



工业和信息化普通高等教育“十二五”规划教材
21世纪高等教育计算机规划教材



SQL Server 2005 数据库应用技术

Application of SQL Server 2005
Database

■ 刘卫国 刘泽星 主编

- 坚持面向应用——实例详实，体现基本知识体系
- 适度突出原理——体现本质概念和思维能力训练
- 符合教学需要——体现课程教学要求和技术发展趋势



 人民邮电出版社
POSTS & TELECOM PRESS



工业和信息化普通高等教育“十二五”规划教材
21世纪高等教育计算机规划教材

COMPUTER

SQL Server 2005 数据库应用技术

Application of SQL Server 2005
Database

刘卫国 刘泽星 主编



人民邮电出版社

北京

图书在版编目 (C I P) 数据

SQL Server 2005数据库应用技术 / 刘卫国, 刘泽星
主编. -- 北京: 人民邮电出版社, 2013. 12
21世纪高等教育计算机规划教材
ISBN 978-7-115-33447-3

I. ①S… II. ①刘… ②刘… 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(2013)第286505号

内 容 提 要

本书以 SQL Server 2005 为实践环境, 系统介绍了数据库技术的基本原理、数据库的操作以及数据库应用系统的开发方法。全书共分为 10 章, 包括数据库技术概论、创建和管理数据库、创建和管理表、索引与数据完整性、查询与视图、T-SQL 程序设计、存储过程与触发器、事务和锁、SQL Server 安全管理、数据库应用系统开发等内容。全书以 SQL Server 2005 数据库操作为主线, 体现 SQL Server 2005 的基本知识体系; 同时适度突出关系数据库的基本原理, 体现关系数据库的概念和应用要求。

本书参照教育部高等学校计算机基础课程教学指导委员会提出的数据库相关课程的教学基本要求, 既可作为高等院校数据库应用课程的教材, 又可供社会各类计算机应用人员及参加各类计算机等级考试的读者学习参考。

-
- ◆ 主 编 刘卫国 刘泽星
责任编辑 邹文波
责任印制 彭志环 杨林杰
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京昌平百善印刷厂印刷
 - ◆ 开本: 787×1092 1/16
印张: 18.5 2013 年 12 第 1 版
字数: 487 千字 2013 年 12 月北京第 1 次印刷
-

定价: 39.00 元

读者服务热线: (010)81055256 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京崇工商广字第 0021 号

前言

数据库技术在 20 世纪 60 年代末产生并发展起来,它在计算机应用中的地位 and 作用日益重要。目前,数据处理已成为计算机应用的主要领域,采用数据库技术进行数据处理是当今的主流技术,其核心是建立、管理和使用数据库。使用数据库技术减少了不必要的多余数据,可以为多种应用服务,而且数据的存储独立于使用这些数据的应用程序。数据库技术已成为了信息系统的核心技术和基础,它不仅是计算机学科的一个重要分支,而且与人们的现实生活息息相关。数据库技术是管理信息系统、决策支持系统、企业资源规划、客户关系管理、数据仓库和数据挖掘等应用的重要支撑。

在数据库系统中,通过数据库管理系统来对数据进行统一管理。为了能开发出适用的数据库应用系统,就需要熟悉和掌握一种数据库管理系统。SQL Server 2005 是 Microsoft 公司推出的一个全面的企业级数据库平台产品,用于大规模联机事务处理、数据仓库和电子商务应用的数据库和数据分析;它使用集成的商业智能(Business Intelligence, BI)工具,结合了分析、报表、集成和通知功能;它的安全性、稳定性和可靠性,使得企业能更可靠地管理来自关键业务的信息,更有效地运行复杂的商业应用,从而在很大程度上帮助企业根据数据做出更快、更好的决策。

本书以 SQL Server 2005 为实践环境,介绍数据库技术的基本原理、数据库的操作以及数据库应用系统的开发方法。全书以“教学管理”数据库贯穿始终,围绕“教学管理”数据库设计编排了大量实例。书中实例新颖、系统,具有启发性;并且案例内容与本书内容相互呼应,便于读者学习、巩固和提高。全书共分为 10 章,包括数据库技术概论、创建和管理数据库、创建和管理表、索引与数据完整性、查询与视图、T-SQL 程序设计、存储过程与触发器、事务和锁、SQL Server 安全管理、数据库应用系统开发等内容。全书以 SQL Server 2005 数据库操作为主线,体现 SQL Server 2005 的基本知识体系;同时适度突出关系数据库的基本原理,体现关系数据库的概念和应用要求。

本书参照教育部高等学校计算机基础课程教学指导委员会提出的数据库相关课程的教学基本要求,既可作为高等院校数据库应用课程的教材,又可供社会各类计算机应用人员及参加各类计算机等级考试的读者学习参考。

为了方便教学和读者上机操作练习,编者还编写了《SQL Server 2005 数据库应用技术实验指导与习题选解》一书,作为与本书配套的实验教材。另外,还有与本书配套的教学课件、各章习题答案、实例数据库等教学资源,可从人民邮电出版社教学资源与服务网(<http://www.ptpedu.com.cn>)下载使用。

本书第 1 章~第 4 章由刘卫国编写,第 5 章~第 8 章由刘泽星编写,第 9 章、第 10 章由熊拥军编写。全书由刘卫国、刘泽星主编并定稿。此外,参与编写工作的还有文碧望、石玉、欧鹏杰、刘苏洲、伍敏、欧阳佳、胡勇刚等。

