



普通高等教育“十二五”应用型本科规划教材

数据库理论与应用

臧文科 乔 鸿 许文杰 编著

配电子课件



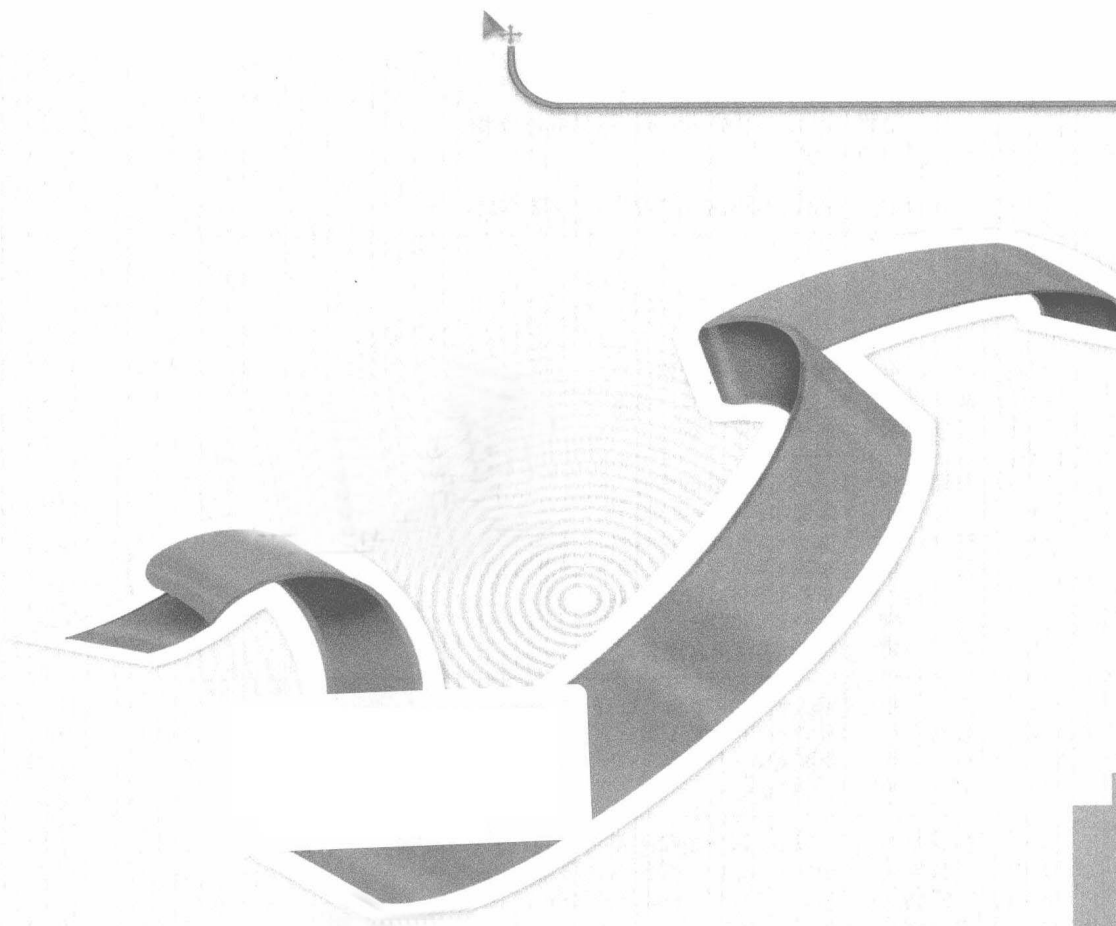
西安交通大学出版社
XI'AN JIAOTONG UNIVERSITY PRESS



普通高等教育“十二五”应用型本科规划教材

数据库理论与应用

臧文科 乔 鸿 许文杰 编著



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

内容简介

数据库系统是进行软件开发不可或缺的数据存储软件,主流软件的开发都离不开数据库的支撑。本书系统全面地阐述了数据库系统的基础理论、基本技术和基本应用。全书分为10章,主要讲解了数据库基本知识,关系数据库理论,SQL语言,数据库保护的概念和方法,数据库的设计方法与技术,Oracle数据库和SQL Server数据库及其使用。针对时下流行的ASP.NET开发,从Web开发角度做了重点阐述,希望能够从中获得对Web数据库开发的基本认识。讲解条理清楚,内容深浅适中。

