

面向21世纪信息管理与信息系统专业
东南大学出版社 核心课程系列教材



数据库技术 及其应用

刘秋生 / 主编

数据库技术及其应用

主 编 刘秋生

副主编 何有世 张海斌 陆丽丹

东南大学出版社

·南京·

内容提要

全书分两大部分共 12 章:第一部分主要介绍数据库的基础理论与基本概念;第二部分围绕数据库系统的开发工具,介绍表、数据库的建立和维护,视图和查询设计,表单设计,报表和标签设计,项目设计等数据库操作过程、操作命令和应用软件设计。每章配有习题。全书围绕数据库的基本概念、基本操作和信息处理的基本功能展开讨论,既有完整的理论性体系,又有很强的实用性,便于教学和学习。

本书可作为高等院校信息管理与信息系统专业本科和信息化方向研究生的技术基础课教材,也可以作为高职高专、职工大学、业余大学、夜大学、函授大学、成人教育学院等经济类、管理类本专科等层次的数据库技术课程的教材,还可以作为广大信息系统研发爱好者及数据员的自学用书。

图书在版编目(CIP)数据

数据库技术及其应用/刘秋生主编. —南京:东南大学出版社,2003.6

ISBN 7-81089-274-6

I. 数... II. 刘... 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(2003)第 041922 号

东南大学出版社出版发行
(南京四牌楼 2 号 邮编 210096)

出版人:宋增民

江苏省新华书店经销 南京京新印刷厂印刷

开本:B5 印张:20 字数:437 千字

2003 年 7 月第 1 版 2003 年 7 月第 1 次印刷

印数:1—4000 定价:28.00 元

(凡因印装质量问题,可直接向发行科调换。电话:025-3795801)

面向 21 世纪信息管理与信息系统专业

核心课程教材建设委员会

孙建军（南京大学信息管理系教授、博导）

吴清烈（东南大学经济管理学院副教授、博士）

武 忠（东南大学经济管理学院副教授、博士）

史田华（南京理工大学信息管理系教授）

王曰芬（南京理工大学信息管理系副教授、博士）

郑会颂（南京邮电学院管理工程系教授）

何有世（江苏大学工商管理学院教授、博士）

刘秋生（江苏大学工商管理学院副教授、博士）

周建屏（苏州大学图书馆副研究馆员）

成 颖（南京大学信息管理系）

前 言

数据库系统概念的形成至今尚不到 50 年时间,但是,数据库系统随着信息技术的发展日新月异,理论上更加完善,应用极为广泛。数据库系统软件十分丰富,特别是数据库管理系统层出不穷。数据库系统应用软件即信息系统已经成为企业管理的必备工具。数据库技术已成为当今信息社会的基础技术,是管理类学生必须掌握的基础知识。

本书作为信息管理与信息系统专业的核心课程教材,是在编者多年从事数据库技术及应用的教学实践基础上编写而成的。

本书的主要特点是:

(1) 系统性强。从数据库系统的应用着手,认识数据库在信息系统中的作用,全面地介绍数据的收集、传输、存储、加工、维护和使用的的基础知识和基本操作。

(2) 重点突出。全书围绕 3 个中心,分别重点介绍数据库的基础知识、数据库的基本操作和数据库系统应用编程设计。

(3) 实用性强。强调了理论与实践相结合。将学生容易理解掌握的学生学籍管理作为典例,并采用图文并茂的方式,便于阅读理解。

(4) 内容精练。在内容上作了精心的安排,以操作简便、容易掌握、实用性强、应用面广的面向对象的关系型数据库管理系统 Visual FoxPro 作为数据库系统的开发工具。围绕 Visual FoxPro 的基本概念、基本操作,由浅入深地、系统地介绍关系型数据库管理系统的功能和数据处理的方法。

(5) 适用面广。本书是面向理工类专业学生的教学用书,也可以作为计算机专业学生和其他工程技术人员的自学用书。

全书共分 12 章。建议总课时为 90 学时,其中上机实验课为 30 学时。在有条件的情况下,安排多媒体教室上课 30 学时,否则适当增加上机实验课时,同时还应该适当安排学生利用课余时间独立上机完成数据库操作练习。各院校可以根据实际情况按上述比例压缩或增加学时。

本书的构思由江苏大学刘秋生副教授完成,并编写了第 1~4 章及第 7、第 8、第 9 章。其他章节由江苏大学何有世教授、张海斌老师和常熟广播电视大学的陆丽丹老师编写。江苏大学赵广凤老师,吴红平、金海燕等同志为组稿、复核、数据处理等工作付出了大量的精力,在此一并表示衷心感谢!