由于编者水平有限,书中难免存在疏漏或不妥之处,恳请广大读者批评指正。

编者

2013 年 12 月

第 1 章 数据库技术概论	1
1.1 数据库技术的产生与发展	1
1.1.1 数据与数据处理	1
1.1.2 数据管理技术的 4 个发展阶段	2
1.2 数据库系统	5
1.2.1 数据库系统的组成	5
1.2.2 数据库的结构体系	6
1.2.3 数据库系统的特点	8
1.3 数据模型	9
1.3.1 数据模型的组成要素	9
1.3.2 数据抽象的过程	9
1.3.3 概念模型	10
1.3.4 逻辑模型	12
1.4 关系数据库基础知识	14
1.4.1 关系数据库的基本概念	14
1.4.2 关系运算	16
1.4.3 关系的完整性约束	18
1.5 关系的规范化理论	19
1.5.1 关系模式的数据冗余和 操作异常问题	19
1.5.2 函数依赖的基本概念	21
1.5.3 关系模式的范式	22
1.5.4 关系模式的分解	24
1.6 数据库的设计方法	27
1.6.1 数据库设计的基本步骤	27
1.6.2 E-R 模型到关系模型的转化	28
1.6.3 关系数据库设计实例	30
1.7 SQL Server 数据库概述	31
1.7.1 SQL Server 的发展	31
1.7.2 SQL Server 2005 的安装	31

1.7.3 SQL Server 2005 的常用管理工具	35
1.8 SQL 和 T-SQL 概述	42
1.8.1 SQL 语言的发展与特点	42
1.8.2 T-SQL 语言简介	42
习题	43
第 2 章 创建和管理数据库	46
2.1 SQL Server 2005 数据库的基本概念	46
2.1.1 SQL Server 2005 数据库类型	46
2.1.2 数据库文件和文件组	47
2.1.3 数据库对象及其标识符	48
2.2 创建数据库	49
2.2.1 使用管理工具创建数据库	49
2.2.2 使用 T-SQL 语句创建数据库	52
2.3 管理数据库	55
2.3.1 查看和修改数据库	55
2.3.2 删除数据库	59
2.4 备份与还原数据库	60
2.4.1 备份数据库	61
2.4.2 还原数据库	67
2.5 分离与附加数据库	73
2.5.1 分离数据库	73
2.5.2 附加数据库	74
2.6 导入与导出数据库	76
2.6.1 导入数据	76
2.6.2 导出数据	78
习题	81
第 3 章 创建和管理表	83
3.1 SQL Server 2005 表的基本知识	83
3.1.1 表的类型	83
3.1.2 表的设计——数据类型	84

3.2 创建表	88	5.3.1 连接概述	135
3.2.1 使用管理工具创建表	88	5.3.2 内连接	136
3.2.2 使用 T-SQL 语句创建表	91	5.3.3 外连接	138
3.3 管理表	92	5.3.4 交叉连接	141
3.3.1 查看表	92	5.4 创建视图	141
3.3.2 修改表	93	5.4.1 视图的概念	141
3.3.3 删除表	95	5.4.2 创建视图的方法	142
3.4 维护表中数据	96	5.5 视图的管理	143
3.4.1 数据插入	96	5.5.1 查看和修改视图	143
3.4.2 数据更新	98	5.5.2 删除视图	144
3.4.3 数据删除	99	5.5.3 视图的应用	145
习题	100	习题	146
第 4 章 索引与数据完整性	102	第 6 章 T-SQL 程序设计	150
4.1 索引概述	102	6.1 数据与表达式	150
4.1.1 索引的概念	102	6.1.1 用户定义数据类型	150
4.1.2 索引的分类	103	6.1.2 常量与变量	151
4.2 索引的操作	103	6.1.3 运算符与表达式	155
4.2.1 创建索引	103	6.2 函数	158
4.2.2 查看与修改索引	105	6.2.1 常用函数	158
4.2.3 删除索引	107	6.2.2 用户定义函数	163
4.3 实施数据完整性	108	6.3 程序控制流语句	168
4.3.1 使用规则实施数据完整性	109	6.3.1 语句块和注释	168
4.3.2 使用默认值实施数据完整性	111	6.3.2 选择控制	170
4.3.3 使用约束实施数据完整性	113	6.3.3 循环控制	174
习题	119	6.3.4 批处理	176
第 5 章 查询与视图	121	6.4 游标管理与应用	177
5.1 基本查询	121	6.4.1 游标概述	177
5.1.1 简单查询	121	6.4.2 声明游标	178
5.1.2 带条件查询	124	6.4.3 使用与管理游标	180
5.1.3 查询结果处理	127	6.4.4 游标的综合应用示例	183
5.2 嵌套查询	132	习题	184
5.2.1 单值嵌套查询	132	第 7 章 存储过程与触发器	187
5.2.2 多值嵌套查询	133	7.1 存储过程概述	187
5.3 连接查询	135	7.2 创建与管理存储过程	189

