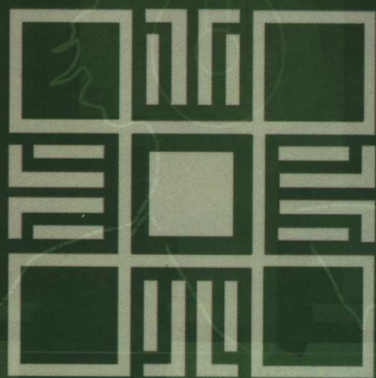


21

世纪高等学校规划教材·计算机基础教育系列

数据库及其应用

陈维默 编著



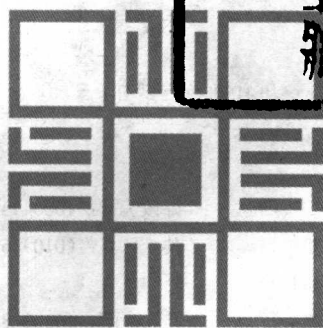
中国电力出版社

www.infopower.com.cn

数据库及其应用

陈维默 编著

江苏工业学院图书馆
藏书章



中国电力出版社

www.infopower.com.cn

内容提要

本书描述了数据库的基本知识,结合 Access 2003 数据库介绍了数据库设计过程,详细阐述了利用 Visual Basic 6.0 编程开发数据库系统技术,简单介绍了工程数据库的基本知识。

全书共 8 章内容。主要包括了数据库的基本知识;操纵数据库的结构化查询语言 SQL; Access 2003 数据库设计、窗体设计和面向对象的程序设计语言 VBA;运用 Visual Basic 6.0 中的数据控件、数据感知控件和数据库访问对象(DAO)技术操作数据库;工程数据库的定义、特点、特性及其结构分析。

本书的主要思路是让读者掌握一种 DBMS 的技术,并能开发一个简单的工程领域中的数据库管理系统。本书各章均附有习题,最后还附上机操作题,便于学习者进一步理解和掌握各章的知识。因此,本书特别适合作为普通高等院校工程类学生的“数据库及其应用”课程的通用教材,也可作为工程领域中科技人员及其他有关人员的参考书。

图书在版编目(CIP)数据

数据库及其应用 / 陈维默主编. —北京: 中国电力出版社, 2005.10

21 世纪高等学校规划教材系列

ISBN 7-5083-3868-5

I.数... II.陈... III.数据库系统—高等学校—教材 IV.F

中国版本图书馆 CIP 数据核字(2005)第 121968 号

责任编辑: 危锦怡

丛 书 名: 21 世纪高等学校规划教材·计算机基础教育系列

书 名: 数据库及其应用

出版发行: 中国电力出版社

地 址: 北京市三里河路 6 号

邮政编码: 100044

电 话: (010) 68358031 (总机)

传 真: (010) 68316497, 88383619

本书如有印装质量问题,我社负责退换

服务电话: (010) 88515918 (总机)

传 真: (010) 88518169

E-mail: infopower@cepp.com.cn

印 刷: 北京同江印刷厂

开本尺寸: 185×233

印 张: 12.5

字 数: 297 千字

书 号: ISBN 7-5083-3868-5

版 次: 2006 年 1 月北京第 1 版

印 次: 2006 年 1 月第 1 次印刷

印 数: 0001—4000 册

定 价: 19.00 元

版权所有,翻印必究

前 言

由于现代科学技术特别是电子计算机技术的迅猛发展,数据库技术在各个领域中的应用日趋普遍。为适应 21 世纪人才培养的需要,并结合多年的教学实践,在为工程类本科生讲授的《数据库及其应用》讲义的基础上,针对数据库技术在工程领域应用所需的基本理论知识和操作技能,编写了这本教材。本书以传授知识与培养能力并重,它具有如下 3 个特点。

(1) 注重实用性兼顾理论性。从工程实用性出发,本书所举例子均精选于数据库技术应用于工程领域中的实例。

(2) 力求深入浅出。对于一些较为抽象,难以理解的概念、理论,一方面从内容的编排上尽量循序渐进并逐步深入;另一方面尽可能地选用具体的实例且予以解释。

(3) 注重能力培养。学习《数据库及其应用》这门课程,对于工程类的学生重点在于如何应用数据库技术解决生产实践中的问题。因此,掌握一种 DBMS 技术,开发一个简单的工程领域中的数据库管理系统,成为本书的主线。这可使学习者获得较为完整的开发数据库管理系统技术及技能的训练,从而可以明显地提高学习者应用计算机解决问题的能力,且达到掌握知识与能力培养并重的目的。

