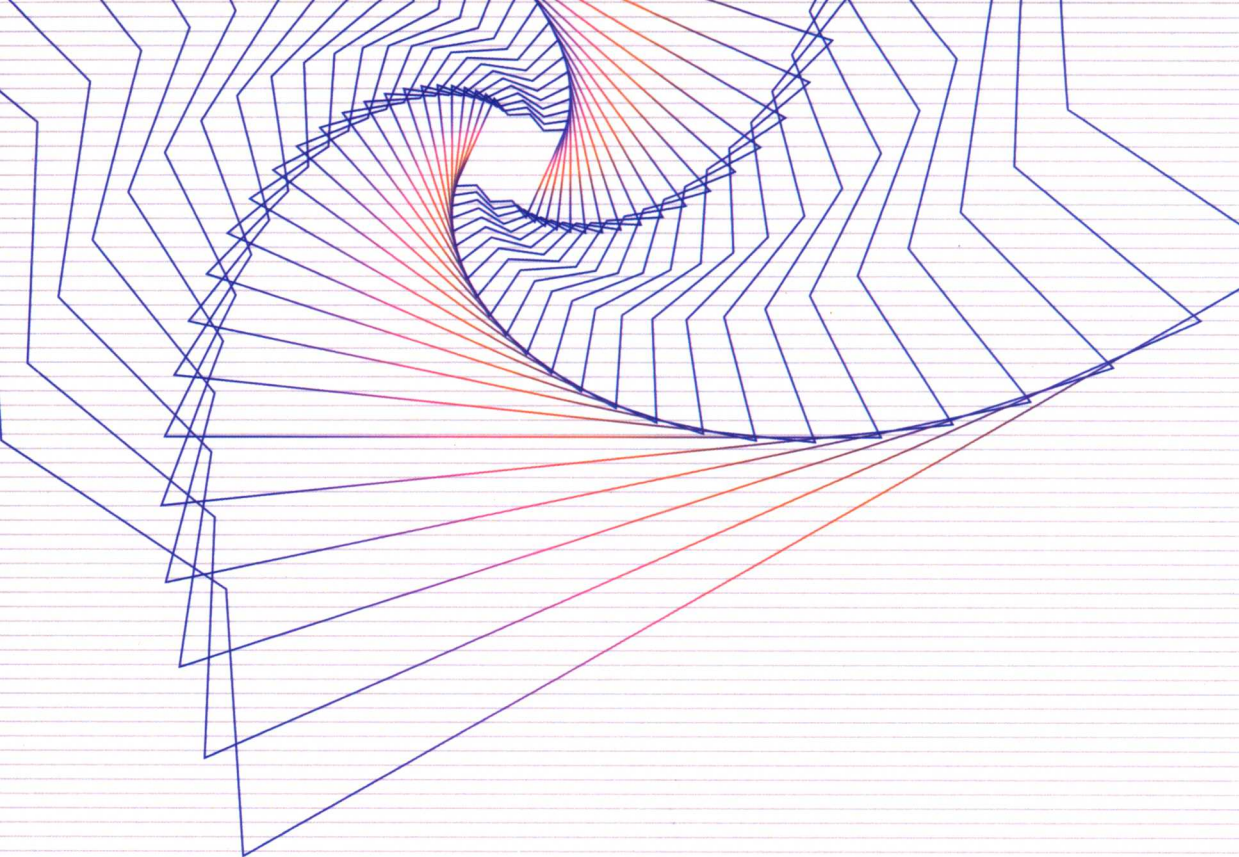


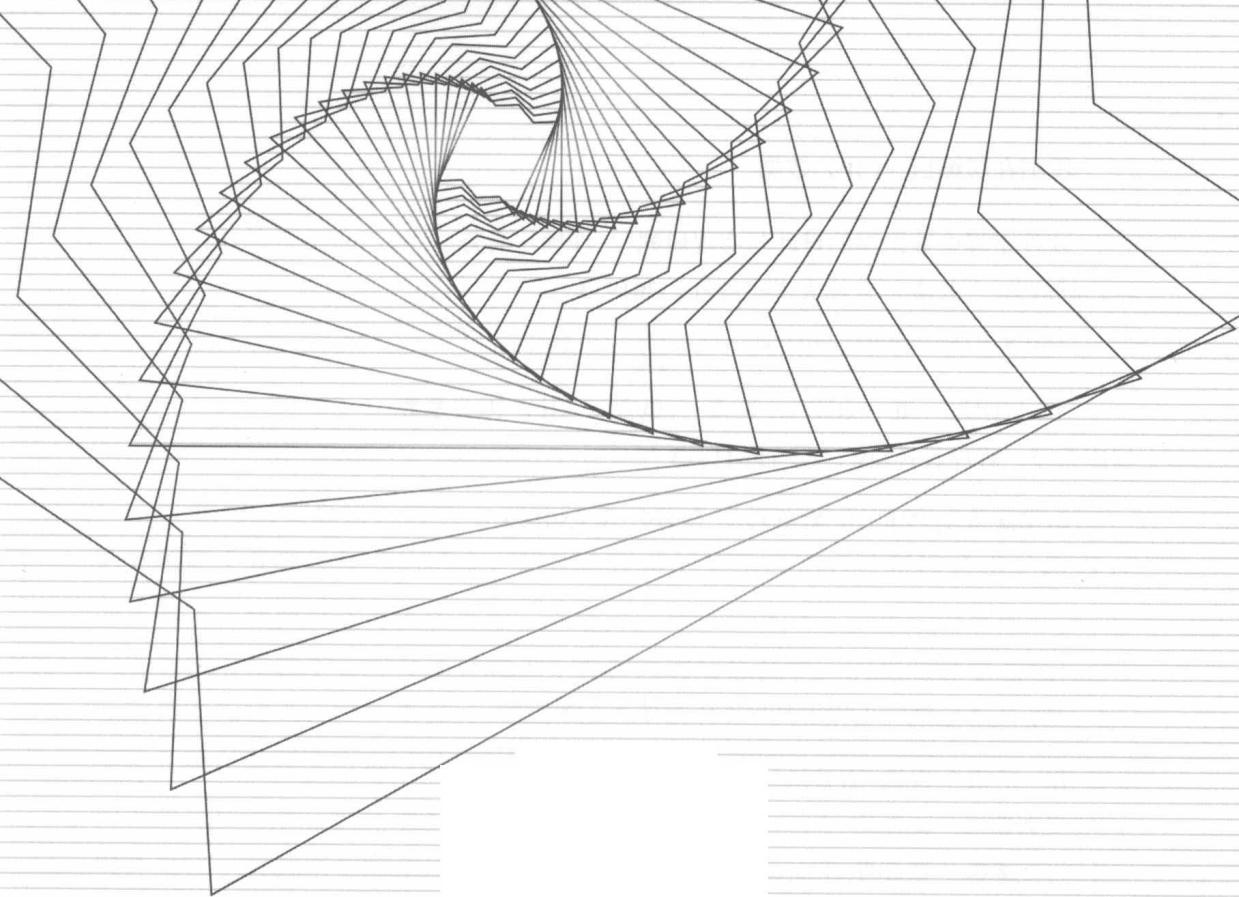


软件工程中数据库的 设计与实现研究

◎ 许楠 高秀艳 赵滨 著

非外借

 吉林大学出版社



软件工程中数据库的 设计与实现研究

© 许楠 高秀艳 赵滨 著

 吉林大学出版社

图书在版编目 (CIP) 数据

软件工程中数据库的设计与实现研究 / 许楠, 高秀艳, 赵滨著. — 长春: 吉林大学出版社, 2019. 3

ISBN 978-7-5692-4463-2

I. ①软… II. ①许… ②高… ③赵… III. ①关系数据库系统—程序设计 IV. ①TP311.13

中国版本图书馆 CIP 数据核字 (2019) 第 050199 号

书 名 软件工程中数据库的设计与实现研究

作 者 许楠 高秀艳 赵滨 著

策划编辑 魏丹丹

责任编辑 魏丹丹

责任校对 严悦铭

装帧设计 凯祥文化

出版发行 吉林大学出版社

社 址 长春市人民大街 4059 号

邮政编码 130021

发行电话 0431-89580028/29/21

网 址 <http://www.jlup.com.cn>

电子邮箱 jdcbs@jlu.edu.cn

印 刷 河北纪元数字印刷有限公司

开 本 787mm×1092mm 1/16

印 张 13.5

字 数 231 千字

版 次 2019 年 3 月 第 1 版

印 次 2019 年 3 月 第 1 次

书 号 ISBN 978-7-5692-4463-2

定 价 55.00 元

前 言

21 世纪是信息化的时代，也是互联网技术飞速发展的时代。互联网技术的发展离不开软件技术的进步，而软件技术的进步离不开数据库技术的提高。因此从事互联网行业的工作人员，或多或少都需要学习一定量的数据库知识。

