



高等学校教材

计算机 信息管理基础

董丕明 宋旭东 倪承普 编著



中国铁道出版社

高等学校教材

计算机信息管理基础

(关系数据库 ACCESS)

董丕明 宋旭东 倪承普 编著

中国铁道出版社

1999年·北京

(京)新登字 063 号

内 容 简 介

本书是依据《工科非计算机基础教学指南》中计算机信息管理基础和《普通高等学校计算机教育基本要求(本科)》中数据库的要求编写的。其内容包括两大部分。第一部分(第一至第三章)概述介绍了数据库理论及 SQL 语言;第二部分(第四至第十一章)较详细地介绍了 Office 97 中文版的 Access 部分。

本书为高等院校教材,也可作为短训班教材。

图书在版编目(CIP)数据

计算机信息管理基础/董丕明编著. —北京:中国铁道出版社,1999 高等学校教材
ISBN 7-113-03430-6

I. 计… II. 董… III. 电子计算机-信息管理-高等学校-教材 IV. TP3

中国版本图书馆 CIP 数据核字(1999)第 30683 号

书 名:计算机信息管理基础

作 者:董丕明 宋旭东 倪承普

出版发行:中国铁道出版社(100054,北京市宣武区右安门西街 8 号)

责任编辑:郭 宇

封面设计:薛小卉

印 刷:中国铁道出版社印刷厂

开 本:787×1092 1/18 印张:16.75 字数:423 千

版 本:1999 年 8 月第 1 版 1999 年 8 月第 1 次印刷

印 数:1~4000 册

书 号:ISBN7-113-03430-6/TP·384

定 价:21.60 元

版权所有 盗印必究

凡购买铁道版的图书,如有缺页、倒页、脱页者,请与本社发行部调换。

前 言

本书系“面向 21 世纪铁路高等教育教学内容和课程体系改革计划项目”——“非计算机专业计算机系列课程教学内容和课程设置改革的研究与实践”课题组计划编写的系列教材之一。根据 1997 年 10 月大连会议和 1998 年 5 月上海会议精神,本书突出了面向应用、重视实践、通俗易懂的原则。在课题组的领导下,依据国家教委教高司[1997]155 号文件和教育部教高司[1998]74 号文件,关于《工科非计算机基础教学指南》中计算机信息管理基础和《普通高等学校计算机基础教育教学基本要求(本科)》中数据库的要求,以及课题组的研究成果编写了此教材。

本书由两大部分组成:第一至第三章概述数据库基础理论及 SQL 语言。第四至十一章详述 Office97 中文版的 Access 部分。按课题组的要求,与之配套的“上机指导和习题”正在编辑中,将陆续出版。

Access 97 是一种关系型数据库管理系统,运行于 Windows 95 工作平台,在国内外已非常流行。其功能强大、使用方便、且易于扩充,是中小型数据库应用系统首选的开发平台。它还是一种多媒体数据库管理系统,具有处理多种媒体的数据信息的功能。

Access 97 具有全面的数据共享、信息交流等功能,支持 OLE、DDE 技术,具有 ODBC 特性,能够导入或导出 dBASE、Paradox、FoxPro、Lotus1-2-3、SQL 服务器、Oracle 和 Btrieve 等流行的以字符为基础的数据库格式的文件,同时也能链接到这些数据库及其数据上。Windows 中的所有优点都反映于 Access 中,用户可在任意 Windows 应用程序和 Access 之间相互剪切、复制和粘贴。在 Windows 95 中,如果用户使用 Microsoft Office 产品(Excel、Word、PowerPoint),则能通过集成这些软件把 Access 变成真正的数据库操作环境。

为了方便用户使用 Access 97,它提供了大量“向导”及多种“生成器”,通过使用这些“向导”和“生成器”,可以快速开发出功能强大、性能完善的数据库应用系统,而不需具有专业的数据库应用系统开发知识。

