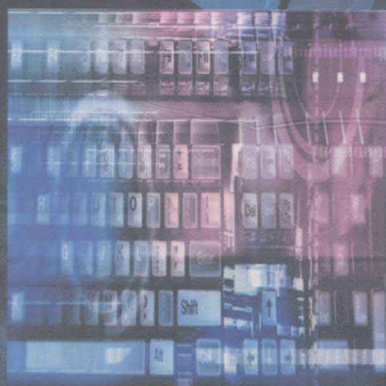


数据库系统

原理与应用

主编 刘德春



湖北人民出版社

数据库系统 原理与应用

主 编：刘德春

编写人员：殷海娜 孙晓莹

郭东恩 吴绍兴

江苏工业学院图书馆
藏书章

湖北人民出版社

鄂新登字 01 号

内 容 提 要

本书以关系数据库为核心,完整地论述了数据库系统的基本概念、基本原理和应用技术。主要包括:数据库系统基本概念、数据模型与概念模型、关系数据库标准语言SQL、关系数据库设计理论、数据库设计、数据库保护、数据库技术新进展、典型关系数据库及SQL标准介绍、Access数据库实例分析。

本书内容循序渐进、深入浅出,通过实例分析力图使读者能对数据库系统有一个全面、深入、系统的了解。同时在每章后面都附有相应的习题,便于读者巩固所学的知识。

本书可作为高等院校计算机及相关专业的数据库课程教材,也可供从事计算机软件工作的科技人员、工程技术人员以及其他有关人员参阅。

数据库系统原理与应用

刘德春 主编

出版: 湖北人民出版社
发行:

地址:武汉市雄楚大街 268 号
邮编:430070

印刷:南阳市寰宇印务有限公司

开本:787 毫米×1092 毫米 1/16

字数:431 千字

版次:2003 年 8 月第 1 版

印数:1—5 000

书号:ISBN 7—216—03739—1/G·1058

印张:17.5

插页:2

印次:2003 年 8 月第 1 次印刷

定价:28.00 元

本社网址:<http://www.hbpp.com.cn>

前 言

数据库技术是 20 世纪后期产生和发展起来的一项计算机数据管理技术,是计算机科学技术中发展最快的领域之一,它的出现和发展使计算机应用渗透到工农业生产、商业、行政管理、科学研究、教育、工程技术和国防军事等人类社会的各个领域,而且已经围绕着数据库技术形成了一个巨大的相关软件产业。本教材是为适应目前数据库技术迅猛发展的形势,以及为使数据库教学能够体现这些新的发展目标而编写的。

本书是一本适合高等院校计算机专业或相关专业的数据库课程教学用书,它以关系数据库系统为核心,完整地论述了数据库系统的基本概念、基本原理、应用技术和最新发展技术,并结合 Access 的应用实践,使读者能对数据库系统有一个全面、深入、系统的理解和掌握,为进一步从事数据库系统的研究、开发和应用奠定坚实的基础。

本书的主要内容包括:数据库系统的基本概念、数据模型与概念模型、关系数据库标准语言 SQL、关系数据库设计理论、数据库的设计、数据库的保护、数据库技术新进展、典型关系数据库产品及 SQL 标准介绍和 Access 2000 数据库实例分析等。

本书的编写以关系数据库系统为核心,在系统论述数据库的基础上,重点讨论关系数据库的原理与实现,其中对关系数据模型、关系数据库体系结构、数据库的设计理论与设计方法、数据库保护等进行了较为详细系统的介绍,对层次数据库和网状数据库等传统数据库内容进行了简介。在本书的编写过程中始终注重理论联系实际,结合数据库应用举例介绍,加强对数据库理论的学习。为适应信息技术发展的需要,书中还对数据库设计方法、开发工具、标准 SQL 语言等数据库应用技术和反映当前数据库领域发展趋势的新产品新技术进行了较全面的论述,为读者开发数据库应用系统或维护管理大中型数据库系统打下基础。

本书在内容选取、章节安排、难易程度等方面充分考虑教学的需要,力求使概念准确、清晰,重点明确,内容广泛,便于取舍,每章均配有习题便于教学使用。

本书由刘德春、殷海娜、孙晓莹、吴绍兴、郭东恩编写,刘德春任主编,其中刘德春编写第 1 章、第 6 章、第 8 章,吴绍兴编写第 2 章、第 7 章,孙晓莹编写第 3 章,殷海娜编写第 5 章,郭东恩编写第 4 章、第 9 章,全书由刘德春统稿审定。

