

国家精品课“数据库原理与应用”配套教材
卓越系列·国家示范性高等职业院校重点建设专业教材
(计算机类)

数据库原理与应用

主编 孙 锋

参编 王庆桦



内 容 简 介

为了更好地适应高职高专课程教学的需要,本书强调“理论带动实训,实训推动理论”的教学方式。本书分为两部分共 17 个单元,系统地介绍了数据库的基础理论知识和数据库的应用技术。其中,理论篇——数据库原理,主要讲述数据库系统导论、数据模型、关系数据库的基本理论、结构化查询语言 SQL、关系数据库规范理论、数据库设计与维护以及数据库的保护与安全等内容;实训篇——数据库应用,通过微软最新的 Visual FoxPro 9.0 开发平台,以一个已经实际应用的实训项目“单位联系管理系统”为主线,逐步介绍数据库和数据表的创建、表单与报表、结构化程序设计、菜单设计以及应用程序的发布等内容。

本书不仅可以作为高职高专计算机及相关专业“数据库原理与应用”课程教材,而且还可以供相关技术人员参考。

图书在版编目(CIP)数据

数据库原理与应用 / 孙锋主编. —天津:天津大学出版社,2008.6

ISBN 978-7-5618-2683-6

I. 数… II. 孙… III. 数据库系统-高等学校-教材
IV. TP311.13

中国版本图书馆 CIP 数据核字(2008)第 075509 号

出版发行 天津大学出版社

出版人 杨欢

地 址 天津市卫津路 92 号天津大学内(邮编:300072)

电 话 发行部:022—27403647 邮购部:022—27402742

印 刷 廊坊市长虹印刷有限公司

经 销 全国各地新华书店

开 本 169mm×239mm

印 张 21.75

字 数 464 千

版 次 2008 年 6 月第 1 版

印 次 2008 年 6 月第 1 次

印 数 1—3 000

定 价 36.00 元

凡购本书,如有缺页、倒页、脱页等质量问题,烦请向我社发行部门联系调换。

版权所有 侵权必究

前 言

数据库技术是近年来计算机科学技术领域发展速度最快的方向之一,已经成为计算机信息系统与应用系统的核心技术和重要基础。数据库技术在各行业得到了广泛的应用,成为存储、使用和更新信息资源的主要手段。

数据库技术包括数据库系统原理和数据库应用系统两大部分,前者是后者的理论基础,只有在正确理论指导下才能设计出合理、适用的数据库系统。本教材首先给出必要的理论基础,然后在理论指导下进行实践。全书分为两部分,共 17 个单元。

第一部分 理论篇——数据库原理(第 1 章到第 7 章)

第 1 章 数据库系统导论,主要讲述数据管理技术的发展、数据库系统的结构和数据库管理系统。

第 2 章 数据模型,主要讲述概念模型和数据模型。

第 3 章 关系数据库的基本理论,主要讲述关系模型的基本概念、关系代数的基本运算和关系的完整性。

第 4 章 结构化查询语言 SQL,主要讲述 SQL 概述、数据定义语句、数据查询语句、数据操纵语句、视图、数据控制语句和嵌入式 SQL。

第 5 章 关系数据库规范理论,主要讲述问题提出、函数依赖、关系模式的范式和关系模式的规范化。

第 6 章 数据库设计与维护,主要讲述数据库设计概述、需求分析、概念结构设计、逻辑结构设计、数据库的物理设计、数据库的实施和数据库的运行与维护。

第 7 章 数据库的保护与安全,主要讲述数据库的并发控制技术、数据库的安全性、数据库的完整性控制和数据备份与恢复技术。

第二部分 实训篇——数据库应用(分为四大模块,实训单元 1 到实训单元 10)

