

普通高等教育“十二五”规划教材

数据库 原理与应用

张丽娜 杜益虹 刘丽娜 主编



化学工业出版社

013063215

TP311.13-43

217

普通高等教育“十二五”规划教材

数据库原理与应用

张丽娜 杜益虹 刘丽娜 主 编

涂嘉庆 周 苏 卢淑静 副主编



TP311.13-43

217



化学工业出版社



北航

C1671323

30.00元

013083512

本书全面阐述了数据库系统的基本概念、理论、方法和技术。全书共 10 章, 主要内容包括数据库系统基础、数据模型和概念模型、关系数据库、关系数据库标准语言——SQL、SQL Server 关系数据库、关系数据库理论、数据库设计、数据库保护、数据库技术的新发展, 以及用于配合课堂教学的课内实践。

本书以培养具有高素质、高技能的应用型人才为目标, 结合计算机、信息管理等相关专业课程群建设的专业培养目标, 以职业能力培养为重点, 从岗位出发, 围绕“工学结合”的思路, 结合本课程既有理论知识又有实用基础的特点, 在内容上重点突出了“职业性、实践性、开放性”特点。本书集中了本课程组诸位老师近十余年的教学经验和教学成果, 配有课后习题和实践部分的参考答案。

本书适用于计算机类专业的本科生、高职高专、专升本的学生使用, 也可以作为大学各专业公共课教材和全国计算机等级考试——数据库原理与应用的参考书。

数据库原理与应用

张丽娜 杜益虹 刘丽娜 主编

张丽娜 杜益虹 刘丽娜 主编

图书在版编目(CIP)数据

数据库原理与应用 / 张丽娜, 杜益虹, 刘丽娜主编.

北京: 化学工业出版社, 2013.6

普通高等教育“十二五”规划教材

ISBN 978-7-122-17355-3

I. ①数… II. ①张… ②杜… ③刘… 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 (2013) 第 101187 号

责任编辑: 王听讲

装帧设计: 关 飞

责任校对: 宋 夏

出版发行: 化学工业出版社 (北京市东城区青年湖南街 13 号 邮政编码 100011)

印 装: 大厂聚鑫刷有限责任公司

787mm×1092mm 1/16 印张 13 $\frac{3}{4}$ 字数 339 千字 2013 年 8 月北京第 1 版第 1 次印刷

购书咨询: 010-64518888 (传真: 010-64519686) 售后服务: 010-64518899

网 址: <http://www.cip.com.cn>

凡购买本书, 如有缺损质量问题, 本社销售中心负责调换。

定 价: 30.00 元

版权所有 违者必究

前 言

数据库原理作为信息类专业的专业基础课程,如何从教材层面体现高素质应用型人才培养目标,成为了当前迫切需要解决的问题。为此,我们邀请长期工作在教学第一线的课程专业教师,在内容的选择、设计和实验文档的组织等方面都做了精心考虑和安排。参与本书编写的教师结合长期的教学实践,针对应用能力培养的目标,把实验实践环节和理论教学相融合,以实践能力培养促进理论知识的学习,有效地提高了课程的教学效果,所以本教材也是“数据库原理”精品课程组成员十余年教学改革和工作成果的总结。

本书遵循应用型人才培养目标,围绕“工学结合”教学理念,从计算机相关专业岗位群的调研出发,确定了数据库方面的专业技能,即数据库设计能力、应用能力和维护能力,归纳出了数据库技术的两条工作流:一是数据库的应用维护工作流,二是数据库的设计工作流,提取了工作过程中的典型工作步骤,界定了数据库的具体技能,从数据库的具体技能归纳出必须具备的数据库方面的原理知识和应用知识,以“原理够用,应用为先”为原则,对教学内容进行了设置。对原理知识部分中不常用又晦涩难懂的知识,诸如层次模型、网状模型、关系演算、关系模式的分解算法等内容进行了删减,增加了应用知识的篇幅,比如强化对关系数据库管理系统 SQL Server 和 SQL 语言的认识及运用。最终形成了以原理和应用两条主线共同支撑的教学内容,主要涵盖了数据库原理和应用两部分教学内容。

原理部分包括数据库的基本知识、概念模型和关系模型、关系数据库及理论、新型数据库技术发展。应用部分包括关系数据库标准语言(SQL)、数据库保护技术、SQL Server 关系数据库系统、数据库设计方法。

我们将为使用本书的教师免费提供电子教案和教学资源,需要者可以到化学工业出版社的教学资源网站 <http://www.cipedu.com.cn> 免费下载使用。欢迎登录本课程的精品课程网站 <http://jpkc.wucc.cn/sql> 下载相关教学资源和学习资料,本书还配有课后习题和实践部分的参考答案。

本书可作为信息类相关专业的本专科教材,讲授学时为 50~80 学时,教师可以根据学时、专业、培养目标和学生的实际情况选择。本书文字简明易懂、侧重工程实践、便于学生自学,也可以供从事数据库相关工作的科技人员参考学习。

