



普通高等教育“十二五”规划教材

数据库实用教程 (SQL Server 2008 + C#)

罗贤缙 秦金磊 张锋奇 谢萍 编著



中国电力出版社
CHINA ELECTRIC POWER PRESS



普通高等教育“十二五”规划教材

数据库实用教程 (SQL Server 2008 + C#)

编 著 罗贤缙 秦金磊 张锋奇 谢 萍
主 审 王敬敏



中国电力出版社
CHINA ELECTRIC POWER PRESS

内 容 提 要

本书为普通高等教育“十二五”规划教材。本书是作者在教学实践基础上，针对工科院校数据库原理与应用课程学时短、实践性强的教学需要编写而成。本书以 Microsoft SQL Server 2008 中文版和 Visual Studio 2010 为教学平台，介绍了数据库基本概念，包括数据库系统概述、数据模型、关系数据库、关系数据库规范化理论和数据库设计、数据库及数据表的创建、数据查询与视图；然后介绍了基于 C# 的数据库应用开发，包括 C# 基础知识、Windows 窗体编程、基于 ADO.NET 的数据库开发。本书包含大量例题和课后习题，语言通俗、结构合理、图文并茂，具有较强的实用性。

本书可作为普通高等学校非计算机专业数据库相关课程教材和成人函授相关专业、社会培训教材，也可供广大从事数据库开发工作的专业技术人员和管理人员参考。

图书在版编目 (CIP) 数据

数据库实用教程: SQL Server 2008+C#/罗贤缙等编著. —北京: 中国电力出版社, 2015.4

普通高等教育“十二五”规划教材

ISBN 978-7-5123-7373-0

I. ①数… II. ①罗… 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 (2015) 第 050070 号

中国电力出版社出版、发行

(北京市东城区北京站西街 19 号 100005 <http://www.cepp.sgcc.com.cn>)

航远印刷有限公司印刷

各地新华书店经售

*

2015 年 4 月第一版 2015 年 4 月北京第一次印刷

787 毫米×1092 毫米 16 开本 16.25 印张 393 千字

定价 33.00 元

敬告读者

本书封底贴有防伪标签，刮开涂层可查询真伪
本书如有印装质量问题，我社发行部负责退换

版 权 专 有 翻 印 必 究

前言

数据库技术自 20 世纪 60 年代末产生至今已有 50 多年的历史,在商业、医疗保健、教育、政府组织、图书馆、军事、工业控制等领域都取得了广泛应用,已成为信息管理、电子商务、网络服务等系统的核心技术和重要基础,也是计算机、控制、信息等相关工科专业的工程技术人员所必须具备的专业知识。开设数据库相关课程的目的是使学生全面地学习数据库系统的概念和应用技术,让学生在数据库基础知识、数据库操作能力、数据库管理能力和数据库应用设计能力等方面达到一定水平。

Microsoft SQL Server 是目前流行的大中型关系型数据库管理系统,SQL Server 2008 是其重大的产品版本,它推出了许多新的特性和关键改进,使其能应对数据爆炸的新挑战及下一代数据驱动应用程序的需求。Microsoft Visual Studio 是目前流行的 Windows 应用程序开发环境,功能强大、使用简单方便。

全书共有 14 章,从内容上分为数据库原理、SQL Server 2008 和基于 C#的数据库应用开发三部分。第一部分的主要内容是数据库基础理论,包括第 1~5 章,介绍数据管理技术发展、数据库系统的组成与结构、数据模型、关系数据库基础知识及规范化理论、数据库设计的方法和步骤;第二部分的主要内容是数据库应用,包括第 6~9 章,介绍在 SQL Server 2008 环境下如何完成数据库和数据表等数据对象的创建及操作、表数据的操纵、T-SQL 语言、数据的查询等;第三部分主要内容是在 Visual Studio 2010 环境下利用 C#进行数据库系统的开发,包括第 10~14 章,介绍 C#面向对象编程基础、Windows 窗体编程及基于 ADO.NET 的数据库开发。全书在保证概念准确的基础上力求通俗易懂、简洁明了,同时配以丰富的图表深入浅出地进行阐述,所选例题具有代表性,能帮助学生迅速掌握简单数据库系统的开发,为将来更深入的学习和应用打下基础。

本书主要由罗贤缙、秦金磊、张锋奇编写,参与编写的还有谢萍,博士生导师王敬敏教授主审,在编写过程中得到了许多老师的支持,提出了很多宝贵意见,在此一并表示感谢!