本书共 8 章,第 1、2 章介绍数据库的基本概念和术语,第 3 章介绍操纵数据库的结构化查询语言 SQL,第 4、5 章介绍 Access 2003 数据库的设计方法及窗体设计,第 6 章介绍了面向对象程序设计语言 VBA,第 7 章介绍 VB 开发数据库技术,第 8 章简单介绍工程数据库的发展、定义和特点。书末还附有结合各章内容的上机操作练习题。

限于编著者对数据库原理和技术的把握水平和经验,书中难免有不妥和疏漏之处,恳请广大读者批评指正。

作 者
2005 年 9 月

目 录

前 言

第 1 章 数据库的基本概念	1
1.1 数据库简史	1
1.2 信息、数据与数据处理	2
1.3 数据库	4
1.4 数据库系统	7
1.5 数据模型	14
习题一	18
第 2 章 关系数据库基础	20
2.1 关系模型的基本概念	20
2.2 关系数据语言	25
习题二	31
第 3 章 操纵数据库的结构化查询语言——SQL	33
3.1 数据定义	34
3.2 数据查询	35
3.3 数据更新	43
3.4 视图	44
3.5 数据控制	46
习题三	47
第 4 章 Access 2003 数据库设计	51
4.1 创建数据库	51
4.2 创建数据表	55
4.3 建立表间关系	63
4.4 修改表结构	66
4.5 在数据库窗口操作表	68
4.6 对数据表的操作	69
习题四	80
第 5 章 Access 2003 窗体设计	86
5.1 创建窗体	86
5.2 窗体控件及其属性	88

5.3 窗体属性	115
习题五	118
第 6 章 面向对象程序设计语言 VBA	122
6.1 VBA 概述	122
6.2 模块、函数及程序	122
6.3 Visual Basic 的数据类型与数据库对象	123
6.4 变量、常量的有关概念	125
6.5 程序控制流程	128
6.6 处理执行时的错误	137
习题六	138
第 7 章 Visual Basic 数据库开发技术	140
7.1 利用数据感知控件访问数据库	140
7.2 利用数据库访问对象技术操作数据库	156
习题七	169
第 8 章 工程数据库简介	171
8.1 工程数据库的历史沿革及其定义	171
8.2 工程数据库与商用数据库的特点分析	173
8.3 工程数据库的特性	177
8.4 工程数据库结构	180
习题八	186
附录 A 上机操作题	187
参考文献	194

第 1 章 数据库的基本概念

随着计算机技术的蓬勃发展, 计算机应用从科学计算、过程控制进入数据处理, 使计算机从少数科学家手中的珍品成为人们日常工作中处理数据的得力助手和有力工具。当今世界, 在计算机的三大主要应用 (科学计算、过程控制和数据处理) 领域中, 数据处理迅速上升为计算机应用的主要方面。在本章中, 我们从数据处理开始介绍数据处理的发展、数据模型和数据库以及所涉及的基本概念, 这将作为后面各章学习的基础。

1.1 数据库简史

20 世纪 60 年代前, 计算机主要用于科学计算, 进入 60 年代后, 在社会大系统中出现了巨大的信息流以及与之相伴随的宏大的数据流, 为了适应和满足社会发展的需要, 必须采用新的技术和手段, 对这些数据进行收集、存储、加工、检索、分类、统计和传输等。谁能担当此重任, 这就是数据库技术产生的历史背景。

20 世纪 60 年代, 美国系统发展公司在为美国海军基地研制数据库时, 首先采用了“Data Base”一词。1968 年 IBM 公司在数据库管理系统方面率先研制成功了 IWS 了系统, 1969 年 10 月 CODASYL 数据库研制者提出了网络模型数据库系统规范报告 DBTG, 使数据库系统开始走向规范化和标准化。因此, 许多专家认为数据库技术诞生于 20 世纪 60 年代末。

数据库技术的产生来源于社会的实际需要, 而数据库技术的实现必须有理论作为指导, 系统的开发和应用又不断促进数据库理论的发展和完善。1970 年 E.F.CODD 发表了著名论文《大型共享数据库的数据关系模型》, 成功地奠定了关系数据库理论的基石, 为此, 1981 年他获得了计算机科学的最高奖——ACM 图林奖。1971 年美国数据系统语言协会在正式发表的 DBTG 报告中, 提出了三级抽象模式, 解决了数据独立性问题。1974 年 IBM SAN JOSC 实验室成功研制了关系数据库管理系统 System R, 并随后投放软件市场, 1976 年美籍华人陈平山提出了数据库逻辑设计的实体——联系方法。1978 年新奥尔良发表了 DBDWD 报告, 他把数据库系统全设计的过程分为 4 个阶段: 需求分析、信息分析与定义、逻辑设计和物理设计。1980 年 J.D.UHMAN 所著的《数据库系统原理》一书出版, 1984 年 David Maier 所著的《关系数据库理论》标志着数据库理论上的成熟。