软件工程技术在现今来看属于一类新兴的信息技术，但是有许多领域都涉及软件工程技术。数据库设计中运用软件工程技术非常常见，在提高计算机水平及发展科学技术方面，软件工程技术有着很大的推动作用，它的出现为人类的生产生活带来了诸多便利。本书主要根据软件工程技术数据库设计中的应用展开研究，对其产生的作用给予分析，为广大互联网技术人员学习数据库相关知识提供依据。

全书共八章。第一章对数据库的基本概念及其相关知识进行了论述，为了解数据库打下了理论基础。第二章探究了 .NET 数据库的开发实践，主要介绍了使用 ADO.NET 对后台数据库进行数据访问的各种方法和数据绑定技术。第三章首先阐述了软件工程技术的使用寿命、使用规格以及技术价值，然后整理了软件工程技术在数据库设计中的作用，最后研究了软件工程中的数据库编程技术。第四章介绍了数据库管理系统 MySQL、软件工程中的物理表设计、数据库故障判断与排除以及数据库实施等内容。第五章整理了怎样利用视图优化管理系统，存储过程在管理系统中的应用方法和触发器的基本概念、语法及其在管理系统中的一些具体应用。第六章探讨了 MySQL 数据库的基本操作及其对表、索引和触发器的基本操作。第七章主要分析了当前数据库存在的一些问题，并对这些问题进行分析，总结出优化数据库的方法。第八章主要举例说明了数据库设计在驾校系统中的具体应用。

本书由许楠、高秀艳、赵滨共同编写完成，具体分工如下：许楠负责第一章至第四章内容的编写和全书的统稿工作，高秀艳负责第五章、第六章内容的编写工作，赵滨负责第七章、第八章内容的编写工作。