本书部分内容加以“*”标注,供教师讲授和学生学时选用。

由于编者水平和经验有限,书中错误在所难免,恳请广大读者批评指正。

编 者

2014 年 12 月

目 录

前言

第1章 数据库系统概述	1
1.1 信息、数据和数据处理	1
1.2 数据管理技术的发展	2
1.3 数据库系统的组成与结构	4
习题1	8
第2章 数据模型	10
2.1 现实世界的抽象过程	10
2.2 概念模型	10
2.3 数据模型及组成要素	15
2.4 常见的数据库类型	15
习题2	18
第3章 关系数据库	20
3.1 关系数据库的基本概念	20
3.2 关系的数学定义	21
3.3 关系代数	24
习题3	28
第4章 关系数据库规范化理论	32
4.1 关系数据库规范化理论问题的提出	32
*4.2 函数依赖	33
4.3 范式	34
习题4	39
第5章 数据库设计	41
5.1 数据库设计概述	41
5.2 数据库设计实例	42
习题5	48
第6章 Microsoft SQL Server 2008 概述	51
6.1 Microsoft SQL Server 2008 简介	51
6.2 安装 SQL Server 2008	52
6.3 使用 SQL Server Management Studio	61
习题6	62

第 7 章 SQL Server 2008 数据库管理	63
7.1 SQL Server 2008 数据库结构	63
7.2 创建数据库	66
7.3 修改数据库	69
7.4 删除数据库	70
7.5 分离与附加数据库	71
习题 7	73
第 8 章 SQL Server 2008 数据表管理	74
8.1 SQL Server 2008 基础	74
8.2 数据表的创建	80
8.3 数据表的修改和删除	89
8.4 表数据的操纵	92
8.5 索引	99
习题 8	103
第 9 章 数据查询与视图	106
9.1 SELECT 语句概述	106
9.2 单表查询	107
9.3 多表查询	116
9.4 嵌套查询	118
9.5 视图	120
习题 9	126
第 10 章 Microsoft.NET 及其开发工具	128
10.1 Microsoft.NET 概述	128
10.2 Microsoft .NET Framework 介绍	129
10.3 Visual Studio .NET IDE 简介	132
习题 10	138
第 11 章 C#编程基础	139
11.1 标识符和关键字	139
11.2 变量和常量	140
11.3 数据类型	141
11.4 运算符和表达式	148
11.5 流程控制语句	154
11.6 异常处理	159
习题 11	161
第 12 章 C#面向对象基础	162
12.1 面向对象编程概述	162
12.2 类和对象	163
12.3 属性和索引器	165
12.4 方法	168

12.5	接口	171
12.6	继承与多态	172
12.7	委托与事件	175
习题 12		180
第 13 章	Windows 窗体编程	181
13.1	窗体编程概述	181
13.2	用 C# 创建 Windows 窗体程序实例	182
13.3	Windows 窗体简介	189
13.4	常用公共控件	192
13.5	定时器组件	209
13.6	消息框	211
习题 13		214
第 14 章	使用 C# 开发数据库	217
14.1	数据库的连接	217
14.2	ADO.NET 概述	222
14.3	ADO.NET 应用实例	241
习题 14		248
参考文献		251

第1章 数据库系统概述

数据库技术是一门研究数据管理的技术,始于20世纪60年代末,经过几十年的发展,数据库技术成为计算机技术的一个重要分支,也成为日常生活中不可缺少的一部分。数据库技术主要研究如何存储、管理和使用数据。本章主要介绍数据管理技术的发展和数据库系统的基本概念,为后面各章的学习奠定基础。

1.1 信息、数据和数据处理

计算机的出现,开辟了计算机处理的新纪元。数据处理的基本问题是数据的组织、存储、检索、维护和加工利用。

数据是数据库系统研究和处理的对象,数据与信息是分不开的,它们既有联系又有区别,因此首先介绍信息与数据的概念。

1.1.1 信息与数据

1. 信息

信息是对现实世界事物存在方式或运动状态的反映。具体地说,信息是一种已经被加工为特定形式的数据,这种数据形式对接受者来说是有意义的,并且对当前和将来的决策具有实际价值。

信息有如下一些重要特征:

- (1) 信息传递需要物质载体,信息的获取和传输需要消耗能量。
- (2) 信息是可以感知的,不同的信息有不同的感知方式(如感觉器官、仪器或传感器等)。
- (3) 信息是可以存储、压缩、加工、传递、共享、再生和增值的。

信息是资源,人类进行各项社会活动,不仅要考虑物质条件,还要研究和利用信息。正因为如此,人类将能源、物质和信息并列为人类社会活动的三大因素。

2. 数据

为了了解世界、研究世界和交流信息,人们需要描述各种事物。用自然语言来描述虽然很直接,但过于繁琐,不便于形式化,而且也不利于计算机表达。因此,人们常常只抽取那些感兴趣的事物特征或属性来描述事物。例如一个学生可以用如下属性描述:

(张伟, 201402000130, 男, 1996.5, 动力系, 热能)

这种描述事物的符号称为数据,数据的含义是数据的语义,数据和其语义是密不可分的。上述数据的语义是:

(学生姓名、学号、性别、出生年月、所在系、专业)

人们通过解释、推论、归纳、分析和综合等方法,从数据中获取有意义的内容称为信息。因此,数据是信息存在的一种形式,只有通过解释或处理才能成为有用的信息。

那么,我们可以得到如下信息:张伟是动力系热能专业的男生,学号 201402000130, 1996 年 5 月出生。

3. 信息与数据

信息与数据是两个既有联系又有区别的概念。数据是信息的载体,而信息是数据的内涵。同一信息可以由不同的数据表示,同一数据可以有不同的解释。因此在许多场合下,对它们不做严格区分,可互换使用。例如,通常说的“信息处理”与“数据处理”具有同义性。

1.1.2 数据处理

数据处理是人们直接或利用机器加工数据的过程。对数据进行查找、统计、分类、交换等都属于加工。

在数据处理的一系列活动中,数据收集、存储、分类传输等基本操作环节称为数据管理,而加工、计算、输出等操作是千变万化的,不同业务有不同的处理。

数据管理技术是解决上述基本环节的,而其他环节是通过应用程序实现的。

1.2 数据管理技术的发展

随着计算机软硬件技术的发展,数据管理技术的发展大致经历了人工管理、文件管理和数据库系统三个阶段。

1.2.1 人工管理阶段

20 世纪 50 年代中期以前,计算机的外部设备只有磁带、卡片和纸带,没有磁盘等可直接存储的设备,没有操作系统和数据管理软件,数据处理方式是批处理,所有数据完全由人工进行管理,因此这个阶段称为人工管理阶段,主要有以下几个特点。

(1) 一组数据对应于一个应用程序,应用程序与其处理的数据合成一个整体。在进行计算时,系统将程序与数据一并装入内存,用完后就将它们撤销、释放,其数据管理模型如图 1.1 所示。

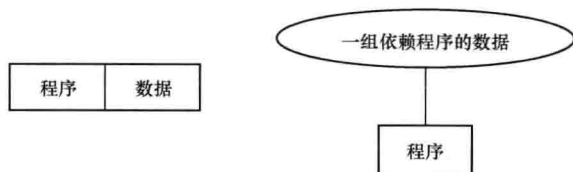


图 1.1 人工管理阶段数据管理模型

(2) 没有文件概念,数据的组织方式由应用程序开发人员自行设计。

(3) 没有对数据进行管理的软件,程序员不仅要设计数据的逻辑结构,还要设计数据的物理结构,如存储结构、存取方法、输入/输出方式等,因此数据与程序不具有独立性,如果数据存储改变了,程序员就必须修改程序。

(4) 数据面向应用,即使两个程序使用同样的数据,也必须各自定义自己的数据存储和存取方式,不能共享相同的数据定义,因此,造成了程序之间大量的冗余数据。

1.2.2 文件系统阶段

20 世纪 50 年代末至 60 年代中期,计算机外存已经有了磁鼓、磁盘等存储设备,软件有了操作系统。人们在操作系统的支持下,设计开发了文件系统。这时,计算机不仅用于科学计算,也已大量用于数据处理,其特点如下:

(1) 数据以文件的形式长期保存,在文件系统中,按一定的规则将数据组织为一个文件,长期存放在外存储器中。

(2) 程序员只需要用文件名与数据打交道,不必关心数据的物理位置,可由文件系统提供的读/写方法去读/写数据。

(3) 文件形式多样化。为了方便数据的存储和查找,人们研究了许多文件类型,如索引文件、链接文件、顺序文件和倒排文件等,数据的存取基本上是以记录为单位的。

(4) 数据与程序之间具有一定的独立性。应用程序通过文件系统对数据文件中的数据进行存取和加工,因此处理数据时,程序不必过多地考虑数据的物理存储的细节,文件系统充当应用程序与数据之间的一种接口,可使应用程序和数据都具有一定的独立性,这样,程序员可以集中精力于算法,而不必过多地考虑物理细节。并且,数据在存储上的改变不一定反映在程序上,这可以大大节省维护程序的工作。其数据管理模型如图 1.2 所示。