Access 97 提供了对 Internet 和 Intranet 的支持。还提供了与 Visual Basic 的接口,使得较为复杂的过程可以通过 Visual Basic 语言来处理,这更加强了 Access 的处理功能。

本书的第一、二、三章由倪承普编写,第四、五、六、七章由董丕明编写,第八、九、十、十一章由宋旭东编写,全书由董丕明统稿,段凡丁和苏斌主审。

感谢铁道部科教司、上海铁道大学陶树平先生、西南交通大学景红、北方交通大学

王移芝、长沙铁道学院刘卫国、大连铁道学院领导及同行们的关怀和支持。

因时间仓促,水平有限,错误及遗漏之处在所难免,敬请读者不吝赐教。

作者
1999年3月

目 录

第一章 概 论	1
1.1 数据的概念	1
1.2 信息的概念	1
1.3 信息组织与数据库系统	2
第二章 基本概念	8
2.1 常用术语的说明	8
2.2 数据模型	8
2.3 关系模型	13
2.4 数据库系统结构	21
2.5 关系数据库	34
第三章 结构化查询语言——SQL	41
3.1 SQL 查询	41
3.2 SQL 数据定义	50
3.3 SQL 数据操纵	54
3.4 SQL 数据控制	57
3.5 数据字典	58
第四章 Access 数据库综述	60
4.1 Access 的基本概念	60
4.2 Access 的运行环境与启动	61
4.3 浏览 Access 窗口	64
4.4 Access 的帮助系统	65
第五章 Access 数据库基本操作	68
5.1 设计 Access 数据库	68
5.2 创建 Access 数据库	70
5.3 操作 Access 数据库	79
5.4 维护 Access 数据库	84
第六章 表	88
6.1 建立数据库表	88
6.2 在数据库表中输入和编辑数据	101

6.3	记录的索引查找排序和筛选	106
6.4	建立表间的关系	109
6.5	打印表数据	113
第七章	查 询	115
7.1	查询的基本概念	115
7.2	使用查询向导建立查询	117
7.3	查询设计视图和选择查询	130
7.4	操作查询	135
7.5	SQL 特定查询	140
第八章	窗 体	145
8.1	基本概念	145
8.2	使用向导创建窗体	150
8.3	在设计视图中创建窗体	159
8.4	利用窗体操作数据	175
8.5	窗体的基本操作	175
第九章	报 表	178
9.1	基本概念	178
9.2	使用向导创建报表	182
9.3	在设计视图中创建报表	190
9.4	报表中操作数据	195
9.5	报表的基本操作	199
第十章	宏	202
10.1	基本概念	202
10.2	创建宏	223
10.3	调试宏	228
10.4	执行宏	230
10.5	宏的应用	231
第十一章	模 块	235
11.1	Access 模块及模块基础	235
11.2	VBA 程序设计简介	243
主要参考文献		262

第一章 概 论

当今我们生活在一个飞速发展的信息时代。各行各业的工作都要求人们提高处理信息的能力。这种能力实际上不仅限于使用计算机的能力,还包括有效地运用各种通信交流途径,通过媒体收集和发送有用的信息,并运用信息处理工具从大量信息中挑选出有价值的材料,设法把它们改造成适合我们使用的方式。无论一个人现在或将来从事任何职业,他都会越来越强烈地感受到信息社会对人们的能力所提出的挑战,同时也会觉察到计算机所发挥的巨大作用,计算机正在越来越深入到我们日常生活的各个角落。可以说计算机的踪迹无处不在,无时不有,无法离开。

人类的一切活动都离不开数据,离不开信息。但是在不同的领域里,信息的含义有所不同。一般认为信息(information)是数据、消息中所包含的意义。数据和信息有时可以混用,例如,数据处理也称为信息处理;有时必须分清,例如,不能把信息系统称为数据系统。