本书的全体编著人员结合科研成果和教学经验,突出重点、突出实用,反复提炼、精益求精,以通俗、简练的语言,阐述较先进的技术。

欢迎读者对本书中存在的缺点和不足提出批评指正。

编 者

于 2003 年 4 月

目 录

1 概 述	(1)	2.4.2 关系代数运算	(20)
1.1 数据库技术的发展	(1)	2.4.3 关系数据库理论	(25)
1.1.1 人工管理阶段	(1)	习题 2	(28)
1.1.2 文件管理阶段	(1)	3 数据库设计	(30)
1.1.3 数据库管理阶段	(2)	3.1 需求分析	(30)
1.2 数据库基础知识	(2)	3.1.1 用户类型	(30)
1.2.1 数据与信息	(2)	3.1.2 需求分析任务	(30)
1.2.2 记录与数据表	(4)	3.1.3 需求分析方法	(31)
1.2.3 数据库	(5)	3.1.4 需求说明书	(31)
1.2.4 数据库技术新概念	(6)	3.2 收集、筛选数据	(32)
1.3 数据库系统的研究与应用领域	(6)	3.2.1 数据收集	(32)
1.3.1 数据库技术的研究范围	(6)	3.2.2 数据筛选	(33)
1.3.2 数据库技术应用领域	(7)	3.3 数据定义	(34)
1.4 数据库管理系统	(8)	3.3.1 数据定义的依据	(34)
1.4.1 数据库管理系统的功能	(9)	3.3.2 数据定义的内容	(34)
1.4.2 几种典型的数据库管理系统	(10)	3.3.3 建立数据字典	(36)
1.5 数据库系统	(11)	3.4 数据库设计	(37)
1.5.1 数据库系统的组成	(11)	3.4.1 面向对象的数据库设计的基本概 念	(37)
1.5.2 数据库系统的特征	(13)	3.4.2 约束和触发	(38)
1.5.3 数据库系统的结构	(13)	3.4.3 数据库建模	(40)
习题 1	(14)	3.4.4 子类	(43)
2 数据模型	(15)	3.4.5 对约束的建模	(44)
2.1 E-R 模型	(15)	3.4.6 ODL 设计转换成关系设计	(44)
2.1.1 E-R 模型的基本概念	(15)	习题 3	(46)
2.1.2 E-R 图	(16)	4 RDBMS 标准语言 SQL 简介	(47)
2.2 层次型数据模型	(17)	4.1 SQL 的数据定义功能	(48)
2.3 网状型数据模型	(17)	4.1.1 创建和维护数据表	(48)
2.4 关系型数据模型	(18)	4.1.2 建立和删除数据表的索引	(50)
2.4.1 关系型数据模型的概念	(18)	4.2 SQL 的数据操纵功能	(50)

4.2.1 查询语句	(50)	5.5.4 表达式	(83)
4.2.2 更新语句	(57)	5.5.5 函数	(86)
4.3 视图	(60)	5.6 Visual FoxPro 的基本操作类型	(94)
4.3.1 创建视图	(60)	习题 5	(96)
4.3.2 修改视图	(61)	6 数据表、库操作	(98)
4.3.3 使用视图	(62)	6.1 创建自由表	(98)
4.4 SQL 数据控制功能	(63)	6.1.1 使用表设计器建立新表	(98)
4.4.1 使用权的授予	(63)	6.1.2 使用表向导创建表	(100)
4.4.2 使用权的收回	(64)	6.1.3 使用命令创建表	(102)
4.5 常用数据库管理系统	(64)	6.2 打开和关闭表	(102)
4.5.1 Oracle 数据库管理系统	(64)	6.2.1 打开表	(103)
4.5.2 DB2 数据库管理系统	(65)	6.2.2 关闭表	(103)
4.5.3 Sybase 数据库管理系统	(66)	6.3 查看表的内容	(103)
4.5.4 PowerBuilder 数据库管理系统	(66)	6.3.1 浏览窗口	(103)
4.5.5 MS - Access 数据库管理系统	(67)	6.3.2 编辑窗口	(104)
4.5.6 Informix 数据库管理系统	(67)	6.3.3 记录定位	(104)
4.5.7 MS Server 2000 数据库管理系统	(67)	6.3.4 调整浏览窗口	(104)
习题 4	(68)	6.4 增加、修改表内容	(105)
5 Visual FoxPro 基础	(70)	6.4.1 在表中加入新记录	(105)
5.1 Visual FoxPro 安装、启动	(70)	6.4.2 在表中编辑字段	(106)
5.1.1 安装前的准备	(70)	6.4.3 记录的删除	(106)
5.1.2 安装过程	(70)	6.4.4 成批修改表内容	(108)
5.1.3 启动 Visual FoxPro	(71)	6.5 修改表结构	(109)
5.2 Visual FoxPro 界面主菜单及操作方法	(72)	6.6 索引	(109)
5.3 Visual FoxPro 操作方式、命令书写规定	(76)	6.6.1 索引的种类	(109)
5.4 Visual FoxPro 技术指标	(77)	6.6.2 索引的建立	(110)
5.5 Visual FoxPro 数据类型、常量、变量、表达式	(78)	6.6.3 打开索引	(111)
5.5.1 Visual FoxPro 数据类型	(78)	6.6.4 索引的维护	(112)
5.5.2 常量	(79)	6.7 数据库的创建	(113)
5.5.3 变量	(81)	6.8 数据库的打开和关闭	(114)
		6.8.1 打开数据库	(114)
		6.8.2 关闭数据库	(114)
		6.9 在数据库中创建、添加、移去表	(114)
		6.9.1 在数据库中创建新表	(115)
		6.9.2 添加数据库表	(116)