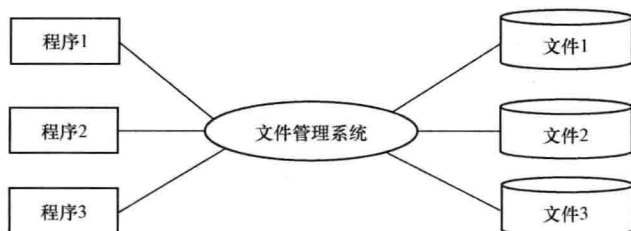


图 1.2 文件系统阶段数据管理模型

尽管文件系统有上述优点,但是这些数据在文件中只是简单地存放,文件中的数据没有结构,文件之间没有有机的联系,仍不能表示复杂的数据结构;数据的存放仍然依赖于应用程序的使用

方法,基本上是一个数据文件对应于一个或几个应用程序;数据面向应用,数据独立性差,仍然出现重复存储、冗余度大、一致性差(同一数据在不同文件中的值不一样)等问题。

1.2.3 数据库系统阶段

随着计算机软硬件技术的发展,数据处理规模的扩大,20 世纪 60 年代后期出现了数据库技术。从不同的角度去定义数据库可能差别较大,但是对数据库所应具有的特点,其认识上大体是一致的。数据库技术的若干特点如下:

(1) 数据结构化。数据库是存储在磁盘等外部直接存取设备上的数据集合,按一定的数据结构组织起来。与文件系统相比,文件系统文件之间不存在联系,因而从整体上看是没有结构的;而数据库中的文件是相互联系着的,并在总体上遵从一定的结构形式。这是文件系统与数据库系统最大的区别。数据库系统正是通过文件之间的联系反映现实世界事物间的自然联系的。

(2) 数据共享。数据库中的数据是考虑所有用户的数据需求、面向整个系统组织的,因此数据库中包含了所有用户的数据成分,但每个用户通常只用到其中一部分数据。不同用户所使用的数据可以重叠,同一部分数据也可为多个用户共享。

(3) 减少了数据冗余。在数据库方式下,用户不是自建文件,而是取自数据库中的某个子集,它并非独立存在,而是靠数据库管理系统(Database Management System,简称 DBMS)从数据库中映像出来的,所以叫做逻辑文件。如图 1.3 所示,用户使用的是逻辑文件,因此尽管一个数据可能出现在不同的逻辑文件中,但实际上的物理存储只出现一次,这就减少了数据冗余。

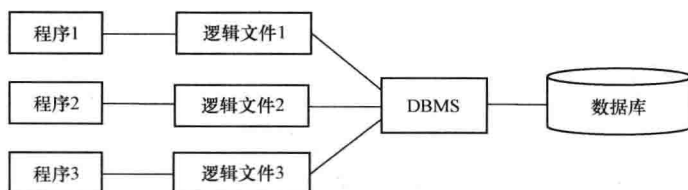


图 1.3 数据库系统阶段数据管理模型

(4) 有较高的数据独立性。数据独立的好处是数据存储方式的改变不会影响应用程序。数据独立又有两个含义,即物理数据独立性和逻辑数据独立性。所谓物理数据独立性是指数据库物理结构(包括数据的组织和存储、存取方法及外部存储设备等)发生改变时,不会影响到逻辑结构,而用户使用的是逻辑数据,所以不必改动程序;所谓逻辑数据独立性是指数据库全局逻辑发生改变时,用户也不必改动程序,就像数据库并没有发生变化一样。这是因为用户仅使用数据库的一个子集,全局变化与否与具体用户无关,只要能从数据库中导出所要的数据即可。

(5) 用户接口。在数据库系统中,数据库管理系统为用户与数据库的接口提供了数据库定义、数据库运行、数据库维护和数据安全性、完整性等控制功能。此外,还支持某些程序设计语言,并有专门的数据操纵语言,为用户编程提供了方便。

从文件系统管理发展到数据库系统管理是信息处理领域的重大变化,人们由传统的关注系统功能设计转向关注数据的结构设计,数据的结构设计成为信息系统的中心问题。

1.3 数据库系统的组成与结构

通常把引进了数据库技术的计算机称为数据库系统,它的目的是存储和产生所需要的有用信息。这些有用的信息可以是对于使用该系统的个人或组织有意义的任何事情。

1.3.1 数据库系统的组成

数据库系统(Database System, DBS)是数据库应用系统的简称,指计算机系统中引入数据库之后,用来组织和存取大量数据的管理系统。数据库系统由计算机系统、数据库、数据库管理系统、应用程序和用户组成。数据库系统的各个组件之间的关系如图 1.4 所示。