1.1 数据的概念

所谓数据,通常指用符号记录下来的可加以鉴别的信息。数据的概念包括两个方面:其一,数据内容是事物特性的反映或描述;其二,数据是符号的集合。

由于记录和描述事物的特性必须借助一定的符号,这些符号就是数据形式。例如,本书的写作日期是“1998年7月27日”,当然也可以将以上的形式用“07/27/98”来表示。

数据的概念在数据处理领域中比在科学计算领域中已经大大地拓宽了。所谓用“符号”,不仅仅指数字、字母、文字和其它特殊字符,而且还包括图形、图象、声音等多媒体数据;所谓“记录下来”也不仅是指印在纸上,而且包括记录在磁介质、光介质上、半导体存储器里。

数据在空间上的传递称为通信,在时间上的传递称为存储。

1.2 信息的概念

信息是关于现实世界事物的存在方式或运动形态反映的综合,是人们进行各种活动所需要的知识。数据与信息既有联系又有区别。数据是载荷信息的物理符号或称为

载体。数据能表示信息,但并非任何数据都能表示信息,正如人们常说的“如果计算机输入的是垃圾,输出的也会是垃圾”。同一数据也可能有不同的解释。因此,信息只是人们消化理解了的数据。信息是抽象的,不随数据设备所决定的数据形式而改变;而数据表示方式却具有可选性。

信息是反映客观现实世界的知识,用不同的数据形式可以表示同样的信息。例如,同一条新闻信息可用文字在报纸上刊登,在电台上用声音广播,在电视上用图象放映以及在计算机网络上传播,其信息内容可以相同。计算机只能处理二进制形式的数字信息,客观世界的信息表现形式却是丰富多彩的,在计算机里它们的形式得到了统一。

数据处理是指将数据转换成信息的过程。广义地讲,它包括对数据的收集、存储、加工、分类、检索、传播等一系列活动。狭义地讲,它是指对所输入的数据进行加工整理。其基本目的是从大量的、已知的数据出发,根据事物之间的固有联系和运动规律,通过分析归纳、演绎推导等手段,萃取出对人们有价值、有意义的信息,作为决策的依据。由此可见,信息是一种被加工成特定形式的数据,这种数据形式对于数据接收者来说是有意义的。对数据加工可以相对比较简单也可以相当复杂。简单加工包括组织、编码、分类、排序等;复杂加工可以复杂到使用统计学方法、数学模型等对数据进行深层次的加工。

数据是原料,是输入,而信息是产出,是输出结果。当两个或两个以上数据处理过程前后相继时,前一过程称为预处理。预处理的输出作为二次数据,成为后面处理过程的输入,此时信息和数据的概念就产生了交叉,表现出相对性。如图 1.1 所示。人们有时说“信息处理”,其真正含义应该是为了产生信息而处理数据。

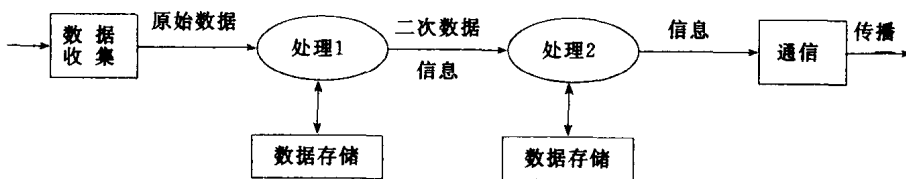


图 1.1 数据与信息的关系

1.3 信息组织与数据库系统

随着时代的发展,人们越来越认识到信息是一种宝贵的资源,而要发展这种资源的作用,使它能够更多的人所共享,有效地存储和管理信息越来越成为一个必须解决的重要问题。如果缺乏有效的管理和维护,数据的价值就好像新鲜水果一样会随着时间的推移而腐烂变坏。为了解决新信息的存储、管理和维护问题,人们进行了许多研究,建立了各种模型和各种实际的信息管理系统。首先人们考虑的是如何把文件内容组织好,以便能够较好地存储各种复杂数据。随着信息管理研究和应用的深入,人们逐渐开发出一

