21世纪应用型本科系列教材

数据库原理及SQL Server应用

主　编　张水平
副主编　张凤琴
主　审　沈钧毅

西安交通大学出版社
XI'AN JIAOTONG UNIVERSITY PRESS

需要全本请访问 www.ertongbook.com

前　言

随着计算机技术的不断提高，计算机的应用面在不断拓宽。尤其是在信息技术大发展的时代，信息的存储、信息的管理、信息的获取已成为各个领域适应市场发展的基本需求。因此数据库技术已渗透到各种应用中。并已成为各信息管理系统、决策支持系统、地理信息系统、办公信息系统等各类信息处理应用系统的核心技术。

数据库技术与各类边缘技术互相渗透，互相结合，使得数据库技术的应用前景更为广阔。微软的 SQL Server 数据库管理系统正是应这一市场需求而开发的。

SQL Server 数据库管理系统是目前市场上非常流行的一种支持 C/S 和 B/S 模式的数据库产品。它有入门容易、开发成本低、运行和维护费用低、技术更新快、兼容性好、功能完整等特点，在电子商务、数据仓库、信息管理系统等小中型应用系统解决方案中扮演着重要的角色，为企业的信息管理提供了强大的技术支持，近几年来具有很高的市场占有量。

本书在介绍数据库基本概念、数据库应用系统设计方法基础上，以教学管理数据库应用系统为背景，讨论了应用 SQL Server 提供的语言和工具，设计完成数据库应用系统的基本方法和思想。

鉴于多年的教学实践，考虑教学特点以及目前人才市场对数据库应用系统开发者能力的要求，我们在书中将理论基础和应用实践的比例进行了最佳的调整。

书中的内容包括两部分：理论篇——数据库原理；应用篇——SQL Server 技术。二者的比例基本是 4∶6。

其中理论篇包括：数据库系统基础、关系运算、关系的规范化理论及数据库系统设计方法等；应用篇包括：SQL Server 基础、物理数据库的设计、数据库的操作、数据的查询与统计、视图、存储过程、触发器、Transact-SQL 语言编程方法、数据的备份与恢复等。

本书结构清晰合理，系统性强，实用性好。各章涉及到的重要概念，均有相关的例题以帮助读者理解和掌握。

为了便于教师课程教学的安排，本书附有教学学时分配指南；为了便于学生课后复习，各章后附有实习实例、内容小结、习题等。

本书的第 1、2 章由张凤琴同志编写；第 3、4 章由张青凤同志编写；第 5、8 章由李淑玲同志编写；第 6、7、9 章由张水平、王蓉同志编写；第 10 章由詹金珍同志编写。全书由沈钧毅教授审阅，并提出宝贵的修改意见，在此表示感谢。

本书可作为计算机专业的教学用书，也可作为参加相关计算机专业自学考试用书及相关技术人员的参考用书。

本书按照由浅入深，循序渐进的原则编写。各章内容既有联系，又有相对独立性，可根据不同教学要求有重点、有选择地组织教学，其中带“*”的章节为可选内容。附表给出了根据不同学时要求各章节的学时分配，供教学时参考。

本书编写内容不足之处在所难免，恳请专家和读者批评指正。

附表　各层次教学学时分配指南

章　节	学　时　分　配							
	70 学时		60 学时		50 学时		40 学时	
总学时	理论	实习	理论	实习	理论	实习	理论	实习
	46	24	40	20	30	20	24	16
第 1 章 数据库系统概述	4	0	4	0	2	0	4	0
第 2 章 关系运算	4		2		2	0		
第 3 章 关系的规范化理论	4		2		2	0		
第 4 章 数据库设计方法	4		4		2	0		
第 5 章 SQL Server 概述	2	2	2	2	2	2	2	2
第 6 章 物理数据库创建与维护	4	4	4	4	4	4	4	4
第 7 章 数据库的查询操作	8	6	8	6	8	6	6	6
第 8 章 Transact-SQL 语言	6	6	6	4	4	6	4	2
第 9 章 存储过程及触发器	4	4	4	2	4	4	4	2
第 10 章 数据库的维护	2	2	2	2	2	2	0	
复习	2		2		2		2	

作者

2007 年 11 月

理论篇　数据库原理

第 1 章　数据库系统概述

数据库技术已成为计算机应用领域的重要技术之一，是软件学科的一个独立分支。它主要研究如何存储、使用和管理数据，是计算机管理技术发展的最新阶段。几十年来，数据库无论在理论上还是实现技术上均得到很大的发展，由最初的仅针对事务处理中大量数据管理，进一步应用到情报检索、人工智能、专家系统、计算机辅助设计、决策支持等方面。在这个充满信息的社会里，数据库的建设规模、数据库信息量的大小和使用频度已成为衡量一个国家信息化程度的重要标志。