在我国, 数据库技术起步较晚, 20 世纪 70 年代后期才开始引进数据库管理系统, 但发展十分迅速, 并取得了丰硕的成果, 特别是关系数据库在我国最为流行, 应用最广。深为人知的 Xbase 和本书将要介绍的 Access 2003 是最典型的代表。

计算机科学的迅猛发展, 特别是微型计算机的出现和普及, 为数据库技术的应用提供了广阔的舞台。微型计算机冲破了计算机神秘世界的高墙深院, 进入了社会的各个角落乃至千家万户, 由此微机数据库受到了人们的普遍重视, 应用水平不断得到提高。

1.2 信息、数据与数据处理

人类的一切活动都离不开数据、信息。在不同的领域里，信息的含义有所不同。一般认为信息是数据和消息中所包含的意义。本节对数据、信息和它们之间的关系作简要的介绍。

1.2.1 信息

所谓信息泛指通过各种方式传播的，可被感受的声音、文字、图像、符号等所表征的某一特定事物的消息、情报或知识。换句话说，信息是对客观事物的反映，是为某一特定目的而提供的决策依据。

在现实世界中，人们经常接触各种各样的信息，并根据这些信息来制定决策。例如，人们根据汛情通报决定防汛措施；当我们走进商店欲购买某种商品时，首先要了解该商品的价格、款式或花色，然后根据这些信息决定是否要购买；人们根据电视节目预告来决定是否要收看；工厂生产的商品是否对路，就必须了解市场动态和社会需求，并根据这些信息和自身的实际情况来决定生产哪些畅销对路的产品等。

总之，现实世界是一个充满信息的世界。当然，在处理实际问题时，人们只关心对其有用的信息。

信息有如下一些重要特征：

- 信息传递需要物质载体，信息的获取和传递需要消耗能量。
- 信息是可以感知的。
- 信息是可以存储、压缩、加工、传递、共享、扩散、再生和增值的。

1.2.2 数据

所谓数据是指表达信息的某种符号，是信息的一种量化表示。数据反映信息，而信息依靠数据来表达。表达信息的符号可以是数字、文字和图形。

计算机只能存放和处理数据。因此，必须人为地把信息转换成可以被计算机接受的数据，但不管是什么数据，都是以二进制形式存储在计算机内并被计算机加工处理的。

数据与信息是密切相关联的。信息是向人们（或机器）提供关于现实世界有关事物的知识，数据则是载荷信息的物理符号，二者是不可分离而又有一定区别的概念。例如，“今年大学招生录取线为 450 分，张红的高考成绩为 500 分。”这段文字（数据）提供了张红可能被某一个大学录取的信息。但是，在一些不很严格的场合下，对它们并不做严格的区分，而当作同义词互换使用。譬如，信息处理与数据处理、信息资源与数据资源、信息采集与数据采集等。

1.2.3 数据处理

所谓数据处理泛指对原始数据进行收集、整理、存储、分类、排序、检索、维护、加工、统计和传输等一系列活动的总称。数据处理的目的是获得我们所需要的资料和提取有用的数据成份并将其作为决策的依据。

当今时代是一个信息爆炸的时代,宏大的数据流单靠手工和简单的工具已无法进行处理,更远远赶不上社会的实际需要。电子计算机的出现使数据处理发生了划时代的变革,而数据库技术的发展,则使数据处理跨入了一个崭新的阶段。在数据处理中,数据的管理技术是重要的组成部分,它的发展大致经历了3个阶段:手工管理方式、文件管理方式、数据库系统方式。

1. 手工管理方式

手工管理方式又称自由管理方式。早期的电子计算机,没有必要的软件支持。用户直接在裸机上作业,虽然用户有绝对的自由权,但必须确定数据的物理存放位置,因而用户程序不仅要设计数据处理的方法,而且还要记住数据在内外存储器中的地址,这使得程序高度依赖于数据存储地址。

这种管理方式,迫使用户程序直接与物理地址打交道,使用户的负担极重,而且这种数据管理方式既不灵活,又不安全,编程效率极低。

2. 文件管理方式

文件管理方式是有关数据组织成文件。这种文件脱离程序而独立存在,并容易命名,而应用程序是通过文件名来存取文件中的数据。

这些数据文件,由一个专门的软件系统——文件管理系统实施统一管理。文件的组织方式是按照统一的规则和方法来组织和存取的。文件管理系统是一个独立的系统软件,它是应用程序与数据文件之间的一个接口,应用程序通过文件管理系统建立和存取文件,如图 1-1 所示。