7.2.1 创建存储过程.....	189	9.3 管理数据库用户.....	223
7.2.2 执行存储过程.....	191	9.4 管理 SQL Server 角色.....	224
7.2.3 修改存储过程.....	191	9.4.1 SQL Server 角色的类型.....	225
7.2.4 删除存储过程.....	193	9.4.2 固定服务器角色管理.....	226
7.2.5 存储过程参数与状态值.....	193	9.4.3 数据库角色管理.....	228
7.3 触发器概述.....	196	9.4.4 用户定义数据库角色.....	230
7.4 创建与管理触发器.....	196	9.5 管理 SQL Server 权限.....	231
7.4.1 创建触发器.....	197	9.5.1 权限的种类.....	231
7.4.2 修改触发器.....	199	9.5.2 授予权限.....	232
7.4.3 删除触发器.....	200	9.5.3 禁止与撤消权限.....	235
习题.....	201	9.5.4 查看权限.....	237
第 8 章 事务和锁.....	203	习题.....	237
8.1 事务.....	203	第 10 章 数据库应用系统开发....	239
8.1.1 事务的概念.....	203	10.1 数据库应用系统的开发过程.....	239
8.1.2 事务管理.....	204	10.2 用 VB.NET 访问 SQL Server	
8.2 锁.....	208	数据库.....	241
8.2.1 锁的概念.....	208	10.2.1 VB.NET 程序设计概述.....	241
8.2.2 隔离级别.....	209	10.2.2 VB.NET 程序设计基础知识.....	244
8.2.3 查看和终止锁.....	210	10.2.3 VB.NET 数据库应用	
8.3 死锁及其处理.....	212	程序开发.....	250
8.3.1 死锁的发生.....	212	10.3 数据库系统开发案例——教学	
8.3.2 死锁的处理.....	213	信息管理系统.....	256
8.3.3 死锁的避免.....	214	10.3.1 系统需求分析.....	256
习题.....	214	10.3.2 系统功能设计.....	256
第 9 章 SQL Server 安全管理....	216	10.3.3 数据库设计.....	257
9.1 SQL Server 身份验证.....	216	10.3.4 系统主窗体的创建.....	260
9.1.1 身份验证.....	216	10.3.5 系统管理模块功能.....	262
9.1.2 身份验证模式的设置.....	217	10.3.6 学籍信息管理模块的创建.....	265
9.2 管理登录账号.....	218	10.3.7 课程信息管理模块的创建.....	273
9.2.1 创建登录账户.....	218	10.3.8 成绩信息管理模块的创建.....	280
9.2.2 修改登录账户.....	222	习题.....	287
9.2.3 删除登录账户.....	222	参考文献.....	289

第 1 章

数据库技术概论

本章学习目标:

- 了解数据库技术的产生背景与发展概况。
- 掌握数据库系统的组成与特点。
- 理解数据模型的概念及基本知识。
- 掌握关系数据库的基本理论及设计方法。
- 了解 SQL Server 2005 常用管理工具与 SQL 语言的基本概念。

数据库技术是作为一门数据处理技术发展起来的,在计算机应用中的地位 and 作用日益重要。目前,数据处理已成为计算机应用的主要领域,采用数据库技术进行数据处理是当今的主流技术,其核心是建立、管理和使用数据库。使用数据库技术减少了不必要的多余数据、可以为多种应用服务,而且数据的存储独立于使用这些数据的应用程序。在数据库系统中,通过数据库管理系统来对数据进行统一管理,为了能开发出适用的数据库应用系统,就需要熟悉和掌握一种数据库管理系统。SQL Server 是目前广为使用的数据库管理系统,本书以 SQL Server 2005 为实践环境,介绍数据库的基本操作和数据库应用系统开发的方法。

1.1 数据库技术的产生与发展

数据库技术是一门研究如何存储、使用和管理数据的技术,是计算机数据管理技术的最新发展阶段,它能把大量的数据按照一定的结构存储起来,在数据库管理系统的集中管理下实现数据共享。

1.1.1 数据与数据处理

人类在长期的生产社会实践中会产生大量数据,如何对数据进行分类、组织、存储、检索和维护成为迫切的需要,只有在计算机成为数据处理的工具之后,才使数据处理现代化成为可能。数据库应用涉及到数据、信息、数据处理和数据管理等基本概念。

1. 数据和信息

数据(Data)和信息(Information)是数据处理中的两个基本概念,有时可以混用,如平时讲数据处理就是信息处理,但有时必须分清。一般认为,数据是对客观事物的某些特征及相互联系的一种抽象化、符号化的表示,即数据是人们用于记录事物情况的物理符号。为了描述客观事物而用到的数字、字符及所有能输入到计算机中并能被计算机处理的符号都可以看作是数据。在

实际应用中有两种基本形式的数据，一种是可以参与数值运算的数值型数据，如表示成绩、工资的数据；另一种是由字符组成、不能参与数值运算的字符型数据，如表示姓名、职称的数据。此外，还有图形、图像、声音、动画和视频等多媒体数据，如照片、商标等。

信息是指数据中所包含的意义。通俗地讲，信息是经过加工处理并对人类社会实践和生产活动产生决策影响的数据。例如，“谢博涵”、“湖南”、“575”只是单纯的数据，而“谢博涵同学来自湖南，入学成绩为 575 分”就是一条有意义的信息。因此，不经过加工处理的数据只是一种原始材料，对人类活动产生不了决策作用，它的价值只是在于记录了客观世界的事实；只有经过提炼和加工，原始数据发生了质的变化，才能给人们以新的知识和智慧。

数据与信息既有区别，又有联系。数据是用来表示信息的，是承载信息的物理符号；信息是加工处理后的数据，是数据所表达的内容。另一方面，信息不随表示它的数据形式而改变，它是反映客观现实世界的知识；而数据则具有任意性，用不同的数据形式可以表示同样的信息。例如，一个城市的天气预报情况是一条信息，而描述该信息的数据形式可以是文字、图像或声音等。

2. 数据处理

数据处理是指将数据转换成信息的过程，其基本目的是从大量的、杂乱无章的、难以理解的数据中整理出对人们有价值、有意义的数据（即信息）作为决策的依据。例如，全体考生各门课程的考试成绩记录了考生的考试情况，这属于原始数据；接下来对考试成绩进行分析和处理，如按成绩从高到低顺序排列、统计各分数段的人数等，进而可以根据招生人数确定录取分数线，此时输出的数据即包含了丰富的信息。

数据管理是指数据的收集、组织、存储、检索和维护等操作，这些操作是数据处理的中心环节，是任何数据处理业务中不可缺少的部分。数据管理的基本目的是为了实现数据共享、降低数据冗余，提高数据的独立性、安全性和完整性，从而能更加有效地管理和使用数据资源。