本书讲解详尽,内容的取舍和安排循序渐进,通俗易懂,实例丰富,并注重培养解决实际问题的能力。本书可以作为各高等学校计算机及相关专业学习数据库课程的教材,也可作为广大计算机爱好者的自学用书。

图书在版编目(CIP)数据

数据库理论与应用/臧文科编著. —西安:西安交通大学出版社,2015.8
ISBN 978-7-5605-7657-2

I. ①数… II. ①臧… III. ①数据库系统
IV. ①TP311.13

中国版本图书馆CIP数据核字(2015)第162805号

书 名 数据库理论与应用
编 著 臧文科 乔 鸿 许文杰
责任编辑 毛 帆 曹 旻

出版发行 西安交通大学出版社
(西安市兴庆南路10号 邮政编码 710049)
网 址 <http://www.xjtpress.com>
电 话 (029)82668357 82667874(发行中心)
(029)82668315(总编办)
传 真 (029)82668280
印 刷 虎彩印艺股份有限公司

开 本 787mm×1092mm 1/16 印张 12.75 字数 303千字
版次印次 2015年10月第1版 2015年10月第1次印刷
书 号 ISBN 978-7-5605-7657-2/TP·677
定 价 31.00元

读者购书、书店添货、如发现印装质量问题,请与本社发行中心联系、调换。

订购热线:(029)82665248 (029)8266524

投稿热线:(029)82669097 QQ:8377981

读者信箱:lg_book@163.com

版权所有 侵权必究

前言

FOREWORD

数据库是现代化数据管理的最重要、最广泛、最先进的技术,是计算机科学的重要分支。本书为计算机及相关的众多学科提供了利用计算机技术进行数据管理的基本理论知识,是计算机科学与技术、软件工程及其相关专业学科的专业必修课。

本书主要介绍数据库的基本理论和数据库管理系统的基本应用技术。全书共 10 章。第 1 章数据库概论,主要介绍数据库的产生与发展,包括数据库、数据库管理系统和数据库系统的概念,数据模型与数据库体系结构等内容。第 2 章关系数据库,主要介绍关系数据模型、关系模型的完整性规则、关系代数的基本运算等。第 3 章关系数据理论,主要介绍实体类型的属性关系、数据的函数依赖、关系数据库模型的规范化理论、关系模式的分解算法等。第 4 章 SQL 语言,主要介绍标准 SQL 概述,SQL 的数据定义、数据查询、数据更新等。第 5 章数据库保护,主要介绍数据库的安全性、完整性、并发控制、备份与恢复等。第 6 章数据库设计,包括数据库设计概述、需求分析、概念结构设计、逻辑结构设计、物理结构设计、数据库的实施和维护。第 7 章和第 8 章介绍了 Oracle 数据库和 SQL Server 数据库管理系统的基本操作。第 9 章数据库编程,主要介绍了存储过程、触发器、事务等。第 10 章中结合了 SQL Server 和 ASP.NET 应用的开发过程,阐述了数据库在开发过程中的应用。

参加本书编写的人员还有任丽艳、毕雪、毕莹、孙佃良、马凤芳、李丽君等,他们也认真地校对了书中的内容;在试用过程中,许多老师和读者也对本书的编写提出了许多宝贵建议和修改意见,在此表示感谢。由于时间仓促,加之编者水平有限,书中疏漏和不当之处难免,恳请读者批评指正。

编 者
2015 年 9 月

第1章 数据库概论

- 1.1 基础概念 / 001
 - 1.1.1 发展历史 / 001
 - 1.1.2 基本概念 / 001
 - 1.1.3 数据管理技术的发展 / 003
- 1.2 数据描述 / 005
 - 1.2.1 三种设计 / 006
 - 1.2.2 数据联系 / 008
- 1.3 数据模型 / 010
 - 1.3.1 模型概念 / 010
 - 1.3.2 结构模型 / 011
 - 1.3.3 实体联系模型 / 011
 - 1.3.4 关系模型 / 014
- 1.4 体系结构 / 016
 - 1.4.1 数据库模式 / 016
 - 1.4.2 三级模式结构 / 016
 - 1.4.3 二级映像 / 018
 - 1.4.4 二级数据独立 / 019
- 1.5 DBMS / 020
 - 1.5.1 DBMS 的组成 / 020
 - 1.5.2 DBMS 的主要功能 / 020
 - 1.5.3 用户访问数据库的过程 / 020