实训单元 1 数据库和数据表的创建,主要讲述 Visual FoxPro 9.0 简介、项目管理器、数据库操作和数据表操作。

实训单元 2 数据表的进一步应用,主要讲述编辑记录中的数据、表的排序与索引、表的数据统计、表间的关系、参照完整性和表间操作命令。

实训单元 3 查询和视图,主要讲述使用命令方式查询、查询和视图。

实训单元 4 设计表单,主要讲述创建表单、表单数据环境和运行、表单控件和管理表单。

实训单元 5 报表,主要讲述创建报表、报表设计器、报表控件、打印报表和标签。

实训单元 6 结构化程序设计,主要讲述 Visual FoxPro 编程方式、程序文件的使用

用、常用命令、Visual FoxPro 的语言基础、Visual FoxPro 的基本语句和过程与函数。

实训单元 7 面向对象程序设计(一),主要讲述标签、命令按钮与命令按钮组、文本框、选项按钮组、复选框、列表框、组合框和页框。

实训单元 8 面向对象程序设计(二),主要讲述编辑框、表格控件、计时器、微调器和其他控件。

实训单元 9 联系单位管理系统开发,主要讲述系统集成和系统发布。

实训单元 10 产品生产销售管理系统开发,主要讲述菜单设计器。

本书具有以下特色。

1. 完整的结构设计。本书既包括数据库原理也包括数据库应用,适合按层次组织教学的需要。

2. 涉及内容适度。在数据库原理部分中,将数据库原理的最基础的部分提炼出来,做了深入浅出的论述,以指导应用系统的设计;在数据库应用部分中,以应用较广的 Visual FoxPro 9.0 为工具,讲解了数据库应用系统的基本操作和简单管理信息系统的开发。

3. 适用性与针对性强。本书在编写上突出高等职业教育的特点,强调理论与实训充分结合,突出职业技能训练。

4. 完整的实训项目制作步骤。本书的实训项目实践性较强,实际应用价值高。本书设计了一个完整的实训示例贯穿整个实训部分,由浅入深介绍管理信息系统的开发步骤和 Visual FoxPro 9.0 的使用方法,并且与理论教学内容相呼应,起到了活学活用的目的。另外,每个实训单元后配有实训项目环节,提出了实训要求,并提供了实训题目和必要的指导。整个实训项目完成后将得到一个完整的实用开发项目。

本书由 2006 年国家级精品课程“数据库原理与应用”课题组教师独立编写,作者都是长期在第一线从事高职高专院校计算机技术教学的专业教师,对高职高专学生的基础、特点和认知规律有深入的研究,在教学实践中积累了丰富的经验。本书由孙锋副教授主编,王庆桦参编。其中,第 5 章到第 7 章、实训单元 1 到实训单元 5 由孙锋编写;第 1 章到第 4 章、实训单元 6 到实训单元 10 由王庆桦编写。

由于水平有限、时间仓促,书中错漏之处在所难免,恳请读者批评指正。

编者

2007 年 12 月

第一部分

理论篇——数据库原理

第 1 章 数据库系统导论

随着信息技术的不断进步,用于管理信息资源的数据库技术也得到了快速的发展,其应用领域也越来越广泛,涉及办公自动化系统、管理信息系统、专家系统、过程控制、联机分析处理、计算机辅助设计与制造等领域。因此,数据库技术是近年来计算机科学技术中发展最快的方向之一,它已成为计算机信息系统与应用系统的核心技术和重要基础。

数据库技术就是研究如何对数据进行科学地组织、管理和处理,以便提供可共享的、安全的、可靠的数据信息的技术。

1.1 数据管理技术的发展

1.1.1 信息与数据

在计算机应用技术中,信息与数据这两个概念有很多相似之处,但其表述的具体内容是有区别的。

数据(Data)是数据库中存储的基本对象,通常指描述事物的符号。这些符号具有不同的数据类型,它可以是数字、文字,也可以是图形、图像、声音、说明性信息等。例如,定义学生的年龄是“18”岁,学生性别是“男”,当然也可以将“男”的文字形式改为用字母“F”表示,这里的“18”、“男”和“F”都是数据。因此,数据代表真实世界的客观事实。

信息(Information)是经过加工处理后具有一定含义的数据集合,它具有超出事实数据本身之外的额外价值。信息是标识复杂客观实体的数据,是人们进行各种活动所需要的知识。例如,可以将学生年龄是“18”岁、性别为“男”的两组相对独立的数据组合在一起形成一条表示学生基本情况的信息。