本书的编写工作得到了温州市“数据库原理”精品课程建设项目的资助和支持,参与精品课程项目的张丽娜、杜益虹、刘丽娜、卢淑静、翁正秋、周苏、涂嘉庆等老师,都积极参与了本书的编写工作。由于编者水平所限,书中难免存在一些缺点,殷切希望广大读者批评指正。

编 者

2013 年 6 月

目 录

第一章 数据库系统基础	1	第二节 关系代数	24
第一节 基本概念	1	一、传统的集合运算	24
一、信息与数据	1	二、专门的关系运算	26
二、数据库	2	三、关系代数表示检索的实例	31
三、数据库管理系统	2	习题	32
四、数据库系统	2	第四章 关系数据库标准语言——SQL	35
五、数据库技术	3	第一节 SQL 概述	35
第二节 数据库技术及发展	3	一、SQL 的发展	35
一、数据库技术的发展	3	二、SQL 的优点	36
二、当代数据库研究的内容	6	三、SQL Server 的硬件要求	36
第三节 数据库系统的结构	7	四、SQL 的应用情况	37
一、数据库系统的体系结构	7	五、SQL 的主要功能	37
二、数据库管理系统的功能结构	8	六、SQL 语句的执行	38
三、数据库系统的模式结构	9	第二节 创建数据库	39
习题	10	一、数据库文件	40
第二章 数据模型和概念模型	12	二、数据库对象	40
第一节 信息的 3 个世界	12	三、创建数据库	41
一、现实世界	12	第三节 创建表	42
二、信息世界	12	一、创建表	42
三、机器世界	13	二、修改表的结构	43
四、现实世界、信息世界和机器世界的关系	13	三、删除表	44
第二节 概念模型	14	四、索引	45
一、概念模型涉及的基本概念	14	第四节 数据查询	46
二、实体联系的类型	14	一、SELECT 语句基础	46
三、概念模型的表示方法	15	二、选择列	46
第三节 数据模型	16	三、选择行	48
一、数据模型的三要素	16	四、聚合函数	50
二、常见的数据模型	17	五、连接查询	51
习题	18	六、嵌套查询	53
第三章 关系数据库	20	七、数据分组	57
第一节 关系模型及其三要素	20	八、排序	59
一、数据结构	20	九、集合运算	59
二、关系操作	23	第五节 表数据操作	60
三、关系的完整性	23	一、插入数据	61
		二、删除数据	62

三、修改数据	62	一、关系模式分解的算法基础	110
第六节 视图	63	二、模式分解	113
一、定义视图	63	习题	116
二、使用和修改视图	64	第七章 数据库设计	117
三、删除视图	65	第一节 数据库设计概述	117
第七节 游标	66	一、数据库设计的内容	117
一、声明游标	66	二、数据库设计的基本步骤	118
二、打开游标	66	第二节 系统需求分析	118
三、读取数据	67	一、需求分析的任务	118
四、关闭游标	67	二、系统需求分析方法	119
五、删除游标	67	三、数据流图	120
习题	68	四、数据字典	121
第五章 SQL Server 关系数据库	73	第三节 概念结构设计	124
第一节 SQL Server 的系统概述	73	一、设计方法概述	124
一、客户机/服务器系统	74	二、设计步骤	125
二、SQL Server 的数据库对象	75	三、实体与属性划分的原则	126
三、SQL Server 的管理工具和工作环境	76	四、数据抽象方法	126
第二节 Transact-SQL	82	五、各分 E-R 图之间的冲突及解决办法	128
一、默认值约束	83	第四节 逻辑结构设计	130
二、规则对象	85	一、基本 E-R 图转换为关系模型的基本方法	131
三、存储过程	87	二、用户子模式的设计	134
四、触发器	91	第五节 数据库的物理设计	134
习题	100	一、设计数据库物理结构要求设计人员了解的内容	134
第六章 关系数据库理论	101	二、设计数据库物理结构的步骤	135
第一节 关系模式的评价	101	三、评价数据库的物理结构	136
一、关系模式	101	第六节 数据库的实施和维护	136
二、关系模式的评价	102	一、数据库实施的主要工作	136
第二节 函数依赖	103	二、数据的载入	136
一、数据依赖	103	三、数据库的试运行	136
二、函数依赖	103	四、数据库的运行与恢复	137
三、函数依赖的几种特例	104	习题	141
第三节 范式	104	第八章 数据库保护	142
一、第一范式	105	第一节 数据库的安全性	142
二、第二范式	105	一、数据库的安全措施	142
三、第三范式	106	二、存取控制	144
四、BC 范式	107	三、其他数据库安全性手段	147
五、多值依赖及 4NF	107	四、SQL Server 的安全体系结构	147
六、第五范式	108	五、SQL Server 的安全认证模式	148
七、关系模式的规范化	109		
第四节 关系模式的分解算法	110		