套通用的数据管理模式,建立起各种复杂的数据库,开发了为建立数据库所必须的数据库管理系统,并进一步发展了许多相关技术,开发了许多实际的软件。让我们首先对普通文件中数据的组织做一个基本的介绍。

1.3.1 文件中的数据组织和操作

最简单的文件结构是顺序文件结构,普通的字符序列文件和执行的程序文件都具有顺序文件结构。这种结构的特点是简单,例如文本文件可以由顺序存放和读取的一个个字符构成,可执行程序文件里存放的是一串机器代码。对这些文件,其内容的顺序存放很大程度上决定了它的使用方式。

顺序文件结构是最基本的文件结构,是其他各种复杂文件结构的基础。顺序文件的基本使用方式有两种,它们是:

第一种方式称为顺序方式。按这种方式使用文件无须借助其他结构,对文件内容的访问按照记录在文件中存放的次序,一个记录一个记录地顺序进行。如果在实际应用中对文件内容的处理总以同样顺序进行,这种方式最简单而且有效。该方式的缺点也很明显,如果需要查找文件中某一个特定的记录,那么没有其他办法,只能从头开始顺序地找。当实际应用要求快速查找特定记录时(这是文件内容的另一种使用方式,称为随机访问方式),简单的顺序文件往往无法满足需要。

要提高对记录进行随机访问的速度,可以为顺序文件建立另外的结构,为文件里的记录建立一种索引(index)。这种情况就像书籍的最前面有章节目录,通过这种目录可以找到具体章节页码;许多书籍的最后有索引,通过索引可以找到某个词出现的位置。文件的索引本身就是一个表格,其中每个表目包括一个记录的关键码和该记录的存储位置信息。有了索引,在需要访问一个记录时首先查找索引,通过索引项找到记录的存放位置,这样随机访问文件记录的速度大大加快了。索引表本身通常也存放在单独的磁盘文件里,这样的文件称为索引文件,文件中的索引项同样顺序排列。对于建立了索引的文件,当文件的内容增加、减少或改变位置时,都需要相应更新索引文件中的信息,保证它们的一致性,这样才能确保索引的有效性。

对于这种由记录作为基本单元构成的文件,需要做的操作主要是增加、修改和删除记录,更新记录内容等。这些操作统称为文件更新(update)或文件维护(maintain)。从上面的讨论可以看出,文件更新是一种相当复杂的工作。当系统管理的数据非常多,数据之间的关系复杂时,在文件更新过程中保持数据的正确性、一致性,保证更新的效率就更加困难。对于数据管理中这些问题的研究和实践的产物就是计算机科学技术中重要的部分——数据库和数据库管理系统。

1.3.2 数据库系统和数据库管理系统

数据库管理系统软件是为建立、使用和维护各种应用数据库而设立的通用软件。为

了理解它的特点,让我们先说明数据库的基本结构。数据库从逻辑上说其目的就是为了对大量数据进行严格有效的管理,同时又要提供灵活方便的使用方式。例如:数据库就好像一种文件柜,具有一层层由粗到细的框架结构。其结构如下:

档案柜→抽屉(放资料卡)→

卡片(多项栏目资料组成)→卡片上的各项栏目资料

其中箭头右边的项是左边项的组成成员。数据库的结构可以相应地对照叙述为:

数据库(database)→关系表(relation tables)

→数据记录(dataRecords)→数据项(data items)

数据库由一个或多个“关系表”组成;一个关系表是多个“数据记录”的集合。例如一个学生班级的成绩记录表是一个关系表,每个学生的成绩记录(按照一定的数据格式)是关系表中的一行,而这些成绩记录的全部构成该班级的成绩记录表,即该关系表的内容。每个学生成绩记录称为一个数据记录,它由若干具有明确格式的数据项组成,例如一个学生的某门课的成绩(是一个数据项)取值应是在 0~100 之间的一个整数。讨论数据库的逻辑结构,主要就是指这些“关系表”的组成结构,它是建立在刚才介绍的文件系统基础之上的。