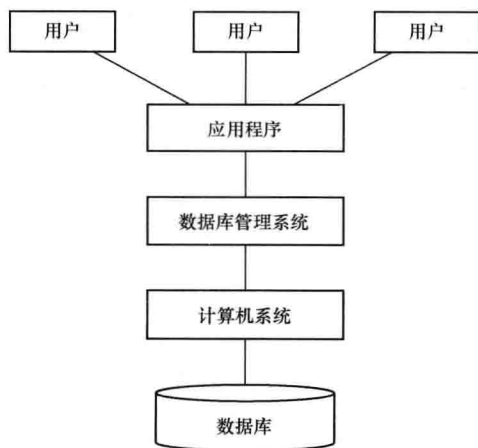


图 1.4 数据库系统的组成

1. 计算机系统

计算机系统由硬件和必要的软件组成。

(1) 硬件:指存储数据库和运行数据库管理系统 DBMS(包括操作系统)的硬件资源,它包括物理存储数据库的磁盘、光盘等外存储器及其附属设备、控制器、I/O 通道、内存、CPU 等其他外部设备。

(2) 软件:指计算机正常运行所需要的操作系统和各种驱动程序等。

2. 数据库

数据库是指数据库系统中集中存储的一批数据的集合。它是数据库系统的工作对象。

特别需要指出的是,数据库中存储的数据是“集

成的”和“共享的”。

“集成”,是指把某特定应用环境中的与各种应用相关的数据及数据之间的联系(联系也是一种数据)全部集中,并按照一定的结构形式进行存储,在文件之间局部或全部地消除了冗余,这使数据库系统具有整体数据结构化和数据冗余小的特点。

“共享”,是指数据库中的数据可为多个不同的用户所共享。即多个不同的用户,使用多

种不同的语言，为了不同的应用目的，而同时存取数据库，甚至同时存取同一块数据。共享实际上是基于数据库是“集成的”这一事实的结果。

3. 数据库管理系统

数据库管理系统 (Database Management System, DBMS) 是数据库系统关键组成部分，是数据库系统的核心软件。任何数据操作，包括数据库定义、数据查询、数据维护、数据库运行控制等都是在 DBMS 管理下进行的。DBMS 是用户与数据库之间的接口，应用程序只有通过 DBMS 才能和数据库打交道。目前流行的数据库管理系统有 Oracle、SQL Server、Access 等。通常，DBMS 的主要功能包括以下几个方面。

(1) 数据定义功能。DBMS 提供数据定义语言 (Data Definition Language, DDL) 来定义数据库的三级模式，也就是为数据库构建数据框架。

(2) 数据存取功能。DBMS 提供数据操纵语言 (Data Manipulation Language, DML) 实现对数据库的数据的基本存取操作：检索、插入、删除和修改。检索就是查询，是最重要最常用的一类操作，所以有些系统把 DML 称为查询语言。插入、修改和删除有时也称为更新操作。

DML 有两类：一类是交互式语言，语法简单，可独立使用，所以称为自主型或自含型；另一类是把数据库存取语句嵌入主语言 (Host Language) 中，如嵌入 FORTRAN、PASCAL、或 C 等高级语言中使用，这类 DML 本身不能独立使用，因此称为宿主型。现代的 DBMS 多含有自主型与宿主型两类 DML，允许用户选择使用。

DBMS 执行 DML 语句，完成对数据库的操作。对于自主型语言，DBMS 通常用解释执行的方式；对于宿主型语言，DBMS 一般提供以下两种方法。

1) 预编译方法。DBMS 提供预编译程序，对源程序进行扫描，识别嵌入的 DML 语句并把它们转换成主语言代码 (或主语言调用语句)，以便被主语言编译程序接受和执行。

2) 修改、扩充主语言编译程序的方法 (亦称增强型编译方法)。

DDL 和 DML 是数据库的用户在设计应用程序时必须用的程序设计语言的一个子集，被称为数据子语言。一种非常典型的数据子语言是 IBM 公司开发的 SQL 语言，它包含查询、操纵、定义和控制四个方面，是一种综合、通用、功能强大的关系数据库语言。

(3) 数据库运行管理功能。DBMS 提供数据控制功能 (Data Control Language, DCL)，即数据库的安全性、完整性和并发控制等，对数据库运行进行有效的控制和管理，以确保数据库数据的正确有效和数据库系统的有效运行。

1) 数据库的安全性 (Security) 控制，指采取一定安全保密措施确保数据库数据不被非法用户存取。DBMS 提供口令检查或其他手段检查用户身份，只存合法用户才能进入数据库系统；提供用户级数据存取权限的定义机制，系统自动检查用户能否执行这些操作，只有检查通过才能执行这些操作。