6.9.3 从数据库中移去表	(116)	8.4.2 表单设计时控件的选择和引用	(168)
6.10 数据表设置	(117)	8.4.3 表单上控件的应用	(168)
6.10.1 创建表之间的关系	(117)	8.5 类的设计与应用	(175)
6.10.2 设置字段属性	(118)	8.5.1 类	(175)
6.10.3 设置表属性	(121)	8.5.2 Visual FoxPro 的基类	(176)
6.10.4 设置参照完整性	(122)	8.5.3 类的设计与应用	(177)
习题 6	(124)	习题 8	(181)
7 视图与查询设计	(126)	9 报表和标签设计	(182)
7.1 视图设计	(126)	9.1 报表与标签的设计方法	(182)
7.1.1 视图设计过程	(126)	9.2 报表设计与创建报表文件	(184)
7.1.2 创建视图	(126)	9.2.1 报表设计	(184)
7.1.3 视图格式修改	(131)	9.2.2 创建报表格式文件	(186)
7.2 视图调用	(131)	9.3 报表格式文件的调用与修改	(194)
7.2.1 打开视图	(132)	9.3.1 报表格式文件的调用	(194)
7.2.2 输出视图数据	(132)	9.3.2 报表格式文件的修改	(195)
7.2.3 视图数据的修改	(133)	9.4 标签设计与创建标签格式文件	(195)
7.2.4 视图的关闭	(133)	9.4.1 标签设计	(195)
7.3 查询设计	(133)	9.4.2 创建标签格式文件	(195)
7.3.1 查询设计的概念	(133)	9.5 标签的调用与修改	(197)
7.3.2 查询设计过程	(134)	9.5.1 标签的调用	(197)
7.3.3 创建查询文件	(134)	9.5.2 标签格式文件的修改	(198)
7.4 查询文件的调用	(145)	9.6 数值统计	(199)
7.5 查询修改	(146)	9.6.1 计数命令	(199)
习题 7	(146)	9.6.2 求和命令	(199)
8 表单、类设计	(148)	9.6.3 求平均值命令	(200)
8.1 对象	(148)	9.6.4 分类求和	(201)
8.1.1 对象	(148)	习题 9	(201)
8.1.2 Visual FoxPro 的容器与控件	(150)	10 菜单设计和工具栏定制	(203)
8.2 表单设计方法	(153)	10.1 菜单设计	(203)
8.2.1 表单设计步骤	(154)	10.1.1 菜单的分类	(203)
8.2.2 表单设计实例	(155)	10.1.2 菜单的组成	(204)
8.3 表单的创建和维护	(156)	10.1.3 菜单的布局	(205)
8.3.1 用向导创建表单	(156)	10.2 创建菜单系统	(208)
8.3.2 用表单设计器创建表单	(160)		
8.4 表单控件的应用	(167)		
8.4.1 控件与数据的关系	(167)		