1.1.2 数据管理技术的 4 个发展阶段

数据库系统的核心任务是数据管理，但并不是一开始就有数据库技术，它的产生与发展是随着数据管理技术的不断发展而逐步形成的。数据管理技术经历了人工管理、文件管理、数据库管理和新型数据库系统 4 个发展阶段。

1. 人工管理阶段

20 世纪 50 年代中期以前，计算机主要应用于科学计算，虽然此时也有数据管理的问题，但这时的数据管理是以人工管理方式进行的。在硬件方面，外存储器只有磁带、卡片和纸带等，没有磁盘等直接存取的外存储器。在软件方面，只有汇编语言，没有操作系统，没有对数据进行管理的软件。数据处理方式基本上是批处理。在人工管理阶段，数据管理的特点如下。

（1）数据不保存

人工管理阶段处理的数据量较少，一般不需要将数据长期保存，只是在计算时将数据随程序一起输入，计算完后将结果输出，而数据和程序一起从内存中被释放。若再要计算，则需重新输入数据和程序。

（2）由应用程序管理数据

系统没有专用的软件对数据进行管理，数据需要由应用程序自行管理。每个应用程序不仅要规定数据的逻辑结构，而且要设计数据的存储结构及输入输出方法等，程序设计任务繁重。

（3）数据有冗余，无法实现共享

程序与数据是一个整体，一个程序中的数据无法被其他程序使用，因此程序与程序之间存在

大量的重复数据，数据无法实现共享。

（4）数据对程序不具有独立性

由于程序对数据的依赖性，数据的逻辑结构或存储结构一旦有所改变，则必须修改相应程序，这就进一步加重了程序设计的负担。

2. 文件管理阶段

20 世纪 50 年代后期至 60 年代后期，计算机开始大量用于数据管理。硬件上出现了直接存取的大容量外存储器，如磁盘、磁鼓等，这为计算机数据管理提供了物质基础。软件方面，出现了高级语言和操作系统。操作系统中的文件系统专门用于管理数据，这又为数据管理提供了技术支持。数据处理方式上不仅有批处理，而且有联机实时处理。

数据处理应用程序利用操作系统的文件管理功能，将相关数据按一定的规则构成文件，通过文件系统对文件中的数据进行存取和管理，实现数据的文件管理方式。其特点可以概括为如下两点。

（1）数据可以长期保存

文件系统为程序和数据之间提供了一个公共接口，使应用程序采用统一的存取方法来存取和操作数据。数据可以组织成文件，能够长期保存、反复使用。

（2）数据对程序有一定独立性

程序和数据不再是一个整体，而是通过文件系统把数据组织成一个独立的数据文件，由文件系统对数据的存取进行管理，程序员只需通过文件名来访问数据文件，不必过多考虑数据的物理存储细节，因此程序员可集中精力进行算法设计，大大减少了程序维护的工作量。

文件管理使计算机在数据管理方面有了长足的进步。时至今日，文件系统仍是一般高级语言普遍采用的数据管理方式。然而，当数据量增加、使用数据的用户越来越多时，文件管理便不能适应更有效地使用数据的需要了，其症结表现在 3 个方面。

（1）数据的共享性差、冗余度大，容易造成数据不一致

由于数据文件是根据应用程序的需要而建立的，当不同的应用程序所使用的数据有相同部分时，也必须建立各自的数据文件，即数据不能共享，造成大量数据重复。这样不仅浪费存储空间，而且使数据修改变得非常困难，容易产生数据不一致，即同样的数据在不同的文件中所存储的数值不同，造成矛盾。

（2）数据独立性差

在文件系统中，尽管数据和程序有一定的独立性，但这种独立性主要是针对某一特定应用而言的，就整个应用系统而言，文件系统还未能彻底体现数据逻辑结构独立于数据存储的物理结构的要求。在文件系统中，数据和应用程序是互相依赖的，即程序的编写与数据组织方式有关，如果改变数据的组织方式，就必须修改有关应用程序。而应用程序发生变化，如改用另一种程序设计语言来编写程序，也需修改文件的数据结构。

（3）数据之间缺乏有机的联系，缺乏对数据的统一控制和管理

文件系统中各数据文件之间是相互独立的，没有从整体上反映现实世界事物之间的内在联系，因此很难对数据进行合理的组织以适应不同应用的需要。在同一个应用项目中的各个数据文件没有统一的管理机构，数据完整性和安全性很难得到保证。

3. 数据库管理阶段

20 世纪 60 年代后期，计算机用于数据管理的规模更加庞大，数据量急剧增加，数据共享性要求更加强烈。同时，计算机硬件价格下降，而软件价格上升，编制和维护软件所需成本相对增加，其中维护成本更高。这些成为数据管理技术在文件管理的基础上发展到数据库管理的原动力。

数据库 (Database, DB) 是按一定的组织方式存储起来的、相互关联的数据集合。在数据库管理阶段, 由一种叫作数据库管理系统 (Database Management System, 数据库管理系统) 的系统软件来对数据进行统一的控制和管理, 把所有应用程序中使用的相关数据汇集起来, 按统一的数据模型存储在数据库中, 为各个应用程序所使用。在应用程序和数据库之间保持较高的独立性, 数据具有完整性、一致性和安全性高等特点, 并且具有充分的共享性, 有效地减少了数据冗余。