本书结合了大量代码实例讲述具体知识，设计思路简单实用，实践步骤清晰，可以触类旁通。全书遵循教学规律，由浅入深、循序渐进。以“教务管理系统”项目作为主线，将相关理论知识与实践操作紧密结合，力争使读者掌握完整的开发项目的设计流程、设计方法、操作方法和应用技巧。本书具有如下特点：本书采用详细的图文结合，对于读者易犯的错误，以“提示”等信息方式列出，便于读者透彻理解；以主流开发环境和开发工具为例，更具有代表性，后台数据库选择了 MySQL、SQL Server 等数据库产品，前台开发工具选择了 .NET 等产品；尽量减少繁多的前台开发语言的干扰，使读者能快速聚焦于数据库模式设计和优化、存储过程、触发器设计等相关知识，具有知识全面、实例精彩、指导性强等优点。

在本书的撰写过程中，借鉴和吸收了许多专家学者的研究成果，在此向他们表示诚挚的感谢。由于笔者知识水平有限，书中难免出现纰漏，敬请广大读者批评指正。

许楠

2018 年 11 月

目 录

第一章 数据库概述	1
第一节 数据库基础知识	1
第二节 关系数据库概述	9
第三节 数据库设计基础	12
第四节 主流数据库管理系统	16
第二章 .NET 数据库开发实践	18
第一节 ADO.NET 简介	18
第二节 基于 .NET Framework 数据提供者编写 数据访问程序	24
第三节 使用 DataReader 与 DataSet 存取数据	33
第四节 ADO.NET 数据库基本操作	41
第五节 数据绑定技术	56
第三章 软件工程与数据库	72
第一节 数据库设计中的软件工程技术价值	72
第二节 数据库设计中软件工程技术的作用	74
第三节 软件工程中的数据库编程技术	76
第四章 软件工程中的数据库物理设计与实现	79
第一节 数据库管理系统——MySQL 简介	79
第二节 软件工程中的物理表的设计	93
第三节 软件工程中的数据库故障判断及排除	118
第四节 软件工程中的数据库实施	121