数据与信息既有联系又有区别。数据是表示信息的,但并非任何数据都表示信息;信息是加工处理后的数据,是数据所表达的内容。同时,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相同的信息。

将数据转换成信息的过程称为数据处理,它包括对各种类型的数据进行收集、储存、分类、加工和传输等一系列活动,具体讲就是对所输入的数据进行加工整理。其目的是从大量的、已知的数据出发,根据事物之间的固有联系和运动规律推导、抽取有

价值的、有意义的信息。可以用下面的表达式简单地表示出信息与数据之间的关系：

$$\text{信息} = \text{数据} + \text{处理}$$

这里,数据可以看作原料,是输入;而信息是产出,是输出结果。由此可见,信息是一种被加工成特定形式的数据。

1.1.2 数据管理技术的发展

从数据本身来讲,数据管理是指收集数据、组织数据、存储数据和维护数据等几个方面。随着计算机硬件和软件技术的发展,计算机数据管理技术也在不断改进,大致经历了3个阶段,即人工管理阶段、文件系统阶段和数据库系统阶段。

1. 人工管理阶段

20世纪50年代中期及以前,计算机主要用于科学计算,数据量较少,一般不需要长期保存。硬件方面,外部存储器只有卡片、磁带和纸带,还没有磁盘等直接存取的存储设备;软件方面,没有专门管理数据的软件,数据处理方式基本是批处理。此阶段数据与应用程序之间的关系是一一对应的关系,如图1-1-1所示。

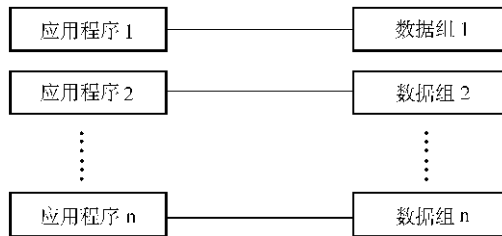


图 1-1-1 人工管理阶段程序与数据之间的关系

这一阶段数据管理有如下特点。

(1)数据面向具体应用,不共享。一组数据只能对应一组应用程序,如果数据的类型、格式或者数据的存取方法、输入/输出方式等改变了,程序必须做相应的修改。这使得数据不能共享,即使两个应用程序涉及某些相同数据也必须各自定义,无法互相利用。因此,程序与程序之间存在大量的冗余。

(2)数据不单独保存。由于应用程序与数据之间结合得非常紧密,每处理一批数据都要特地为这批数据编制相应的应用程序。数据只为本程序所使用,无法被其他应用程序利用。因此,程序的数据均不能单独保存。

(3)没有软件系统对数据进行管理。数据管理任务包括数据存储结构、存取方法、输入/输出方式等,完全由程序开发人员全面负责,没有专门的软件加以管理。一旦数据发生改变,就必须修改程序,这就给应用程序开发人员增加了很大的负担。

(4)没有文件的概念。这个阶段只有程序的概念,没有文件的概念。数据的组织方式必须由程序员自行设计。

2. 文件系统阶段

20世纪50年代后期至20世纪60年代中后期,计算机不仅用于科学计算,而且还用于信息管理。硬件方面,外存储器有了磁盘、磁鼓等直接存取的存储设备;软件方面,操作系统中已经有了专门的管理外存的数据软件,一般称为文件系统。数据处理方式不仅有批处理,而且还有联机实时处理。此阶段数据与应用程序之间的关系如图1-1-2所示。

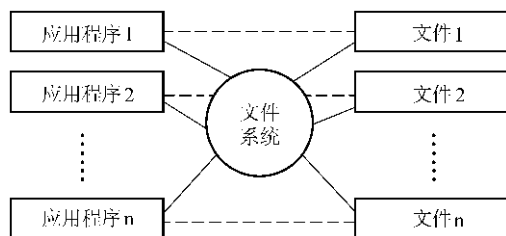


图 1-1-2 文件系统阶段程序与数据之间的关系

这一阶段数据管理有如下特点。

(1)程序与数据分开存储,数据以“文件”形式可长期保存在外部存储器上,并可对文件进行多次查询、修改、插入和删除等操作。

(2)有专门的文件系统进行数据管理,程序和数据之间通过文件系统提供存取方法进行转换。因此,程序和数据之间具有一定的独立性,程序只需用文件名访问数据,不必关心数据的物理位置。数据的存取以记录为单位,并出现了多种文件组织形式,如索引文件、随机文件和直接存取文件等。