本章主要介绍数据库的基本概念，使读者了解数据库技术的发展、数据库系统组成、数据库管理系统功能、数据库体系结构以及数据模型等知识。

1.1　数据库技术的发展

信息资源是当今社会最重要和最宝贵的资源之一，作为电子信息资源系统的核心和基础的数据库管理技术，经历了三个阶段：人工管理阶段、文件系统管理阶段和数据库系统管理阶段。

1.1.1　相关术语

在介绍数据库技术发展之前，首先介绍一下相关的术语。

1. 数据

数据(Data)是描述现实世界事物属性的物理符号，它可以是数字、文字、图形、图像、声音、动画等形式。这些数据的多种表现形式都可以经过数字化后存入计算机，用于记录现实世界事物。

现实世界的一切事物均可以用数据来表示。在使用计算机存储和管理数据时，需要抽象出事物的一些特性(即属性)。例如在某学校的学生档案管理系统中，学生的基本信息包括学号、姓名、性别、出生日期、所在系、专业等数据集合表示。对于某一个学生来说，描述他的一条数据应为：

(0502112，张凌，男，1987.06.26，计算机系，软件工程专业)

对于这条数据，了解学生特性的人会理解为：学号为 0502112 的学生，是姓名为张凌的男同学，1987.06.26 出生，现在是本校计算机系软件工程专业的学生。但对于不了解学生特性的人则无法理解其中的全部含义。因此，数据的本身并不完全能表达其内容，还需要经过数据的属性解释，所以数据的描述有“型”和“值”之分。其中数据的“型”是指数据的结构，数据的

需要全本请访问 www.ertongbook.com

“值”是指数据的具体取值。

2. 信息

信息(Information)是现实世界事物的存在方式或运动形态反映的总和,是人们进行各种活动所需要的知识。信息具有可感知性、可存储性、可加工性、可传递性和可再生性等自然属性。它是社会上各行各业不可缺少的资源。信息与数据是两个概念,但又具有密切关系。信息是数据的内涵,是数据的语义解释。因此,在某些不需要严格区分的场合,对两者不加区别地使用,如信息处理也说成数据处理。例如,500 是数据,而本年计算机系的招生人数为 500 是信息。

3. 数据管理

数据管理(Data Management)是指数据的收集、整理、组织、存储、维护、检索、传送等操作,这部分操作是数据处理业务的基本环节,而且是任何数据处理业务中必不可少的共有部分。

4. 数据库

数据库(DB,Database)是存放数据的仓库,它是指长期储存在计算机内的、有组织的、可共享的数据集合。数据库中的数据按一定的数学模型组织、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。例如,数据库中存放的学生基本信息如表 1-1 所示。

表 1-1　学生基本信息

学号	姓名	性别	出生日期	所在的系	专业
0502112	张凌	男	1987.06.26	计算机系	软件工程
0502213	刘江	男	1988.01.20	计算机系	计算机科学与技术
0503214	宁尚	女	1987.01.12	通信工程系	通信工程
⋮					

5. 数据库管理系统

数据库管理系统(DBMS,Database Management System)是位于用户与操作系统(OS)之间的一层数据管理软件。DBMS 为用户或应用程序提供访问 DB 的方法,包括 DB 的建立、查询、更新及各种数据控制。

6. 数据库系统

数据库系统(DBS,Database System)是实现有组织地、动态地存储大量关联数据、方便多用户访问的计算机硬件、软件和数据资源组成的系统。一般由计算机硬件、操作系统、DB、DBMS 及开发工具和各种人员(如数据库管理员、用户等)构成。

1.1.2　数据库管理技术的发展

数据库是数据管理的手段和技术。数据库管理技术的发展经历了三个阶段,即人工管理阶段、文件系统管理阶段和数据库系统管理阶段。

1. 人工管理阶段

20 世纪 50 年代中期以前，计算机主要用于科学计算。当时的硬件的外部设备只有纸带、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，没有管理数据库的软件，程序员在程序中不仅要规定数据的逻辑结构，还要设计其物理结构，包括存储结构、存取方法、输入输出方式等等。当数据的物理组织或存储设备改变时，用户程序就必须重新编写。由于数据的组织面向应用，不同的应用程序之间不能共享数据，使得不同的应用之间存在大量的重复数据，应用程序之间数据一致性的维护工作也很难进行。数据处理方式是批处理。人工管理阶段的应用程序和数据之间关系如图 1－1 所示。