第2章 关系数据库

- 2.1 关系数据库 / 023
- 2.2 关系模型 / 023
 - 2.2.1 基本概念 / 024
 - 2.2.2 基本关系的性质 / 025
 - 2.2.3 完整性 / 025
 - 2.2.4 形式定义 / 026

数据库理论与应用

- 2.3 关系代数 / 027
 - 2.3.1 基本操作 / 027
 - 2.3.2 组合操作 / 029
 - 2.3.3 表达式 / 032
- 2.4 关系演算 / 033
 - 2.4.1 元组关系演算 / 034
 - 2.4.2 域关系演算 / 034

第3章 关系数据理论

- 3.1 函数依赖 / 035
 - 3.1.1 各种定义 / 035
 - 3.1.2 码 / 036
 - 3.1.3 推理规则 / 037
 - 3.1.4 闭包 / 038
 - 3.1.5 最小函数依赖集 / 039
- 3.2 范式 / 042
 - 3.2.1 第一范式(1NF) / 042
 - 3.2.2 第二范式(2NF) / 043
 - 3.2.3 第三范式(3NF) / 043
 - 3.2.4 BC 范式(BCNF) / 044
 - 3.2.5 存在问题 / 044
- 3.3 分解 / 045
 - 3.3.1 分解原则 / 045
 - 3.3.2 无损连接性 / 047
 - 3.3.3 模式分解 / 049

第4章 SQL 语言

- 4.1 SQL 概述 / 052
 - 4.1.1 SQL 发展历程 / 052
 - 4.1.2 SQL 体系结构 / 052
 - 4.1.3 SQL 组成 / 057
- 4.2 SQL 定义 / 058
 - 4.2.1 模式定义 / 058

数据库理论与应用

- 4.2.2 表的定义 / 058
- 4.2.3 视图的定义 / 061
- 4.2.4 索引的定义 / 062
- 4.3 SQL 查询 / 062
 - 4.3.1 单表查询 / 063
 - 4.3.2 连接查询 / 067
 - 4.3.3 嵌套查询 / 067
- 4.4 SQL 更新 / 069
 - 4.4.1 插入 / 069
 - 4.4.2 修改 / 069
 - 4.4.3 删除 / 069

第5章 数据库保护

- 5.1 数据库安全性 / 071
 - 5.1.1 权限 / 071
 - 5.1.2 授权与回收 / 072
 - 5.1.3 SQL Server 的安全性机制 / 074
 - 5.1.4 数据库角色 / 074
- 5.2 数据库完整性 / 075
 - 5.2.1 几类完整性 / 076
 - 5.2.2 SQL Server 的完整性 / 076
- 5.3 数据库恢复 / 079
 - 5.3.1 事务 / 079
 - 5.3.2 恢复的原则与方法 / 081
 - 5.3.3 SQL 中的恢复操作 / 081
- 5.4 数据库并发性控制 / 081
 - 5.4.1 并发操作问题 / 081
 - 5.4.2 封锁 / 083
 - 5.4.3 死锁与活锁 / 084
 - 5.4.4 两段封锁预防 / 085

第6章 数据库设计

- 6.1 设计概述 / 088

数据库理论与应用

- 6.1.1 软件方法学 / 088
- 6.1.2 数据库设计方法学 / 088
- 6.1.3 数据库设计步骤 / 090
- 6.2 需求分析 / 092
 - 6.2.1 需求描述 / 092
 - 6.2.2 需求分析方法 / 092
 - 6.2.3 需求分析步骤 / 093
 - 6.2.4 数据字典 / 095
- 6.3 概念结构设计 / 096
 - 6.3.1 概念模型 / 096
 - 6.3.2 概念设计步骤 / 097
 - 6.3.3 数据抽象 / 099
 - 6.3.4 局部 E-R 模型设计 / 100
 - 6.3.5 全局概念结构设计 / 101
 - 6.3.6 实例 / 102
- 6.4 逻辑结构设计 / 107
 - 6.4.1 逻辑结构设计步骤 / 107
 - 6.4.2 E-R 模型向关系模型的转换 / 107
 - 6.4.3 数据模型的优化 / 108
- 6.5 物理设计 / 109
 - 6.5.1 物理设计环境 / 109
 - 6.5.2 物理设计步骤 / 109
- 6.6 实施与维护 / 110
 - 6.6.1 数据库实现 / 110
 - 6.6.2 试运行 / 110
 - 6.6.3 运行与维护 / 111