4. 新型数据库系统

数据库技术的发展先后经历了层次数据库、网状数据库和关系数据库。层次数据库和网状数据库可以看作是第一代数据库系统, 关系数据库可以看作是第二代数据库系统。自 20 世纪 70 年代提出关系数据模型和关系数据库后, 数据库技术得到了蓬勃发展, 应用也越来越广泛。但随着应用的不断深入, 占主导地位的关系数据库系统已不能满足新的应用领域的需求。例如, 在实际应用中, 除了需要处理数字、字符数据的简单应用之外, 还需要存储并检索复杂的复合数据 (如集合、数组、结构)、多媒体数据、计算机辅助设计绘制的工程图纸和地理信息系统 (Geographic Information System, GIS) 提供的空间数据等。对于这些复杂数据, 关系数据库无法实现对它们的管理。正是实际应用中涌现出的许多问题, 促使数据库技术不断向前发展, 出现了许多不同类型的新型数据库系统, 下面概要性地做一下介绍。

(1) 分布式数据库系统

分布式数据库系统 (Distributed Database System, DDBS) 是数据库技术与计算机网络技术、分布式处理技术相结合的产物。分布式数据库系统是将系统中的数据地理上分布在计算机网络的不同结点, 但逻辑上属于一个整体的数据库系统, 它不同于将数据存储在服务器上供用户共享存取的网络数据库系统, 分布式数据库系统不仅能支持局部应用 (访问本地数据库), 而且能支持全局应用 (访问异地数据库)。

分布式数据库系统的主要特点如下。

① 数据是分布的。数据库中的数据分布在计算机网络的不同结点上, 而不是集中在一个结点, 区别于数据存放在服务器上由各用户共享的网络数据库系统。

② 数据是逻辑相关的。分布在不同结点的数据逻辑上属于同一数据库系统, 数据间存在相互关联, 区别于由计算机网络连接的多个独立数据库系统。

③ 结点的自治性。每个结点都有自己的计算机软、硬件资源, 包括数据库、数据库管理系统等, 因而能够独立地管理局部数据库。局部数据库中的数据可以仅供本结点用户存取使用, 也可供其他结点上的用户存取使用, 提供全局应用。

(2) 面向对象数据库系统

面向对象数据库系统 (Object-Oriented Database System, OODBS) 是将面向对象的模型、方法和机制, 与先进的数据库技术有机地结合而形成的新型数据库系统。它从关系模型中脱离出来, 强调在数据库框架中发展类型、数据抽象、继承和持久性。它的基本设计思想是, 一方面把面向对象语言向数据库方向扩展, 使应用程序能够存取并处理对象; 另一方面扩展数据库系统, 使其具有面向对象的特征, 提供一种综合的语义数据建模概念集, 以便对现实世界中复杂应用的实体和联系建模。因此, 面向对象数据库系统首先是一个数据库系统, 具备数据库系统的基本功能; 其次是一个面向对象的系统, 针对面向对象的程序设计语言的永久性对象存储管理而设计, 充分支持完整的面向对象概念和机制。

(3) 多媒体数据库系统

多媒体数据库系统 (Multimedia Database System, MDBS) 是数据库技术与多媒体技术相结

合的产物。随着信息技术的发展,数据库应用从传统的企业信息管理扩展到计算机辅助设计(Computer Aided Design, CAD)、计算机辅助制造(Computer Aided Manufacture, CAM)、办公自动化(Office Automation, OA)、人工智能(Artificial Intelligence, AI)等多种应用领域。这些领域中要求处理的数据不仅包括传统的数字、字符等格式化数据,还包括大量的多媒体形式的非格式化数据,如图形、图像、声音等。这种能存储和管理多媒体的数据库称为多媒体数据库。

多媒体数据库与传统格式化数据库的结构和操作有很大差别。现有数据库管理系统无论从模型的语义描述能力、系统功能、数据操作,还是存储管理、存储方法上都不能适应非格式化数据的处理要求。综合程序设计语言、人工智能和数据库领域的研究成果,设计支持多媒体数据管理的数据库管理系统已成为数据库领域中一个新的重要研究方向。

在多媒体信息管理环境中,不仅数据本身的结构和存储形式各不相同,而且不同领域对数据处理的要求也比一般事务管理复杂得多,因而对数据库管理系统提出了更高的功能要求。

(4) 数据仓库技术

随着信息技术的高速发展,数据库应用的规模、范围和深度不断扩大,一般的事务处理已不能满足应用的需要,企业界需要在大量数据基础上的决策支持,数据仓库(Data Warehouse, DW)技术的兴起满足了这一需求。数据仓库作为决策支持系统(Decision Support System, DSS)的有效解决方案,涉及 3 方面的技术内容,即数据仓库技术、联机分析处理(On-Line Analysis Processing, OLAP)技术和数据挖掘(Data Mining, DM)技术。

数据仓库、OLAP 和数据挖掘是作为 3 种独立的数据处理技术出现的。数据仓库用于数据的存储和组织,OLAP 集中于数据的分析,数据挖掘则致力于知识的自动发现。它们都可以分别应用到信息系统的设计和实现中,以提高相应部分的处理能力。但是,由于这 3 种技术内在的联系性和互补性,将它们结合起来即是一种新的 DSS 架构。这一架构是以数据库中的大量数据为基础,其系统由数据驱动。

1.2 数据库系统

数据库系统(Database System, DBS)是指基于数据库的计算机应用系统。和一般的应用系统相比,数据库系统有其自身的特点,它涉及到一些相互联系而又有区别的基本概念。

1.2.1 数据库系统的组成

数据库系统是一个计算机应用系统,它是把有关计算机硬件、软件、数据和人员组合起来为用户提供信息服务的系统。因此,数据库系统是由计算机系统、数据库及其描述机构、数据库管理系统和有关人员组成,是具有高度组织性的整体。