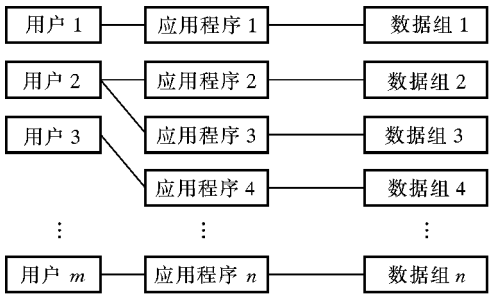


图 1－1　应用程序和数据的关系

这个阶段有以下几个特点：

(1) 数据不保存。在该阶段计算机主要用于科学计算，一般不需要将数据长期保存，只在计算一个题目时，将数据输入计算机，算完题，得到计算结果即可。

(2) 数据管理由应用程序自己管理。应用程序需要自己规定数据的逻辑结构、物理结构。

(3) 数据不共享，冗余度极大。数据面向应用，即一组数据对应一个程序。

(4) 数据不具有独立性，完整性差。当数据的物理组织或存储设备改变时，用户就需要相应地修改自己的应用程序。

2. 文件系统管理阶段

20 世纪 50 年代到 60 年代中期，计算机的应用范围逐渐扩大，计算机不再只用于科学计算的单一任务，而可以做一些非数值数据的处理。又由于大容量的磁盘等辅助存储设备的出现，使得专门管理辅助存储设备上的数据的文件系统应运而生，它是操作系统中的一个子系统。在文件系统中，按一定的规则将数据组织成为一个文件，应用程序通过文件系统对文件中的数据进行存取和加工。文件系统对数据的管理，实际上是通过应用程序和数据之间的一种接口实现的，如图 1－2 所示。

文件系统的最大特点是解决了应用程序和数据之间的一个公共接口问题，使得应用程序采用统一的存取方法来操作数据。在文件系统阶段中数据管理的特点如下：

(1) 数据可以长期保存。由于有了直接存取设备，数据可以长期保存在外存中进行反复的查询、修改、插入和删除等操作。

(2) 数据由文件系统管理。文件系统把数据组织成为相互独立的数据文件，应用程序按照“按文件名访问，按记录进行存取”管理技术，实现数据的更新。文件系统的记录内部有结

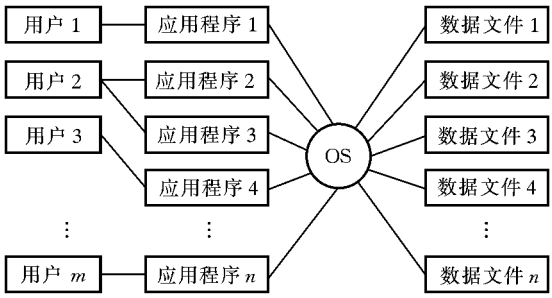


图 1-2　应用程序与文件的关系

构，但整体无结构。例如，使用 C 语言定义学生档案管理系统中的学生基本信息、选课和成绩的结构如下：

学生基本信息结构：

```
struct StudentInfo {
   int nStuId;
   char sStuName[10];
   char sSex;
   char sDept[30];
}
```

学生选课结构：

```
struct Course {
   int nCourseId;
   char sCourseName[30];
   char sDept[30];
}
```

学生成绩结构：

```
struct Score {
   int nStuId;
   int nCourseId;
   int nScore;
}
```

学生档案管理系统中学生基本信息、选课和成绩三类数据分别按照其定义存储在三个文件中。此时的应用程序和数据文件之间由文件系统来提供存取方法，实现了“按文件名访问，按记录进行存取”，使应用程序与数据文件之间有了一定的独立性，程序员不必过多地考虑物理的细节，而将精力集中在算法上。但是文件系统只是简单地存取数据，数据文件相互之间并没有有机的联系，即数据存取依赖于应用程序的使用方法，不同的应用程序仍然很难共享同一数据文件。

(3) 数据共享性差，冗余度大。在文件系统中，一个文件基本上对应一个应用程序。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，而不能共享相同数据，因此数据冗余度仍然较大，浪费许多的存储空间。

(4) 数据独立性差。文件系统中文件是为某一特定应用服务的。文件的逻辑结构对应程序来说是优化的，但要对现有的数据增加一些应用会比较困难，因此系统难以扩充。另外，一旦数据的逻辑结构改变，修改文件系统的定义，也必须修改应用程序，因此数据独立性也较差。