第 7 章 Oracle 数据库

- 7.1 介绍 / 113
 - 7.1.1 Oracle 的成长历程 / 113
 - 7.1.2 Oracle 与 SQL Server / 113
 - 7.1.3 Oracle 的产品版本 / 114

数据库理论与应用

- 7.1.4 Oracle 的最新版本 / 115
- 7.1.5 Oracle 12c 的新特性 / 115
- 7.2 安装 / 117
 - 7.2.1 软件下载 / 117
 - 7.2.2 使用 Oracle 11g 的基本条件 / 118
 - 7.2.3 在 Windows 系统中安装 Oracle 11g / 118
 - 7.2.4 软件配置 / 123
- 7.3 使用 / 126
 - 7.3.1 自带工具 / 126
 - 7.3.2 辅助工具 / 128
 - 7.3.3 建立数据库 / 131
- 7.4 应用 / 137
 - 7.4.1 创建数据库 / 137
 - 7.4.2 数据的查询、修改和删除 / 138
 - 7.4.3 视图 / 141

第 8 章 SQL Server 数据库

- 8.1 介绍 / 142
 - 8.1.1 SQL Server 的成长历程 / 142
 - 8.1.2 SQL Server 的产品版本 / 142
 - 8.1.3 SQL Server 2012 / 144
- 8.2 安装 / 145
 - 8.2.1 SQL Server 2012 数据库安装 / 145
 - 8.2.2 安装的系统要求 / 145
 - 8.2.3 安装步骤 / 145
- 8.3 使用 / 155
 - 8.3.1 SQL Server 工具 / 155
 - 8.3.2 建立数据库——建立你的第一个数据库和表 / 156

第 9 章 数据库编程

- 9.1 存储过程 / 160
 - 9.1.1 介绍 / 160
 - 9.1.2 定义 / 161

数据库理论与应用

- 9.1.3 调用 / 163
- 9.2 游标 / 165
 - 9.2.1 介绍 / 165
 - 9.2.2 定义 / 166
- 9.3 触发器 / 168
 - 9.3.1 介绍 / 168
 - 9.3.2 定义 / 169
 - 9.3.3 使用 / 171
- 9.4 事务 / 173
 - 9.4.1 定义 / 173
 - 9.4.2 使用 / 174

第 10 章 数据库应用

- 10.1 ADO.NET 概述 / 178
 - 10.1.1 ADO.NET 简介 / 178
 - 10.1.2 ADO.NET 组件 / 179
 - 10.1.3 ADO.NET 数据提供者 / 179
- 10.2 数据库访问 / 179
 - 10.2.1 数据库类型映射 / 179
 - 10.2.2 数据库连接 / 180
 - 10.2.3 数据库操作 / 181
- 10.3 数据库绑定 / 184
- 10.4 存储过程调用 / 186
 - 10.4.1 返回单一记录集 / 186
 - 10.4.2 没有输入输出 / 187
 - 10.4.3 有返回值 / 188
 - 10.4.4 有输入参数和输出参数 / 189
- 10.5 数据库应用 / 190

1.1 基础概念

1.1.1 发展历史

在信息化高速发展的今天,数据库对于每个人来说都是生活中不可缺少的一部分。许多信息系统都是以数据库为基础建立起来的,数据库本身存在的重要性的应用的广泛性决定了其在数字化的今天的地位。数据库技术是研究数据库的结构、存储、设计、管理和使用的一门软件学科,是信息化社会的重要基础技术之一,是计算机科学领域中发展最为迅速的一个重要分支。它是计算机信息系统与应用系统的核心技术和重要基础,是人们存储数据、管理信息和共享资源的最常用、最先进的技术,在科学、技术、经济、文化和军事等各个领域都发挥着重要的作用。

20 世纪 50 年代,美国为了战争的需要,把各种情报集中存储在计算机中,被称为 Database,数据库这个概念由此诞生。

20 世纪 60 年代末,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,数据库技术作为数据处理的一种新手段迅速发展起来。

20 世纪 70 年代和 80 年代是数据库蓬勃发展的时期,在这一时期,出现了一些层次模型数据库系统和网状模型数据库系统,围绕关系数据模型的大量研究和开发工作也使关系数据库理论和关系模型数据库系统日益完善。后来关系模型数据库系统凭借自身优越性,逐渐取代网状数据库系统和层次模型数据库系统。至今,关系模型数据库系统仍具有重要地位。

