用户在同样的数据上操作，而且数据对所有部门具有相同的含义。数据的冗余减到最少，消除了数据的不一致性。

⑤ 加快应用系统开发速度：程序员和系统分析员可以集中全部精力于应用处理设计，而不必关心数据操纵和文件设计的细节，后援和恢复问题均由系统保证。

⑥ 程序设计方面：数据库方法使系统中的程序数目减少而又不过分增加程序的复杂性，由于数据管理语言命令功能强，应用程序编写起来较快，进一步提高了程序员的生产效率。

⑦ 修改方便：数据独立性使得修改数据库结构时尽量不损害已有的应用程序，使程序维护工作量大为减少。

⑧ 标准化：数据库方法能促进建立整个企业的数据库一致性和用法的标准化工作。

1.2 数据库管理系统 (DBMS)

数据库管理系统 (DBMS) 是数据库系统中对数据进行管理的软件，是数据库系统的核心组成部分。对数据库的一切操作，包括定义、查询、更新及各种控制等都是通过 DBMS 进行的。DBMS 是用户与数据库的接口。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、概念级，再导向内部级，进而操纵存储器中的数据。

DBMS 是针对某种数据模型设计的，可以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，即使同种数据模型的 DBMS，它们在用户接口和系统功能等方面也常常是不相同的。

1.2.1 DBMS 的主要功能

1. 数据库的定义功能

DBMS 提供数据定义语言 (Data Definition Language, DDL) 定义数据库的结构，包括外模式、内模式及其相互之间的映像，定义数据的完整性约束、保密限制等约束条件，又称数据描述语言。定义工作是由 DBA 完成的。在 DBMS 中有 DDL 的编译程序，负责将 DDL 编写的各种模式编译成相应的目标模式。这些目标模式是对数据库的描述，不是数据本身，是数据库的框架 (即结构)，并被保存在数据字典中，供以后进行数据操纵或数据控制时查阅使用。

2. 数据库操纵功能

DBMS 提供数据操纵语言 (Data Manipulation Language, DML) 实现对数据库的操作。基本的数据操作有四种：检索、插入、删除和修改。DML 有两类：一类是嵌入在宿主语言中使用，例如嵌入在 COBOL、FORTRAN、C 等高级语言中，这类 DML 称为宿主型 DML；另一类是可以独立交互使用的 DML，称为自主型或自含型 DML。因而 DBMS 中必须包括 DML 的编译程序或解释程序。

3. 数据库运行控制功能

DBMS 对数据库的控制主要通过四方面实现：数据安全性控制、数据完整性控制、多用户环境下的并发控制和数据库的恢复。

① 数据库安全性的控制是对数据库的一种保护。使用数据的用户必须向 DBMS 标识自己，由系统确定是否可以对指定的数据进行存取。它的作用是防止未经授权的用户修改数据库中的数据，以免数据泄露、更改或破坏，使企业蒙受巨大的损失。

② 数据完整性控制是 DBMS 对数据库提供保护的另一个重要方面。其目的是保持进入数据库中存储数据的语义的正确性和有效性，防止对数据造成违反其语义的操作。因此，DBMS 都允许对数据库中各类数据定义若干语义完整性约束，由 DBMS 强制实行。

③ 并发控制是 DBMS 的第三类控制机制。数据库技术的一个优点是数据的共享性，但多个应用程序同时对数据库进行操作可能会破坏数据的正确性，或者在数据库中存储了错误的数据，或者用户读取了不正确的数据（称为脏数据）。并发控制机制能防止上述情况发生，正确处理好多用户、多任务环境下的并发操作。

④ 数据库的恢复机制是保护数据库的又一个重要方面。在对数据库进行操作的过程中，可能会出现各种故障，如停电、软/硬件各种错误、人为破坏等，导致数据库损坏或者数据不正确。此时，DBMS 的恢复机制有能力把数据库恢复至最近某个正确的状态。为了保证恢复工作的正常进行，系统要经常为数据库建立若干备份副本（一般放在磁带上）。

4. 数据库的维护功能

数据库的维护功能包括数据库初始数据的载入、转换、转储，数据库的重组和性能监视、分析等。这些功能由各个实用程序完成，如装配程序（装配数据库）、重组程序（重新组织数据库）、日志程序（用于更新操作和数据库的恢复）、统计分析程序等。

5. 数据字典

数据字典（Data Dictionary, DD）中存放着数据库三级结构的描述，对于数据库的操作要通过查阅 DD 进行。现在在部分大型系统中，把 DD 单独抽出来成一个系统，成为一个软件工具，使得 DD 成为一个比 DBMS 更高级的用户和数据库之间的接口。

数据字典的任务是管理有关数据的信息，又称“数据库的数据库”。它的主要任务如下：

① 描述数据库系统的所有对象，并确定其属性。例如，一个模式中包含的记录型与一个记录型包含的数据项；用户的标识、口令；物理文件名称、物理位置及其文件组织方式等。数据字典在描述时赋给每个对象一个唯一标识。