10.2.1	创建定制菜单·····	(208)	11.5.2	子程序调用·····	(238)
10.2.2	菜单设计器的组成和使用方法 ·····	(209)	11.5.3	数据传递·····	(241)
10.2.3	创建快捷菜单·····	(210)	习题 11	·····	(241)
10.2.4	创建 SDI 菜单·····	(210)	12	项目设计及其应用·····	(243)
10.2.5	创建用户开发的菜单系统····· ·····	(211)	12.1	项目管理器的基本功能····· ·····	(243)
10.2.6	定制菜单·····	(216)	12.1.1	程序和表单的运行·····	(243)
10.3	执行菜单文件·····	(217)	12.1.2	报表预览·····	(244)
10.4	菜单文件的维护·····	(218)	12.1.3	打开数据库和浏览表格····· ·····	(244)
10.5	工具栏应用·····	(218)	12.1.4	使用项目信息对话框·····	(244)
10.5.1	创建工具栏·····	(218)	12.1.5	折叠和展开项目管理器····· ·····	(247)
10.5.2	工具栏应用于表单·····	(220)	12.2	Visual FoxPro 的 .app 文件和 .exe文件的建立·····	(248)
习题 10	·····	(221)	12.2.1	项目连编·····	(248)
11	程序、过程设计·····	(221)	12.2.2	设置主文件·····	(249)
11.1	程序设计基础·····	(221)	12.2.3	加入项目文件·····	(252)
11.1.1	建立与编辑程序文件·····	(222)	12.2.4	引用可修改的文件·····	(253)
11.1.2	利用设计工具建立程序代码 ·····	(222)	12.2.5	从命令窗口建立项目·····	(253)
11.1.3	程序文件的调用·····	(222)	12.2.6	定制应用程序的运行状态····· ·····	(256)
11.2	顺序程序设计·····	(225)	习题 12	·····	(258)
11.3	分支程序设计·····	(225)	附录	·····	(259)
11.3.1	IF···ENDIF 结构·····	(227)	附录 A	函数表·····	(259)
11.3.2	DO···CASE 结构·····	(230)	附录 B	文件类型·····	(273)
11.3.3	如何使用 IF 语句和 DO CASE 语句·····	(230)	附录 C	控件和对象表·····	(274)
11.4	循环程序设计·····	(231)	附录 D	事件表·····	(276)
11.4.1	使用 DO WHILE···ENDDO 结构 ·····	(232)	附录 E	方法表·····	(279)
11.4.2	使用 FOR···ENDFOR 结构····· ·····	(233)	附录 F	属性表·····	(282)
11.4.3	使用 SCAN···ENDSCAN 结构····· ·····	(234)	附录 G	环境配置·····	(293)
11.5	执行外部程序·····	(234)	附录 H	系统变量·····	(296)
11.5.1	子程序格式·····	(235)	附录 I	系统常用命令·····	(299)
			参考文献	·····	(310)

1 概 述

数据库概念的提出至今不到 50 年的历程,但是已经形成了坚实的理论基础和独特的数据技术,其内涵不断深入,应用广泛,发展飞速。数据库技术推动了现代管理技术的进步,成为现代化管理的基石,深刻地影响着人们的生活、工作,改变着人们的思想观念。管理现代化的需求给数据库技术提供了广泛的应用途径,也给数据库技术提出了新的课题。数据库技术、管理技术、信息系统相互渗透,促进了企业信息化、全球数字化、资源一体化的发展。

1.1 数据库技术的发展

在数据处理中,通常计算比较简单,但对数据管理的要求较高,包括数据的收集、整理、组织、存储、维护、检索、统计、传输等一系列的工作。利用计算机对数据进行处理,一般来说分为如下 5 个基本环节:

- (1) 原始数据的收集。
- (2) 数据的规范化及其编码。
- (3) 数据输入。
- (4) 数据处理。
- (5) 数据输出。

随着数据库技术的发展,以及数据处理量的增长,计算机数据库管理技术也在不断地发展。根据提供的数据独立性、数据共享性、数据完整性、数据存取方式等水平的高低,数据处理技术的发展可以划分为 3 个阶段:人工管理阶段、文件管理阶段以及数据库管理阶段。

1.1.1 人工管理阶段

20 世纪以前,数据处理都是通过手工进行的,这时的计算机主要用于科学计算,既没有专用的软件,也没有大容量的存储设备。在这种环境下,数据处理的数量少、精度低、速度慢,即使部分数据可通过计算机处理,但应用程序与数据之间的依赖性太强,数据不能从程序中独立出来,数据组之间可能造成许多重复和数据冗余。

1.1.2 文件管理阶段

20 世纪 50 年代后期至 60 年代中期为文件管理阶段,这时计算机处理的数据不仅是数值,而且也包含文字,数据与应用程序分别组织存储,由专用程序实现程序与数据之间的统一接口。解决了应用程序与数据之间的一个公共接口问题,使得应用程序采用统一的存取方法操作数据,但只能简单地存放数据,文件之间没有有机的联系,数据的存放依赖于应用程序的使用方法,不同的应用程序仍然很难共