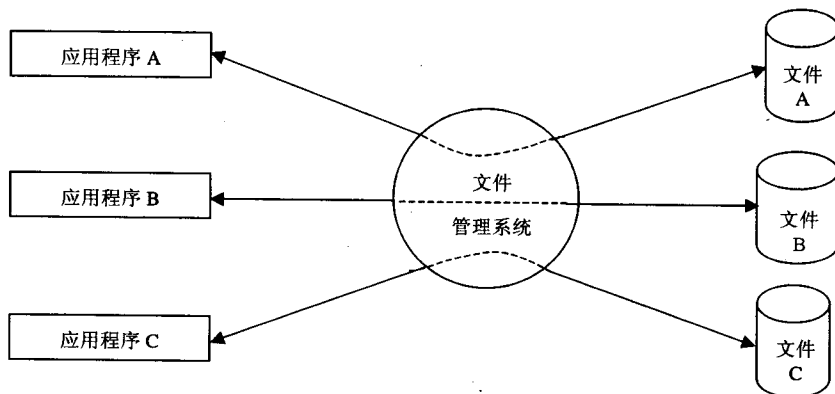


图 1-1 文件管理方式示意图

文件管理方式具有以下弱点:

(1) 尽管数据以文件方式独立存放,但程序与数据紧密相关,一旦数据文件离开了使用它的应用程序,便失去了存在的价值。

(2) 由于不同应用程序各自建立相应的数据文件,从而造成了数据冗余,使空间利用率大为降低。

(3) 由于同一数据存放在不同的数据文件中,很容易造成数据的不一致性。

(4) 文件管理方式不能反映信息之间的联系。

为了克服文件管理方式的上述弱点，又产生了数据库系统方式。

3. 数据库系统方式

数据库系统方式与文件管理方式不同，其数据组织面向整个系统，即用整体观点规划数据，形成一个数据中心，构成一个数据仓库。库中的数据能满足所有用户的不同要求，供不同的用户所共享。这时，应用程序不再与一个孤立的数据文件相对应，而是将取自整体数据集的某个子集作为逻辑文件且与应用程序相对应，并通过一个系统软件——数据库管理系统（DBMS）来实现逻辑文件与物理数据之间的映射。

1.3 数 据 库

自 1968 年第一个商品化的数据库管理系统问世以来，数据库技术得到了迅猛发展。随着计算机应用的不断深入，数据库的重要性日益被人们所认识，它已成为信息管理、办公自动化和计算机辅助设计等主要软件工具之一。

1.3.1 数据库的概念

数据库一词的英文写法为 DataBase 或 Data Base，简称 DB。Data 的中文含义是数据，base 是基地，DataBase 意指供给数据的基地即为数据库。顾名思义，数据库是计算机存放数据的“仓库”。仓库建立在外存储媒介中，如磁盘、光盘等。

在我们的日常生活中，有各种各样的库，例如书库、粮库、汽车配件库等。这些库都具有这样几个共同特点：仓库由一到多个库房组成，可以存放现实世界中一定的东西，这些东西有条理、有组织、合理地存放在仓库中某个库房的货位或货架上，且这些东西由仓库管理人员进行管理，管理人员有一套管理体制，即系统。人们到仓库中取货，必须通过仓库管理员。

数据库系统实质上就是一个记录保存系统。它可以将各种数据很方便地保存到数据库中，数据库中的数据由一个称为数据库管理系统的软件进行管理，对数据库的访问必须通过数据库管理系统。

例如，在计算机上有一个学生情况数据库 STUDENT，它帮助我们记录学生信息。现在如果我们想查询有关管理信息系学生的学号、姓名、性别、年龄情况，可输入下面的命令：

SELECT 学号、姓名、性别、年龄 FROM STUDENT WHERE 系名=“管理信息系”

于是，在计算机屏幕上就会立即显示所需要的学生库中管理信息系全部学生的信息，如图 1-2 所示。

学 号	姓 名	性 别	系 名	年 龄
970001	李 一	男	管理信息系	16
970002	王 二	男	管理信息系	17
970003	张小丽	女	管理信息系	15
970004	赵 芳	女	管理信息系	17
...				...

图 1-2 显示查询结果

当记录的数据发生变化时，我们就要对数据库中的数据进行更新。例如，每过一年，学生年龄增长一岁，要在学生数据库 STUDENT 中反映出来，就要更新数据库中所有的记录，命令如下：

```
UPDATE STUDENT SET 年龄 = 年龄+1
```

我们可以将数据库中没有必要保留的记录删除掉。例如，学号为“97002”的学生已退学，用命令将该学生的记录从库文件 STUDENT 中删除掉，命令如下：

```
DELETE FROM STUDENT WHERE 学号=“930002”
```