(3)数据不只对应某个应用程序,可以被重复使用。但程序还是基于特定的物理结构和存取方法,因此数据结构与程序之间的依赖关系仍然存在。

虽然这一阶段较人工管理阶段有了很大的改进,但仍显露出很多缺点,主要以下两个方面。

(1)数据冗余度大。文件系统中数据文件结构的设计仍然对应于某个应用程序,也就是说,数据还是面向应用的。当不同的应用程序所需要的数据有部分相同时,也必须建立各自的文件,而不能共享部分相同的数据。因此,出现大量重复数据,浪费了存储空间。

(2)数据独立性差。文件系统中数据文件是为某一特定要求设计的,数据与程序相互依赖。如果改变数据的逻辑结构或文件的组织方式,必须修改相应的应用程序;而应用程序的改变,比如说应用程序的编程语言改变了,也将影响数据文件结构的改变。

因此,文件系统是一个不具有弹性的、无结构的数据集合,即文件之间是独立的,不能反映现实世界事物之间的内在联系。

3. 数据库系统阶段

20 世纪 60 年代后期以来,计算机用于管理的范围越来越广泛,数据量也急剧增加。硬件技术方面,开始出现了大容量、价格低廉的磁盘;软件技术方面,操作系统更加成熟,程序设计语言的功能更加强大。在数据处理方式上,联机实时处理要求更多,另外提出分布式数据处理方式,用于解决多用户、多应用共享数据的要求。在这样的背景下,数据库技术应运而生,它主要是解决数据的独立性,实现数据的统一管理,达到数据共享的目的,也因此出现了统一管理数据的专门软件系统即数据库管理系统(DataBase Management System,简称 DBMS)。这一阶段程序与数据之间的关系如图 1-1-3 所示。

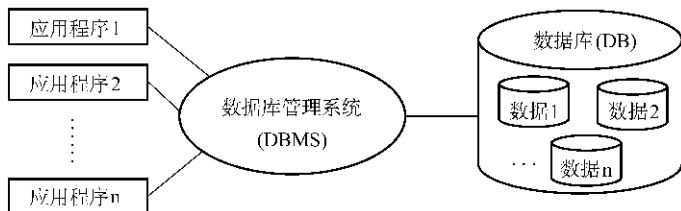


图 1-1-3 数据库系统阶段程序与数据之间的关系

数据库系统阶段的数据管理具有以下特点。

(1) 数据结构化。数据结构化是数据库系统与文件系统的根本区别,是数据库系统的主要特征之一。传统文件的最简单形式是等长同格式的记录集合。在文件系统中,相互独立的文件的记录内部是有结构的,类似于属性之间的联系,而记录之间是没有结构的、孤立的。

例如,有 3 个文件,学生(学号、姓名、年龄、性别、出生日期、专业、住址);课程(课程号、课程名称、授课教师);成绩(学号、课程号、成绩),要想查找某人选修的全部课程的课程名称和对应成绩,如果采用文件系统管理方式,则必须编写一段很复杂的程序来实现。

而数据库系统采用数据模型来表示复杂的数据结构,数据模型不仅表示数据本身的联系,还表示数据之间的联系。只要定义好数据模型,上述询问可以非常容易地联机查到。

(2) 数据的冗余度低,共享性高,易扩充。数据库系统从整体角度看待和描述数据,数据不再面向某个应用而是面向整个系统,因此一个数据可以被多个用户、多个应用共享使用。这样就大大减少了数据冗余,提高了共享性,节约了存储空间。数据共享还能够避免数据之间的不相容性与不一致性。