享一个数据文件,数据的冗余性较大。

1.1.3 数据库管理阶段

20世纪60年代后期,进入数据库管理阶段,实现了有组织地、动态地存储大量关联数据,方便了多用户访问,可以使数据充分地共享,同时与应用程序高度独立。

数据库技术的形成过程中,主要有以下3件事奠定了数据库技术的基本理论和方法。

(1) 1969年美国IBM公司的数据管理系统IMS(Information Management System)研制成功,并得到成功应用。

(2) 同年,美国CODASYL委员会公布了他们的成果DBTG(Data Base Task Group)报告。

(3) 从1970年起,IBM公司的高级研究员E. F. Codd连续发表了一系列论文,奠定了关系型数据库理论的基础。

1978年,美国ANSI/X3/SPARC(美国计算机与信息处理国家标准化委员会规划与需求委员会)DBMS(DataBase Management System)研究组的研究报告“ANSI/X3/SPARC DBMS Framework of the Study Group on DBMS”(简称SPARC报告)发表,它标志着数据库技术进入了成熟阶段。SPARC报告提出了一个标准化的数据库系统模型,对数据库系统的总体结构、特征、各个组成部分以及相应接口,都作了较明确的规定。迄今为止,它仍然是数据库技术影响最大的重要文件。SPARC报告中的一个重要思想是采用内部级、概念级、外部级3级结构,在最大程度上实现了数据独立性,提出了脱离具体数据库管理系统的数据库模型的概念,使用数据字典(目录)以支持系统内部的数据管理,以及映像功能的一些明确规定等,这在DBMS软件开发中起到了关键性的作用。

20世纪90年代以后,数据库技术有了突飞猛进的发展,数据库结构从单一模式发展成为多种模式复合的异构体数据库,提出了数据库仓库、数据挖掘等新的概念,其应用领域不断扩展。

1.2 数据库基础知识

1.2.1 数据与信息

1) 数据(数据项或字段)

数据是描述事物特征的特定的符号(也称为数据项或字段),是数据库组织和数据处理中最基本的单元。它不仅是我们日常工作中所熟悉的数字,而且还包含了我们在描述事物过程中经常采用的文字、图像、图形、声音等形式,这些也属于数据库中的数据范畴。数据是人们传达思想、进行信息交流的载体。

通过数据将事物的信息及时、正确、全面地描述或记录下来是数据处理过程中的关键。描述一个事物往往涉及许多方面的概念和理论。

描述事物的用途直接影响着事物属性的定义。对于不同的用户,需求不同,侧重点不同,需要的信息自然不同,因此,在描述事物时使用的属性也不相同。例如,我们把人作为描述的对象,首先要确定人的用途是什么,若描述学生,则将会涉及学生的学习情况、政治思想表现、工作能力和学习能力等方面的基本内容;若描述一个职工,则将会涉及职工的工资、保险、工作能力和技术等级等。当我们要描述一个学生的学习情况时,针对的是一个较具体的数据处理需求,可以通过学生的姓名、学号、专业、年级、开课时间、课程名称、考试类别和考试成绩等属性反映每一位学生的每一门课程的学习成果,这些属性通过采用人们日常交流中所约定的符号而被转换成数据。

数据来源于我们的日常工作和生活,存储于各种媒体中,经过加工、传送,为人们的工作、生活服务。

2) 数据的描述

数据的描述是从客观事物出发,经过概念、规则或逻辑推理转换成数据,这一过程经历了3个领域:现实世界、概念世界和数据世界。

(1) 现实世界。现实世界是存在于人们头脑之外的客观世界,是不以人的意志为转移的客观实体。事物可分成“对象”与“性质”两大类,又可分为“特殊事物”与“共同事物”两个重要级别。

(2) 概念世界。概念世界是通过人们对现实世界中事物的规范、约定在人们头脑中的反映;由一切定义、定理、规则等组成,也称为逻辑世界。事物往往通过各种属性来表达其概念。

(3) 数据世界。数据世界是概念世界中信息的数据化,现实世界中的事物及联系在这里用数据模型描述。

3) 信息

信息至今还没有一个精确的定义,通俗的意义是有用的消息,是客观世界的正确反映,是对数据的解释。例如:“50”是一个数据,“50元人民币”就成了信息。信息通常被定义为:经过加工后的数据,对接收者存在着直接或潜在的价值。这反映了信息的事实性、价值性、可传播性和等级性。由于信息自身的特性,目前无法全面地定义信息。