3. 数据库系统管理阶段

20 世纪 60 年代以后，随着计算机在数据管理领域的普遍应用，数据量的急剧增长，人们对数据管理技术提出了更高的要求，希望面向企业或部门以数据为中心组织数据，减少数据的冗余，提供更高的数据共享能力，同时要求程序和数据具有较强的独立性。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，以降低应用程序研制与维护的费用。数据库技术正是在这样的应用需求的基础上发展起来的。在数据库系统管理阶段，数据与应用程序之间的关系如图 1－3 所示。

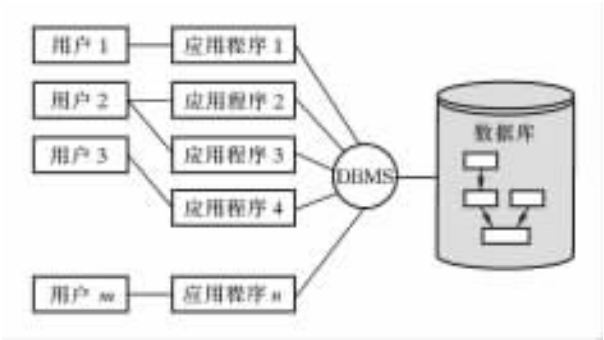


图 1－3　应用程序与数据库的关系

数据库系统管理阶段有如下特点：

(1) 数据结构化。

数据结构化是数据库管理阶段与文件管理阶段的根本区别。在文件系统阶段，相互独立的文件记录内部是有结构的，但记录之间没有联系。在数据库系统管理阶段，数据结构化模型不仅描述数据本身的特点，还描述数据之间的联系。例如，学生档案管理系统中的学生信息包括学生情况记录，如表 1－2 所示。

表 1－2　学生情况记录

学号	姓名	性别	出生年月	民族	政治面貌	班级	所在系	成绩情况	家庭成员

表中“成绩情况”和“家庭成员”栏目内各个学生所具有的信息量的多少变化很大。若以等长记录形式存储时，每个学生记录的长度必须等于信息量最多的记录的长度，这样会造成存储空间的大量浪费。因此，最好采用不等长记录(变长记录)或主记录与详细记录相结合的方法建立文件。

在学生信息记录中，从“学号”到“所在系”8 个数据项是任何一个学生必须具有而且是等长的，将这一部分作为主记录，而将“成绩情况”和“家庭成员”作为详细记录分别存储在学生成绩和家庭成员中。这样，一个学生的基本信息记录的结构如图 1－4 所示。图 1－4(a)表示每个记录的记录格式，而图 1－4(b)表示学生基本概况记录的某一个记录值，这样由一个主记录附加若干个值的详细记录的方法增加了灵活性，节省了存储空间。

(2) 数据共享性高、冗余度低、易扩充。

数据库从整体观点来看就是描述数据，数据不再是面向某一应用，而是面向整个系统，这可以大大减小数据的冗余度，既节约存储空间、减少存取时间，又可避免数据之间的不相容性

学号	姓名	性别	出生年月	民族	政治面貌	班级	所在系

↓成绩情况

学号	课程	成绩

↓家庭情况

关系	姓名	单位	联系电话

(a) 主记录及其详细记录格式示例

061102	张芳	女	1986.10	汉	团员	21011	计算机科学与技术

↓学号为 061102 成绩情况

061102	数据库	89
061102	软件工程	92
061102	C 语言	86

↓学号为 061102 家庭情况

061102	父亲	张恒	微电子研究所	88491114
061102	母亲	姜佳	西安市第一中学	88492114

(b) 某个学生记录示例

图 1-4　主记录及其详细记录

和不一致性。

由于数据库中的数据面向整个系统，是有结构的数据，不仅可以被多个应用共享，而且容易增加新的应用，可以适应各种应用需求。当应用需求改变或增加时，只要更新选取整体数据的不同子集便可以满足新的要求，这就使得数据库系统具有弹性大，易扩充的特点。

(3) 数据独立性强。

数据独立性是指应用程序和数据之间相互独立，彼此之间不受影响。数据独立性分为数据的物理独立性和数据的逻辑独立性，它是数据库系统通过提供两个层次的映像功能，使数据的物理结构独立于全局逻辑结构，使数据的全局逻辑结构独立于应用程序。当数据的物理组织或存储设备改变时，不影响全局逻辑结构。而当数据的全局逻辑结构改变时，有可能不用改变应用程序。

(4) 数据的管理由 DBMS 管理和控制。

数据库的管理由专门的系统软件数据库管理系统(DBMS，DataBase Management System)来实现。DBMS 是位于用户和操作系统之间的一种数据管理软件，它能够科学地组织和存储数据，高效地获取和维护数据。DBMS 是以数据库的方式存储数据的，是数据的一种更高级的组织形式，在应用程序和数据库之间由 DBMS 负责数据的存取。DBMS 对数据的处理方式和文件系统不同，它把所有应用程序中使用的数据以及数据间的联系汇集在一起，以便于应用程序查询和使用。