20 世纪 90 年代,关系模型数据库技术又有了进一步的发展。数据库技术与面向对象技术、网络技术相互渗透,产生了面向对象的数据库和网络数据库。

21 世纪以来,面向对象数据库和网络数据库技术逐渐成熟并得到了广泛的应用。

数据库从出现到现在仅仅几十年的时间,就已经形成了坚实的理论基础,成熟的商业产品和广泛的应用领域,它的诞生和发展给计算机信息管理带来了一场巨大的革命。数据库系统的出现使信息系统从以加工数据的程序为中心转向围绕共享的数据为中心的新阶段。

1.1.2 基本概念

1. 信息 (information)

信息泛指以各种方式传播的、可被感知的数字、文字、图像、声音等符号所表征的某一事物的消息、情报和知识,它是观念性的东西,是人们头脑对现实世界的抽象反映。

2. 数据(data)

数据,实际上就是描述事物的符号记录,通常指用符号记录下来的、可以识别的信息。数据是所有计算机系统所要处理的对象,是数据库中存储的基本对象。

数字是最简单的一种数据,但数据不仅仅包括数字,数据有多种形式,如文字、符号、图形、图像以及声音等,它们都可以经过数字化存入计算机。例如,当用书号、书名、单价、作者、出版社来描述一本书时(000001,数据库理论与应用,49.00元,张刚,中国人民大学出版社),这些描述就是这本书的数据,我们可以从这一数据的含义中得到这本书的有关信息。

数据有一定的格式,这些格式的规定就是数据的语法,而数据的含义就是数据的语义。通过解释、推理、归纳、分析和综合等,从数据中所获得的有意义的内容称为信息。

信息与数据之间存在着固有的联系:数据是信息的符号表示或载体;信息则是数据的内涵,是对数据语义的解释。例如在“一个和尚挑水喝,两个和尚抬水喝,三个和尚没水喝”这句谚语中,一个和尚、两个和尚和三个和尚是数据;一个和尚挑水喝,两个和尚抬水喝,三个和尚没水喝是信息。

3. 数据库(DataBase, 简称 DB)

数据库是长期存储在计算机内、有组织的、统一管理的相关数据的集合。简单来说,数据库就是存放数据的仓库。数据库能为各种用户共享,具有冗余度小、数据间联系紧密和数据独立性高等特点。例如,图书馆可能同时有描述图书的数据(图书编号,书名,作者,单价,出版社,出版日期)和图书借阅数据(图书编号,书名,单价,借阅者,借阅天数),两者中图书编号、书名、单价是冗余数据。在构造数据库时,由于数据可共享,只保存一套数据就可以,从而消除数据冗余。

4. 数据库管理系统(DataBase Management System, 简称 DBMS)

数据库管理系统是一种系统软件,负责数据库中的数据组织、数据操纵、数据维护、控制及保护和数据服务等,是数据库的核心。它对数据库进行统一的管理和控制,以保证数据库的安全性和完整性。用户通过 DBMS 访问数据库中的数据,数据库管理员也通过 DBMS 进行数据库的维护工作,它可以使多个应用程序和用户用不同的方法在同一时刻或不同时刻去建立、修改和访问数据库。它的功能我们将在以后章节中作更加详细的描述。

数据库管理系统是数据库系统的一个重要组成部分,在一个数据库管理系统中可以有多个数据库。

5. 关系型数据库管理系统(Relational DataBase Management System, 简称 RDBMS)

在数据库管理系统的基础上增加关系即为关系型数据库管理系统,它通过数据、关系和由对数据库的约束组成的数据模型来存放和管理数据。

6. 数据库系统(DataBase System, 简称 DBS)

数据库系统是由数据库及其管理软件组成的系统。它是为适应数据处理的需要而发展起来的一种较为理想的数据处理的核心机构。它是一个实际可运行的集存储、维护和为应用系统提供数据功能为一体的软件系统,是存储介质、处理对象和管理系统的集合体。