信息与数据有着密切的联系,也存在着明显的差异。数据是信息的载体,信息是对数据的解释。在信息子系统之间通信时,上级子系统的输出对上级子系统而言是信息;当经过通信设备到达下级子系统时,下级子系统接受的是数据,如图1.2.1所示。数据与信息的关系如表1.2.1所示。

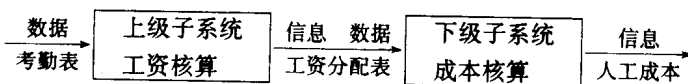


图 1.2.1 数据与信息的关系

表 1.2.1 数据与信息区别

项 目	区别 1	区别 2
数据	未加工的原始材料	用以装载信息的物理符号、信息的载体
信息	加工了的数据	数据的解释

1.2.2 记录与数据表

1) 记录

记录(Record)是按一定次序排列起来的数据项的集合,一个记录(数据表中的一行数据)由若干个数据项(数据表中的列)组成,引用一个记录实际就是引用它所包含的一组数据项。记录的作用是为了某种用途来描述实体属性。例如,一个人的基本档案资料、一个学生的成绩、一个客户档案等都可以是一个记录。实体集合的总描述,称为记录型,即记录型是相应于实体型的。记录集合的描述是由各个数据项的描述组成,即记录型数据项的命名集合。记录型又叫记录格式。

2) 数据表

数据表(Data Table)是按一定形式组织起来的相关记录的命名集合(特定的二维表),即描述同质总体的数据集合,它相应于实集,是存放在存储器上的一组记录,一般地,一个数据表所含有的记录(数据行)内容可以有一个或多个记录,一个数据表可以与相关的其他文件联系。每一个数据表都有自己的一个名称。

3) 关键字

关键字(Key)是用来识别记录的一个或一组数据项,它是识别记录和数据表中查找的标志,也是文件的基本数据。例如,每个学生具有惟一的学号,因此,学号可以作为描述学生的关键字。

关键字可以由数据表中的一个或多个数据项组成的,用于惟一标识数据表中的记录。在数据处理过程中关键字可以分成:

超关键字 在数据表中能惟一确定记录的一列或多列的数据组称为“超关键字(Super Key)”。显然,数据表中至少有一个超关键字(全体数据项构成)。超关键字虽然能惟一地标识记录,但是,在超关键字中可能会有多余的数据项。希望用最少的数据项来惟一地标识记录。如果用一列数据项构成关键字,则称为“单一关键字(Single Key)”;如果用两列或两列以上的数据项构成关键字,则称为“合成关键字(Composite Key)”。

候选关键字 如果从一个超关键字中去掉其中的一列数据项后不能惟一地标识数据表中的记录,则这样的关键字称为“候选关键字(Candidate Key)”。候选关键字是超关键字中数据项最少的关键字。一个数据表中可能存在多个候选关键字,但至少有一个候选关键字。在数据处理过程中,创建数据库的数据表时,候选关键字可以转成主关键字。

主关键字 从数据表的候选关键字中,挑选出其中的一个作为主关键字

(Primary Key)。由于关键字必须惟一地标识记录,因此当数据表的数据为一列或多列组成后,则该列或组成关键字的指定列在数据表中是不能有重复的记录的。必须小心使用空值,否则将无法输入数据。

外部关键字 当一个数据表(表 A)的主关键字被包含在另一个数据表(表 B)中时,表 A 中主关键字被称为表 B 的外部关键字(Foreign Key)。例如,学生表中,“学号”是主关键字,而在成绩表中,“学号”则是外部关键字。

1.2.3 数据库

1) 数据库的含义

数据库是有用数据的有序集合,从形式上看数据库是数据的仓库,对数据进行组织、存储和管理。

数据库中数据组织的最小单元是数据项,这是独立的不可分割的处理单元,描述了事物的某一属性。描述一个事物某一用途的全部属性的数据集称为记录,如一个学生一门课程或一个学期的学习成绩等,即记录是由数据项组成;描述一类相同属性事物的记录集称为数据文件,如学生成绩表,在数据文件中数据存取的最小单位是记录,数据文件是由记录组成的,在数据库组织中数据文件被形象地称为表;数据库是由数据文件组成的,在数据库中记载了各张表的特征和表间的联系。一个数据库可以记录、加工、传递知识,网络教育、电子商务、信息系统等无不建立在数据库的基础之上。