数据库(DataBase,通常简记为 DB)是为了满足一定范围里许多用户的需要,在计算机里建立的一组相互关联的数据集合。与普通基于文件的信息存储方式相比,数据库系统的主要特点包括:采用一个称为“数据库管理系统”的系统软件,管理和维护数据库里的数据文件,对数据的存储、更新、检索(查找)操作采用统一的处理和控制方式;增强对数据库中数据的完整性和一致性的保证(统一管理各种数据、控制其使用方式、并对数据更新的方式和过程加以控制);能够尽可能地减少数据冗余,在数据库中尽量消除信息的重复存储(减少数据的重复拷贝,争取最小数据冗余量);尽量使存储的数据能够同时为多个应用程序和用户服务(对数据共享的控制、管理和服务)。以学校的信息管理为例,学校里的许多部门都需要使用有关学生的基本信息,如学校学籍管理部门、教务部门、各个系、宿舍管理部门、学生会和共青团组织等。如果各部门不考虑信息的共享问题,而是在各自的计算机里分别存储有关各个学生的信息,那么其中的许多信息必然要有大量重复存储,这样不仅浪费存储资源(要消费大量外存储器的空间),更重要的是数据的准确性和统一性根本无法保证,一个部门对于某项信息的改动无法反映到其他部门的计算机里。使用数据库系统就可能解决这些问题,这时可以设法保证一个学生各方面的信息只有一份,提供给学校各个部门共享,而数据的管理工作由系统统一处理。这样做,自然能够更好地满足实际工作的需要。

在实际工作中,不同的人对各种数据的使用权限不应该是一样的。例如,教师有权提供学生考试的成绩;教务部门保存、管理这些数据,但他们没有权力对成绩数据进行修改;学生可以查阅自己的有关信息,但不能增添或修改自己的记录;等等。在数据库系统中可以根据需要为各种数据项设置不同级别的安全访问限制,以满足人们在建立数

据库安全方面的需要。每个或每类用户只能在限定范围中使用系统里的数据。

1.3.3 数据库管理系统

数据库管理系统(DataBase Management Systems,通常简记为 DBMS)是支持建立和管理数据库的系统软件,它负责对数据库做严格的管理和控制,以保证数据库安全和数据库内数据的完整。用户通过 DBMS 的支持,访问数据库中的数据;数据库的管理员也通过 DBMS 的支持进行数据库的维护工作。DBMS 是一类非常复杂的软件系统,其主要功能包括:定义数据库结构和所存储的格式(这些构成数据库的“数据字典”),规定数据在外存储器的存储安排方式,负责各种与数据有关的控制和管理任务,包括在数据库系统运行过程中维护数据库数据的一致性、完整性,保证数据安全性等。数据库管理系统还提供支持数据库建立和维护的各种工具,例如对数据库的数据装入、数据库的重新组织、数据库的恢复以及数据库的工作性能监测等。DBMS 还要为数据库的工作提供下列支持:

(1)DBMS 提供称为查询语言的数据库使用接口。用户可以直接给出用查询语言描述的数据库使用命令,而 DBMS 软件负责对命令进行解释,执行命令,并把查询的结果提供给用户。

(2)提供对数据存取的多层次接口。为了有效地支持对大量数据的组织和存取,DBMS 通常还提供了与操作系统,特别是与文件系统的数据存取接口(支持在磁盘等外围设备上进行有效的数据读写)。为了改善数据库存取,DBMS 建立了查询索引结构,它的使用可以有效地提高信息查询速度。