数据库系统通常由软件、数据库和数据管理员组成。其软件主要包括操作系统、各种宿主语言、实用程序以及数据库管理系统。

数据库系统的目标是解决数据冗余问题,实现数据独立性,实现数据共享并解决由于数

据共享而带来的数据完整性、安全性及并发控制等一系列问题。为实现这一目标,数据库的运行必须有一个软件系统来控制,这个系统软件即上面提到的数据库管理系统 DBMS,因此我们说数据库管理系统是数据库系统的重要组成部分。

数据库系统可以用图 1.1 表示。

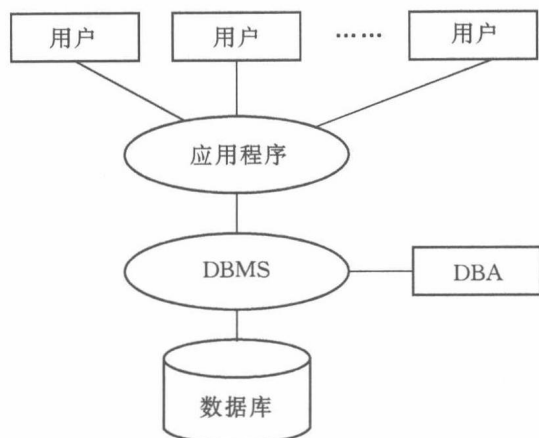


图 1.1 数据库系统

1.1.3 数据管理技术的发展

随着计算机技术的发展,特别是在计算机软件、硬件与网络技术发展的前提下,人们的数据处理要求不断提高,对数据处理速度及规模的需求远远超出了过去人工或机械方式的能力范围,在此情况下,数据管理技术也不断改进。数据管理具体就是指人们对数据进行收集、组织、存储、加工、传播和利用的一系列活动总和,经历了人工管理、文件管理、数据库管理三个阶段,每个阶段的发展以数据存储冗余不断减小、数据独立性不断增强、数据操作更加方便和简单为标志,各有各的特点。

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。硬件方面,计算机的外存只有磁带、卡片、纸带,没有磁盘等直接存取的存储设备,存储量非常小;软件方面,没有操作系统,没有高级语言,数据处理的方式是批处理,即机器一次处理一批数据,直到运算完成为止,然后才能进行另外一批数据的处理,中间不能被打断。原因是此时的外存如磁带、卡片等只能顺序输入。

人工管理阶段的数据具有以下几个特点:

- ①数据不保存。
- ②数据不具有独立性。
- ③数据不共享。数据是面向应用程序的,一组数据对应一个程序。不同应用的数据之间是相互独立、彼此无关的,即使两个不同应用涉及到相同的数据,也必须各自定义,无法相互利用,互相参照,因此程序与程序之间有大量的冗余数据。如图 1.2 所示。
- ④由应用程序管理数据。数据没有专门的软件进行管理,需要应用程序自己进行管理,

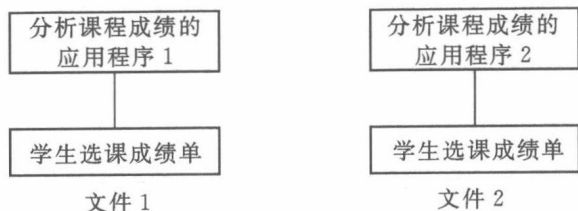


图 1.2 两个应用程序各自使用同一数据

应用程序中要规定数据的逻辑结构并设计物理结构(包括存储结构、存取方法、输入/输出方式等)。

2. 文件系统阶段

20 世纪 50 年代后期到 60 年代中期,数据管理发展到文件系统阶段。此时的计算机不仅用于科学计算,还大量用于管理。外存储器有了磁盘等直接存取的存储设备;在软件方面,操作系统中已有了专门的管理数据软件,称为文件系统。从处理方式上讲,不仅能够进行文件批处理,而且能够联机实时处理,联机实时处理是指在需要的时候随时从存储设备中查询、修改或更新,因为操作系统的文件管理功能提供了这种可能。

文件系统阶段如图 1.3 所示。



图 1.3 文件系统阶段

文件系统阶段具有以下几个特点:

①数据长期保存。

②程序和数据具有一定的独立性,程序和数据分开存储。比如要保存数据时,只需给出保存指令,而不用程序员精心设计一套程序,控制计算机物理地实现保存数据。在读取数据时,只要给出文件名,而不必知道文件的具体的存放地址。文件的逻辑结构和物理存储结构由系统进行转换,程序与数据有了一定的独立性。数据的改变不一定要引起程序的改变,例如保存的文件中有 10 条记录,使用某一个查询程序;当文件中有 10000 条记录时,仍然使用保留的这一个查询程序。