我们可以向数据库中插入 (INSERT) 新记录。例如，向库中增加一个新学生“周丽同学”，命令如下：

```
INSERT INTO STUDENT (学号, 姓名, 性别, 系名, 宿舍, 年龄, 家庭住址 )  
VALUES (“970040”, “周丽”, “女”, “计算机”, “学 2202”, 16, “北京海淀”)
```

在上述例子中。我们对数据库进行了 4 种基本操作：检索、更新、删除和插入。这 4 种操作通常称为数据库操作，它们是最常用的操作。

一个数据库系统的主要功能之一就是允许数据操作，但是数据库系统绝非仅此而已，还需要有许多附加功能（例如屏幕格式定义、菜单定义、打印输出控制等）和数据定义（即建库功能）功能。上述这些功能，说准确一点，就是支持一种数据库语言，用户可以用这种语言编写程序去处理数据。本章的例子是采用数据库标准语言 SQL 来表达的。

每一种数据库系统都用它自己支持的语言（命令）来实现对数据的操作，这有点类似于地球上不同的国家和民族使用各自的语言来表达各自的思想。但是，语言的不同给不同的民族之间的交流带来了困难。数据库的语言同样如此，因而国际标准化组织 (ISO) 选定了 SQL 语言作为数据库的标准语言。关于 SQL 语言和其他数据库语言，我们在后面的章节将做专门介绍。

从上述例子中我们可以看出，数据库看起来就像一张大二维表，对数据库的各种操作就像在表上按条件插入几行或删除几行，或从表中选取几行和几列组成新的表。我们还可看出，数据库中对数据的存取不一定以记录为单位（文件系统对数据的存取是以记录为单位的），也可以是字段。事实上，今天我们用的数据库就是一个或几个彼此相关的表的集合，数据记录就是以表中行的形式存放在库中的。

由此可见，用数据库对数据管理比传统的手工对数据管理优越得多，主要表现在：紧凑、迅速、简洁和及时。

由于数据库是一个很复杂的系统，涉及面很广，因而许多专家从不同角度给数据库下过定义。总的来说，数据库是存储在一起的相关数据集合，数据被结构化。这些数据去掉了有害的或不必要的冗余，能为多种应用服务，数据的存储独立于使用它的程序；对数据库插入新的数据，修改和检索原有的数据，均可按一种公用的可控制方式进行。

1.3.2 数据库的特点

数据库系统与文件系统相比，克服了文件系统的缺陷。用数据库技术管理数据，主要有

以下5个方面的特点。

1. 数据共享

这是数据库系统区别于文件系统的最大特点之一，也是数据库系统技术先进性的重要体现。共享是指多用户、多种应用、多种语言互相覆盖地共享数据集合，所有用户可同时存取数据库中的数据。

2. 面向全组织的数据结构化

在数据库中，数据不再像文件系统那样从属于特定的应用，而是按照某种数据模型组织成为一个结构化的整体。它不仅描述了数据本身的特性，而且也描述了数据与数据之间的种种联系，这使数据库具备了复杂的结构。

数据结构化，有利于实现数据共享。比如，一个学校，可以把学校所有的各个应用（人事、学籍、科研、财务、后勤等）的数据组织到一个数据库中，并且使其结构化。数据实现集中、统一的存储与管理，各种应用存取各自相关的数据子集，以满足各种应用要求，实现数据共享。

3. 数据独立性

文件系统管理中，应用程序严重依赖于数据文件，如果把应用程序使用的磁带顺序文件改为磁盘索引文件，则必须对应用程序进行修改。而数据库技术的重要特征就是数据独立于应用程序而存在，数据与程序相互独立，互不依赖，不因一方的改变而改变另一方，这大大简化了应用程序的设计与维护的工作量。

数据库是面向整个系统的，以最优的方式服务于一个或多个应用程序（用户），以实现数据共享。

4. 可控数据冗余度

数据共享、结构化和数据独立性的优点可使数据存储不必重复，这不仅可以节省存储空间，而且从根本上保证了数据的一致性，并且这也是有别文件系统的重要特征。

从理论上讲，数据存储完全不必重复，即冗余度为零，但有时为了提高检索速度，常有意安排若干冗余。这种冗余由用户自己控制，被称为可控冗余度。可控冗余要求任何一个冗余的改变都能自动地对其余冗余加以改变，这个过程叫做传播更新。

5. 统一数据控制功能

数据库是系统中各用户的共享资源，因而计算机的共享一般是并发的，即多个用户同时使用数据库。因此，数据库管理系统必须提供以下4个方面的数据控制功能，以保证整个系统的正常运转。

（1）数据安全性控制。数据的安全性是指采取一定安全保密措施以确保数据库中的数据不被非法用户存取，从而避免数据的泄密和破坏。

（2）数据完整性控制。数据的完整性指数据的正确性、有效性与相容性。系统要提供必要的功能，以保证数据库中的数据在输入、修改过程中始终符合原来的定义和规定。

（3）并发控制。当多个用户并发进程同时存取、修改数据库中数据时，可能会发生互相干扰而得到错误结果，并使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。