第五章 软件工程中的数据库业务逻辑设计与实现.....	124
第一节 利用视图优化管理系统.....	124
第二节 存储过程在管理系统中的应用.....	134
第三节 触发器在管理系统中的应用.....	144
第六章 MySQL 数据库对象操作	149
第一节 数据库的基本操作.....	149
第二节 表的基本操作.....	160
第三节 数据库索引.....	172
第四节 触发器常用操作.....	174
第七章 数据库优化研究.....	177
第一节 数据库优化概述.....	177
第二节 影响数据库性能的要素.....	178
第三节 数据库优化的原则、步骤及建议.....	179
第四节 数据库的优化.....	183
第八章 数据库设计实例.....	188
第一节 数据库系统功能.....	188
第二节 数据库设计.....	191
第三节 数据库测试.....	202
参考文献.....	204
后记.....	207

第一章 数据库概述

本章将介绍数据库的基础知识，包括数据库的产生和发展、关系数据库概述、数据库设计基础、主流数据库管理系统等。

第一节 数据库基础知识

20 世纪 60 年代后期，数据库技术开始发展。20 世纪 70 年代以来，数据库技术进入了一个崭新的发展阶段，并得到了广泛的应用。对于计算机科学与技术领域来说，它是一项非常重要的新技术。^①

信息 (Information) 的本质就是人们观察事物的属性和运动状态，并给予一定的反映。人们可以通过信息了解某一事物的两个不同方面：一是事物以什么方式存在；二是在某一时刻事物的运动状态。信息是人们进行多样化活动之后的产物，其中主要的活动有三种：第一种是社会活动；第二种是经济活动；第三种是生产活动。

信息对人们来说是非常重要的存在，人们可以通过接收信息去感知事物，通过信息处理工具对信息进行处理，使信息再生或者增值。如今全球都在朝着信息化的方向发展，信息的重要性也就不言而喻了，它对任何一个国家来说，都是一种不可或缺的资源。

对任何事物来说，数据 (Data) 都是具有一定意义的。数据不仅可以反映

^① 徐孝凯. 数据库技术基础教程 [M]. 北京：清华大学出版社，2004.

各种各样的事物存在的方式，而且还可以记录不同事物的运动状态，是信息的载体。数据可以使用不同的方式去展示信息，如数字方式、文字符号、图形或图像、声音等。

数据与信息在概念上是有区别的。从信息处理角度看，任何事物的存在方式和运动状态都可以通过数据来表示，数据经过加工处理后，使其具有知识性并对人类活动产生作用，从而形成信息。从计算机的角度看，数据泛指那些可以被计算机接受并处理的符号，是数据库中存储的基本对象。

一、有关数据库的概念

我们首先要了解数据库中相关词汇的概念，之后再具体阐述数据库技术。有关的基本概念主要包括以下五种。

第一，数据（Data）：它是用来描述并记录事物的符号。

第二，数据库（Database，DB）：存储在计算机存储设备中的、结构化的相关数据的集合。它不仅包括描述事物的数据本身，也包括描述不同事物之间的联系。

第三，数据库系统（Database System，DBS）：是为适应数据处理的需要而发展起来的一种较为理想的数据处理系统，也是一个为实际可运行的存储、维护和应用系统提供数据的软件系统，是存储介质、处理对象和管理系统的集合体。

第四，数据库应用系统（Database Application System，DBAS）：系统开发人员利用数据库系统资源开发的面向某一类实际应用的软件系统。

第五，数据库管理系统（Database Management System，DBMS）：它是指一些操纵和管理数据库的软件，如 Visual FoxPro 和 Access 等。

二、数据库

数据库是数据库系统的核心部分，是数据库系统的管理对象。

所谓数据库，从表面进行理解，就是海量数据的集合；从深层次理解，就是相关数据之间被有组织地联系在一起，然后存放在计算机中形成的集合。数据库中的数据可以被多个用户同时访问、下载，被数据库管理系统统一管理，并独立于不同应用程序。