在本书的编写过程中,北京航空航天大学的周孝宽教授给予了许多帮助,同时南阳理工学院网络中心给予了大力协助支持,在此一并表示诚挚的谢意。由于编者水平和条件所限,本书难免会有诸多不足之处,恳请专家和读者提出宝贵意见并批评指正。

编 者
2003 年 7 月

目 录

第 1 章 数据库系统基本概念	(1)
1.1 数据库的作用及相关概念	(1)
1.1.1 信息与数据的特征及关系	(1)
1.1.2 信息的三种世界及其描述	(2)
1.1.3 数据管理与数据库的作用及特点	(5)
1.1.4 数据库管理系统与信息管理系统	(7)
1.2 数据库系统的组成和结构	(9)
1.2.1 数据库系统的组成	(9)
1.2.2 数据库管理系统的功能结构	(12)
1.2.3 数据库系统的三级数据模式结构	(13)
1.3 数据库技术及其发展	(15)
1.3.1 手工管理数据阶段及特点	(16)
1.3.2 文件系统阶段及特点	(16)
1.3.3 数据库系统阶段及特点	(17)
1.4 数据库系统的特点	(20)
1.4.1 实现数据的集中化控制	(20)
1.4.2 数据的冗余小	(21)
1.4.3 采用一定的数据模型实现数据结构化	(21)
1.4.4 避免了数据的不一致性	(21)
1.4.5 实现数据共享	(21)
1.4.6 提供数据库保护	(22)
1.4.7 数据独立性	(22)
1.4.8 数据由 DBMS 统一管理和控制	(22)
习题	(23)
第 2 章 数据模型与概念模型	(24)
2.1 数据模型	(24)
2.1.1 信息世界中的基本概念	(24)
2.1.2 表示概念模型的实体—联系方法	(25)
2.1.3 实体联系类型	(25)
2.1.4 概念模型的表示方法	(27)
2.1.5 数据模型	(28)
2.2 数据模型的三要素	(29)
2.3 层次模型	(29)
2.3.1 层次模型的数据结构	(30)
2.3.2 层次模型中多对多联系的表示	(32)
2.3.3 层次模型的数据操作和完整性的约束条件	(33)

2.3.4 层次模型的存储结构·····	(33)
2.4 网状模型及特点 ·····	(35)
2.4.1 网状模型的数据结构·····	(35)
2.4.2 网状模型的完整性约束条件·····	(36)
2.4.3 网状模型的存储结构·····	(36)
2.4.4 网状模型与层次模型的比较·····	(36)
2.5 关系数据模型 ·····	(37)
2.5.1 关系数据库的概述·····	(37)
2.5.2 关系数据结构·····	(38)
2.5.3 关系的完整性·····	(40)
2.5.4 关系代数·····	(41)
2.5.5 关系演算·····	(43)
2.6 关系数据库管理系统 ·····	(51)
2.7 关系模型的优缺点 ·····	(51)
2.8 关系模型与非关系模型比较 ·····	(52)
习题 ·····	(52)
第3章 关系数据库标准语言 SQL ·····	(54)
3.1 SQL语言概述·····	(54)
3.2 SQL语言特点·····	(54)
3.2.1 SQL基本功能·····	(54)
3.2.2 SQL语言的特点 ·····	(55)
3.2.3 SQL语言分类简介·····	(55)
3.2.4 本章示例说明·····	(56)
3.3 数据定义语言 ·····	(57)
3.3.1 定义、删除基本表·····	(58)
3.3.2 索引的定义与维护 ·····	(61)
3.4 数据查询语句 ·····	(62)
3.4.1 Select 语句概述及语句语法 ·····	(62)
3.4.2 数据查询中关系运算符的执行 ·····	(63)
3.4.3 运算符的使用 ·····	(67)
3.4.4 数据排序与分组 ·····	(70)
3.5 数据更新 ·····	(72)
3.5.1 插入数据 ·····	(73)
3.5.2 修改数据 ·····	(73)
3.5.3 删除数据 ·····	(74)
3.6 子查询和集合运算 ·····	(75)
3.6.1 子查询·····	(75)
3.6.2 集合运算·····	(80)
3.7 数据控制 ·····	(81)
3.7.1 数据库安全控制·····	(82)