（4）数据恢复。当系统因发生故障造成数据错误或当对数据库数据的操作发生错误时，

系统能进行应急处理，并把数据库恢复到正确状态。

数据库系统中应用程序与数据库之间的关系如图 1-3 所示。

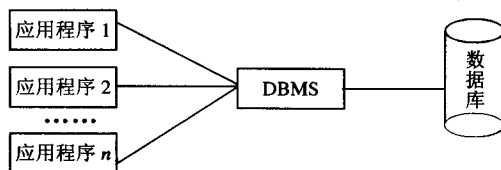


图 1-3 数据库系统中应用程序与数据库之间的关系

1.4 数据库系统

数据库系统（Data Base System, DBS）是一个复杂的系统，它是由硬件、软件（操作系统、数据库管理系统和编译系统等）、数据库和用户构成的系统。

1.4.1 数据库系统

数据库是数据库系统（DBS）的核心和管理对象。数据库系统已经不仅仅是一个对数据进行管理的软件，也不仅仅是一个数据库，数据库系统是一个实际运行的，按照数据库方式存储、维护和向应用系统提供数据支持的系统。数据库系统的层次如图 1-4 所示。

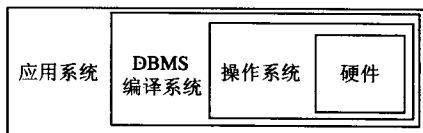


图 1-4 数据库系统层次

1. 硬件

数据库系统建立在计算机系统上。计算机的硬件资源包括 CPU、内存、显示器、磁盘、磁带及其他外部设备。数据库系统对计算机系统硬件资源有如下特殊的要求。

(1) 需要足够大的内存来存放操作系统、数据库管理系统核心模块、数据库数据缓冲区和应用程序。

(2) 需要足够大的磁盘等直接存取设备来存储数据库中庞大的数据。

(3) 为了安全，需要足够的磁盘、磁带、光盘等存储介质来制作数据库的备份。

(4) 要求具有较高的通道能力，以提高数据传送率。

(5) 要求系统支持联网，以实现数据共享。

2. 软件

数据库系统的软件主要是指负责数据库存取、维护和管理的软件系统。该软件系统通常叫做数据库管理系统（Data Base Management System, DBMS）。数据库系统各类用户对数据

库的各种操作请求,都由 DBMS 来完成,它是数据库系统的核心软件。DBMS 提供一种超出硬件层之上可以对数据库进行观察的功能,并支持用较高级的语言来表达用户的操作,使用户不受硬件层细节的影响。DBMS 是在操作系统支持下工作的。

数据库系统的其他软件还包括各种高级语言处理程序(编译或解释程序)、应用开发工具软件和应用程序等。

3. 数据库

数据库是数据库系统的核心和管理对象,库中的数据是集成的、共享的、最小冗余的,且能为多种应用服务。其中“集成”是指某特定应用环境中的各种应用的数据及其数据之间的联系(联系也是一种数据)全部集中地按照一定的结构形式进行存储。“共享”是指数据库中的一块块数据可为多个不同的用户使用多种不同的语言,为了不同的目的而同时存取数据,甚至同时存取同一块数据。

4. 数据库用户

数据库系统中存在一组管理、开发、使用数据库的用户。有的用户需要从数据库中查询信息,有的用户要为查询信息的用户编写处理程序,有的用户专门来管理、维护数据库等。通常情况下,根据他们的工作职能将这些用户分成 3 类,说明如下:

(1) 终端用户。这类用户使用数据库系统提供的终端命令语言或者菜单驱动、表格驱动、图形显示和报表生成等对话方式来存取数据库中的数据。这类人员一般为不怎么精通计算机和程序设计的各级管理人员。

(2) 应用程序员。这类人员是负责设计和编制应用程序的人员。他们通常使用 C 语言、数据库语言或 4GL(第四代语言)等来设计和编写应用程序,并对数据库进行存取操作,所编制的应用程序供终端用户使用。

(3) 数据库管理员。这类人员是指全面负责数据库系统的管理维护和正常使用的人员。由于数据库是整个企业的数据资源,它和人力、物力、设备、财产、资金一样,是企业中的基本资源,应由专门人员或机构加以管理。对于大型数据库系统,数据库管理员(Data Base Administrator, DBA)极为重要,常专门设置 DBA 办公室。DBA 是数据库系统能否正常运转的关键。DBA 不仅要求具有较高的技术专长,而且还要具备较深的资历,并熟悉各部门全部数据的性质和用途,兼有系统程序员、系统分析员的能力。DBA 的具体职责是:

- 1) 决定数据库的内容和结构。
- 2) 决定数据库的存储结构和存取策略,使数据的存储空间利用率和存取效率均较优。
- 3) 定义数据的安全性要求和完整性约束条件。
- 4) 监督并控制数据库的使用和运行。
- 5) 改进和重组数据库系统。

DBA 可利用工具来做上述工作,其中有两个重要的工具是必需的,一个是系列的实用程序(DBMS 的装配、重组、日志、恢复、统计等程序),另一个工具是数据字典(存储数据库结构的定义、记录类型和字段定义等信息)。数据库系统创建初期,DBA 还必须和系统分析员一起负责应用系统的需求分析和规范说明,以及确定系统的软硬件配置并参与数据库的设计。

需要注意的是,鉴于不同规模的数据库系统,对用户的人员配置是不相同的。只有大型

数据库系统才配备应用程序员和数据库管理员。对于常见的微机数据库系统，通常只有一个用户，可以兼终端用户、应用程序员和数据库管理员。

为了便于读者更直观地了解数据库系统，我们不妨把它与图书馆做一比较，见图 1-5。大家都知道，图书馆是存放和借阅图书的部门，而数据库系统则是存储数据并负责用户访问数据的机构。正像图书馆不能简单地与书库等同起来一样，我们也不能把数据库系统仅仅理解成存储数据的集合，而应该将它理解成一个系统。

1.4.2 数据库管理系统 DBMS

在文件系统中，用户对其所使用的数据文件的物理组织和存储细节全要进行安排和处理。这给用户带来很大不便，而数据库系统的目标之一就是要解决这个问题。数据库系统把对“存储数据”的管理、维护和使用的复杂性都转嫁到 DBMS 身上。因此，DBMS 是一种非常复杂的、综合性的、在数据库系统对数据进行管理的大型系统软件，它是数据库系统的核心组成部分，在操作系统（OS）支持下进行工作。在确保数据安全可靠的同时，DBMS 大大提高了用户使用数据的简明性和方便性，用户在数据库系统中的一切操作，包括数据定义、查询、更新及各种控制，都是通过 DBMS 进行的。DBMS 允许用户抽象地、逻辑地处理数据而不必关心这些数据在计算机中的布局 and 物理位置，这给数据库用户带来了极大的好处。

DBMS 总是基于某种数据模型。根据采用的数据模型的不同，DBMS 可以分成网状、层次、关系、面向对象等。



图 1-5 数据库系统与图书馆比较

1.4.3 数据库管理系统的抽象层次

在数据库系统中，对于同一意义下的数据，比如学生数据，从计算机中处理的二进制数据到用户处理的诸如学生、姓名、年龄等概念的数据之间，存在着许多层抽象和转换。

由图 1-6 可知，为了提高数据库的独立性，DBMS 把数据库从逻辑上分为三层（三级），面向用户或程序员的用户层（或称为外部级），面向数据库管理员和系统分析员的概念层（或称为概念级）和内层（或称为内部级），它们反映了看待数据库的三种不同的角度。

尽管目前世界上大多数数据库管理系统种类不同，支持的数据模型不同，使用的语言不

同, 存储结构不同, 但它们的结构基本是相同的, 都支持数据库系统结构的三级抽象描述。

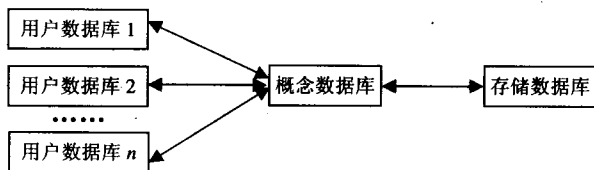


图 1-6 数据库的 3 个抽象层次

在数据库系统中, 用户看到的数据与计算机中存放的数据是两回事。用户只能看到外层(外视图), 而其他两层是看不到的。它们之间实际上要经过两次抽象和转换: 一次是把所有用户的数据综合抽象成一个统一的全局数据的共同视图, 即中间层(概念层), 以实现数据共享和减少数据冗余; 第二次是把全局数据的共同视图抽象为计算机数据的实际存储, 即把全局数据按照存储的最优组织形式进行物理存放, 以提高性能和存放效率。当计算机向用户提供数据时则做相反的变换。因此, 对于一个数据库系统而言, 实际上只有物理数据库是存在的, 概念层数据库只不过是物理数据库的一种抽象描述, 用户层数据库则是用户与数据库的接口。

数据库系统采用三级结构的优点是保证了数据的独立性, 简化了用户接口, 有利于数据共享和数据的安全保密。这也正是数据库系统优于文件系统之处。图 1-7 给出了内视图、概念视图、外视图的三级结构以及它们之间的两级映像, 还给出了 DBMS 和数据库管理员与三级结构的关系。在进一步讨论内视图、概念视图、外视图以及映像等概念之前, 先对视图和模式进行定义。