第二节 数据库的完整性	163	第十章 课内实践	197
一、完整性约束条件及完整性控制	163	说明	197
二、SQL Server 的数据完整性及实现		实验一 SQL Server 的可视化操作	198
方法	167	一、实验目的	198
第三节 数据库的并发控制	168	二、实验指南	198
一、事务	168	三、实验内容	203
二、并发控制	171	四、思考题	203
第四节 数据库恢复技术	174	实验二 单表查询	204
一、故障的种类	174	一、实验目的	204
二、数据恢复的实现技术	174	二、实验指南	204
三、SQL Server 的数据备份和恢复	175	三、实验内容	206
习题	184	四、思考题	207
第九章 数据库技术的新发展	187	实验三 连接查询和嵌套查询	207
第一节 新一代数据库系统及应用	187	一、实验目的	207
第二节 扩展的关系数据库系统	188	二、实验指南	207
一、基于逻辑的关系数据库系统	188	三、实验内容	207
二、基于嵌套关系模型的关系数据库		四、思考题	208
系统	188	实验四 综合查询	208
三、专家数据库系统	189	一、实验目的	208
第三节 面向对象的数据库系统	189	二、实验指南	208
一、面向对象的数据模型	189	三、实验内容	208
二、面向对象的数据库管理系统	189	四、思考题	209
第四节 分布式数据库系统	190	实验五 用 SQL 实现数据库的建立与维护	209
一、分布式数据库设计概述	190	一、实验目的	209
二、分布式数据库系统的体系结构	190	二、实验指南	209
第五节 数据仓库与联机分析处理技术	191	三、实验内容	209
一、数据仓库的概念及特点	191	四、思考题	210
二、数据仓库与数据库的关系及比较	192	实验六 存储过程和触发器	210
三、联机分析处理技术	192	一、目的和要求	210
第六节 其他数据库新技术	193	二、实验准备	210
一、多媒体数据库技术	193	三、实验内容	210
二、时态数据库技术	193	四、思考题	211
三、移动数据库技术	194	实验七 数据库保护	211
四、主动数据库技术	194	一、实验目的	211
五、数据流技术	194	二、实验准备	211
六、基于 Web 数据库的问技术	195	三、实验内容	211
习题	196	四、思考题	211
		参考文献	212

第一章 数据库系统基础

【知识目标】

- 了解数据库技术的发展;
- 理解三级模式结构;
- 掌握数据库、数据库管理系统、数据库系统等基本概念;
- 掌握数据库管理系统的功能与组成。

【能力目标】

- 能说出数据库技术发展的整个过程;
- 能表达出数据、数据库、数据库管理系统、数据库系统概念;
- 能够使用数据库系统的体系结构。

第一节 基本概念

一、信息与数据

数据和信息是数据处理中的两个基本概念,它们具有不同的含义。

1. 数据

说起数据(Data),人们首先想到的就是数字。其实数字只是最简单的一种数据。数据的种类很多,在日常生活中数据无处不在,文字、图形、图像、声音……,这些都是数据。

为了认识世界,交流信息,人们需要描述事物。数据实际上是描述事物的符号记录。在日常生活中人们直接用自然语言(如汉语)描述事物。在计算机中,为了存储和处理这些事物,就要抽出对这些事物感兴趣的特征组成一个记录来描述。例如,在学生档案中,如果人们最感兴趣的是学生的姓名、性别、出生年月、籍贯、所在系别、入学时间,那么可以如下形式描述成雷明,男,1982,江苏,计算机系,2000。

数据与其语义是不可分的。对于上面这条学生记录,了解其语义的人会得到如下信息:雷明是个大学生,1982年出生,江苏人,2000年考入计算机系,而不了解其语义的人则无法理解其含义。可见,数据的形式本身并不能完全表达其内容,需要经过语义解释。

2. 信息

信息是一种重要的资源,一般认为,信息是关于现实世界事物的存在方式或运动状态反映的综合。例如,学生在多媒体教室上课,台风级别是7级等。

3. 数据处理

数据处理又称信息处理,是将数据转换成信息的过程,包括对数据的收集、存储、加工、检索和传输等一系列活动,其目的是从大量的原始数据中抽取和推导出有价值的信息,做各种应用。

我们可简单地用下列式子表示信息、数据与数据处理的关系。

信息=数据+数据处理

分析：数据可以形象地比喻为原料——输入；信息就像产品——输出；而数据处理是原料变成产品的过程。从这种角度看，“数据处理”的真正含义应该是为了产生信息而处理数据。

二、数据库

所谓数据库（DataBase，DB）就是存储数据的仓库。一般定义为：长期储存在计算机内、有组织的、可共享的、统一管理的数据集合。数据库中的数据按一定的数据模型组织、描述和储存，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。

三、数据库管理系统

数据库管理系统（DataBase Management System，DBMS）是数据库系统中对数据进行管理的一组大型软件系统，它是数据库系统的核心组成部分。数据库的一切操作，包括查询、更新及各种控制，都是通过数据库管理系统进行的。

数据库管理系统是位于用户与操作系统之间的一层数据管理软件。

数据库在建立、运用和维护时由数据库管理系统统一管理、统一控制。数据库管理系统使用户能方便地定义数据和操纵数据，并能够保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后的系统恢复。