3.7.2	数据操作权限的设置	(84)
3.8	视图	(85)
3.8.1	视图简介	(85)
3.8.2	视图的定义与使用	(86)
3.9	嵌入式 SQL 语言	(87)
3.9.1	嵌入式 SQL 语言的简介	(87)
3.9.2	嵌入式 SQL 语言的一般形式	(88)
3.9.3	嵌入式 SQL 语言与主语言之间的通信	(88)
3.9.4	动态 SQL 语句	(92)
	习题	(92)
第 4 章	关系数据库规范化理论	(95)
4.1	为何引入规范化理论	(95)
4.1.1	关系模式基本要求	(95)
4.1.2	关系模式中的数据依赖	(96)
4.2	函数依赖有关概念	(99)
4.3	范式	(101)
4.3.1	第一范式(1NF)	(102)
4.3.2	第二范式(2NF)	(103)
4.3.3	第三范式(3NF)	(104)
4.3.4	BC 范式(BCNF)	(106)
4.3.5	多值依赖与第四范式(4NF)	(107)
4.3.6	连接依赖与关系的第五范式	(110)
4.4	数据依赖公理系统	(112)
4.4.1	函数依赖公理系统	(112)
4.4.2	闭包	(113)
4.4.3	函数依赖的覆盖	(113)
4.4.4	依赖集的最小化及其在数据库中应用	(114)
4.5	关系模式的规范化	(115)
4.5.1	关系模式规范化的步骤	(116)
4.5.2	关系模式的分解	(117)
4.5.3	模式分解算法	(119)
	习题	(120)
第 5 章	数据库设计	(122)
5.1	数据库的设计方法	(122)
5.2	数据库的设计步骤	(122)
5.3	需求分析	(125)
5.3.1	需求分析的任务	(125)
5.3.2	需求分析的方法	(125)
5.3.3	数据字典	(130)
5.4	概念结构设计	(134)

5.4.1	概念结构设计的方法步骤	(134)
5.4.2	数据抽象、局部视图的设计和视图集成	(135)
5.5	逻辑结构设计	(147)
5.5.1	E-R 图到关系模式的转换	(147)
5.5.2	关系模式的优化	(149)
5.4.3	用户子模式设计	(149)
5.6	数据库物理设计	(151)
5.6.1	数据库物理设计内容方法	(151)
5.6.2	评价物理结构	(153)
5.7	数据库的实施	(153)
5.8	数据库的运行和维护	(155)
习题		(156)
第 6 章	数据库安全保护	(159)
6.1	数据库的安全性	(160)
6.1.1	数据库安全性的含义	(160)
6.1.2	安全性控制的一般方法	(160)
6.2	完整性控制	(165)
6.2.1	数据库完整性的含义	(165)
6.2.2	完整性规则的组成	(165)
6.2.3	完整性约束条件的分类	(166)
6.3	并发控制与封锁	(167)
6.3.1	数据库并发性的含义	(167)
6.3.2	事务	(168)
6.3.3	并发操作与数据的不致性	(169)
6.3.4	封锁	(172)
6.4	数据库的恢复	(178)
6.4.1	数据库恢复的含义	(178)
6.4.2	数据库恢复的原理及其实现技术	(178)
6.4.3	数据库的故障和恢复的策略	(180)
6.5	Access 的保护方法	(182)
6.5.1	Access 安全管理	(182)
6.5.2	保护 Access 数据库	(189)
习题		(190)
第 7 章	数据库技术新进展	(191)
7.1	数据库技术发展概述	(191)
7.2	数据库模型及数据库系统的发展	(191)
7.2.1	第一代数据库系统	(192)
7.2.2	第二代数据库系统	(192)
7.2.3	第三代数据库系统	(194)
7.2.4	Web 数据库的发展	(195)

7.3	面向对象的数据技术	(197)
7.3.1	面向对象的模型的核心概念	(198)
7.3.2	面向对象的数据库系统的特点	(200)
7.3.3	面向对象的数据库语言	(201)
7.3.4	面向对象的数据库的研究内容	(201)
7.4	数据库技术与其他相关技术相结合	(202)
7.4.1	分布式数据库	(202)
7.4.2	主动数据库	(204)
7.4.3	多媒体数据库	(206)
7.4.4	数据库技术新进展与科学数据库工程	(210)
7.5	Web 与数据库的集成技术	(211)
7.5.1	Web 与数据库集成的一般结构和方法	(211)
7.5.2	主要厂家技术及产品介绍	(213)
习题		(218)
第 8 章	典型关系数据库及 SQL 标准介绍	(219)
8.1	关系数据库管理系统产品概述	(219)
8.2	典型数据库	(221)
8.2.1	ORACLE	(221)
8.2.2	SYBASE	(222)
8.2.3	INFORMIX	(222)
8.2.4	DB2	(222)
8.2.5	INGRES	(223)
8.3	SQL 标准介绍	(223)
8.3.1	概述	(223)
8.3.2	SQL-92	(224)
8.3.3	SQL:1999	(226)
习题		(229)
第 9 章	Access 数据库实例分析	(230)
9.1	系统任务的提出和设计思路	(233)
9.1.1	系统任务的提出	(233)
9.1.2	设计思路	(233)
9.2	系统各部分功能介绍	(234)
9.2.1	数据库的启动	(234)
9.2.2	数据库的组成	(234)
9.2.3	系统各部分功能介绍	(234)
9.3	建立数据库	(235)
9.4	表的建立	(236)
9.4.1	表的组成结构	(237)
9.4.2	表的建立	(238)
9.4.3	建立表间关系	(244)