2) 数据库的特性

数据库技术自产生以来形成了较完善的理论体系和强大的数据处理功能,数据库具有如下特点。

(1) 数据共享性。一个数据库可以供多种不同的用户使用,如学生成绩数据表,可以供学生、学校和用人单位等多种用户使用。在数据库里,数据与程序独立,提高了数据的使用价值,同时简化了程序设计,提高了程序的灵活性,方便了用户的操作。

(2) 数据的一体化和结构化。数据库按某种模型组织、存储和处理数据,不仅使内部数据之间彼此相关,而且文件之间在结构上也有有机地联系在一起,整个数据库形成一个整体,即数据库的一体化,这样使数据库具有较大的适应性,易于维护与扩充,应用数据灵活方便。

(3) 较少的冗余度。数据库的数据组织是从描述事物的整体出发,数据的冗余大大减少,在数据文件中除了作为表间联系的关键字和为了数据安全、可靠所采取的备份副本之外,存储的数据冗余度保持在尽可能小的程度。

(4) 数据独立性好。数据库系统提供了数据的映射功能,当需要改变存储结构时,逻辑结构可以不变,从而避免了不必要的程序修改工作。

(5) 对数据进行集中统一的控制。系统提供统一的数据定义、预处理、查询以及维护等手段,并统一控制数据的安全性、完整性、保密性和并发性,使得对数据的应用更加有效和可靠。

1.2.4 数据库技术新概念

数据仓库(Data Warehouse) 20世纪80年代中期,“数据仓库之父”William H. Inmon先生在其《建立数据仓库》一书中定义了数据仓库的概念,随后又给出了更为精确的定义:数据仓库是在企业管理和决策中面向主题的、集成的、与时间相关的、不可修改的数据集合。与其他数据库应用不同的是,数据仓库更像一种过程,一种对分布在企业内部各处的业务数据的整合、加工和分析的过程,而不是一种可以购买的产品。

数据集市(Data Mart) 或者叫做“小数据仓库”。如果说数据仓库是建立在企业级的数据模型之上的话,那么数据集市就是企业级数据仓库的一个子集,它主要面向部门级业务,并且只面向某个特定的主题。数据集市可以在一定程度上缓解访问数据仓库的瓶颈。

数据挖掘(Data Mining) 是一种决策支持过程,它主要基于AI、机器学习、统计学等技术,高度自动化地分析企业原有的数据,做出归纳性的推理,从中挖掘出潜在的模式,预测客户的行为,帮助企业的决策者调整市场策略,减少风险,做出正确的决策。

数据抽取(Extract)、转换(Transform)、清洗(Cleaning)、装载(Load) 是构建数据仓库的重要一环,用户从数据源抽取所需的数据,经过数据清洗,最终按照预先定义好的数据仓库模型,将数据加载到数据仓库中去。

1.3 数据库系统的研究与应用领域

数据库技术的研究与应用范围十分广泛,而且正在不断地高速向前发展。

1.3.1 数据库技术的研究范围

数据库技术的研究范围从总体上可分成3个主要领域。

1) 数据库管理系统软件研制

数据库管理系统(DBMS)是数据库系统的基础,研制DBMS的基本目标是扩大功能、提高性能、增强可用性。研制以DBMS为核心的一组相互联系的软件系统已经成为当前数据库软件产品的方向,这些在DBMS基础上运行的软件系统有:数据通信软件、表格软件、数据字典、报表书写、图形系统等。

由于数据库应用领域的不断扩大,数据库不仅广泛应用于管理,而且已经应用到工程设计、图形图像处理 and 声音多媒体介质的复杂数据处理、控制和计算机辅助设计等领域。数据库应用领域的不断扩大,给DBMS软件的研制带来了新的课题,这不仅涉及应用系统的设计方法,而且涉及数据库系统模型,以及实现技术等多种新的问题。出现了面向对象的数据库、多媒体数据库、工程数据库、语音数据库、图形图像数据库、数据仓库等研究方向,这些新的研究方向都来源于用户的需求、信息化的发展和数据库技术的应用。

2) 数据库设计

在数据库管理系统的支持下,按照企业的实际需要,设计出一个结构良好、使用方便、数据处理效率高的数据库及其应用系统,是数据库设计的主要任务。在这个领域内,主要研究内容是数据库设计方法学,数据模型方法的研究,数据库设计的计算机辅助方法的研究,数据库设计规范和研究,以及数据设计工具的开发等。

3) 数据库理论

数据库理论研究的主要内容集中于关系规范化理论、关系数据库理论、数据库与人工智能结合的专家库、数据库与逻辑、逻辑演绎推理结合的知识库等方面。此外,演绎数据库、面向对象的数据库、知识数据库系统的研究也都属于这一研究方向。例如,IBM公司的E.F.Codd研究的关系数据模型等。