数据库依据组织结构所支持的模型类型分为不同的类型。如果一个数据库中数据的组织结构所支持的模型是关系模型,则该数据库为关系数据库;如果一个数据库中数据的组织结构所支持的模型是面向对象模型,则该数据库为对象数据库。例如,在 Access 数据库管理系统中,数据的组织结构所支持的是关系模型,所以 Access 数据库为关系数据库。

数据库管理系统必须提供的数据控制功能包含以下四种:第一,要保证数据安全;第二,要保证数据完整;第三,要实现数据的并发控制;第四,要保证数据库可以恢复。

三、计算机数据管理的发展

数据库的出现对数据处理的作用非常大,首先就解决了有关数据存储的问题。数据库系统能将所有数据集中到一起,进行统一管理、调用,并且使每个用户都能够获取数据信息。

从数据的本质来说,数据处理就是对各种不同的信息进行加工。通过计算机技术对数据进行一系列操作,这些操作可以按照以下步骤进行:第一步,对相关数据进行采集;第二步,对采集后的数据进行整理;第三步,将整理后的数据进行存储;第四步,将存储后的数据进行分类;第五步,将分类后的数据进行排序;第六步,将排序后的数据进行检索;第七步,将检索后的数据进行加工;第八步,将加工后的数据进行统计;第九步,将统计后的数据进行传输。

对数据执行上述操作的目的是,把海量的数据转换成有用的信息。信息可以让人们更加直观地了解数据的意义,因为信息每一个人都理解,但是数据只有那些专业的人士才能理解。计算机通过一些专门的数据分析软件对数据进行分析处理,得出对人们有用的信息,人们将依据所得信息做出选择并使用。

人们使用计算机外存储器来存储数据,通过计算机软件来管理数据,通过应用程序来对数据进行加工处理。

数据处理的最重要的部分是数据管理。数据管理中包含六部分:第一部分是数据的分类;第二部分是数据的组织;第三部分是数据的编码;第四部分是数据的存储;第五部分是数据的检索;第六部分是数据的维护。

随着计算机硬件和软件的发展,数据管理经历了以下三个发展阶段。

(一) 人工管理阶段

20 世纪 50 年代中期以前, 计算机的应用还比较简单, 主要是用来进行科学计算。因为那时计算机处于起步阶段, 所以硬件和软件都存在很多不足之处。从硬件的角度来说, 内存空间较小, 存储量非常有限, 计算速度很慢, 外存储器只有磁带、卡片以及纸带等设备, 磁盘还没有出现; 从软件的角度来说, 因为当时没有更好的操作系统, 所以数据处理非常粗糙, 都是成批进行数据的处理, 更严重的是, 根本没有数据管理软件。

(二) 文件系统阶段

1958—1965 年间, 计算机的应用功能多了一些, 不再局限于科学计算, 而是被广泛应用于一些事务的处理上, 计算机硬件和软件的功能都有了很大的提升。

从硬件角度看: 计算机不再像过去一样, 只是用磁带、卡片等存储器, 人们制造出了磁盘用于存储, 计算速度也有了一定的提升。

从软件角度看: 计算机不仅配备了操作系统, 而且操作系统的性能有了很大的提高, 同时用于数据管理的软件也被制造出来。

从数据处理方式看: 处理数据的方式也不再单一, 不仅可以进行成批处理, 同时也可以进行联机处理, 而且联机处理数据是实时的。这个时期的计算机有三大明显的改变: 数据可以进行更加长期的保存; 数据可以通过软件进行直接管理; 数据结构也朝着多样化趋势发展。

(三) 数据库系统阶段

计算机发展到今天还在不断进步, 目前的计算机应用范围可以说覆盖了 70% 的领域, 数据量更是在不断增长。这一阶段的计算机, 从硬件角度来说, 应用了大量的集成电路, 使内存空间进一步扩大, 计算速度更快, 磁盘容量在不断扩充, 增加了磁鼓及光盘等存储设备; 从软件角度来说, 随着硬件价格下降, 软件价格上涨, 编制和维护系统软件 and 应用程序所需的成本相对增加。为了满足多用户、多应用共享数据的需求, 专门针对数据进行管理的数据库技术出现了。