从文件系统发展到数据库技术在信息领域中具有里程碑的意义。在文件系统阶段，人们在信息处理中关注的中心问题是系统功能的设计，因此程序设计占主导地位；而数据库方式下，数据开始占中心地位，数据的结构设计成为信息系统首要关心的问题，而应用程序则以既定的数据结构为基础进行设计。

数据管理三个阶段的比较如表 1-3 所示。

表 1-3 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	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 DBMS 提供数据安全性和完整性等

1.2 数据库系统组成

1.2.1 DBS 的组成

数据库系统(DBS)是由计算机系统、数据库、数据库管理系统、应用程序和相关人员组成的具有高度组织的总体。

1. 计算机系统

计算机系统包括计算机硬件和操作系统。计算机硬件是构成计算机的各种物理设备，包括存储数据所需的外部设备。硬件的配置应满足整个数据库系统的需要。操作系统是控制和管理计算机系统内各种硬件和软件资源、有效地组织多道程序运行的系统软件，是用户与计算机之间的接口。

2. 数据库

数据库既有存放实际数据的物理数据库，也有存放数据逻辑结构的数据字典。

物理数据库是指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的集合，例如学生档案管理系统中，数据库存放着学生、课程、选课、成绩、教师、教材等信息。

数据字典存放有关数据库定义信息，如用户名表及权限、数据库表的定义等。数据字典是数据库管理系统的组成部分之一，它是由数据库管理系统自动生成并维护的一组表和视图。

需要全本请访问 www.ertongbook.com

数据字典是数据库管理系统工作的依据，数据库管理系统通过它对数据库中的数据进行管理与维护。

3. 数据库管理系统 DBMS

DBMS 是一个通用的软件系统。它能对数据库进行有效的管理，包括安全性管理、存储管理、完整性管理等，为数据的访问和保户提供强大的处理功能，如查询处理、并发控制、故障恢复等，同时为用户提供一个应用、操作和管理的平台，使其更方便地创建、检索、存储和处理数据库中的数据。

数据库管理系统是数据库系统的核心软件，有关详细内容在 1.2.2 中讲解。

4. 应用程序

应用程序主要指实现业务逻辑的程序。它要为用户提供一个友好的操作数据的图形用户界面，通过数据库语言或相应的数据访问接口，存取数据库中的数据。

5. 相关人员

数据库管理系统中相关人员主要包括系统分析员和数据库设计人员、数据库管理员、应用程序员和最终用户。

(1) 系统分析员和数据库设计人员：系统分析员负责应用系统的需求分析和规范说明，他们和用户及数据库管理员一起确定系统的硬件配置，并参与数据库系统的概要设计。数据库设计人员负责数据库中数据的确定、数据库各级模式的设计。

(2) 数据库管理员(DBA，Data Base Administrator)：负责管理和监控数据库系统，解决应用中系统出现的问题。DBA 的具体职责包括：负责数据库系统软件的安装和维护；参与数据库中的设计；决定数据库的存储结构和存取策略；定义数据库的安全性要求和完整性约束条件；监控数据库的使用和运行；参与数据库的性能改进、数据库的重组和重构，以提高系统的性能。

(3) 应用程序员：负责编写使用数据库的应用程序，这些应用程序可对数据进行检索、建立、删除或改变现存的信息。

(4) 最终用户：用户是指应用系统的最终使用者，他们通过应用程序的操作界面使用数据库，完成日常业务处理。

1.2.2　DBMS

数据库管理系统 DBMS 是对数据进行管理的软件系统，是数据库系统的核心软件。数据库系统的一切操作，包括创建各种数据库对象，如表、视图、存储过程等，以及应用程序对这些对象的操作(如插入数据到表中，对表中原有数据的检索、修改、删除等)，都是通过 DBMS 进行的。

带有数据库的计算机系统硬软件层次如图 1－5 所示。

数据库管理系统主要提供以下功能。

1. 数据库定义功能

DBMS 提供数据定义语言(DDL，Data Definition Language)。数据库设计人员通过 DDL 来描述和定义数据库中的对象，例如定义数据库的结构(包括数据库的模式、子模式、存储模式以及相互之间的映像)，定义数据库的完整性约束、保密限制等约束条件。这些定义存储在数