1. 计算机硬件

计算机硬件系统是数据库系统的物质基础,是存储数据库及运行数据库管理系统的硬件资源,主要包括计算机主机、存储设备、输入输出设备及计算机网络环境。

2. 计算机软件

数据库系统中的软件包括操作系统、数据库管理系统及数据库应用系统等。

数据库管理系统(DBMS)是数据库系统的核心软件之一,它提供数据定义、数据操纵、数据库管理、数据库建立和维护以及通信等功能。数据库管理系统提供对数据库中数据资源进行统

一管理和控制的功能，将用户、应用程序与数据库数据相互隔离，是数据库系统的核心，其功能的强弱是衡量数据库系统性能优劣的主要指标。数据库管理系统必须运行在相应的系统平台上，有操作系统和相关系统软件的支持。

数据库管理系统功能的强弱随系统而异，大系统的功能较强、较全，小系统的功能较弱、较少。目前较流行的数据库管理系统有 Access、Visual FoxPro、SQL Server、Oracle、Sybase 等。

数据库应用系统是指系统开发人员利用数据库系统资源开发出来的、面向某一类实际应用的应用软件系统。从实现技术角度而言，它是以数据库技术为基础的计算机应用系统。

3. 数据库

数据库（DB）是指数据库系统中按照一定的方式组织的、存储在外部存储设备上的、能为多个用户共享的、与应用程序相互独立的相关数据集合。它不仅包括描述事物的数据本身，而且还包括相关事物之间的联系。

数据库中的数据往往不是像文件系统那样，只面向某一项特定应用，而是面向多种应用，可以被多个用户、多个应用程序共享。其数据结构独立于使用数据的程序，对于数据的增加、删除、修改和检索由数据库管理系统进行统一管理和控制，用户对数据库进行的各种操作都是由数据库管理系统实现的。

4. 数据库系统的有关人员

数据库系统的有关人员主要有 3 类：最终用户、数据库应用系统开发人员和数据库管理员（Database Administrator，DBA）。最终用户指通过应用系统的用户界面使用数据库的人员，他们一般对数据库知识了解不多。数据库应用系统开发人员包括系统分析员、系统设计员和程序员。系统分析员负责应用系统的分析，他们和用户、数据库管理员相配合，参与系统分析；系统设计员负责应用系统设计和数据库设计；程序员则根据设计要求进行编码。数据库管理员是数据管理机构的一组人员，他们负责对整个数据库系统进行总体控制和维护，以保证数据库系统的正常运行。

综上所述，数据库中包含的数据是存储在存储介质上的数据文件的集合；每个用户均可使用其中的数据，不同用户使用的数据可以重叠，同一组数据可以为多个用户共享；数据库管理系统为用户提供对数据的存储组织、操作管理功能；用户通过数据库管理系统和应用程序实现数据库系统的操作与应用。

1.2.2 数据库的结构体系

为了有效地组织、管理数据，提高数据库的逻辑独立性和物理独立性，人们为数据库设计了一个严谨的结构体系，数据库领域公认的标准结构是三级模式结构及二级映射。三级模式包括外模式、概念模式和内模式，二级映射则分别是概念模式/内模式的映射以及外模式/概念模式的映射。这种三级模式与二级映射构成了数据库的结构体系，如图 1-1 所示。

1. 数据库的三级模式

美国国家标准协会（American National Standards Institute，ANSI）的数据库管理系统研究小组于 1978 年提出了标准化的建议，将数据库结构体系分为三级，即面向用户或应用程序员的用户级、面向建立和维护数据库人员的概念级、面向系统程序员的物理级。用户级对应外模式，概念级对应概念模式，物理级对应内模式，使不同级别的用户对数据库形成不同的视图。所谓视图，就是指观察、认识和理解数据的范围、角度和方法，是数据库在用户眼中的反映。很显然，不同层次（级别）用户所看到的数据库是不相同的。

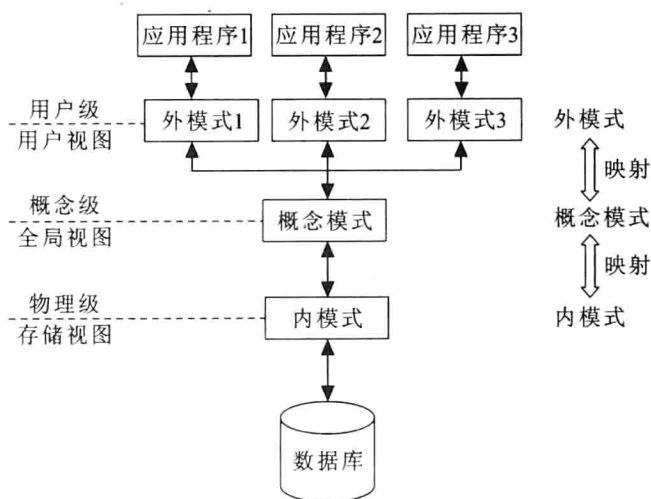


图 1-1 数据库的三级模式以及二级映射

（1）概念模式

概念模式又称逻辑模式，或简称为模式，对应于概念级。它是由数据库设计者综合所有用户的数据，按照统一的观点构造的全局逻辑结构，是对数据库中全部数据的逻辑结构和特征的总体描述，是所有用户的公共数据视图（全局视图）。它是由数据库系统提供的数据库定义语言（Data Definition Language, DDL）来描述、定义的，体现并反映了数据库系统的整体观。

（2）外模式

外模式又称子模式或用户模式，对应于用户级。它是某个或某几个用户所看到的数据库的数据视图，是与某一应用有关的数据的逻辑表示。外模式是从概念模式导出的一个子集，包含概念模式中允许特定用户使用的那部分数据。用户可以通过外模式定义语言（外模式 DDL）来描述、定义对应于用户的数据记录（外模式），也可以利用数据操纵语言（Data Manipulation Language, DML）对这些数据记录进行操作。外模式反映了数据库的用户观。