四、数据库系统

在计算机中引入数据库之后所形成的系统就是数据库系统。数据库系统会

解决一些数据的相关问题,如数据冗余问题、数据共享问题等。

(一) 数据库系统的特点

数据库系统具有如下四个方面的特点:①实现数据共享,减少数据冗余。②整体数据结构化。③具有较高的数据独立性。④有统一的数据控制功能。

(二) 数据库系统的结构

在实际应用中的数据库系统软件种类是非常丰富的,尽管数据库系统软件的种类多种多样,但是它们对数据的定义模式几乎相同,从整体上可以划分为两类,一类是三级模式,另一类是二级影射模式。数据定义包括定义构成数据库结构的外模式、模式及内模式,定义外模式与模式之间的映射,定义模式与内模式之间的映射,定义有关的约束条件等。

1. 三级模式

数据库的基本结构就是三级模式,也就是所谓的外模式、模式和内模式结构。这种模式并不对数据进行具体的组织,而是通过 DBMS 去管理组织,这样对用户比较有利。首先用户可以更加简单地处理数据,其次用户不仅可以忽略数据的表示,同时也可以忽略数据的存储方式。

(1) 外模式。外模式也叫作子模式,通俗理解为用户模式。数据库用户想要看见和使用局部数据的逻辑结构就得通过外模式来实现,数据库用户特征的描述同样也离不开外模式,同时数据库用户的数据视图也是通过外模式来实现的,如在某一应用中,与这个应用相关的数据逻辑表示就是通过外模式来实现的。一般情况下,模式与外模式之间是包含关系,一个数据库可以有一个外模式,也可以有多个外模式。

(2) 模式。模式也称为逻辑模式,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。在数据库系统模式结构中,模式位于中间层地位,这就意味着它不会涉及数据的物理存储细节,也不会涉及硬件环境,与某些具体的应用程序没有任何关系,与程序设计语言之间同样没有任何关系。简单来说,模式就是将数据库中的数据在逻辑级上体现出来,也就是逻辑视图。

(3) 内模式。内模式也叫作存储模式。虽然一个数据库中可以有多个外模式,但内模式只能有一个。内模式既可以描述数据物理结构,同时也可以描述

数据的存储方式。在数据库内部，数据就是通过内模式来表示的，而用于内模式的描述语言是由 DBMS 所定义的，其中一种是内模式 DDL，或称存储模式 DDL (Data Definition Language)。

2. 二级映射

在三级模式之间存在着一定的联系，这种联系可以通过二级映射来实现。

(1) 外模式/模式的映射。所谓外模式/模式的映射，简单来说，就是某一个外模式与模式之间的对应关系。映射关系中，最典型的特点就是一对多，这也就意味着一个模式可以对应多个不同的外模式。在不同的外模式中，它的映射定义是不同的，当模式发生变化时，映射关系也会随之做出相应的改变。

(2) 模式/内模式的映射。模式/内模式的映射同样是用于定义对应关系的。这种对应关系是指数据逻辑结构和存储结构之间的关系。通过这种对应关系可以知道逻辑记录与字段在内部的呈现，为了保证模式不发生改变，当数据库的存储结构发生一定变动时，模式/内模式的映射也要做出相应的修改。

五、数据模型

数据库需要根据应用系统中数据的性质、内在联系，按照管理的要求来设计和组织。数据模型指的是，将真实世界的环境通过计算机技术虚拟呈现出来的一个框架。主要步骤有：首先，人们要将真实世界中的事物感知到大脑中，通过大脑对事物进行整理；然后，将事物抽象为一个概念模型，但这个模型既不依赖于计算机系统，也不依赖于 DBMS 所支持的概念模型；最后，把这个事物的抽象模型进行转换，使之成为 DBMS 能够支持的数据模型。

(一) 相关概念

(1) 实体。所谓的实体指的就是事物，事物需要满足两个特点，第一是客观存在，第二是相互区别。对于实体来说，它既可以是具体的事物，也可以是某些抽象的事物。例如，某个人或者是某个东西就是具体的事物；而空间、时间等就是抽象的事物。