目前常用的 DBMS 有 Oracle、MySql、Microsoft SQL Server、DB2、Sybase、FoxPro 和 Access 等。

四、数据库系统

数据库系统（DataBase System）是指在计算机系统中引入数据库后的系统构成，一般由数据库、数据库管理系统（及其开发工具）、应用系统、数据库管理员和用户构成。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 远远不够，还要有专门的人员来完成，这些人称为数据库管理员（Database Administrator，DBA）。数据库系统如图 1-1 所示。

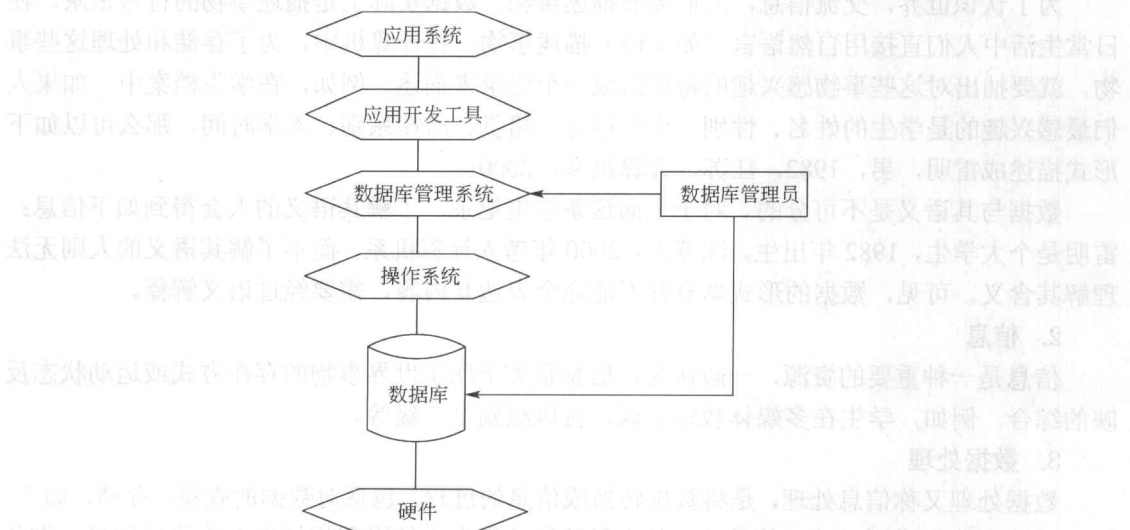


图 1-1 数据库系统

在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。

数据库管理员（DataBase Administrator, DBA）在数据库管理中是极其重要的，是所谓的超级用户。DBA 全面负责管理、控制和维护数据库，使数据能被任何有使用权限的人有效使用。DBA 可以是一个人，也可以是几个人组成的小组。DBA 主要有以下职责。

- ① 参与数据库设计的全过程，决定整个数据库的结构和信息内容。
- ② 帮助终端用户使用数据库，如培训终端用户，解答终端用户在日常使用数据库系统时遇到的问题等。
- ③ 定义数据的安全性和完整性，负责分配用户对数据库的使用权和口令管理等，制定数据库访问策略。
- ④ 监督控制数据库的使用和运行，改进和重新构造数据库系统。当数据库受到损坏时，负责恢复数据库；当数据库的结构需要改变时，完成对数据结构的改变。

DBA 不仅要有较高的技术水平和较深的资历，还应具有了解和阐明管理要求的能力。特别对于大型数据库系统，DBA 极为重要。而常见的微型计算机系统往往只有一个用户，没有必要专门设置专职的 DBA，DBA 通常由应用程序员和终端用户兼任。

五、数据库技术

数据库技术是使用计算机管理数据的一门最新技术。数据库技术所研究的问题是如何科学地组织和存储数据，如何高效地处理数据以获取其内在的信息。使用数据库对数据进行管理是计算机应用的一个重要而广阔的领域。

第二节 数据库技术及发展

一、数据库技术的发展

数据管理是指如何对数据进行分类、组织、编码、储存、检索和维护，是数据处理的中心问题。随着计算机硬件和软件的发展，数据管理经历了人工管理、文件系统、数据库系统三个初级阶段和高级数据库阶段。数据库管理初级阶段的比较见表 1-1。

表 1-1 数据管理三个初级阶段的比较

		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	无直接存储设备	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	数据的管理者	人	文件系统	数据库管理系统
	数据面向的对象	某一应用程序	某一应用程序	现实世界
	数据的共享程度	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，完全依赖于应用程序	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
	数据的控制能力	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. 人工管理阶段

在 20 世纪 50 年代中期以前，计算机主要用于科学计算。当时的硬件状况是，外存只有

纸带、卡片、磁带，没有磁盘等直接存取的存储设备；软件状况是，没有操作系统，没有管理数据的软件；数据处理方式是批处理。

人工管理数据具有如下特点。

(1) 数据不保存

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。不仅对用户数据如此处置，对系统软件有时也是这样。

(2) 数据需要由应用程序自己管理

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。没有相应的软件系统负责数据的管理工作，因此程序员负担很重。