② 描述数据库系统对象之间的各种交叉联系。例如，哪个用户使用哪个外模式，哪些模式或记录型分配在哪些区域及对应于哪些物理文件、存储在何种物理设备上。

③ 登记所有对象的完整性及安全性限制等。

④ 对数据字典本身的维护、保护、查询与输出。

数据字典的主要作用是：供数据库管理系统快速查找有关对象的信息。数据库管理系统在处理用户存取时，要经常查阅数据字典中的用户表、外模式表和模式表；供数据库管理员查询，以掌握整个系统的运行情况；支持数据库设计与系统分析。

上述是一般 DBMS 所具备的功能。通常在大、中型机上实现的 DBMS 功能较强、较全，在微机上实现的 DBMS 功能较弱。

还应指出,用宿主语言编写的应用程序并不属于 DBMS 的范围。应用程序是用宿主语言和 DML 编写的。程序中的 DML 语句是由 DBMS 解释执行的,而其余部分仍由宿主语言编译系统去编译。

1.2.2 DBMS 的组成

DBMS 通常由三部分组成:数据描述语言及其翻译程序、数据操纵语言及其处理程序和数据库管理的例行程序。

1. 数据描述语言

数据描述语言 (DDL) 对应数据库系统的三级模式 (外模式、模式和内模式) 分别由三种不同的 DDL (外模式 DDL、模式 DDL 和内模式 DDL) 实现,它们是专门提供给 DBA 使用的,一般用户不必去关心。

① 外模式 DDL 是专门定义某一用户的局部逻辑结构。

② 模式 DDL 用来描述数据库的全局逻辑结构。它包括数据库中所有元素的名称、特征及其相互关系的描述,并包括数据的安全保密性和完整性以及存储安排、存取路径等信息。

③ 内模式 DDL 用来定义物理结构的数据描述语言。它有存储记录和块的概念,但不受任何存储设备和设备规格 (如柱面大小、磁道容量等) 的限制。它包括对存储记录类型、索引方法等方面的描述。

2. 数据操纵语言

数据操纵语言 (DML) 是用户与 DBMS 之间的接口,是用户用于存储、控制、检索和更新数据库的工具。

DML 由一组命令组成,这些语句可分为四类:

① 存储语句。用户使用存储语句向数据库中存放数据。系统给出新增数据库记录的数据库码,并分配相应的存储空间。

② 控制语句。用户通过这类语句向 DBMS 发出使用数据库的命令,使数据库置于可用状态。操作结束后,必须使用关闭数据库的命令,以便对数据库的数据进行保护。

③ 检索语句。用户通过这类语句把需要检索的数据从数据库中选择出来传至内存,交给应用程序处理。

④ 更新语句。用户通过这组更新语句完成对数据库的插入、删除和修改数据的操作。

3. 数据库管理的例行程序

数据库管理的例行程序随系统而异。一般说来,它通常由如下三部分组成:

① 语言翻译处理程序:包括 DDL 翻译程序、DML 处理程序、终端查询语言解释程序、数据库控制语言的翻译程序等。

② 公用程序:定义公用程序和维护公用程序。定义公用程序包括信息格式定义、模式定义、外模式定义和保密定义公用程序等。维护公用程序包括数据装入、数据库更新、重组、重构、恢复、统计分析、工作日记、转储和打印公用程序等。

③ 系统运行控制程序:包括数据存取、更新、有效性检验、完整性保护程序、并发控制、数据库管理、通信控制程序等。

1.2.3 DBMS 的工作过程

下面通过应用程序 A 调用 DBMS 读取数据库中的一个记录的全过程(见图 1-1)来了解 DBMS 的工作过程。在应用程序 A 运行时, DBMS 首先开辟一个数据库的系统缓冲区, 用于输入/输出数据。三级模式的定义存放在数据字典中, 具体过程如下:

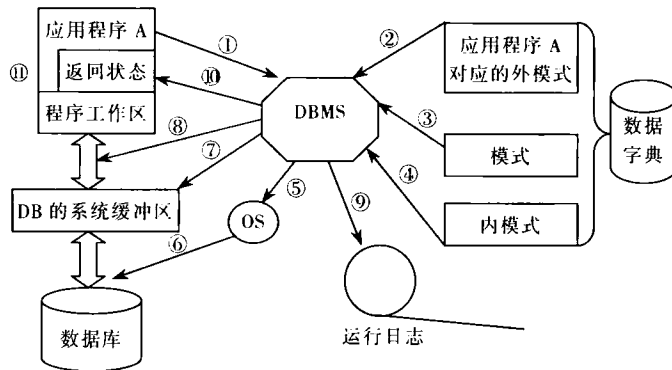


图 1-1 用户访问数据库的过程

① 应用程序 A 中有一条读记录的 DML 语句。该语句给出涉及外模式中记录类型名及想要读记录的键值。当计算机执行该 DML 语句时, 立即启动 DBMS, 并向 DBMS 发出读记录的命令。