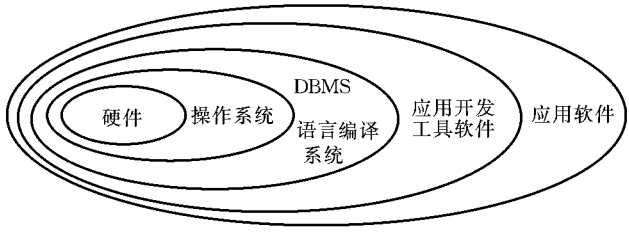


图 1-5 DBMS 在计算机系统中的地位

据字典中,是 DBMS 运行的基本依据。

2. 数据操作功能

DBMS 提供数据操作语言(DML,Data Manipulation Language)。用户使用 DML 语言实现对数据库中数据的基本操作,如检索、插入、修改和删除。

DML 分为两类:一种是嵌入某种主语言(如 C、PASCAL、COBOL 等)中的数据库操作语言,称为宿主型 DML;一种是可以独立使用的数据库操作语言,称为自主型或自含型 DML,例如 SQL Server 中的 Transact-SQL 语言等。

3. 数据库运行控制功能

DBMS 提供数据控制语言(DCL,Data Control Language)。用户使用 DCL 语言实现以下几方面的数据控制功能。

(1) 数据的安全控制:

数据的安全控制是对数据库的一种保护,即防止不合法的用户造成数据库中数据的泄密或破坏。控制每个用户的使用权限,控制某些数据的使用和处理的方式。

(2) 数据的完整性控制:

数据的完整性控制是对数据库中数据的一种保护措施,以保证进入数据库中的存储数据的语义正确、有效,防止出现违反其语义的任何操作。

(3) 数据的并发控制:

在多个应用程序同时对数据库进行存取、修改等操作时,可能会发生相互干扰而得到错误的结果,或在数据库中保存了错误的数据。DBMS 数据并发控制功能能够对多用户的并发操作加以控制和协调,防止出现错误的结果,正确处理多用户、多任务环境下的并发操作。

(4) 数据的数据恢复:

数据库在运行的过程中,可能会出现各种故障,如硬件故障、软件故障、操作员的误操作以及有意的破坏等,都可能造成数据库的损坏或运行结果不正确。DBMS 数据的数据恢复就是把数据库从被破坏的、不正确的状态恢复到正常状态。

4. 数据库的建立和维护

数据库的建立和维护包括数据库的初始建立、数据的转换、数据库的转储和恢复、数据库的重组和重构、性能监测和分析等等。

1.3　数据库系统体系结构

数据库系统的结构从数据库管理系统的角度来看，通常分为三级模式结构，这是数据库管理系统的内部体系结构。这是本节介绍的主要内容。数据库系统的结构从最终用户的角度来看，通常分为集中式结构（又可分为单用户结构和主从结构）、分布式结构和并行结构，这些是数据库管理系统的外部系统结构。

数据库系统的体系结构是数据库系统的一个总体框架。实际应用中数据库产品多种多样，各支持不同的数据模型和不同的数据库语言，而且数据的存储结构也各不相同，但是绝大多数数据库系统在总体结构上都具有三级模式的结构特点。

1.3.1　数据库系统的三级模式

数据库系统的三级模式结构是由外模式、概念模式和内模式构成，如图 1－6 所示。三级模式结构之间形成两级映射，所以通常又称数据库的体系结构为三级模式两级映射。数据库系统设计员可在视图层、逻辑层和物理层对数据抽象，通过外模式、模式和内模式来描述不同层次上的数据特性。

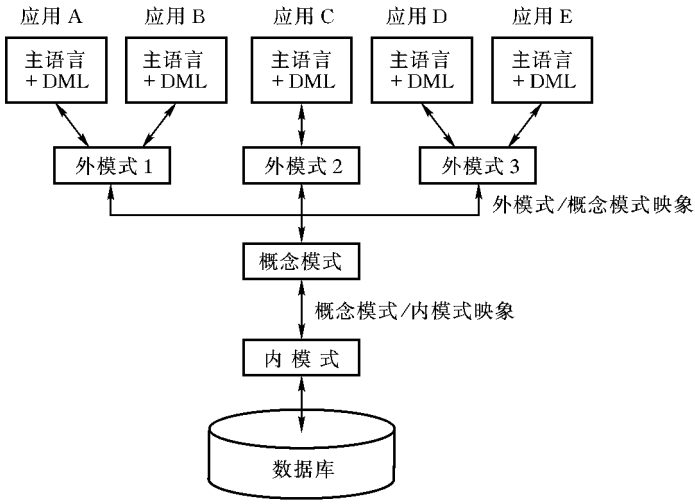


图 1－6　数据库系统结构的三级模式二级映像

1. 概念模式

概念模式也称模式或逻辑模式，是数据库中全部数据的逻辑结构和特征的描述，它由若干个记录类型组成，只涉及到行的描述，不涉及到具体的值。概念模式的一个具体值称为模式的一个实例，同一个模式可以有很多实例。例如，在学生档案管理系统中的学生情况记录结构定义为（学号，姓名，性别，出生年月，民族，政治面貌，班级，所在系），这是记录型，而（061102，“张芳”，“女”，1986.10，“汉”，“团员”，21011，“计算机科学与技术”）则是该记录型的一个记录值，换句话说是该记录类型的一个实例。