(3) 数据不共享

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。

(4) 数据不具有独立性

数据的逻辑结构或物理结构发生变化后，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。

人工管理阶段应用程序与数据之间的对应关系如图 1-2 所示。



图 1-2 人工管理阶段

2. 文件系统阶段

20 世纪 50 年代后期到 60 年代中期，计算机的应用范围逐渐扩大，计算机不仅用于科学计算，而且还大量用于管理。这时硬件上已有了磁盘、磁鼓等直接存取存储设备；软件方面，操作系统中已经有了专门的数据管理软件，一般称为文件系统；处理方式上不仅有了文件批处理，而且能够联机实时处理，用文件系统管理数据具有如下特点。

(1) 数据可以长期保存

由于计算机大量用于数据处理，数据需要长期保留在外部存储器上，反复进行查询、修改、插入和删除等操作。

(2) 由专门的软件即文件系统进行数据管理

程序和数据之间由软件提供的存取方法进行转换，使应用程序与数据之间有了一定的独立性，程序员可以不必过多地考虑物理细节，将精力集中于算法。而且数据在存储上的改变不一定反映在程序上，大大节省了维护程序的工作量。

(3) 数据共享性差

在文件系统中，一个文件基本上对应于一个应用程序，即文件仍然是面向应用的。当不同的应用程序具有部分相同的数据时，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，给数据的修改和维护带来了困难，容易造成数据的不一致。

(4) 数据独立性低

文件系统中的文件是为某一特定应用服务的，文件的逻辑结构对该应用程序来说是优化的，因此要想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改变，必须修改应用程序，修改文件结构的定义。而应用程序的改变，例如，应用程

序改用不同的高级语言等,也将引起文件的数据结构的改变。因此数据与程序之间仍缺乏独立性。可见,文件系统仍然是一个不具有弹性的无结构的数据集合,即文件之间是孤立的,不能反映现实世界事物之间的内在联系。

文件系统阶段应用程序与数据之间的关系如图 1-3 所示。

3. 数据库阶段

20 世纪 60 年代后期以来,计算机用于管理的规模更为庞大,应用越来越广泛,数据量急剧增长,同时多种应用、多种语言互相覆盖地共享数据集合的要求越来越强烈。这时硬件已有大容量磁盘,硬件价格下降,软件价格上升,为编制和维护系统软件及应用程序所需的成本相对增加;在处理方式上,联机实时处理的要求更多,并开始提出和考虑分布处理。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,使数据为尽可能多地应用服务,就出现了数据库技术,出现了统一管理数据的专门软件系统——数据库管理系统。用数据库系统来管理数据具有如下特点。

(1) 数据结构化

数据结构化是数据库与文件系统的根本区别。在文件系统中,相互独立的文件的记录内部是有结构的。传统文件的最简单形式是等长同格式的记录集合。例如,一个学生人事记录文件,每个记录都有如图 1-4 所示的记录格式。

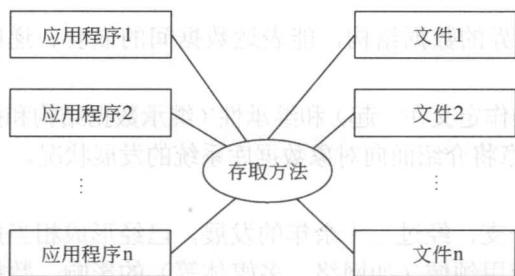


图 1-3 文件系统阶段应用程序与数据之间的对应关系

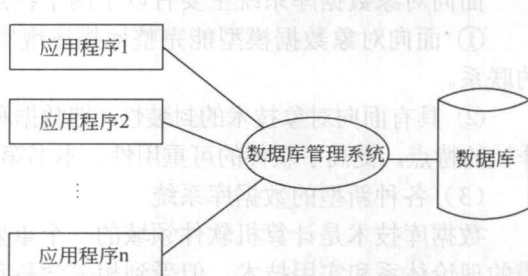


图 1-4 数据库系统阶段程序与数据的关系

(2) 较高的数据共享性

数据共享是指允许多个用户同时存取数据而互不影响,该特征正是数据库技术先进性的体现。主要包括 3 个方面:所有用户可以同时存取数据;数据库不仅可以为当前的用户服务,也可以为将来的新用户服务;可以使用多种语言完成与数据库的接口。

(3) 较高的数据独立性

所谓数据独立是指数据与应用程序之间的彼此独立,它们之间不存在相互依赖的关系。应用程序不随数据存储结构的变化而变化,这是数据库一个最基本的优点。

在数据库管理系统中,对数据的定义和管理已经从应用程序中分离出来,通过数据库管理系统统一控制。

(4) 数据由 DBMS 统一管理和控制

DBMS 加入了安全保密机制,可以防止对数据的非法存取。由于进行集中控制,因此有利于控制数据的完整性。数据系统采取了并发访问控制,保证了数据的正确性。另外,数据库系统还采取了一系列措施,实现了对数据库破坏后的恢复。