(2) 实体的属性。所谓实体的属性就是用于描述实体的相关特性。例如，一个动物实体可以用它的外观或者是吃的食物来描述，当人们提起熊猫时，首先会想到它的外观是黑白相间，吃的食物是竹子。

(3) 实体型和实体集。对于一个任意实体来说，它的属性可以有很多个，

这些属性集合到一起表示的是这个实体的所属类型，这种类型也就是实体型。而一个实体是通过属性值来体现的，按类型进行划分的实体集合，叫作实体集。例如，关于手机，它的生产时间、编号等就是一个实体型；而对于长颈鹿来说，所有长颈鹿的集合就是一个实体集。

(4) 实体间的联系及种类。在实体与实体之间是存在对应关系的，而这种对应关系就称为联系，这就意味着现实世界的事物之间是相互关联的。而实体间联系的种类是指在某个实体集中所包含的每个实体与另一个实体集中的某些实体存在的联系。一般情况下，两个实体集之间的联系可以分为三种类型。

第一，一对一联系。假设有两个实体集 A 和 B。在 A 中的任意一个实体最多可以与 B 中的某一个实体有联系；而在 B 中的任意一个实体同样最多与 A 中的某一个实体有联系，则称它们之间的联系为 1:1 联系。

第二，一对多联系。同样假设有两个实体集 A 和 B。在 A 中的任意一个实体可以与 B 中的某几个实体有联系；而在 B 中的任意一个实体最多与 A 中的某一个实体有联系，则称 A 对 B 的联系为 1:n 联系。

第三，多对多联系。同样假设有两个实体集 A 和 B。在 A 中的任意一个实体可以与 B 中的某几个实体有联系；而在 B 中的任意一个实体同样也可以与 A 中的某几个实体有联系，则称 A 对 B 的联系为 m:n 联系。

(二) 数据模型类型

数据既要反映事物本身，同时又要反映事物之间的联系，这就要求数据库中的数据不能是散乱的，需要具有一定的结构，这种结构主要通过数据模型来体现。因此数据库不仅需要对数据进行管理，还需要对数据之间的关系进行整理，并用数据模型表现出来。由此可见，在数据库系统中，实体与实体之间的联系是通过数据模型呈现出来的。一个具体的数据模型应当正确地反映出数据之间存在的整体逻辑关系。

一般情况下，数据模型可分为三种：层次数据模型、网状数据模型和关系数据模型。

(1) 层次数据模型。这种模型重点体现的是数据层次，通过树形结构来表示实体之间的联系。在数据库系统中，这种模型是最先出现的，但是并不是所有的数据模型都是层次模型，层次模型需要具备两个最基本的条件：①在树形结构中，需要仅有一个根节点存在，并且这个根节点是独立存在的，没有“双

亲”。②其他节点虽然不是独立存在，但是却只能有一个“双亲”。

若用图来表示，这个图形就像是一颗倒立的树，一棵树有树根、树枝，而根节点就相当于树根，根的孩子就相当于树枝。同时层次也是从根节点算起，根节点被称为“双亲”。如图 1-1 所示。

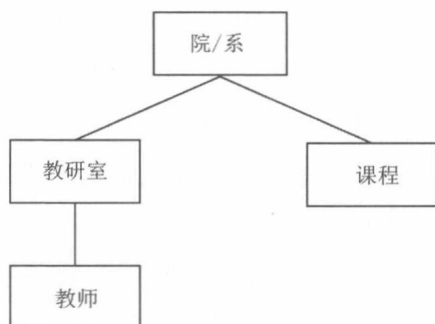


图 1-1 一个系的层次模型

使用层次模型来描述数据之间的联系更加直观易懂，但是层次模型同样存在不足，就是不能够表示实体之间的多对多联系。在 DBMS 系统中，支持层次模型的管理系统称为层次数据库管理系统。

(2) 网状数据模型。从字面上来理解，就是它的模型呈现出网状结构。网状数据模型同样需要具备两个最基本的条件：①在网状数据模型中，可以允许有两个或者是两个以上节点无“双亲”。②一个节点可以有一个以上“双亲”。①

在网状数据模型中，有一个比较典型的系统——DBTG（数据库任务组），这个系统又被称为 CODASYL 系统，是 20 世纪 70 年代数据系统语言协会 CODASYL 下属的数据库任务组提出的一个系统方案。图 1-2 就是一个比较简单的网状数据模型。

① 刘云生. 实时数据库系统 [J]. 计算机科学, 1994, 21 (3): 42-46.