所谓数据的不一致性是指同一数据不同拷贝的值不一样。采用人工管理或文件系统管理时,由于数据被重复存储,当不同的应用、使用和修改不同的拷贝时就很容易造

成数据的不一致。在数据库中数据共享,减少了由于数据冗余造成的不一致现象。

由于数据面向整个系统,是有结构的数据,不但可以被多个应用共享使用,而且容易增加新的应用,这就使得数据库系统弹性大,易于扩充,可以适应各种用户的要求。

(3)数据独立性高。数据独立性包括数据的物理独立性和数据的逻辑独立性。物理独立性是指用户的应用程序与存储在磁盘上的数据库中数据是相互独立的。也就是说,数据在磁盘上的数据库中怎样存储是由数据库管理系统负责管理的,应用程序不需要了解,应用程序要处理的只是数据的逻辑结构。这样当数据的物理结构改变时,可以不影响数据的逻辑结构和应用程序,这就保证了数据的物理独立性。

而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,即当数据的逻辑结构改变了,应用程序却可以保持不变。

(4)数据由数据库管理系统统一管理和控制。数据库系统的共享是并发的(Concurrency)共享,即多个用户可以同时存取数据库中的数据,这个阶段的程序和数据的联系通过 DBMS 来实现。数据库管理系统必须为用户提供存储、检索、更新数据的手段,实现数据库的并发控制、实现数据库的恢复、保证数据完整性和保障数据安全性控制。具体内容将在 1.3 节中介绍。

1.2 数据库系统的结构

1.2.1 数据库系统的组成

数据库系统是指引进了数据库技术后的计算机系统,它能够有组织地、动态地存储大量数据,提供数据处理和数据共享机制,一般由硬件系统、软件系统、数据库和人员组成。数据库系统的组成结构如图 1-1-4 所示。

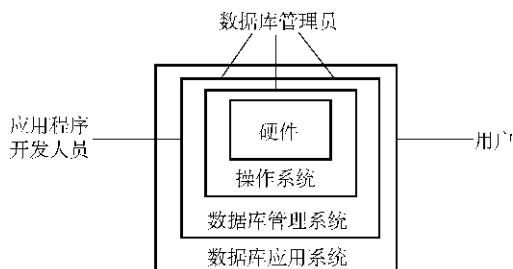


图 1-1-4 数据库系统组成结构

1. 硬件系统

由于数据库系统数据量都很大,加之 DBMS 丰富的功能使得自身的规模也很大,

因此整个数据库系统对硬件资源提出了较高的要求。这些要求包括如下几个方面。

(1)有足够大的内存存放操作系统、DBMS 的核心模块、数据缓冲区和应用程序。

(2)有足够大的磁盘等直接存取设备存放数据库;有足够数量的存储介质(内部存储设备和外部存储设备)作数据备份。

(3)有较高的通信能力,以提高数据传送率。

2. 软件系统

数据库系统的软件主要包括 DBMS 和支持 DBMS 运行的操作系统(Operating System)。DBMS 是整个数据库系统的核心,位于用户与操作系统之间,主要用于数据库的建立、使用和维护,提供对数据库中数据资源进行统一管理和控制,同时将数据库应用程序和数据库中的数据联系起来。另外,数据库系统的软件还包括具有与数据库接口的高级语言 and 应用程序开发工具。

一般来说,一种数据库只支持一种或两种操作系统。然而,近几年来,跨平台作业越来越受到人们的重视,许多大型数据库都同时支持几种操作系统,如 Oracle 数据库等。

应用程序开发工具在这里主要是用来开发与数据库相关的应用程序。现在流行的数据库应用程序开发工具有很多种,如本书所介绍的 Visual FoxPro 9.0 就是一种优秀的工具,它功能齐全,并且处理数据速度较高。

3. 数据库

数据库(DataBase,简称 DB),是存储在计算机存储设备上,结构化的相关数据集。它不仅包括描述事物的数据本身,还包括相关事物之间的联系。

人们收集并抽取出一个应用所需要的大量数据之后,应将其保存起来以供进一步加工处理,进一步抽取有用信息。在科学技术飞速发展的今天,人们的视野越来越广,数据量急剧增加。过去人们把数据存放在文件柜里,现在人们借助计算机和数据库技术科学地保存和管理大量的复杂的数据,以便能方便而充分地利用这些宝贵的信息资源。