9.4.4 建立数据的筛选	(246)
9.5 数据的查询与统计	(247)
9.5.1 查询的定义及种类	(247)
9.5.2 创建查询	(248)
9.5.3 数据的统计	(250)
9.6 窗体的建立	(253)
9.6.1 新建窗体	(254)
9.6.2 创建复杂窗体	(258)
9.6.3 控制窗体的建立	(258)
9.7 宏的设计	(259)
9.7.1 宏的设计	(259)
9.7.2 宏组的制作	(260)
9.7.3 常用宏简介	(261)
9.8 报表的制作	(262)
9.8.1 创建报表	(263)
9.8.2 报表的美化设计	(264)
9.9 数据访问页的制作	(264)
9.9.1 创建空白的数据访问页	(264)
9.9.2 使用字段列表将数据添加到数据访问页	(265)
9.10 系统的优化	(265)
9.10.1 处理窗体的属性	(265)
9.10.2 对命令按钮控件的优化	(266)
9.10.3 优化系统的性能	(267)
习题	(267)
参考文献	(269)

第 1 章 数据库系统的基本概念

数据库(Database),简单地说就是数据的仓库,即数据存放的地方。我们周围有许多数据库的例子,如通讯录是一个小数据库,图书馆则是一个典型的大型数据库。小数据尚可用手工管理,而大型数据库必须由计算机进行管理。在计算机三大应用领域(科学计算、过程控制和数据处理)中,数据处理所占比例约为 70%。20 世纪 60 年代末,数据库技术作为数据处理的最新技术应运而生。

数据库技术是计算机学科中的一个重要的组成部份,它的应用非常广泛,几乎涉及到所有的应用领域。要想掌握好数据库系统技术,必须弄清什么是数据、数据管理、数据库、数据模型和概念模型等专业术语的内涵;了解数据库的发展过程和数据库系统的特点;弄清数据库、数据库管理系统和信息管理系统三者之间的关系。本章将介绍这些数据库系统的基本概念和基础知识。

1.1 数据库的作用及相关概念

1.1.1 信息与数据的特征及关系

1. 信息(Information)、信息特征及作用

在日常生活中,我们经常可以听到“信息”这个名词。什么是信息呢?简单地说,信息就是新的、有用的事实和知识,它具有实效性、有用性和知识性,是客观世界的反映。信息是客观存在的,人类有意识地对信息进行采集并加工、传递,而形成了各种消息、情报、指令、数据及信号等。

信息具有四个基本特征:一是信息的内容是关于客观事物或思想方面的知识,即信息的内容能反映已存在的客观事实、能预测未发生事物的状态和能用于指挥控制事物发展的决策;二是信息是有用的,它是人们活动的必需知识,利用信息能够克服工作中的盲目性,增加主动性和科学性,可以把事情办得更好;三是信息能够在空间和时间上被传递,在空间上传递信息称为信息通信,在时间上传递信息称为信息存储;四是信息需要一定的形式表示,信息与其表现符号不可分离。

信息对于人类社会的发展具有重要意义。它可以提高人们对事物的认识,减少人们活动的盲目性;信息是社会机体进行活动的纽带,社会的各个组织通过信息网相互了解并协同工作,使整个社会协调发展;社会越发展,信息的作用就越突出;信息又是管理活动的核心,要想把事物管理好,就需要掌握更多的信息,并利用信息进行工作。

2. 数据(Data)、数据与信息的关系及数据的特征

数据实际上就是描述事物的符号记录,也可以说是描述事物的信息,信息是通过各种方式传播事物的消息,消息通过数据来表达。如文字、图形、图像、声音、学生的档案记录、货物的运