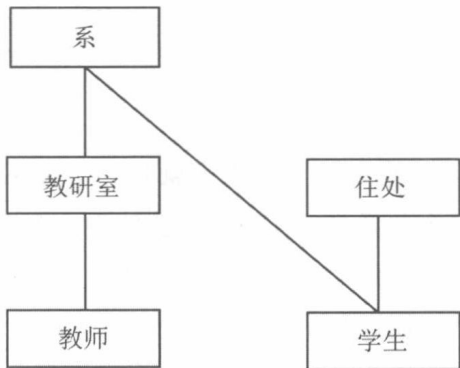


图 1-2 一个系的网状模型

从理论的角度来说，数据实体型之间的联系可以用以上两种数据模型体现，但在自然界中，实体间的联系往往是用网状模型体现的。这是因为大多数实体型间并不存在层次关系，因而实体型间很难形成树形结构，层次模型也体现不出直观的效果，而网状数据模型刚好弥补了层次模型的不足。

(3) 关系数据模型。关系数据模型是指用二维表结构来表示实体及实体之间联系的模型，它是目前最重要的一种模型。关系数据模型以关系数学理论为基础，操作的对象和结果都是二维表，这种二维表就是关系。

对于关系数据模型来说，它有三大优势：一是，它是以数学概念为基础的，是严谨的；二是，它的概念内容比较稳定，不太容易变化；三是，它的存取路径非常直观，任何一个用户都可以看见，并且数据是独立的。

当然，关系数据模型同样也有劣势。它的劣势是，查询效果相较于其他模型来说较弱，因此要改进数据查询请求，这在一定程度上会导致数据库管理系统的开发工作难以快速完成。

第二节 关系数据库概述

在数据库管理系统中普遍使用的数据模型是关系模型，它对数据库的意义毋庸置疑。在关系数据库中，数据所使用的组织方式全部来源于关系模型。最先提出关系模型的人是 IBM 公司的 Edgar Frank Codd（埃德加·弗兰克·科德），自 20 世纪 80 年代以来，各大计算机厂商纷纷推出了支持关系数据模型

的数据库管理系统。

一、关系术语

(1) 关系。在 Access 中，一个二维表就是一个关系，一个关系就是一个二维表；每个关系有一个关系名，一个关系存储为一个表，具有一个表名。

(2) 关系模式。对关系的描述称为关系模式，一个关系模式对应一个关系结构，其格式为关系名（属性名 1，属性名 2，……，属性名 n ）。

(3) 元组。在一个二维表中，水平方向的行称为元组，每一行是一个元组。每一个元组都对应表中的一个具体的记录。

(4) 属性。二维表中垂直方向的列称为属性，每一列有一个属性名，与前面讲的实体属性相同，在 Access 中表示为字段名。每一个字段的数据类型、宽度等在创建表的结构时设定。

(5) 域。域即属性的取值范围，即不同元组对同一个属性的取值所限定的范围。

(6) 关键字。关键字是指其值能够唯一地标识一个元组的属性或属性组合，在 Access 中表示为字段或字段的组合。比如学生表中有学号、学生姓名、班级等信息，因为学生的学号是不能重复的，所以可以起到唯一标识的作用，可以作为关键字。而学生姓名是有可能重复的，所以学生姓名不能作为关键字。主关键字和候选关键字起唯一标识一个元组的作用。

(7) 外部关键字。如果表中的一个字段不是本表的主关键字，而是另外一个表的主关键字和候选关键字，那么，这个字段就称为外部关键字。

从集合论的观点来定义关系，可以将关系定义为元组的集合。关系模式是命名的属性集合，元组是属性值的集合，一个具体的关系模型是若干个有联系的关系模式的集合。

二、关系运算

（一）传统的集合运算

一般来说，集合运算有三种：并运算、差运算、交运算。如图 1-3 所示。