计算机与网络技术的飞速发展,使得应用领域不断增加,信息技术的应用也逐渐从信息查询、数据传送走向网络环境中的海量数据存储、数据挖掘和决策信息支持,因此数据存储的可靠性和可管理性,正日益成为企业信息化进程中的一个关键环节。目前,市场已经不再满足于现有的分散式存储模式,一些有远见的企业用户开始考虑建立集中数据管理模式。因此,现今存储技术与数据挖掘技术的发展已成为全球各商家及企业最为关注的信息技术之一。

在全球经济一体化的驱动下,企业业务将越来越多地发生在世界各地不同的地点、不同的硬件平台和不同的数据库中,企业大量的数据信息不能够实现真正的共享,使得工作位置变得更加复杂,使企业获取有效信息更加困难。因此,基于网络数据库的开发与利用的网络存储,以及网络数据挖掘技术也就有了巨大的市场潜力。技术的不断进步使高伸缩性的基础设施应运而生,网络存储将是一个发展趋势,存储虚拟化、独立于服务器的存储以及模块化存储的技术和智能化的数据挖掘技术,使企业、用户能够更加方便地使用并有效地管理数据。

数据挖掘技术在国内外不断创新与发展,涉及存储设备与软件开发、系统集成、存储服务供应、解决方案提供,还包括金融、证券、电信、保险、邮政、交通、航空、石化、能源等众多行业建立起相互交流的平台。

1.3.2 数据库技术应用领域

信息系统的建设和应用是建立在数据库基础上,信息技术的发展促进了数据库技术的向前发展。这相当于计算机科学与电子技术的内在联系。信息技术日新月异的发展和广泛的应用领域,事实上是以数据库技术为依托。两者必须同步发展,相互依赖。信息技术的发展给数据库技术提出了新的研究课题,同时数据库技术在信息系统中的应用为数据库技术的研究提供了充足的资金来源。

1) 在 MIS 中的应用

管理信息系统(MIS)的研制,几乎全部是数据库系统功能的应用。在管理信息系统中无论是事务处理、综合业务处理,还是决策支持系统,都体现了数据库的作用。信息是从用户需求的角度来描述事物,数据是从计算机处理的角度来描述、传

输、保存、检索、再现事物的属性。

2) 在 CIMS 中的应用

计算机集成制造系统(CIMS),除了需要传统的 DBMS 功能外,还要求一些新的功能,例如:

- 面向工程环境的数据类型,要求能对这些数据进行数据定义和数据操纵,使其设计者可以定义新的数据类型,修改和重新定义自己的数据结构等。

- 能存储和有效地检索图形、工程数据。
- 能管理设计过程中对象的演变历史。
- 具有处理复杂数据对象间语义完整性、一致性约束能力。
- 具有处理长事务的安全性、可恢复性的能力。

3) 在 CASE 中的应用

计算机辅助软件工程(CASE)涉及软件开发的全过程,在使用 CASE 工具进行软件开发的每个过程,都需要数据库的支持。例如:一切与开发有关的活动(分析过程、方案设计等等)都应由 CASE 数据库来支持。

CASE 环境下包括大型程序和文档的版本管理能力,而且这些是传统的数据库所不能提供的。

4) 在 OAS 中的应用

办公自动化系统(OAS)所处理的数据具有多样性,既有图形、文字、声音等多媒体数据,还有智能电话、传真、电视电话等信号数据和网络控制数据,如纸张、录音磁带、录像磁带等,所以要求数据库能处理多介质信息。

5) 在数据挖掘中的应用

数据挖掘是对已经存在的数据资源通过各种模型、工具,将来自不同地理位置、不同数据库的数据进行分类、汇总统计、模拟、评估、预测,从数据分析中获取有用的决策信息。数据库为数据挖掘提供了最原始、最基本的资料。数据挖掘技术的应用主要有如下几方面:

- 数据挖掘技术的行业应用。
- 人工神经网络数据库技术。
- 数据挖掘与企业 CRM。
- 数据挖掘与银行信誉制。
- 数据挖掘在电子政务中的应用。
- 数据挖掘过程的信息安全保护。

1.4 数据库管理系统

数据库管理系统(DataBase Management System,简称为 DBMS)是指数据库系统中对数据进行管理的软件系统,它是数据库系统的核心部分。数据库系统的一切操作,包括查询、维护以及各种控制都是通过 DBMS 进行的。

DBMS 是基于某种数据模型上的,因此,可以把它看成是某种数据模型在计算