输情况……,这些都是数据。数据的形式本身并不能完全表达其内容,需要经过语义解释。数据与其语义是不可分的。

数据是用于载荷信息的物理符号。这就是说,数据是信息的一种表现形式,数据通过能书写的信息编码表示信息。尽管信息有多种表现形式,它可以通过手势、眼神、声音或图形等方式表达,但数据是信息的最佳表现形式,信息是数据的内涵。由于数据能够书写,因而它能够被记录、存储和处理,从中挖掘出更深层的信息。必须指出的是,在许多不严格的情况下,会把“数据”和“信息”两个概念混为一谈,称“数据”为“信息”。其实,数据不等于信息,数据只是信息表达方式中的一种;正确的数据可表达信息,而虚假、错误的信息所表达的是谬误,不是信息。

数据有以下4个特征:

(1)数据有“型”和“值”之分

数据的型是指数据的结构,而数据的值是指数据的具体取值。数据的结构指数据的内部构成和对外联系。例如:学生的数据由“学号”、“姓名”、“年龄”、“性别”、“所在系”等属性构成,其中“学生”为数据名,“学号”、“性别”等为属性名(或称数据项名);课程也是数据,它由“课程编号”、“课程名称”、“课时数”等数据项构成;“学生”和“课程”之间有“选课”的联系。“学生”和“课程”数据的内部构成及其相互联系就是学生课程数据的类型,而一个具体取值,如“03512,李四,22,男,计算机科学与技术系”,就是一个学生数据值。

(2)数据受数据类型和取值范围的约束

数据类型是针对不同的应用场合设计的。数据的类型不同,则数据表示形式、存储方式及数据能进行的操作运算各不相同。在使用计算机处理信息时,我们应当对数据类型特别重视,为数据选择合适的类型,否则会在应用中产生不良的后果。常见的数据类型有数值型、字符串型、日期型和逻辑型等,它们具有不同的特点和用途。数值型数据就是我们通常所说的算术数据,它能够进行加、减、乘、除等算术运算。字符串型数据是最常用的数据,它可以表示姓名、地址、邮政编码及电话号码等数据,字符串型数据能够进行查找字符串、读取字符串和连接字符串的运算操作。日期型数据是表达日期和时间的数据。逻辑型数据是表达真或假、是或非等逻辑数据。

数据的取值范围又称数据的值域,例如学生性别的值域是{“男”,“女”}。为数据设置值域是保证数据的有效性 & 避免数据输入或修改时出现错误的重要措施。

(3)数据有定性表示和定量表示之分

我们在表示职工的年龄时,可以用“老”、“中”、“青”定性表示,也可以用具体岁数定量表示。由于数据的定性表示是带有模糊因素的粗略表示方式,而数据的定量表示是描述事物的精确表示方式,所以在计算机软件设计中,我们对数据应尽可能地采用定量表示方式。

(4)数据应具有载体和多种表现形式

数据是客体(即客观物体或概念)属性的记录,它必须有一定的物理载体。当数据记录在纸上时,纸张是数据的载体;当数据记录在计算机的外存上时,硬盘、软盘或磁带就是数据的载体。数据具有多种表现形式,它可以用报表、图形、语音及不同的语言符号表示。

1.1.2 信息的三种世界及其描述

信息的三种世界是指现实世界、信息世界和计算机世界(也称数据世界)。数据库是模拟现实世界中某些事务活动的信息集合,数据库中所存储的数据,来源于现实世界的信息流。信息

流用来描述现实世界中一些事物的某些方面的特征及事物间的相互联系。在处理信息流前,必须先对其进行分析并用一定的方法加以描述,然后将描述转换成计算机所能接受的数据形式。

1. 信息的现实世界

现实世界泛指存在于人脑之外的客观世界。信息的现实世界是指我们要管理的客观存在的各种事物、事务之间的相互联系及事物的发生、变化过程。通过对现实世界的了解和认识,使得我们对要管理的对象、管理的过程和方法有了概念模型。认识信息的现实世界并用概念模型加以描述的过程称为系统分析。信息的现实世界通过实体、特征、实体集进行划分和认识。

(1) 实体

现实世界中存在的可以相互区分的事物或概念称为实体。例如,一个学生、一台机器、一部汽车等是具体的事物实体,一门课、一个班级等称为概念实体。

(2) 实体的特征

每个实体都有自己的特征,利用实体的特征可以区别不同的实体。例如,学生通过姓名、性别、年龄、身高、体重等许多特征来描述自己。尽管实体具有许多特征,但是我们在研究时,只选择其中对管理及处理有用的或有意义的特征。例如,对于人事管理,职工的特征可选择姓名、性别、年龄、工资、职务等,而在描述一个人健康情况时,可以用职工的身高、体重、血压等特征表示。