（3）内模式

内模式又称存储模式或物理模式，对应于物理级。它是数据库中全体数据的内部表示或底层描述，是数据库最低一级的逻辑描述，它描述了数据在存储介质上的存储方式和物理结构，对应着实际存储在外存储介质上的数据库。内模式是由内模式定义语言（内模式 DDL）来描述、定义的，它是数据库的存储观。

在一个数据库系统中，只有唯一的数据库，因而作为定义、描述数据库存储结构的内模式和定义、描述数据库逻辑结构的模式，也是唯一的。但建立在数据库系统之上的应用则是非常广泛、多样的，所以对应的外模式不是唯一的，也不可能唯一。

2. 三级模式间的二级映射

数据库的三级模式是数据在三个级别（层次）上的抽象，使用户能够逻辑地、抽象地处理数据，而不必关心数据在计算机中的物理表示和存储方式，把数据的具体组织交给数据库管理系统去完成。为了实现这三个抽象级别的联系和转换，数据库管理系统在三级模式之间提供了二级映射，正是这二级映射保证了数据库中的数据具有较高的物理独立性和逻辑独立性。

（1）概念模式/内模式的映射

数据库中的概念模式和内模式都只有一个，所以概念模式/内模式的映射是唯一的。它确定了数据的全局逻辑结构与存储结构之间的对应关系。当存储结构变化时，概念模式/内模式的映射也应有相应的变化，使其概念模式仍保持不变，即把存储结构变化的影响限制在概念模式之下，这

使数据的存储结构和存储方法独立于应用程序，通过映射功能保证数据存储结构的变化不影响数据的全局逻辑结构的改变，从而不必修改应用程序，即确保了数据的物理独立性。

(2) 外模式/概念模式的映射

数据库中的同一概念模式可以有多个外模式，对于每一个外模式，都存在一个外模式/概念模式的映射，用于定义该外模式和概念模式之间的对应关系。当概念模式发生改变时，例如，增加新的属性或改变属性的数据类型等，只要对外模式/概念模式的映射做相应的修改，而外模式（即数据的局部逻辑结构）保持不变。由于应用程序是依据数据的局部逻辑结构编写的，所以应用程序不必修改，即保证了数据与程序间的逻辑独立性。

1.2.3 数据库系统的特点

数据库系统的出现是计算机数据管理技术的重大进步，它克服了文件系统的缺陷，提供了对数据更高级、更有效的管理。

1. 数据结构化

在文件系统中，文件的记录是有结构的。例如，学生数据文件的每个记录是由学号、姓名、性别、出生年月、籍贯、简历等数据项组成的。但这种结构只适用于特定的应用，对其他应用并不适用。

在数据库系统中，每一个数据库都是为某一应用领域服务的。例如，学校信息管理涉及多个方面的应用，包括对学生的学籍管理、课程管理、学生成绩管理等，还包括教工的人事管理、教学管理、科研管理、住房管理和工资管理等，这些应用彼此之间都有着密切的联系。因此在数据库系统中不仅要考虑某个应用的数据结构，还要考虑整个组织（即多个应用）的数据结构。这种数据组织方式使数据结构化了，这就要求在描述数据时不仅要描述数据本身，还要描述数据之间的联系。而在文件系统中，尽管其记录内部已有了某些结构，但记录之间没有联系。数据库系统实现整体数据的结构化，这是数据库的主要特点之一，也是数据库系统与文件系统的本质区别。

2. 数据共享性高、冗余度低

数据共享是指多个用户或应用程序可以访问同一个数据库中的数据，而且数据库管理系统提供并发和协调机制，保证在多个应用程序同时访问、存取和操作数据库数据时，不产生任何冲突，从而保证数据不遭到破坏。

数据冗余既浪费存储空间，又容易产生数据的不一致。在文件系统中，由于每个应用程序都有自己的数据文件，所以数据存在着大量的重复。

数据库从全局观念来组织和存储数据，数据已经根据特定的数据模型结构化，在数据库中用户的逻辑数据文件和具体的物理数据文件不必一一对应，从而有效地节省了存储资源，减少了数据冗余，保证了数据的一致性。

3. 具有较高的数据独立性

数据独立性是指应用程序与数据库的数据结构之间相互独立。在数据库系统中，因为采用了数据库的三级模式结构，保证了数据库中数据的独立性。在数据存储结构改变时，不影响数据的全局逻辑结构，这样保证了数据的物理独立性。在全局逻辑结构改变时，不影响用户的局部逻辑结构以及应用程序，这样就保证了数据的逻辑独立性。

4. 有统一的数据控制功能

在数据库系统中，数据由数据库管理系统进行统一控制和管理。数据库管理系统提供了一套有效的数据控制手段，包括数据安全性控制、数据完整性控制、数据库的并发控制和数据库的恢复等，增强了多用户环境下数据的安全性和一致性保护。

1.3 数据模型

数据库是现实世界中某种应用环境（一个单位或部门）所涉及的数据的集合，它不仅要反映数据本身的内容，而且要反映数据之间的联系。由于计算机不能直接处理现实世界中的具体事物，所以必须将这些具体事物转换成计算机能够处理的数据。在数据库技术中，用数据模型（Data Model）来对现实世界中的数据进行抽象和表示。

1.3.1 数据模型的组成要素

一般而言，数据模型是一种形式化描述数据、数据之间的联系以及有关语义约束规则的方法，这些规则分为3个方面，即描述实体静态特征的数据结构、描述实体动态特征的数据操作规则和描述实体语义要求的数据完整性约束规则。因此，数据结构、数据操作及数据的完整性约束也被称为数据模型的3个组成要素。