(3)提供数据库日常运行管理的功能。如对运行事务(transactions)的支持、日常运行日志监控以及支持数据库管理员对数据库的维护等。

(4)提供数据库的维护功能,包括数据库的恢复机制。数据库中可能存储着非常重要的数据,一旦丢失或破坏造成的影响可能非常严重。设想一个航空公司订票系统,如果数据库中的订票数据记录丢失或破坏而又无法恢复,造成的损失将难以估量。功能强大的数据库管理系统需要有功能全面的数据恢复机制。通常采用的方法是建立有关数据库最新改动情况的记录文件,在其中记录所有对数据的改动操作,及改动前后的有关内容。必要时可以利用这个记录文件进行数据库恢复。

无论哪一种数据库软件,如果我们没有用它来建立数据库,即没有存入数据,就好像我们造了一个柜子,制定了一套周密的管理办法,指定了管理人员,但还没有往柜里藏进什么卡片或簿册那样,是不起什么作用的。因为它还不是一个数据库,而仅仅是一个数据库管理系统。当然,如果没有数据库管理系统,任何数据库都是建立不起来的。这就是说,数据库和数据库管理系统不是一回事,但它们之间有着依存关系。从认识上,我们应当明确它们的区别和联系,而在平时的对话中甚至书本和文章中,把数据库管理信息系统简称数据库,则是常有的事。在本书往后的叙述中,在上下文不会发生歧义的情

况下,为了简便起见,也会把 Microsoft Access 称作 Access 数据库,或者更简单的称之为 Access 或数据库。

1.3.4 数据库查询

数据库查询(query)就是要从数据库所存储的大量记录中找出所需要(符合一定条件的)所有数据记录。“查询语言”(query language)是数据库管理系统提供给用户的、用于描述查询条件的一种语言。人们用查询语言描述所需数据的情况(查询条件),数据库管理系统按查询条件在数据库里查找,取出符合要求的记录,提供给用户。

查询语言的形式多种多样。一种常见的方式称为“用例子查询”(QueryByExample,简记为 QBE)。用这种方式查询时,系统显示一个记录域的表,用户可以给其中的一些域提供一个或一组值,数据库管理系统把库中所有符合要求的记录挑选出来。例如要在学生档案库查出所有家住大连的姓“李”的学生,那么就在记录域表的姓氏一栏里填入“李”,而在居住地一栏填入“大连”。在这里,“姓氏一栏的李和居住地一栏的大连”就是查询条件。目前使用很广泛的一种查询语言是 SQL(Structured Query Language,结构化查询语言),这种语言已经成为一种标准查询语言,在各种数据库管理系统软件中得到成功的应用。

1.3.5 几种微型机上的数据库管理系统

不久以前微型机的硬件性能还比较弱,那时在微机上使用的数据库管理软件功能也比较简单。随着微机功能的日益强大,微机上的数据库管理系统的功能也越来越强大。现在已经有了许多在微型机上可以用的数据库管理系统软件,其中有不少可以在 Windows 系统或其他图形用户界面的操作环境中使用,有的可以在多种操作系统环境里运行。一些数据库管理系统还伴随地提供用于开发应用信息系统的软件开发工具,这些工具的功能也日臻完善。

常见的微机数据库管理系统有在 DOS 环境中工作的 dBASE(较新版本为版本 5)系统、FoxBase 和 FoxPro 等,它们的一些新版本也可以在 Windows 上使用。在 Windows 系统上运行的还有 Microsoft Access 等。目前较大型的、在许多硬件平台和操作系统上,都能用的数据库管理系统主要有:Oracle、Informix 和 SyBase 等,这些系统也大都提供了微机上使用的版本。

1.3.6 数据仓储技术

对于高层管理人员来说,日常业务的大量运行数据常常不能直接供决策使用,数据使用之前需要经过仔细整理和深入分析,才能从中得到真正有用的信息。让我们用一个有趣的例子来说明直接收集到的“数据”和有用“信息”两者之间的关系。美国的超级市场零售业利用“数据仓储技术”从市场销售的大量数据中“采掘”出一条前所未见的“规