③可以实时处理。由于有了直接存取设备,也有了索引文件、链接存取文件、直接存取文件等,所以既可以采用顺序批处理,也可以采用实时处理方式。

虽然较人工管理阶段有了很大的改进,但文件系统仍存在许多问题:数据和程序缺乏足

够的独立性;当不同的应用程序所需的数据有部分相同时,仍需建立各自的独立数据文件,而不能共享相同的数据。因此,数据冗余大,空间浪费严重,而且相同的数据重复存放,各自管理,当相同部分的数据需要修改时比较麻烦,稍有不慎,就造成数据的不一致。例如,学籍管理需要建立包括学生的姓名、班级、学号等数据的文件。这种逻辑结构和学生成绩管理所需的数据结构是不同的。在学生成绩管理系统中,进行学生成绩排列和统计,程序需要建立自己的文件,除了成绩等数据外,还要有姓名、班级等与学籍管理系统的数据文件相同的数据,数据冗余是显而易见的。此外当有学生转学走或转来时,两个文件都要修改,否则,就会出现有某个学生的成绩,却没有该学生的学籍的情况,反之亦然。如果系统庞大,则会牵一发而动全身,一个微小的变动将引起一连串的变动,利用计算机管理的规模越大,问题就越多。

文件管理系统存在的问题严重阻碍了数据处理技术的发展,但正是这些问题的存在,使得人们对数据管理技术展开了更加深入的研究,进而使数据处理逐渐进入了数据库系统阶段。

3. 数据库系统阶段

从 20 世纪 60 年代后期开始,数据管理进入数据库系统阶段。这一时期计算机管理的规模日益庞大,应用越来越广泛,数据量急剧增长,数据共享的需求越来越迫切。这一时期的计算机有了大容量磁盘,计算能力也非常强。硬件价格下降,编制软件和维护软件的费用相对增加。联机实时处理的要求更多,并开始提出和考虑并行处理。在这样的背景下,数据管理技术进入数据库系统阶段。

数据库系统阶段如图 1.4 所示。

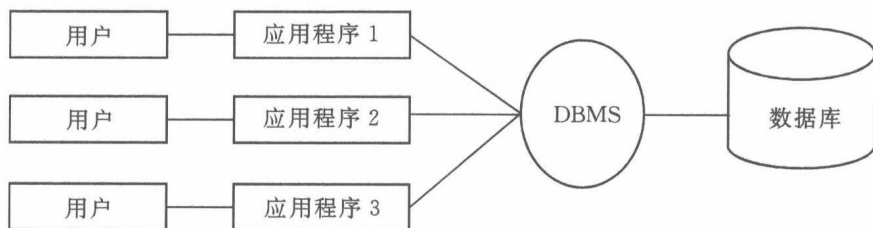


图 1.4 数据库系统阶段

数据库系统阶段具有以下几个特点:

- ①实现数据共享,减少数据冗余;
- ②采用特定的数据模型;
- ③具有较高的数据独立性;
- ④有统一的数据控制功能以及安全性、完整性控制。

1.2 数据描述

在计算机进行数据处理过程中,数据的表示要经历 3 个阶段:现实世界、信息世界和计算机世界的描述。

1.2.1 三种设计

如何将现实中“看得见”、“摸得着”的事物变成计算机能够处理的数据,这中间需要一个复杂的转换过程。比如,如何认识、理解、整理、描述和加工现实生活中的这些事物。从数据转化的顺序来说,数据从现实世界进入到数据库需要经历 3 个阶段:现实世界阶段、信息世界阶段和机器(计算机)世界阶段。

现实世界是指客观存在的世界中的事实及其联系。在这阶段要对现实世界的信息进行收集、分类,并抽象成信息世界的描述形式,然后再将其描述转换成计算机世界中的数据描述。

信息世界是现实世界在人们头脑中的反映,是对客观事物及其联系的一种抽象描述。为了正确直观的反映客观事物及其联系,有必要对所研究的信息世界建立一个抽象的模型,称之为概念模型,它是对现实世界的第一层抽象。目前流行的一种概念模型是实体联系模型(后面会详细介绍),它用实体联系图来描述。在数据库设计中,这一阶段又称为概念设计阶段。