1. 数据结构

数据结构研究数据之间的组织形式（数据的逻辑结构）、数据的存储形式（数据的物理结构）以及数据对象的类型等。存储在数据库中的对象类型的集合是数据库的组成部分。例如在教学管理系统中，要管理的数据对象有学生、课程、选课成绩等，在课程对象集中每门课程包括课程号、课程名、学分等信息，这些基本信息描述了每门课程的特性，构成在数据库中存储的框架，即对象类型。

数据结构用于描述系统的静态特性，是刻画一个数据模型性质最重要的方面。因此，在数据库系统中，通常按照其数据结构类型来命名数据模型。例如层次结构、网状结构和关系结构的数据模型分别命名为层次模型、网状模型和关系模型。

2. 数据操作

数据操作用于描述系统的动态特性，是指对数据库中的各种数据所允许执行的操作的集合，包括操作及有关的操作规则。数据库主要有查询和更新（包括插入、删除和修改等）两大类操作。数据模型必须定义这些操作的确切含义、操作符号、操作规则（如优先级）以及实现操作的语言。

3. 数据的完整性约束

数据的完整性约束是一组完整性规则的集合。完整性规则是给定的数据模型中数据及其联系所具有的约束和依存规则，用以限定符合数据模型的数据库状态以及状态的变化，以保证数据的正确、有效和相容。

数据模型应该反映和规定数据必须遵守的、基本的、通用的完整性约束。此外，数据模型还应该提供定义完整性约束条件的机制，以反映具体所涉及的数据必须遵守的、特定的语义约束条件。例如，在学生信息中的“性别”只能为“男”或“女”，学生选课信息中的“课程号”的值必须取自学校已开设课程的课程号等。

1.3.2 数据抽象的过程

从现实世界中的客观事物到数据库中存储的数据是一个逐步抽象的过程，这个过程经历了现实世界、观念世界和机器世界3个阶段，对应于数据抽象的不同阶段采用不同的数据模型。首先将现实世界的事物及其联系抽象成观念世界的概念模型，然后再转换成机器世界的数据模型。概

念模型并不依赖于具体的计算机系统，它不是数据库管理系统所支持的数据模型，它是现实世界中客观事物的抽象表示。概念模型经过转换成为计算机上某一数据库管理系统支持的数据模型。所以说，数据模型是对现实世界进行抽象和转换的结果，这一过程如图 1-2 所示。

1. 对现实世界的抽象

现实世界就是客观存在的世界，其中存在着各种客观事物及其相互之间的联系，而且每个事物都有自己的特征或性质。计算机处理的对象是现实世界中的客观事物，在对其实施处理的过程中，首先应了解和熟悉现实世界，从对现实世界的调查和观察中抽象出大量描述客观事物的事实，再对这些事实进行整理、分类和规范，进而将规范化的事实数据化，最终实现由数据库系统存储和处理。

2. 观念世界中的概念模型

观念世界是对现实世界的一种抽象，通过对客观事物及其联系的抽象描述，构造出概念模型 (Conceptual Model)。概念模型的特征是按用户需求观点对数据进行建模，表达了数据的全局逻辑结构，是系统用户对整个应用项目涉及的数据的全面描述。概念模型主要用于数据库设计，它独立于实现时的数据库管理系统，也就是说选择何种数据库管理系统都不会影响概念模型的设计。

概念模型的表示方法有很多，目前较常用的是实体联系模型 (Entity Relationship Model)，简称 E-R 模型。E-R 模型主要用 E-R 图来表示。

3. 机器世界中的逻辑模型和物理模型

机器世界是指现实世界在计算机中的体现与反映。现实世界中的客观事物及其联系，在机器世界以逻辑模型 (Logical Model) 描述。在选定数据库管理系统后，就要将 E-R 图表示的概念模型转换为具体的数据库管理系统支持的逻辑模型。逻辑模型的特征是按计算机实现的观点对数据进行建模，表达了数据库的全局逻辑结构，是设计人员对整个应用项目数据库的全面描述，逻辑模型服务于数据库管理系统的应用实现。通常，也把数据的逻辑模型直接称为数据模型。数据库系统中主要的逻辑模型有层次模型、网状模型和关系模型。

物理模型 (Physical Model) 是对数据最低层的抽象，用以描述数据在物理存储介质上的组织结构，与具体的数据库管理系统、操作系统和硬件有关。

从概念模型到逻辑模型的转换是由数据库设计人员完成的，从逻辑模型到物理模型的转换是由数据库管理系统完成的，一般人员不必考虑物理实现细节，因而逻辑模型是数据库系统的基础，也是应用过程中要考虑的核心问题。

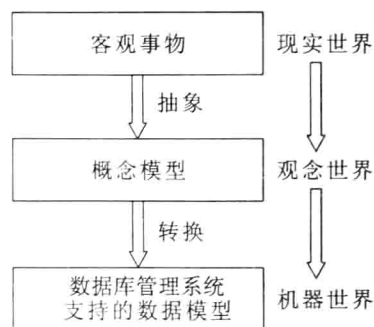


图 1-2 数据抽象的过程

1.3.3 概念模型

当分析某种应用环境所需的数据时，首先要找出涉及的实体及其实体之间的联系，进而得到概念模型，这是数据库设计的先导。

1. 实体与实体集

实体 (Entity) 是现实世界中任何可以相互区分和识别的事物，它可以是能触及的客观对象，例如一位教师、一名学生、一种商品等；还可以是抽象的事件，例如一场足球比赛、一次借书等。

性质相同的同类实体的集合称为实体集 (Entity Set)。例如，一个系的所有教师，2010 年南非世界杯足球赛的全部 64 场比赛等。