概念模式反映的是数据库的结构及其联系，所以是相对稳定的；而实例反映的是数据库某一时刻的状态，所以是相对变动的。

需要说明的是，概念模式不仅要描述概念记录类型，还要描述记录间的联系、操作、数据的完整性、安全性等要求。

描述概念模式的数据定义语言称为“模式 DDL”(Schema Data Definition Language)。

2. 外模式

外模式也称用户模式或子模式，是用户与数据库系统的接口，是用户用到的那部分数据的描述。外模式由概念模式导出，是概念模式的子集，它是用户的数据视图，即与某一应用有关的数据的逻辑表示。

描述外模式的数据定义语言称为“外模式 DDL”。有了外模式后，程序员不必关心概念模式，只与外模式发生联系，按外模式的结构存储和操作数据。

3. 内模式

内模式也称存储模式，是数据物理结构和存储方式的描述，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 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。

描述内模式的数据定义语言称为“内模式 DDL”。

总之，数据按外模式的描述提供给用户，按内模式的描述存储在磁盘上，而概念模式提供了连接这两级模式的相对稳定的中间观点，并使得两级的任意一级的改变都不受另一级的牵制。

数据库系统中的人员包含系统分析员、数据库设计人员、数据库管理员、应用程序员和最终用户等各类人员，不同人员涉及不同级别的数据，具有不同的数据视图。各种人员在数据库体系结构中的数据视图如图 1－7 所示。

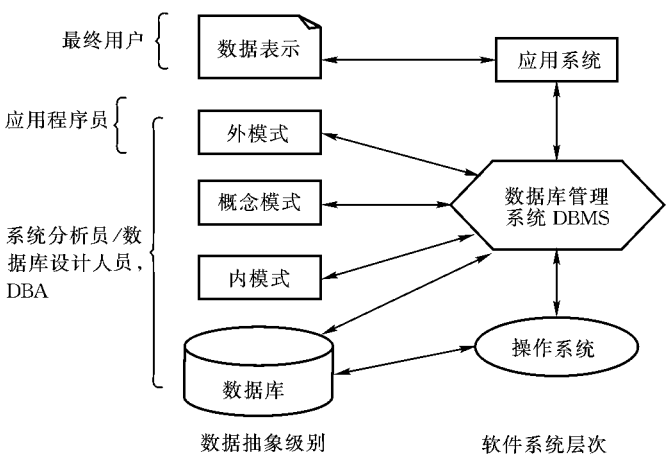


图 1－7　各种人员的数据视图

1.3.2　数据库系统的二级映像和数据独立性

数据库系统的三级模式是对数据的三个抽象级别，数据的具体组织是由 DBMS 进行管理

的，使用户能逻辑地处理数据，不必关心数据在计算机中的表示和存储。为了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中提供了外模式/概念模式和概念模式/内模式二级映像。如图 1－6 所示。

1. 外模式/概念模式映像

对应于同一概念模式，可以有任意多个外模式。外模式/概念模式映像定义某个外模式和概念模式之间的对应关系，这些映像定义通常包含在外模式中。当概念模式改变时，只要将外模式/概念模式映像作相应的改变，就可以保证外模式保持不变，它提供了数据的逻辑独立性。

2. 概念模式/内模式映像

概念模式/内模式映像定义数据逻辑结构和存储结构之间的对应关系。当数据库的存储结构改变时，只要将概念模式/内模式映像作相应的修改，就可以使得概念模式保持不变，它提供了数据的物理独立性。

3. 数据的独立性

数据的独立性是指数据与应用程序相互独立，即将数据的定义和描述从程序中分离出去，由 DBMS 负责数据的存取和管理，用户不必考虑存取的路径和细节，从而简化应用程序，大大减少应用程序编制的工作量。数据的独立性分为数据的逻辑独立性和数据的物理独立性。数据库的体系结构，即三级模式和二级映像保证了数据能够具有较高的逻辑独立性和物理独立性。