4. 高级数据库阶段

20 世纪 70 年代开始,数据库技术又有了很大的发展,并且不断与其他计算机技术相互

渗透。数据库技术与其他学科的内容相结合,是新一代数据库技术的一个显著特征,涌现出各种新型的数据库系统。主要的标志是 20 世纪 80 年代的分布式数据库系统、90 年代的面向对象数据库系统和各种新型数据库系统。

(1) 分布式数据库系统

随着地理上分散的用户对数据共享的要求日益增强,以及计算机网络技术的发展,在传统的集中式数据库系统基础上产生和发展了分布式数据库系统。

分布式数据库系统并不是简单地把集中式数据库安装在不同场地,用网络连接起来以实现(这是分散的数据库系统),而是具有自己的性质和特征。集中式数据库系统中的许多概念和技术,如数据独立性的概念,数据共享和减少冗余的控制策略,并发控制和事务恢复的概念及实现技术等,在分布式数据库中有不同的、更加丰富的内容。

(2) 面向对象数据库系统

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,如多媒体、多维表格数据和 CAD(计算机辅助设计)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。而面向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库技术结合的产物。

面向对象数据库系统主要有以下两个特点。

① 面向对象数据模型能完整地描述现实世界的数据结构,能表达数据间的嵌套、递归的联系。

② 具有面向对象技术的封装性(把数据和操作定义在一起)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。本书第九章将介绍面向对象数据库系统的发展状况。

(3) 各种新型的数据库系统

数据库技术是计算机软件领域的一个重要分支,经过三十余年的发展,已经形成相当规模的理论体系和实用技术。但受到相关学科和应用领域(如网络、多媒体等)的影响,数据库技术的研究并没有停滞,仍在不断发展,并出现许多新的分支。例如,演绎数据库、主动数据库、时态数据库、模糊数据库、并行数据库、多媒体数据库、内存数据库、联邦数据库、 workflow 数据库、空间数据库等。感兴趣的读者可以查阅有关的书籍。

二、当代数据库研究的内容

当前,数据库研究的范围有以下 3 个领域。

1. 数据库管理系统软件的研制

数据库管理系统 DBMS 是数据库系统的基础。DBMS 的研制包括 DBMS 本身以 DBMS 为核心的一组相互联系的软件系统。研制的目标是扩大功能、提高性能和提高用户的生产率。

2. 数据库设计

数据库设计的主要任务是在 DBMS 的支持下,按照应用的要求,为某一部门或组织设计一个结构合理、使用方便、效率较高的数据库及其应用系统。其中主要的研究方向是数据库设计方法学和设计工具,包括数据库设计方法、设计工具和设计理论的研究,数据模型和数据建模的研究,计算机辅助数据库设计方法及其软件系统的研究,数据库设计规范和标准的研究等。

3. 数据库理论

数据库理论的研究主要集中于关系的规范化理论、关系数据理论等。近年来,随着人工

智能与数据库理论的结合以及并行计算机的发展,数据库逻辑演绎和知识推理、并行算法等理论研究,以及演绎数据库系统、知识库系统和数据仓库的研制都已成为新的研究方向。

第三节 数据库系统的结构

考查数据库系统的结构可以从多种不同的角度查看,下面主要从3个方面来讲解数据库系统的体系结构、功能结构及模式结构。

一、数据库系统的体系结构

数据库系统的结构分为单用户结构、主从式结构、分布式结构和客户/服务器结构。

1. 单用户数据库系统

单用户数据库系统是一种早期的、最简单的数据库系统。在单用户系统中,整个数据库系统,包括应用程序、DBMS、数据,都装在一台计算机上,由一个用户独占,不同的机器之间不能共享数据。

例如,一个企业的各个部门都使用本部门的机器来管理本部门的数据,各个部门的机器是独立的。由于不同部门之间不能共享数据,因此企业内部存在大量的冗余数据。例如,人事部门、会计部门、技术部门必须重复存放每一名职工的一些基本信息(职工号、姓名等)。

2. 主从式结构的数据库系统

主从式结构是指一个主机带多个终端的多用户结构。在这种结构中,数据库系统,包括应用程序, DBMS、数据,都集中存放在主机上,所有处理任务都由主机来完成,各个用户通过主机的终端并发地存取数据库,共享数据资源。

主从式结构的优点是简单,数据易于管理与维护。缺点是当终端用户数目增加到一定程度后,主机的任务会过分繁重,成为瓶颈,从而使系统的性能大幅度下降。另外当主机出现故障时,整个系统都不能使用,因此系统的可靠性不高。

3. 分布式结构的数据库系统

分布式结构的数据库系统是指数据库中的数据在逻辑上是一个整体,但物理地分布在计算机网络的不同结点上。网络中的每个结点都可以独立处理本地数据库中的数据,执行局部应用;同时也可以同时存取和处理多个异地数据库中的数据,执行全局应用。