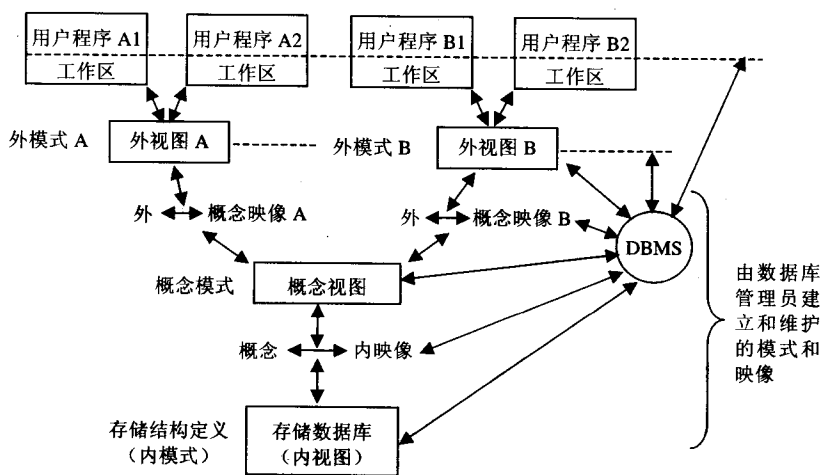


图 1-7 数据库系统的体系结构

1. 视图和模式

数据库结构可用图形和语言描述。图形描述就是数据模型图(简称视图), 是我们看到的数据库结构的图形表示, 非常直观, 但不严格。视图分为外视图、概念视图(也称为视图)

和内视图。模式就是用 DBMS 提供的数据库定义语言 DDL (Data Descriptive Language) 来编写的精确地定义 (描述) 数据库结构 (视图) 的程序。其语言描述严格、准确, 可描述许多细节, 是计算机可接收的方式。对应 3 个不同的视图, 分别用数据库定义语言去定义可以得到 3 个对应的模式, 即外模式 (也称子模式)、概念模式和内模式。

2. 内视图

如前所述, 内视图是处于三级结构的内层, 即与实际存储数据方式有关的一层, 是整个数据库中非常低级的表示。它是由文件和索引 (或其他高效存取文件的结构) 组成的集合。

内视图是借助于“内模式”来描述和定义的。内模式是由一种数据库定义语言——内 DDL 来编写的。

3. 概念视图

概念视图处于三级结构的中间层, 它是存储数据库的抽象表示形式; 是对现实世界的一个抽象, 是现实世界中某应用环境 (企业或单位) 的所有信息内容集合的表示; 也是所有外视图综合起来的结果, 它表示了数据库的整体结构。概念视图用“概念模式”进行描述和定义。概念模式是用另一种数据库定义语言——概念 DDL 来编写的。

4. 外视图

外视图是用户视图, 处于三级结构的外层。外视图是一种局部的逻辑数据结构, 是用户看到和使用的数据库结构。数据库系统中, 用户很多, 由于不同的用户只对数据库中某一部分感兴趣, 因此, 要根据不同用户情况为其设计所需要的不同的数据库结构, 该库结构是从模式中抽象出来的一个虚的库结构。外视图可有多个, 具有相同视图的用户可共用一个用户视图。

从逻辑关系上看, 外模式是概念模式的一部分, 或者说是概念模式的一个逻辑子集 (当然也可以是整个概念模式)。外模式由数据库定义语言——外 DDL 来编写。

5. 二级映像

图 1-7 中给出了二级映像关系, 一个是在外层和概念层之间; 另一个是在概念层和内层之间。

外模式/概念模式的映像定义某个外模式和概念模式之间的对应关系, 当概念模式改变时, DBA 只需要对外模式/概念模式的映像做相应的修改, 可使外模式保持不变, 从而不必修改建立在该外模式上的应用程序, 使程序与数据之间逻辑独立。

概念模式/内模式的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, DBA 只需要对概念模式/内模式的映像做相应的修改, 可使模式保持不变, 从而不必修改应用程序, 使程序与数据之间物理独立。

正是由于二级映像功能, 使得数据库系统中数据具有较高的逻辑独立性和物理独立性。这样可使用户看到的和使用的逻辑数据与实际存储的物理数据完全分开, 从而大大减少了应用程序的维护和开销。

数据库三级结构带来的优点如下:

(1) 保证数据的独立性。将概念模式和内模式分开, 保证了数据物理独立性; 将外模式和概念模式分开, 保证了数据的逻辑独立性。

(2) 简化了用户接口, 方便了用户。用户只要按照外模式编写应用程序或在终端键入操