(3) 实体集

只有相同特征或能用同样特征描述的实体的集合称为实体集。例如学生、工人、汽车等都是实体集。

2. 信息世界

现实世界中的事物反映到人们的头脑里,经过认识、选择、命名、分类等综合分析而形成了印象和概念,从而得到了信息。当事物用信息来描述时,即进入信息世界。在信息世界中,实体的特征在头脑中形成的知识称为属性;实体通过其属性表示称为实例;同类实例的集合称为对象,对象即实体集中的实体集属性表示得出的信息集合。实体与实例是不同的,例如张三是一个实体,而“张三,男,23岁,计算机系学生”是实例,现实世界中的张三除了姓名、性别、年龄和所在系外还有其他的特征,而实例仅对需要的特征通过属性进行了描述。在信息世界中,实体集之间的联系用对象联系表示。

信息世界通过概念模型(也称信息模型)、过程模型和状态模型反映现实世界,它要求对现实世界中的事物、事物间的联系和事物的变化情况准确、如实、全面地表示。概念模型通过 E-R 图中的对象、属性和联系对现实世界的事物及关系给出静态描述;过程模型通过信息流程图和数据字典描述事物的处理方法和信息加工过程;状态模型通过事物状态转换图对事物给出动态描述。数据库主要是根据概念模型设计的,而数据处理方法主要是根据过程模型设计,状态模型对数据库的系统功能设计则有重要的参考价值。

3. 信息的计算机世界

信息世界中的信息,经过数字化处理形成计算机能够处理的数据,就进入了计算机世界,计算机世界也叫机器世界或数据世界。在信息转换为数据的过程中,对计算机硬件和软件(软件主要指数据库管理系统)都有限定,所以信息的表示方法和信息处理能力要受到计算机硬件和

软件限制。也就是说,数据模型应符合具体的计算机系统和 DBMS 的要求。

在计算机世界中用到下列术语:

(1)数据项(Item)

数据项是对象属性的数据表示。数据项有型和值之分,数据项的型是对数据特性的表示,它通过数据项的名称、数据类型、数据宽度和值域等来描述;数据项的值是其具体取值。数据项的型和值都要符合计算机数据的编码要求,即都要符合数据的编码要求。

(2)记录(Record)

记录是实例的数据表示。记录有型和值之分:记录的型是结构,由数据项的型构成;记录的值表示对象中的一个实例,它的分量是数据项值。例如,“姓名,性别,年龄,所在系”是学生数据的记录型,而“张三,男,23,计算机系”是一个学生的记录值,它表示学生对象的一个实例,“张三”、“男”、“23”、“计算机系”都是数据项值。

(3)文件(File)

文件是对象的数据表示,是同类记录的集合。即同一个文件中的记录类型应是一样的。例如将所有学生的登记表组成一个学生数据文件,文件中的每条记录都要按“姓名,性别,年龄,所在系”的结构组织数据项值。

(4)数据模型(DataModel)

现实世界中的事物和相互联系反映到计算机世界中,是文件的记录结构和不同文件间的记录与记录之间的联系,它们数据化的结果就是数据模型。

4. 现实世界、信息世界和计算机世界的关系

现实世界、信息世界和计算机世界这三个领域是由客观到认识、由认识到使用管理的不同层次,后一领域是前一领域的抽象描述。

关于三个领域之间的术语对应关系,可由表 1-1 表示。

表 1-1 三种信息世界的术语对应关系表

现实世界	信息世界	计算机世界
实体	实例	记录
特征	属性	数据项
实体集	对象或实体型	数据或文件
实体间的联系	对象间的联系	数据间的联系
	概念模型	数据模型

现实世界、信息世界和计算机世界的转换关系可以用图 1-1 表示。从图中可以看出,现实世界的事物及联系,通过系统分析成为信息世界的信息模型,而信息模型经过数据化处理转换为数据模型。

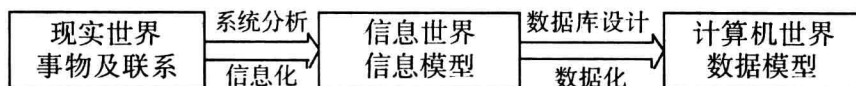


图1-1 信息的三个世界的联系和转换过程

1.1.3 数据管理与数据库的作用及特点

数据管理是数据处理的基础工作,数据库是数据管理的技术和手段。数据库中的数据具有整体性和共享性。

1. 数据处理(Data Processing)及分类

围绕着数据所做的工作均称为数据处理。数据处理是指对数据的收集、组织、整理、加工、存储和传播等工作。我们可以把数据处理分为3类:一类为数据管理,它的主要任务是收集信息、将信息用数据表示并按类别组织保存,其目的是在需要的时候,为各种应用和数据处理提供数据;另一类是数据加工,它的主要任务是对数据进行变换、抽取和运算,通过数据加工会得到更有用的数据,以指导或控制人的行为或事物的变化趋势;最后一类是数据传播,它在空间或时间上以各种形式传播信息,而不改变数据的结构、性质和内容,数据传播会使更多的人得到并理解信息,从而使信息的作用充分发挥出来。