所以,数据库是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。

4. 人员

这里的人员主要是指开发、设计、管理和使用数据库的人员,包括数据库管理员、应用程序开发人员和最终用户。

数据库管理员(DataBase Administrator, DBA),为专设的保证数据库系统正常运行的人员。数据库管理员具体职责包括以下几个方面。

(1)规划数据库的结构及存取策略。DBA 要了解、分析用户的应用需求,创建数据模式,并根据此数据模式决定数据库的内容和结构。同时要 and 数据库设计人员共同决定数据的存储结构和存取策略,以求获得较高的存取效率和存储空间利用率。此外,

DBA 还要负责确定各个用户对数据库的存取权限、数据的保密级别和完整性约束条件。

(2) 监督和控制数据库的使用。DBA 的一个重要职责就是监视数据库系统的运行情况,及时处理运行过程中出现的问题。比如系统发生各种故障时,数据库会因此遭到不同程度的破坏,DBA 必须在最短时间内将数据库恢复到正确状态,并尽可能不影响或少影响系统其他部分的正常运行。

(3) 负责数据库的日常维护。DBA 还负责在系统运行期间的日常维护工作,对运行情况进行记录、统计分析,并可以根据实际情况对数据库加以改进和重组重构。

DBA 的工作十分复杂,尤其是大型数据库的 DBA 一般是由几个人组成的小组协同工作。DBA 的职责十分重要,直接关系到数据库系统的顺利运行。所以,DBA 必须由专业知识和经验较丰富的专业人士来担任。

应用程序开发人员,就是设计数据库管理系统的人员。他们主要负责根据系统的需要分析,使用某种高级语言编写应用程序。应用程序可以对数据库进行访问、修改和存取等操作,并能够将数据库返回的结果按一定的形式显示给出用户。

最终用户,就是从计算机终端与系统交互的用户。最终用户可以通过已经开发好的具有友好界面的应用程序访问数据库,还可以使用数据库系统提供的接口进行联机访问数据库。

1.2.2 数据库系统的模式结构

数据库系统的模式结构就是为了保证数据的独立性,以达到数据统一管理和共享的目的。数据的独立性包括物理独立性和逻辑独立性。物理独立性是指用户的应用程序与存储在磁盘上的数据库中数据的相互独立性。也就是说,在磁盘上数据库中数据的存储是由 DBMS 管理的,用户程序一般不需要了解。应用程序要处理的只是数据的逻辑结构,也就是数据库中的数据,这样在计算机存储设备上的物理存储改变时,应用程序可以不必改变,而由 DBMS 来处理这种改变,这就称为“物理独立性”。有的 DBMS 还提供一些功能使得某些程序上数据库的逻辑结构改变了,用户程序也可以不改变,这又称为“逻辑独立性”。

数据库体系结构是数据库的一个总体框架,是数据库内部的系统结构。1978 年,美国国家标准协会(American National Standard Institute,ANSI)的数据库管理研究小组提出标准化建议,从数据库管理系统角度,将数据库结构分成三级模式和两层映像。