② DBMS 接到命令后, 先从数据字典中调出该程序对应的外模式, 检查该操作是否在合法授权范围内, 若不合法则拒绝执行, 并向应用程序状态返回区发出不成功的状态信息; 若合法则执行下一步。

③ DBMS 调用相应的概念模式描述, 并从外模式映像到概念模式, 也就是把外模式的外部记录格式映像到概念模式记录格式, 决定概念模式应读入哪些记录。

④ DBMS 调用相应的内模式描述, 并把概念模式映像到内模式, 即把概念模式的概念记录格式映像到内模式的内部记录格式, 确定应读入哪些物理记录及具体的地址信息。

⑤ DBMS 向操作系统发出从指定地址读物理记录的命令。

⑥ 操作系统执行读命令, 按指定地址从数据库中把记录读入数据库的系统缓冲区, 并在操作结束后向 DBMS 做出回答。

⑦ DBMS 收到操作系统读操作结束的回答后, 参照模式将读入系统缓冲区中的内容转换成概念记录, 再参照外模式转换成用户要求读取的外部记录。

⑧ 将 DBMS 导出的外部记录从系统缓冲区送到应用程序 A 的“程序工作区”中。

⑨ DBMS 向运行日志数据库发出读一条记录的信息, 以便查询使用数据库的情况。

⑩ DBMS 将操作执行成功与否的状态信息返回给用户。

⑪ DBMS 应用程序根据返回的状态信息决定是否使用工作区中的数据。

如果用户需要修改一个记录内容, 其过程与此类似。这时首先读出目标记录, 并在用户工作区中用主语言的语句进行修改, 然后向 DBMS 发出写回修改记录的命令。DBMS 在系统缓冲区进行必要的转换(转换的过程与读数据时相反)后向操作系统发出写命令, 即可达到修改数据的目的。

1.3 数据模型

DBMS 都是针对数据模型进行设计的, 任何一个数据库都要组织成符合 DBMS 规定的数据库模型。将反映现实世界中有意义的信息转化为能在计算机中表示的数据并能被数据库处理是数据库模型要解决的问题。数据模型不仅要能表示存储了哪些数据, 更重要的是要能以一定的结构形式表示出各种不同数据之间的联系。利用这些联系很快找到相关联的数据, 完成相关数据的运算处理。因此, 数据模型应具有描述数据和数据联系两方面的功能。

1.3.1 信息和数据

数据是数据库系统研究和处理的对象, 而数据与信息是分不开的。信息是对现实世界各种事物的存在特征、运动形态及不同事物间的相互联系等在人脑中的抽象反映, 进而形成概念。数据是对信息的符号化表示, 即用一定的符号表示信息。数据是信息的载体, 而信息是数据的内涵。同一信息可以有不同的数据表示形式; 而同一数据也可能有不同的解释。信息只有通过数据形式表示出来才能被人们理解和接受。尽管数据和信息在概念上不尽相同, 但是通常人们在数据库处理中并不严格区分它们, 数据处理本质上就是信息处理。

1.3.2 数据模型的三个层次

数据模型是对客观事物及其联系的数据描述。从事物的特征到计算机中的数据表示, 对现实世界问题的抽象经历了三个不同层次, 即概念数据模型、逻辑数据模型、物理数据模型。

1. 概念数据模型

概念数据模型又称概念模型, 是现实世界到概念世界的抽象。它是一种与具体的计算机和数据库管理系统无关的, 面向客观世界和用户的模型。侧重于对客观世界复杂事物的结构及它们内在联系的描述, 而将与 DBMS、计算机有关的物理和细节的描述留给其他模型描述。概念模型是整个数据模型的基础。目前较为著名的模型是实体联系模型 (Entity-Relationship Model), 简称 E-R 模型。

2. 逻辑数据模型

逻辑数据模型又称数据模型, 是概念世界的抽象描述到信息世界的转换。它是一种面向数据库系统的模型, 直接与 DBMS 有关。概念模型只有在转换成数据模型后才能在数据库中得以表示。目前最常用的数据模型有层次模型 (Hierarchical Model)、网状模型 (Network Model)、关系模型 (Relational Model)。

这类模型有严格的定义, 包含数据结构、数据操作和数据完整性约束三个要素:

- ① 数据结构是指对实体类型和实体间联系的表达和实现。
- ② 数据操作是指对数据库的检索和更新 (包括插入、删除和修改) 两类操作。
- ③ 数据完整性约束给出数据及其联系应具有制约和依赖规则。

3. 物理数据模型

物理数据模型又称物理模型, 是信息世界模型在机器世界的实现。它是面向计算机物理表