分布式结构的数据库系统是计算机网络发展的必然产物,它适应了地理上分散的公司、团体和组织对于数据库应用的需求。但数据的分布存放,给数据的处理、管理与维护带来困难。此外,当用户需要经常访问远程数据时,系统效率会明显地受到网络交通的制约。

4. 客户/服务器结构的数据库系统

主从式数据库系统中的主机和分布式数据库系统中的每个结点计算机是一个通用计算机,既执行功能又执行应用程序。随着工作站功能的增强和广泛使用,人们开始把DBMS的功能和应用分开,网络中某个(些)结点上的计算机专门用于执行功能,称为数据库服务器,简称服务器,其他结点上的计算机安装DBMS的外围应用开发工具,支持用户的应用,称为客户机,这就是客户/服务器结构的数据库系统。

在客户/服务器结构中,客户端的用户请求被传送到数据库服务器,数据库服务器进行处理后,只将结果返回给用户(而不是整个数据),从而显著减少了网络上的数据传输量,提高了系统的性能、吞吐量和负载能力。

另一方面, 客户/服务器结构的数据库往往更加开放。客户与服务器一般都能在多种不同的硬件和软件平台上运行, 可以使用不同厂商的数据库应用开发工具, 应用程序具有更强的可移植性, 同时也可以减少软件维护开销。

二、数据库管理系统的功能结构

1. DBMS 功能

由于不同的 DBMS 要求的硬件资源、软件环境是不同的, 因此其功能与性能也存在差异, 但一般说来, DBMS 的功能主要包括以下 6 个方面。

(1) 数据定义

数据定义包括定义构成数据库结构的模式、存储模式和外模式, 定义各个外模式与模式之间的映射, 定义模式与存储模式之间的映射, 定义有关的约束条件。例如, 为保证数据库中的数据具有正确语义而定义的完整性规则、为保证数据库安全而定义的用户口令和存取权限等。

(2) 数据操纵

数据操纵包括对数据库数据的检索、插入、修改和删除等基本操作。

(3) 数据库运行管理

对数据库的运行进行管理是 DBMS 运行时的核心部分, 包括对数据库进行并发控制、安全性检查、完整性约束条件的检查和执行、数据库的内部维护 (如索引、数据字典的自动维护) 等。所有访问数据库的操作都要在这些控制程序的统一管理下进行, 以保证数据的安全性、完整性、一致性以及多用户对数据库的并发使用。

(4) 数据组织、存储和管理

数据库中需要存放多种数据, 如数据字典、用户数据、存取路径等, DBMS 负责分门别类地组织、存储和管理这些数据, 确定以何种文件结构和存取方式物理地组织这些数据, 如何实现数据之间的联系, 以便提高存储空间利用率以及提高随机查找、顺序查找、增、删、改等操作的时间效率。

(5) 数据库的建立和维护

建立数据库包括数据库初始数据的输入与数据转换等。维护数据库包括数据库的转储与恢复、数据库的重组织与重构造、性能的监视与分析等。

(6) 数据通信接口

DBMS 需要提供与其他软件系统进行通信的功能。例如, 提供与其他 DBMS 或文件系统的接口, 从而能够将数据转换为另一个 DBMS 或文件系统能够接受的格式, 或者接收其他 DBMS 或文件系统的数据库。

2. DBMS 的组成

DBMS 是许多程序所组成的一个大型软件系统, 每个程序都有自己的功能, 共同完成 DBMS 的一个或几个工作。一个典型的 DBMS 组成模块通常由以下几部分组成。