1. 数据库的三级模式结构

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,它们之间的关系如图 1-1-5 所示。

1) 外模式

外模式也称子模式或用户模式,属于视图层抽象,它是数据库用户(包括应用程序

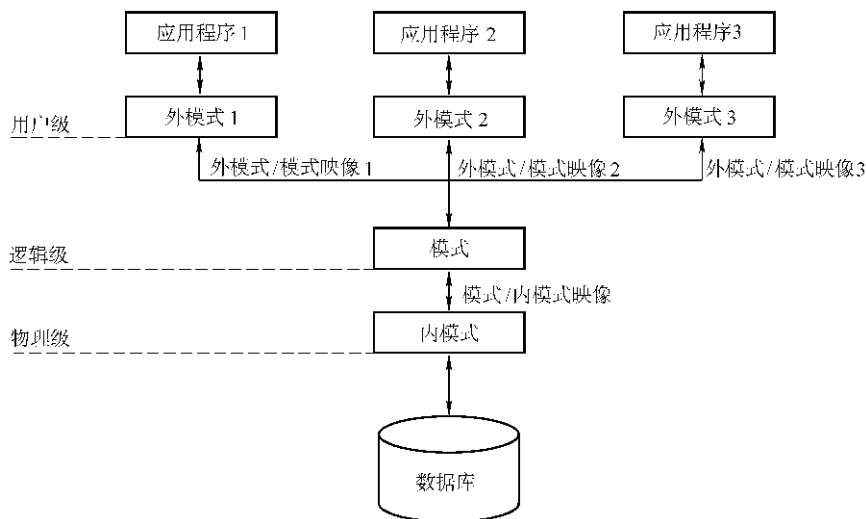


图 1-1-5 数据库系统的三级模式结构

员和最终用户)能够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,如果用户的应用需求、提取数据的方式、对数据保密的要求等方面存在差异,则其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。可见,不同数据库用户的外模式可以不同。

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是不可见的,对于用户来说,外模式就是数据库。这样既实现数据共享,又能保证数据库的安全性。DBMS 提供模式语言(Data Description Language)来严格定义外模式。

2) 模式

模式也称逻辑模式或概念模式,是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,是数据库管理员看到的数据库,属于逻辑层抽象。它介于外模式与内模式之间,既不涉及数据的物理存储细节和硬件环境,也与具体的应用程序无关。

模式实际上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,统一考虑所有用户的需求,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,而且还要定义数据之间的联系,定义与数据有关的安全性、完整性的要求。模式可以减小系统的数据冗余,实现数据共享。DBMS 提供模式描述语言(Data Description Language)来严格地定义模式。

3) 内模式

内模式也称存储模式,是数据在数据库中的内部表示,属于物理层抽象。内模式是数据物理结构和存储方式的描述。一个数据库只有一个内模式,它是 DBMS 管理的最低层。DBMS 提供内模式描述语言(Internal Schema Data Description Language)来严格地定义内模式。

总之,模式描述数据的全局逻辑结构,外模式涉及的是数据的局部逻辑结构,即用户可以直接接触到的数据的逻辑结构,而内模式更多的是由数据库系统内部实现的。

2. 数据库的两层映像与独立性

数据库系统的三级模式是对数据的三个抽象级别,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两层映像:外模式/模式映像和模式/内模式映像。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。

1) 外模式/模式映像

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,数据库系统都提供了一个外模式/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,可由数据库管理员对各个外模式/模式的映像作相应改变,从而保持外模式不变。应用程序是依据数据的外模式编写的,因此应用程序就不必修改了,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) 模式/内模式映像

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 is 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。当数据库的存储结构改变了(例如选用了另一种存储结构),为了保持模式不变,也就是应用程序保持不变,可由数据库管理员对模式/内模式映像作相应改变就可以了。这样,就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,数据库模式即全局逻辑结构是数据库的中心与关键,它独立于数据库的其他层次。因此设计数据库模式结构时应首先确定数据库的逻辑模式。

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所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,该外模式就得做相应改动,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上编制的,它依赖于特定的外模式,与数据库的模式和存储结构独立。不同的应用程序有时可以共用同一个外模式。数据库

的两层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,除非应用需求本身发生变化,否则应用程序一般不需要修改。

数据库的三级模式和两层映像保证了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中分离出去。另外,由于数据的存取由 DBMS 管理,用户不必考虑存取路径等细节,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。

1.3 数据库管理系统

数据库管理系统是处理数据访问的软件系统,也就是位于用户与操作系统之间的一层数据库管理软件。数据库管理系统提供以下几个对数据库资源进行统一管理和控制的功能。

1. 数据定义

DBMS 提供数据定义语言(Data Definition Language,简称 DDL),定义数据库的三级结构,包括外模式、模式和内模式及相互之间的映像,定义数据的完整性、安全控制等约束。各级模式通过 DDL 编译成相应的目标模式并被保存在数据字典中,以便在进行数据操纵和控制时使用。这些定义存于数据字典中,是 DBMS 存储和管理数据的依据。DBMS 根据这些定义,从物理记录导出全局逻辑记录,又从全局逻辑记录导出用户所检索的记录。