律”：“当靠近周末，星期四和星期五两天购买尿布的男人多半是顺便购买一些汽水类饮料”。这究竟是发生于偶然还是有什么更深的原因呢？思索再三，情况原来是这样，孩子的妈妈告诉丈夫，一周所用的尿布已告罄，别忘记回家之前去商场；当男人们购买了尿布之后，已经不便弯腰去选择他们喜爱的其他饮料，只是随手拿了放在高处容易拿到的大瓶可乐饮料一并夹回。这种“规律”性的信息对商场可不是无足轻重的，它和百货架的安放以及促销的策略是有紧密关系的。

概括而言，数据仓储技术的目的是想要把“日常业务的大量运行数据”去粗取精、分析统计，变换为相对稳定且有规律的信息，供决策之用。它能让管理决策者更加聪明。

第二章 基本概念

2.1、常用术语的说明

(1)属性。属性为事物的某一方面特征的抽象描述。如学生这个事物可以通过学生的姓名、学号、性别、年龄和政治面貌等特征来描述,称姓名、学号、性别、年龄和政治面貌为属性。

(2)属性值。属性值为属性的具体取值。如学生王力,其姓名为王力,学号为 80301,性别为男,年龄为 20,政治面貌为共青团员,这些具体值为属性值。

(3)域。域为属性的取值范围,称之为属性的域。如学生的年龄为 16~30 范围内的正整数,其域为(16,30)。

(4)实体。若干个属性的属性值组成的集合,用来表征一个实体。如王力、80301 表示学生王力。同类实体的集合组成了实体集,例如描述全部学生的实体就构成了学生实体集,反映了一个学校全部学生的情况。

(5)关键字。关键字为唯一标识实体的属性或属性组合。例如在学生实体集中,学号可以唯一地标识每个学生实体,所以学号为关键字。在有些实体集中,可以有多个关键字,例如学生实体集,假设学生姓名没有重名,那么属性“姓名”也可以作为关键字。我们通常选定其中一个,被选的那一个关键字称为主关键字,其它的关键字为候选关键字。

(6)次关键字。在实体集中经常选择一些不能唯一标识实体的属性来标识实体,这样的属性叫次关键字。例如“年龄”,“政治面貌”这些属性都可以充当次关键字。

(7)数据项。数据项是实体属性的数据表示。

(8)记录。记录是实体的数据表示。

(9)文件。文件是同类记录的集合,如所有学生的登记表组成一个文件。

(10)数据模型。现实世界中,个体与个体之间总是存在着某些联系。反映到信息世界中是实体与实体之间的联系,由此构成实体模型。反映到计算机世界中是记录与记录之间的联系。实体模型在该领域的数字化表示,就是数据模型。

2.2 数据模型

世间的任何事物都是相互关联的,任何一个实体都不是孤立存在的,因此描述实体

的数据也是相互联系的。通常,这种联系表现在两个方面:一是反映实体内部的联系,反映在数据上便是记录内部各数据项之间的联系,一是实体之间的联系,反映在数据上则是记录之间的联系,这种联系在数据模型中给予描述。

数据系统的一个核心问题,就是研究如何表示和处理实体之间的联系。而我们把表示实体及实体之间联系的数据库的数据结构称为数据模型。换句话说,我们把数据库系统中所包含的所有记录类型,按照它们之间的联系组合在一起,构成一个整体。这个整体的结构就称之为数据库的数据模型,它是数据库中全局逻辑结构的描述。

数据模型是数据库中一个关键的概念,它是实体间联系的一个轮廓视图。数据模型的所有术语都用在模型一级上,整个模型就像一个框架,给它填上具体的数据值就是数据模型的一个实例。