数据是原料,是输入;而信息是产出,是输出结果。“信息处理”的真正含义应该是为了产生信息而处理数据。我们可以用下面表达式简单地表示出信息、数据与数据处理的关系:信息 = 数据 + 处理。

2. 数据管理(Data Management)及内容

在数据处理中,最基本的工作是数据管理工作。数据管理是其他数据处理的核心和基础。具体讲,数据管理工作应包括3项内容:一是组织和保存数据,即将收集到的数据合理地分类组织,将其存储在物理载体上,使数据能够长期地被保存;二是进行数据维护,即根据需要随时进行插入新数据、修改原数据和删除失效数据的操作;三是提供数据查询和数据统计功能,以便快速地得到需要的正确数据,满足各种使用要求。

数据管理在实际工作中的地位很重要。我们周围有许多人从事各种行政管理工作,这些管人、管财、管物或管事(人、财、物和事统称为事务)的工作实际上就是数据管理工作。在事务管理中,事务(人、财、物和事)以数据的形式被记录和保存。例如在财务管理中,财务科通过对各种账目的记账、对账或查账等实现对财务数据的管理。传统的数据管理方法是人工管理方式,即通过手工记账、算账和保管账的方法实现对各种事务的管理。计算机的发展为科学地进行数据管理提供了先进的技术和手段,目前许多数据管理工作利用计算机进行,而数据管理(即信息或事务管理)也成了计算机应用的一个重要分支。

3. 数据库(Database,简称DB)及数据库中的数据性质

数据库简称为DB,它是一个按数据结构来存储和管理数据的计算机软件系统。数据库是长期存储在计算机内有结构的大量的共享的数据集合。它可以供各种用户共享、具有最小冗余度和较高的数据独立性。DBMS在数据库建立、运用和维护时对数据库进行统一控制,以保证数据的完整性、安全性,并在多用户同时使用数据库的时候进行并发控制,在发生故障后对系统进行恢复。

数据库的概念实际上包括两层意思:

(1)数据库是一个实体,它是能够合理保管数据的“仓库”,用户在该“仓库”中存放要管理的

事务数据,“数据”和“库”这两个概念结合成为“数据库”。

(2)数据库是数据管理的新方法和技术,它能够更合理地组织数据、更方便地维护数据、更严密地控制数据和更有效地利用数据。

在数据库技术出现之前,人们采用“数据文件”方法进行数据管理。数据库方法与文件方法相比,具有以下两个明显的进步:

(1)数据库中的数据具有数据整体性

数据库中的数据保持了自身完整的数据结构,该数据结构是从全局观点出发建立的;而文件中的数据一般是不完整的,其数据结构是根据某个局部要求或功能需要建立的。从设计系统的思想方法讲,数据库方法是面向对象的方法,而文件方法是面向过程的方法。数据库要保持数据(即事务)自身的结构完整,强调站在全局的角度设计数据结构,并以数据库为基础进行功能设计;文件系统(用文件方法建立的数据管理系统)则是站在具体要实现的功能角度上考虑数据结构,按各个具体功能需要分别组织数据,数据完全依附于功能需要。让我们通过下面的简单例子来说明数据库的数据整体性特征。

如果按数据库方法设计一个“职工”的数据,应深入到所有使用“职工”数据的部门进行了解,并将得到的信息综合后,才能得出“职工”的数据结构。例如,要到人事处、财务处、校医院、科研处等每个与“职工”数据相关的地方,了解职工的一般情况、工资情况、身体情况及科研情况的综合内容,这种综合内容为“职工”数据的内部组成,它可以用下面的数据结构表示:

职工(职工编号,姓名,性别,出生日期,家庭住址,职务,职称,政治面貌,基本工资,附加工资,身体状况,病史情况,业务特长,主要科研成果。)

如果是按文件方法设计一个“职工”的数据,则需要为人事处、财务处、校医院、科研处等建立不同的“职工”数据文件(职工1, ..., 职工4),以满足各部门对于“职工”数据的要求。这些“职工”数据文件的记录结构为:

职工1(职工编号,姓名,性别,出生日期,家庭住址,职务,职称,政治面貌);