2) 数据的完整性 (Integrity) 控制，指 DBMS 提供必要的功能确保数据库数据的正确性、有效性和相容性。

3) 数据的并发 (Concurrency) 控制，指 DBMS 必须对多用户并发进程同时存取、修改操作进行控制和协调，以防止互相干扰导致错误结果。

(4) 数据库的建立和维护功能。包括数据库初始数据的装入，数据库的转储、恢复、重组织，系统性能监视、分析等功能，这些功能大部分由 DBMS 的实用程序来完成。

4. 应用程序

应用程序介于用户和数据库管理系统之间,是指完成用户操作的程序,该程序将用户的操作(例如,统计学生平均分,查询学生学籍信息等)转换成一系列的命令执行,而这些命令需要对数据库中的数据进行查询、插入、删除和统计,应用程序将这些复杂的数据库操作交由数据库管理系统来完成。

5. 用户

用户是指存储、维护和检索数据库中数据的使用人员,数据库系统中主要有 3 类用户:终端用户、应用程序员和数据库管理员。

(1) 终端用户:从计算机联机终端存取数据库的人员,也可称为最终用户。这类用户使用开发工具提供的终端命令语言、表格或菜单驱动等交互式对话方式来存取数据库中的数据。终端用户一般是指不精通计算机和程序设计的各级管理人员、工程技术人员和科研人员。

(2) 应用程序员:负责设计和编制应用程序的人员,也称为系统开发员。这类用户通常使用 C#、VB、PB 等程序设计语言来设计和编写应用程序,通过程序使用和维护数据库,并对数据库进行存取操作。

(3) 数据库管理员 (Database Administrator, DBA):全面负责数据库系统的“管理、维护和正常使用”的人员,它可以是一个人或一组人。DBA 的主要职责有:参与数据库设计的全过程,与用户、应用程序员、系统分析员紧密合作,设计数据库的结构和内容;决定数据库的存储与存取策略,使数据的存储空间利用率、存取效率均较高;定义数据的安全性和完整性;监督控制数据库的使用和运行,及时处理运行程序中出现的問題;改进和重新构造数据库系统等。

1.3.2 数据库系统体系结构

数据库系统有着严谨的体系结构。目前世界上大量的数据库正在运行中,其类型和规模可能相差很大,但是就其体系结构而言却是大体相同。

1. 数据库系统的三级模式结构

美国国家标准协会 (ANSI) 所属标准计划和要求委员会在 1975 年公布了一个关于数据库标准的报告,突出了数据库的三级模式结构,这就是有名的 SPARC 分级结构。

所谓模式 (Schema) 是数据库中全体数据的逻辑结构和特征描述,关系名及其属性的组合称为这个关系的模式。模式只是对实体的描述,而与具体的值无关。例如,学生(学号,姓名,性别,班号)就是模式。模式的具体值称为实例 (Instance),同一模式可以有多个实例。

从数据库管理的角度看,各数据库的体系结构都具有相同的特征。三级模式结构是从逻辑上对数据库的组织从内到外的三个层次描述,分别称为内模式、概念模式和外模式,如图 1.5 所示。

(1) 内模式,又称存储模式,具体描述了数据如何组织存储在存储介质上。内模式是系统程序员用一定的文件组织起来的一个存储文件和联系手段;也是由他们编制存取程序,实现数据存取的。一个数据库只有一个内模式。

(2) 概念模式,简称模式、概念视图或 DBA 视图,是对数据库的整体逻辑结构和特征的描述,不涉及数据的物理存储细节和硬件环境,与具体的应用程序和使用的应用开发工具无关,它由多个概念记录型组成,还包括记录间的联系、数据的完整性和其他数据控制方面

的要求。一个数据库只有一个概念模式。

(3) 外模式, 外模式通常是模式的一个子集, 故又称为子模式。外模式面向应用, 它是数据库用户能够看到和使用的局部数据结构和特征的描述, 是与某一应用有关的数据的逻辑表示。

综上所述, 概念模式是内模式的逻辑表述, 内模式是概念模式的物理实现, 外模式则是概念模式的部分抽取。三个模式反映了对数据库的三种不同观点: 概念模式表示了概念级的数据库, 体现了对数据库的整体观; 内模式表示了物理级的数据库, 体现了对数据库的存储

观; 外模式表示了用户级数据库, 体现了对数据库的用户观。整体观和存储观只有一个, 而用户观可能有多个, 一个应用对应一个用户观。

2. 三级模式映像

前面谈到的三级模式, 只有内模式才是真正存储数据的, 而概念模式和外模式仅是一种逻辑上表示数据的方法, 这是通过 DBMS 的映像功能实现的。

数据库的三级模式是对数据的三个抽象级别, 它将数据的具体组织留给 DBMS 管理, 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内的具体表示方式与存储方式。为了能够在内部实现这三个抽象层次的联系和转换, 数据库管理系统在这三级模式之间提供了以下两层映像:

(1) 外模式/概念模式映像。

(2) 概念模式/内模式映像。

正是这两层映像保证了数据库系统中数据能够有较高的逻辑独立性和物理独立性。

(1) 外模式/概念模式映像。概念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。对应于一个概念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/概念模式映像, 它定义了该外模式与该概念模式之间的关系。这些映像通常又包含在各自的外模式描述中。

当概念模式改变时 (例如增加新的关系、新的属性等), 由数据库管理员对各个外模式/概念模式的映像作相应改变, 可以使外模式保持不变。应用程序是依据数据的外模式编写的, 从而不必修改应用程序, 保证了数据与程序的逻辑独立性, 简称数据的逻辑独立性。

(2) 概念模式/内模式映像。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/内模式映像是唯一的, 它定义了数据库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概念模式描述中。当数据库的存储结构改变时 (例如选用了另外一种存储结构), 由数据库管理员对概念模式/内模式映像作相应改变, 可以使概念模式保持不变, 从而应用程序也不必改变, 保证了数据与程序的物理独立性, 简称数据的物理独立性。

在数据库的三级模式结构中, 概念模式即全局逻辑结构是数据库的中心与关键, 它独立

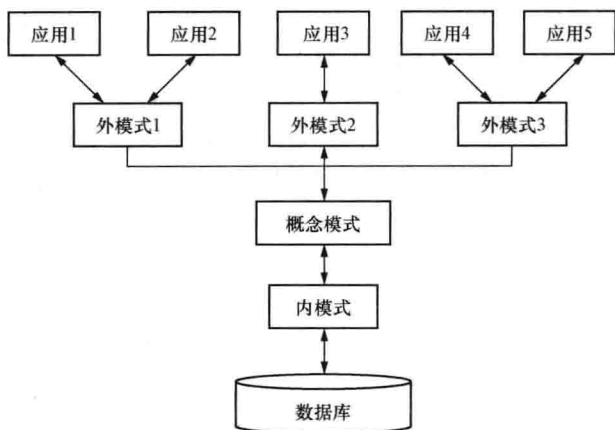


图 1.5 SPARC 分级结构

于数据库的其他层次, 因此设计数据库的模式结构时应首先确定数据库逻辑模式。

数据库的内模式依赖于它的全局逻辑结构, 但独立于数据库的用户视图 (即外模式), 也独立于具体的存储设备。它是将全局逻辑结构中定义的数据结构及其联系按照一定的物理存储策略进行组织, 以达到较好的时间与空间效率。

数据库的外模式面向具体的应用程序, 它定义在逻辑模式之上, 但独立于存储模式和存储设备。当用户需求发生较大变化, 相应外模式不能满足其视图要求时, 该外模式就需要作相应改动, 所以设计外模式时应充分考虑到应用的扩充性。

特定的应用程序是在外模式描述的数据结构之上编制的, 它依赖于特定的外模式, 与数据库的概念结构和存储结构独立, 不同的应用程序有时可以共用同一个外模式。数据库的二级映像保证了数据库外模式的稳定性, 从而从底层保证了应用程序的稳定性, 除非应用程序需求本身发生变化, 一般情况下, 应用程序不必修改。

数据与程序之间的独立性, 使得数据的定义和描述可以从应用程序分离出去。另外, 由于数据的存取由 DBMS 管理, 用户不必考虑存取路径细节, 从而简化了应用程序的编制, 大大减少了应用程序的维护和修改。

三级模式之间的比较如表 1.1 所示。

表 1.1

三级模式之间的比较

	外模式	概念模式	内模式
其他名字	子模式、用户模式	模式、DBA 视图	存储模式、内视图
描述	数据库用户能够看见和使用的局部数据的逻辑结构	数据库中全体数据的逻辑结构	数据物理结构和存储方式的描述
特点	用户与数据库的接口	所有用户的公共数据视图	数据在数据库内部的表示方式
	可以有多个外模式	只有一个概念模式	只有一个内模式
	面向用户程序或最终用户	由 DBA 定义	基本由 DBMS 定义



习 题 1

一、单项选择题

(1) 数据管理技术的发展过程经历了人工管理阶段、文件系统阶段和数据库系统阶段, 其中数据独立性最高的阶段是_____。

- A. 数据库系统 B. 文件系统 C. 人工管理 D. 数据项管理

(2) 数据库系统与文件系统的主要区别是_____。

- A. 数据库系统复杂, 而文件系统简单
B. 文件系统不能解决数据冗余性问题, 而数据库系统可以解决
C. 文件系统只能管理程序文件, 而数据库系统能够管理各种类型的文件
D. 文件系统管理的数据量较少, 而数据库系统可以管理庞大的数据量