2. 数据操纵

DBMS 还提供数据操纵语言(Data Manipulation Language,简称 DML),用户可以使用 DML 操纵数据实现对数据库的基本操作,如存取、检索、插入、删除和修改等。DML 有两类,一类 DML 可以独立交互使用,不依赖于任何程序设计语言称为自主型或自含型语言;另一类 DML 必须嵌入到宿主语言中使用,称为宿主型 DML。在使用高级语言编写的应用程序中,需要使用宿主型 DML 访问数据库中的数据。因此,DBMS 必须包含编译或解释程序。

3. 数据库的运行管理

所有数据库的操作都要在数据库管理系统统一管理和控制下进行,以保证事务的正确运行和数据的安全性、完整性,这也是 DBMS 运行时的核心部分,它包括以下几个方面。

1) 数据的并发(Concurrency)控制

当多个用户的并发进程同时存取、修改或访问数据库时,可能会发生相互干扰而得到错误的结果或使得数据库的完整性遭到破坏,因此必须对多用户的并发操作加以控制和协调。

2) 数据的安全性(Security)保护

数据的安全性保护是指保护数据以防止不合法的使用造成的数据的泄密和破坏,使每个用户只能按规定对某些数据以某些方式进行使用和处理。

3) 数据的完整性(Integrity)控制

数据的完整性控制是指设计一定的完整性规则以确保数据库中数据的正确性、有效性和相容性。例如当输入或修改数据时,不符合数据库定义规定的数据库系统不予接受。

4) 数据库的恢复(Recovery)

计算机系统的硬件故障、软件故障、操作员的失误以及故意的破坏会影响数据库中数据的正确性,甚至会造成数据库部分或全部数据的丢失。DBMS 必须具有将数据库从错误状态恢复到某一已知的正确状态(亦称为完整状态或一致状态)的功能,这就是数据库的恢复功能。

总而言之,数据库是一个通用的综合性的数据集合,它可以供各种用户共享,并且具有最小的冗余度和较高的数据与程序的独立性。

4. 数据字典

数据字典 DD(Data Dictionary)中存放着对实际数据库各级模式所做的定义,也就是对数据库结构的描述。这些数据是数据库系统中有关数据的数据,称之为元数据(Metadata)。因此,数据字典本身也可以看成是一个数据库,只不过它是系统数据库。

数据字典是数据库管理系统存取和管理数据的基本依据,主要由系统管理和使用。DD 描述了对数据库数据的集中管理手段,并且还可以通过查阅数据字典来了解数据库的使用和操作。数据字典经历了人工字典、计算机文件、专用数据字典系统和数据库管理系统与数据字典一体化 4 个发展阶段。专用的数据字典在系统设计、实现、运行和扩充各个阶段是管理和控制数据库的有力信息工具。关于数据字典的具体内容将在第 6 章数据库设计与维护中做详细介绍。

综上所述,数据库管理系统是在建立、运行和维护时对数据库进行统一控制。以确保多种程序并发地使用数据库,并可以及时、有效地处理数据。

本章小结

本章主要介绍了数据库的基本概念、数据管理技术的发展阶段、数据库系统的结构及数据库管理系统的功能。

数据是描述事物的符号,它与信息既有联系又有区别。数据库是存储在计算机存储设备上,结构化的相关数据集合。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。数据库管理系统是处理数据访问的系统软件,它位于用户与操作系统中间。数据库系统是指引进了数据库技术后的计算机系统。

同时,本章简要介绍了数据管理技术的发展大致经历 3 个阶段,即人工管理阶段、文件系统阶段和数据库系统阶段。数据库系统体系结构从数据库管理系统角度看,通常采用三级模式结构,即外模式、模式和内模式。三级模式之间的两层映像,即外模式/模式映像、模式/内模式映像,保证了数据库系统中的数据具有逻辑独立性和物理独立性。

习 题

1. 数据与信息有什么区别与联系？
2. 什么是数据库、数据库系统和数据库管理系统？
3. 计算机数据管理技术发展经历几个阶段,各阶段的特点是什么？
4. 什么是外模式、模式和内模式？
5. 试述数据库系统的两层映像功能。
6. 试述数据库系统的三级模式结构是如何保证数据的独立性的？
7. 简单说明数据库管理系统包含的功能。