计算机世界是信息世界中信息的数据化,即将信息用字符和数值表示,便于计算机识别和处理。在计算机世界中,用逻辑模型来描述现实世界,它是对现实世界的第二层抽象。这一阶段的数据处理在数据库的设计过程中也称为逻辑设计。

计算机世界和信息世界相关术语的对应关系如表 1.1 所示。

表 1.1 计算机世界和信息世界相关术语的对应关系

信息世界	计算机世界
实体	记录
属性	字段
实体集	文件
实体标识符	关键字

下面我们从三个方面介绍数据库中的数据描述:概念设计中的数据描述、逻辑设计中的数据描述、物理数据描述。

1. 概念设计中的数据描述

数据库中概念设计阶段需要根据用户需求设计数据库,设计人员和用户共同参与,要求数据描述简单且易于理解。概念设计中主要术语如下:

(1)实体。现实世界中客观存在并可相互区分的事物叫做实体。一个实体对应了现实世界中的一个事物。实体可以是一个具体的人或物,如一名学生、一朵花等;也可以是抽象的事件或概念,如购买一本图书等。

(2)属性。客观存在的不同事物,具有不同的特性。从客观世界抽象出的不同实体,也具有各自不同的特性。实体的某一特性称为属性,如学生实体有学号、姓名、年龄、性别、系等方面的属性。属性有“型”和“值”之分,“型”即为属性名,如姓名、年龄、性别是属性的型;“值”即为属性的具体内容(如 990001,张三,21,男,信管),这些属性值的集合表示了一个学生实体。

(3)实体型。若干个属性型组成的集合可以表示一个实体的类型,简称实体型。如学生(学号,姓名,年龄,性别,系)就是一个实体型。

(4)实体集。同型实体的集合称为实体集。如所有的学生、所有的课程等。

(5)码。能唯一标识一个实体的属性或属性集称为实体的码,如学生的学号。学生的姓名可能有重名,不能作为学生实体的码。

(6)域。属性值的取值范围称为该属性的域。例如学号的域为6位整数,姓名的域为字符串集合,性别的域为(男,女)。

(7)联系。在现实世界中,事物内部以及事物之间是有联系的,这些联系同样也要抽象和反映到信息世界中来。在信息世界中将被抽象为实体型内部的联系和实体型之间的联系。

2. 逻辑设计中的数据描述

数据库的逻辑设计是根据概念设计所得的概念模型来设计数据库的逻辑结构,也就是DBMS所支持的数据结构。逻辑设计中主要术语如下:

(1)字段:标识实体属性的命名单位称为字段,字段名往往和属性同名。

(2)记录:字段的有序集合称为记录。通常用一个记录描述一个实体。

(3)文件:同一类记录的集合称为文件。文件用来描述实体集。

(4)关键字:能唯一表示文件中每个记录的字段或字段集,称为记录的关键字。

(5)记录型:对应实体型。

概念设计与逻辑设计中术语对应关系如表1.2所示。

表 1.2 概念设计与逻辑设计中术语对应关系

概念设计	实体	实体集	实体型	属性	实体标识符
逻辑设计	记录	文件	记录型	字段(域、数据项)	关键字(键、码)

3. 物理数据描述

在上一节中我们提到的字段、记录、文件等都是数据的重要单位,称为逻辑数据。它们之间的关系可以形象的表述为:数据库是数据文件的集合,数据文件是记录的集合,记录是字段的集合。当把它们存储到计算机的存储介质上时,就称为物理数据。

计算机存储介质是计算机存储器中用于存储某种不连续物理量的媒体。简单地说,存储介质是指存储数据的载体。存储器是计算机系统上的记忆设备,它的主要功能是存储程序和各种数据,并能在计算机运行过程中高速自动地完成程序或数据的存取。

存储器根据不同的分类标准(例如按照存取方式不同,存取介质不同等)可以分为不同类型。根据存储器在计算机系统中的作用划分,计算机系统中存储层次可分为高速缓冲存储器、主存储器、辅助存储器三级。

高速缓冲存储器位于主存储器与CPU之间,用来改善主存储器与CPU的速度匹配问题。特点是存储量较小,价格较高,但存储速度快。

主存储器即内存,用来保存当前正在执行的程序和数据,可以被CPU直接随机地进行读写访问,但这些程序和数据仅用于暂时存放,断电就会导致信息丢失。主存具有一定容