(3) 下列关于数据库系统的叙述中正确的是_____。

- A. 数据库系统减少了数据冗余
B. 数据库系统避免了一切冗余

C. 数据库系统中的一致性是指数据类型一致

D. 数据库系统比文件系统管理更多的数据

(4) 通常所说的数据库系统 (DBS)、数据库管理系统 (DBMS) 和数据库 (DB) 三者之间的关系是_____。

A. DBMS 包含 DB 和 DBS

B. DB 包含 DBMS 和 DBS

C. DBS 包含 DB 和 DBMS

D. 三者无关

(5) 在数据库的三级模式结构当中, 描述数据库全局逻辑结构和特性的是_____。

A. 外模式

B. 内模式

C. 存储模式

D. 模式

(6) 一般的, 一个数据库系统的外模式_____。

A. 只能有一个

B. 最多只能有一个

C. 至少两个

D. 可以有多个

(7) 模式和内模式_____。

A. 只能有一个

B. 最多只能有一个

C. 至少两个

D. 可以有多个

(8) DBMS 是_____。

A. 操作系统的一部分

B. 在操作系统支持下的系统软件

C. 一种编译程序

D. 应用程序系统

(9) 数据库管理系统能实现对数据库中数据的查询、插入、修改和删除, 这类功能称为_____。

A. 数据定义功能

B. 数据管理功能

C. 数据操纵功能

D. 数据控制功能

(10) 数据库系统的数据独立性是指_____。

A. 不会因为数据的数值变化而影响应用程序

B. 不会因为系统数据存储结构与数据逻辑结构的变化而影响应用程序

C. 不会因为存储策略的变化而影响存储结构

D. 不会因为某些存储结构的变化而影响其他的存储结构

(11) 存储在计算机外部存储介质上的结构化的数据集合的英文名字是_____。

A. Data Dictionary

B. Data Base System

C. Data Base

D. Data Base Management System (DBMS)

(12) 数据库系统的核心是_____。

A. 数据库

B. 数据库管理系统

C. 数据模型

D. 软件工具

二、简答题

(1) 文件系统与数据库系统中的文件有何本质上的不同?

(2) 什么是数据独立性? 数据库系统是如何实现数据独立性的?

第2章 数 据 模 型

数据模型是某个数据库的框架,这个框架形式化地描述了数据、数据间的关系及在数据上的约束。同时,数据模型是定义数据库的依据,现有的数据库系统均是基于某种数据模型的。因此,了解数据模型的基本概念是学习数据库的基础。本章介绍数据模型的基本概念和常用的数据模型。

2.1 现实世界的抽象过程

计算机并不能自动识别现实世界,因此需要人们对现实生活中的事物进行抽象,将对现实世界的认识通过一定的方式表示为计算机能支持和识别的数据,整个过程如图 2.1 所示。

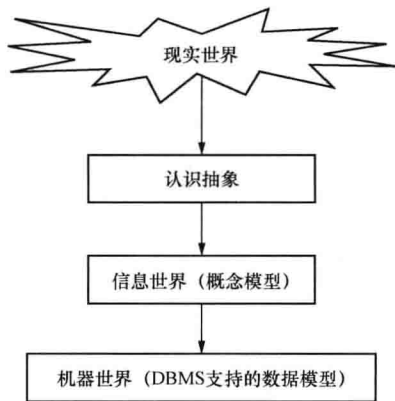


图 2.1 人类认识事物的过程

1. 现实世界

现实世界是存在于人们头脑之外的客观世界。在现实世界中,事物的存在不是孤立的,而是处于相互联系的状态中。现实世界中的事物与事物之间的联系是数据库最原始的信息,例如,学校教学管理中的学生与课程之间的联系,图书管理中图书与读者之间的联系等。

2. 信息世界

信息世界是现实世界在人们头脑中的反映。现实世界中的事物与事物之间的联系经过人们的选择、命名、分类之后进入信息世界,也就是将现实世界中的事物及联系抽象为某一种信息结构,这种信息结构并不依赖于具体的计算机系统,也不是某一种数据库管理系统支持的数据模型,只是概念级的模型,是对现实世界的第一次抽象,是从人的角度看世界。

3. 机器世界

机器世界又称为数据世界,是数据库的处理对象。信息世界的信息经过加工、编码转换为计算机上某一数据库系统能够接受处理的数据形式,即信息的数据化,这是对现实世界的第二次抽象,是从机器的角度看世界。

2.2 概 念 模 型

2.2.1 信息世界中的基本概念

概念模型是现实世界在头脑中的抽象反映,它不依赖于具体的计算机系统,是现实世界到机器世界的一个中间层次。概念模型用于信息世界的建模,是数据库设计人员进行数据库设计的有力工具,也是数据库设计人员和用户之间进行交流的语言。