数据的逻辑独立性是指用户的应用程序与数据库的逻辑结构是相互独立的。数据的逻辑结构发生变化（例如增加新的关系、新的属性、改变属性的数据类型等）后，数据库管理员对各个外模式和概念模式的映像作了相应的改变，可以使外模式保持不变。由于应用程序是依据数据的外模式编写的，当数据的逻辑结构发生变化时可以使外模式保持不变，因此应用程序便不必修改，从而保证数据的逻辑独立性。

数据的物理独立性是指当数据库的内模式发生改变时，数据的逻辑结构不变。由于数据库中只有一个概念模式，也只有一个内模式，所以概念模式和内模式的映像也是唯一的，它定义了数据库的全局逻辑结构与存储结构之间的关系。当数据库的存储模式改变（例如选用了另一种存储结构）时，数据库管理系统对概念模式和内模式的映像作了相应的改变，可以使模式保持不变，从而应用程序也不需要改变。这样就保证了数据的物理独立性。

由于应用程序是在外模式描述的数据结构上编写的，数据库的三级模式二级映像又保证了数据的独立性，因此当数据的逻辑结构和存储结构改变时，应用程序一般不需要修改（应用需求发生了变化的除外）。

1.4　数据模型

模型就是对现实世界特征的模拟和抽象，如一张地图、一组建筑设计沙盘、一架航模飞机等都是具体的模型。数据模型是对现实世界数据特征的抽象，它精确描述了数据、数据之间的联系、数据之间的寓意和完整性约束。很多数据模型还包括一些操作的集合，这些操作用来对数据库进行存取和更新。数据模型应该满足三个方面的要求：一是能真实地模拟现实世界；二

是容易被人们所理解;三是便于在计算机上实现。即涉及到现实世界、信息世界和机器世界三个数据领域,它们之间的关系如图 1-8 所示。

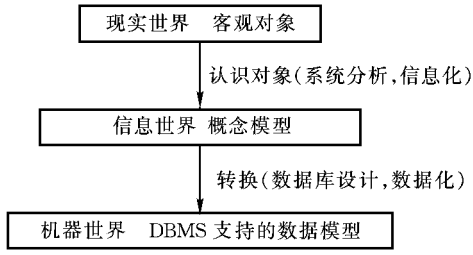


图 1-8 现实世界中客观对象的抽象过程

1.4.1 数据模型的概念

数据模型是数据库系统的重要基础。不同的数据模型提供了不同的模型化数据和信息的工具。这里主要介绍三个数据世界的概念、数据模型的三要素以及数据的抽象。

1. 三个数据世界的概念与描述

(1) 现实世界

现实世界泛指存在于人脑之外的客观世界。现实世界的数据就是客观存在的各种报表、图表和查询格式等原始数据。计算机只能处理数据,所以首先要解决的问题是按用户的观点对数据和信息建模,即人们通过实体、特征、实体集及联系进行分析和认识。

① 实体(Entity)。客观存在并可相互区别的事物或概念称为实体。实体可以是具体的人、事、物,也可以是抽象的概念或联系,例如一个职工、一个学生、一个部门、一门课、学生的一次选课、部门的一次订货,以及老师与系的工作关系等都是实体。

② 实体的特征(Entity Characteristic)。每个实体都有自己的特征,利用它来区别不同的实体。例如,在学生管理系统中的学生实体可以由学号、姓名、性别、年龄、系、专业等特征来描述自己;在人事管理系统中的职工实体可以由职工号、姓名、性别、年龄、工资、职务等特征来描述自己。

③ 实体集(Entity Set)及实体集之间的联系。具有相同特征的或能用相同特征来描述的实体的集合称为实体集。例如,学生、工人、房屋、课程等都是一个实体集。在现实世界中,实体集之间是有着各种各样的联系,例如学生和课程之间有“选课”的联系。

(2) 信息世界

信息世界是现实世界在人们头脑中的反映,经过人们认识、选择、命名、分类等综合分析而形成的印象或概念。在信息世界中,实体特征在人脑中形成的知识成为属性,实体通过其属性表示为实例,同类实例的集合称为对象,对象即实体集中的实体用属性表示得出的信息集合。实体与实例是不同的,例如张芳是一个实体,而“061102,张芳,女,1986.10,汉,团员,21011,计算机科学与技术”是实例。在实体的诸多属性中,能够唯一标识一个实体的属性或属性组成为码。在信息世界中实体集之间的联系是用对象的联系表示。

① 对象(Object)和实例(Instance)。在现实世界中,具有相同性质同类事物或概念的抽

需要全本请访问 www.ertongbook.com