职工2(职工编号,姓名,性别,基本工资,附加工资);

职工3(职工编号,姓名,性别,出生日期,身体状况,病史情况);

职工4(职工编号,姓名,性别,出生日期,职务,职称,业务特长,科研成果)。

从以上例子可以看出,在数据库中使用的“职工”数据全面反映了职工的各个特征,消除了大量的数据冗余;而文件系统中的“职工”数据则是从不同的侧面反映职工的某些特征,尽管它使用了4种数据文件表示“职工”,但无论哪个数据文件都不能完整地表示职工情况。

(2)数据库中的数据具有数据共享性

文件系统的数据库文件是为满足某一个功能模块的使用要求而建立的,数据与功能程序是一一对应的关系。文件系统中的数据与功能程序之间存在着非常紧密的相互依赖关系,即数据离开相关的功能程序就失去了它存在的价值,功能程序如果没有数据支持就无法工作。数据库中的数据是为众多用户所共享其信息而建立的,它已经摆脱了具体程序的限制和制约。数据库的数据共享性表现在两个方面:

1)不同的用户可以按各自的用法使用数据库中的数据。数据库能为用户提供不同的数据视图,以满足个别用户对数据结构、数据命名或约束条件的特殊要求。

2)多个用户可以同时共享数据库中的数据资源,即不同的用户可以同时存取数据库中的同一个数据。

数据共享性不仅满足了各用户对信息内容的要求,同时也满足了各用户之间的信息通信要

求。在上述例子中,数据库中的“职工”数据是为人事处、财务处、校医院、科研处等部门所共同使用,其中人事处可以按“职工 1”、财务处可以按“职工 2”、校医院可以按“职工 3”、科研处可以按“职工 4”的结构形式使用数据,它们使用共同的“职工”数据源。“职工”数据不仅能为现有的各个应用功能提供数据,而且由于其自身结构是完整的,它还可以为今后需要实现的功能或别的应用系统提供相应的信息。

1.1.4 数据库管理系统与信息管理系统

数据库管理系统是提供数据库管理的计算机系统软件,信息管理系统是实现某种具体事物管理功能的应用软件。数据库管理系统为信息管理系统的的设计提供了方法、手段和工具,信息管理系统利用数据库管理系统可以更快、更好地设计和实施。

1. 数据库管理系统(Database Management System,简称 DBMS)

数据库管理系统简称 DBMS,它是专门用于管理数据库的计算机系统软件。数据库管理系统能够为数据库提供数据的定义、建立、维护、查询和统计等操作功能,并完成对数据完整性、安全性进行控制的功能。

数据库管理系统是位于用户与操作系统之间的一层数据管理软件。

数据库在建立、运用和维护时由数据库管理系统统一管理、统一控制。数据库管理系统使用户能方便地定义数据和操纵数据,并能够保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后的系统恢复。

在数据库管理系统的操作功能中,数据定义功能是指为说明库中的数据情况而进行的建立数据库结构的操作,通过数据定义可以建立起数据库的框架;数据库建立功能是指将大批数据录入到数据库的操作,它使得库中含有需要保存的数据记录;数据库维护功能是指对数据的插入、删除和修改操作,其操作能满足库中信息变化或更新的需求;数据查询和统计功能是指通过对数据库的访问,为实际应用提供需要的数据。

数据库管理系统不仅要为数据管理提供数据操作功能,还要为数据库提供必要的数据库控制功能。数据库管理系统的数据库控制主要指对数据安全性和完整性的控制。数据安全性控制是为了保证数据库的数据安全可靠,防止不合法的使用造成数据泄漏和破坏,即避免数据被人偷看、篡改或破坏;数据完整性控制是为了保证数据库中数据的正确、有效和相容,以防止不合语义的错误数据被输入或输出。

数据库管理系统的目标是让用户能够更方便、更有效、更可靠地建立数据库和使用数据库中的信息资源。数据库管理系统不是应用软件,它不能直接用于诸如工资管理、人事管理或资料管理等事务管理工作,但数据库管理系统能够为事务管理提供技术和方法、应用系统的设计平台和设计工具,使相关的事务管理软件很容易设计。也就是说,数据库管理系统是为设计数据管理应用项目提供的计算机软件,利用数据库管理系统设计事务管理系统可以达到事半功倍的效果。我们周围有关数据库管理系统的计算机软件有很多,其中比较著名的系统有 Oracle、Informix、Sybase、SQLServer、Delphi 等。

2. 数据库系统(Database System,简称 DBS)

数据库系统是指在计算机系统中引入数据库后的系统构成,一般由数据库、数据库管理系