通常,数据模型可以表示成由这样两个集合组成:

$$\text{数据模型 DM} = (\text{R}, \text{L})$$

式中 R 表示记录型的集合;L 表示记录型之间联系的集合。

如果按照记录间联系的表示方式,对数据模型加以分类的话,那么数据模型一般可以分成三种:层次模型、网状模型及关系模型。前两种又称为格式化数据模型。

数据模型的好坏,直接影响数据库的性能,数据模型的选择,是设计数据库一项首要任务。

2.2.1 层次模型

在数据库中,把可以用图(这里用有向图)来表示的数据模型,称为格式化数据模型。而格式化数据模型又按照图的特点分为层次模型和网状模型两种。这两种模型都是以记录型作为图中的结点,而图中的边则用来表示记录型之间的联系。层次模型是以记录型为结点的有向树。在树中,把无双亲的记录称为根记录,其它记录称为从属记录。除根记录外,任何记录只有一个父记录。一个父记录可以有多个子记录。从根记录开始,一直到最下一层的记录为止,所具有的层次称为该数据模型的层次,或称层次模型,如图 2.1 所示。

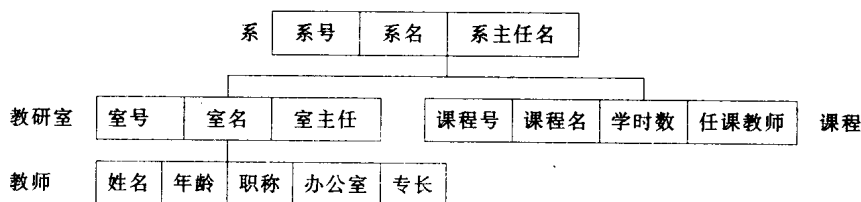


图 2.1 教育层次模型

在层次模型中,每个记录(除根记录外)只有一个双亲,记录之间的联系可以唯一地用双亲表示。在层次模型中,总是双亲记录指向子记录。所以说,记录之间的联系可以

不用命名,只要指出其双亲,就可以找到其子。在层次模型中,从根记录开始,按照双亲—子联系,依次连接的记录序列称为层次路径。在层次模型中,数据是按层次路径存取

的。

目前流行的大型数据库系统中,有许多采用的是层次模型。其中最著名的是 IBM 公司研制的 IMS 系统。除 IMS 系统之外,还有许多采用层次模型的系统,如系统 SYSTEM2000 就是其中的一个。

2.2.2 网状模型

数据库的网状模型,是以记录类型为结点的网状结构。这种结构必须满足如下条件:

- (1)可以有一个以上的结点无双亲。
- (2)至少有一个结点有多于一个的双亲。

图 2.2 所示的是学生的网状模型。

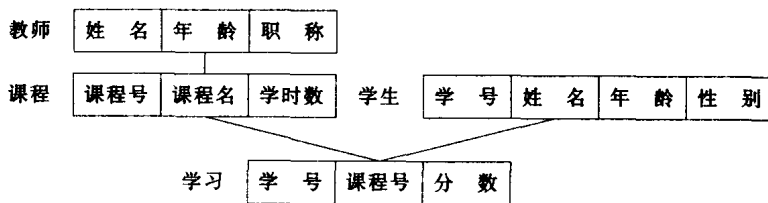


图2.2 学生的网状模型

网状模型和层次模型的差别是:

- (1)一个子结点可以有两个或多个父结点。
- (2)在两个结点之间可以有两种或多种联系。

2.2.3 关系模型

层次模型和网状模型的数据库系统被开发出来之后,在继续开发新型数据库系统的工作中,人们发现层次模型和网状模型缺乏充实的理论基础,难以开展深入的理论研究。于是人们就开始寻求具有较充实的理论基础的数据模型。在这个基础下,IBM 公司的 E. Fcodd 从 1970~1974 年发表了一系列有关关系模型的论文从而奠定了关系数据库的理论基础。

人们常常习惯用表格的形式,表示所关心的现实世界中的信息。例如,表 2.1 就是一个学院各系情况表。

从这个表中,各系的主要信息一目了然。我们将这样一张表,称为一个关系。表中的每一列为关系中的一个属性(数据项),表中的每一行为关系中的一个元组(记录)。实际上关系就相当于一个记录类型,而这个关系的全体元组,就是这个记录的所有值。