(1) 语言编译处理程序

语言编译处理程序包括以下两个程序。

① 数据定义语言 DDL 编译程序。把用 DDL 编写的各级源模式编译成各级目标模式。

这些目标模式是对数据库结构信息的描述, 它们被保存在数据字典中, 供数据操纵控制时使用。

② 数据操纵语言 DML 编译程序。它将应用程序中的 DML 语句转换成可执行程序，实现对数据库的查询、插入、修改等基本操作。

(2) 系统运行控制程序

系统运行控制程序主要包括以下几部分。

① 系统总控程序：用于控制和协调各程序的活动，是 DBMS 运行程序的核心。

② 安全性控制程序：防止未被授权的用户存取数据库中的数据。

③ 完整性控制程序：检查完整性约束条件，确保进入数据库中的数据正确性、有效性和相容性。

④ 并发控制程序：协调多用户、多任务环境下各应用程序对数据库的并发操作，保证数据的一致性。

⑤ 数据存取和更新程序：实施对数据库数据的查询、插入、修改和删除等操作。

⑥ 通信控制程序：实现用户程序与 DBMS 间的通信。

此外，还有事务管理程序、运行日志管理程序等。所有这些程序在数据库系统运行过程中协同操作，监视着对数据库的所有操作，控制、管理数据库资源等。

(3) 系统建立、维护程序

系统建立、维护程序主要包括以下几部分：

① 装配程序：完成初始数据库的装入。

② 重组程序：当数据库系统性能降低时（如查询速度变慢），需要重新组织数据库，重新装入数据。

③ 系统恢复程序：当数据库系统受到破坏时，将数据库系统恢复到以前某个正确的状态。

(4) 数据字典系统程序

管理数据字典，实现数据字典功能。在数据库中，DBMS 与操作系统、应用程序、硬件等协调工作，共同完成数据各种存取操作，其中 DBMS 起着关键的作用。

三、数据库系统的模式结构

1. 数据库系统的三级模式结构

数据库系统的三级模式结构是指数据库系统是由外模式、模式和内模式三级构成，如图 1-5 所示。

(1) 模式

模式也称逻辑模式，是对数据库中全体数据的逻辑结构和特征的描述，是所有用户的公共数据视图，是数据库系统模式结构的中间层，不涉及数据的物理存储

细节和硬件环境，与具体的应用程序，与所使用的应用开发工具及高级程序设计语言（如 C 语言）无关。

实际上，模式是数据库数据在逻辑级上的视图。一个数据库只有一个模式。数据库模式以某一种数据模型为基础，统一综合地考虑了所有用户的需求，并将这些需求有机地结合成一个逻辑整体。定义模式时不仅要定义数据的逻辑结构，例如，数据记录由哪些数据项构成，数据项的名字、类型、取值范围等，而且要定义与数据有关的安全性、完整性要求，以

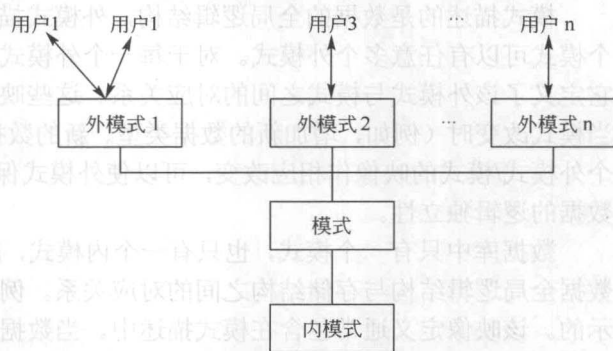


图 1-5 数据库系统的三级模式结构

及定义这些数据之间的联系。

(2) 外模式

外模式也称子模式或用户模式，是对数据库用户（包括应用程序和最终用户）看见和使用的局部数据的逻辑结构和特征的描述，是数据库用户的数据视图，是与某一应用有关的数据的逻辑表示。

外模式通常是模式的子集。一个数据库可以有多个外模式。由于它是各个用户的数据视图，如果不同的用户在不同需求、看待数据的方式、对数据保密的要求等方面存在差异，它们的外模式描述是不同的。即使对模式中同一数据，在外模式中的结构、类型、长度、保密级别等都可以不同。另一方面，同一外模式也可以为某一用户的多个应用系统所使用，但一个应用程序只能使用一个外模式。

外模式是保证数据库安全性的一个有力措施。每个用户只能看见和访问所对应的外模式中的数据，数据库中的其余数据对他们来说是不可见的。

(3) 内模式

内模式也称存储模式，是对数据的物理结构和存储结构的描述，是数据在数据库内部的表示方式。例如，记录的存储方式是顺序存储、按照 B 树结构存储还是按 HASH 方法存储；索引按照什么方式组织；数据是否压缩存储，是否加密；数据的存储记录结构有何规定等。一个数据库只有一个内模式。

2. 数据库的二级映像功能与数据独立性

数据库系统的三级模式是对数据的三个抽象级别，它把数据的具体组织留给 DBMS 管理，使用户能逻辑地、抽象地处理数据，而不必关心数据在计算机中的具体表示方式与存储方式，而为了能够在内部实现这三个抽象层次的联系和转换，数据库系统在这三级模式之间提供了两层映像：外模式/模式映像和模式/内模式映像。正是这两层映像保证了数据库系统中的数据能够具有较高的逻辑独立性和物理独立性。

模式描述的是数据的全局逻辑结构，外模式描述的是数据的局部逻辑结构。对应于同一个模式可以有任意多个外模式。对于每一个外模式，数据库系统都有一个外模式/模式映像，它定义了该外模式与模式之间的对应关系。这些映像定义通常包含在各自外模式的描述中。当模式改变时（例如，增加新的数据类型、新的数据项、新的关系等），由数据库管理员对各个外模式/模式的映像作相应改变，可以使外模式保持不变，从而应用程序不必修改，保证了数据的逻辑独立性。

数据库中只有一个模式，也只有一个内模式，所以模式/内模式映像是唯一的，它定义了数据全局逻辑结构与存储结构之间的对应关系。例如，说明逻辑记录 and 字段在内部是如何表示的。该映像定义通常包含在模式描述中。当数据库的存储结构改变了（例如，采用了更先进的存储结构），由数据库管理员对模式/内模式映像作相应改变，可以使模式保持不变，从而保证了数据的物理独立性。

习 题

一、填空题

1. 数据库的英文缩写是_____，数据库管理系统的英文缩写是_____，数据库系统的英文缩写是_____。