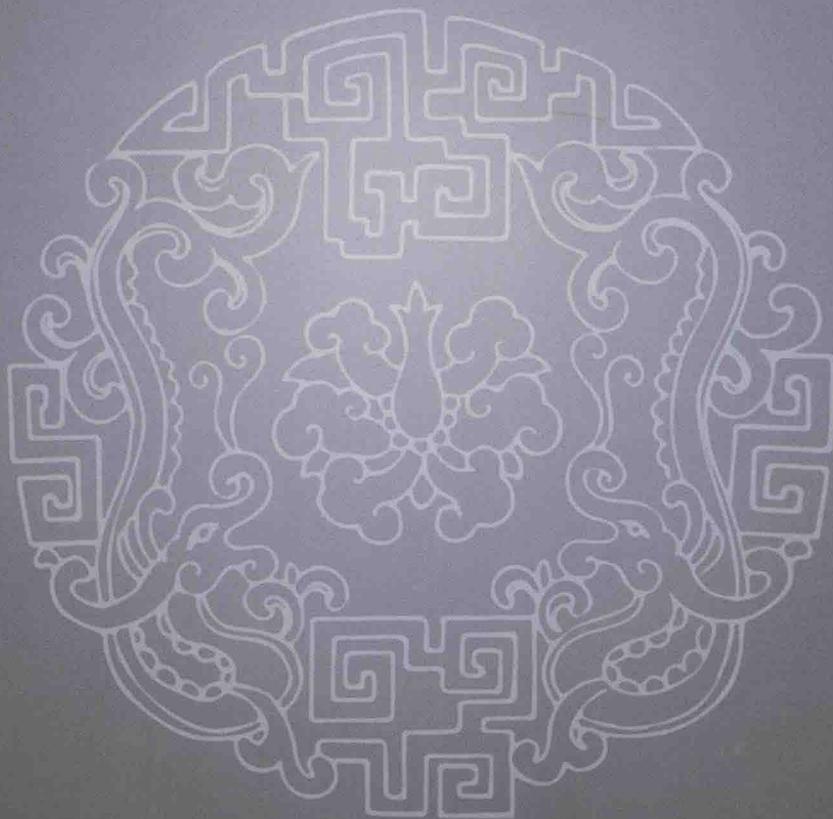


高等学校计算机基础教育课程“十二五”规划教材

# 数据库技术与应用

SHUJUKU JISHU YU YINGYONG

郭永青 主编 焦纯 副主编



中国铁道出版社  
CHINA RAILWAY PUBLISHING HOUSE

高等学校计算机基础教育课程“十二五”规划教材

# 数据库技术与应用

主编 郭永青

副主编 焦 纯

参 编 王路漫 彭 琳 孙 静 赵晓英 孙志勇

## 内 容 简 介

本书以 Access 数据库管理系统为平台, 以实例为主导简明介绍了数据库基础知识、数据库管理技术的具体实现方法和应用。具体内容包括数据库技术概述、数据存储、数据查询、SQL 查询、窗体和报表设计、宏、VBA 程序设计等, 并在最后一章以一个应用案例来巩固所学的知识。

本书面向高等学校非计算机专业, 特别是医药类学校或相关专业的学生, 可作为学习计算机基础课程的教材, 也可以作为了解数据库知识相关人员的参考书。

### 图书在版编目 (CIP) 数据

数据库技术与应用 / 郭永青主编. —北京: 中国

铁道出版社, 2012. 5

高等学校计算机基础教育课程“十二五”规划教材

ISBN 978-7-113-14558-3

I. ①数… II. ①郭… III. ①关系数据库—数据库管理系统, Access—高等学校—教材 IV. ①TP311. 138

中国版本图书馆 CIP 数据核字 (2012) 第 072599 号

书 名: 数据库技术与应用

作 者: 郭永青 主编

---

策 划: 吴宏伟 读者热线: 400-668-0820

责任编辑: 杜 鹏 冯彩茹

封面设计: 付 巍

封面制作: 白 雪

责任印制: 李 佳

---

出版发行: 中国铁道出版社 (100054, 北京市西城区右安门西街 8 号)

网 址: <http://www.51eds.com>

印 刷: 北京海淀五色花印刷厂

版 次: 2012 年 5 月第 1 版 2012 年 5 月第 1 次印刷

开 本: 787mm×1092mm 1/16 印张: 14.75 字数: 349 千

印 数: 1~3 000 册

书 号: ISBN 978-7-113-14558-3

定 价: 29.00 元

---

版权所有 侵权必究

凡购买铁道版图书, 如有印制质量问题, 请与本社教材图书营销部联系调换。电话: (010) 63550836

打击盗版举报电话: (010) 63549504

## 高等学校计算机基础教育课程“十二五”规划教材

主任：童隆正

副主任：郭永青

委员：（按姓氏笔画排列）

王 磊 叶耀辉 田翔华 毕雪华

刘雄鹰 闫朝升 吕格莉 李 丹

李连捷 李祥生 杨长兴 杨国平

周梅红 周智明 赵小龙 赵志坚

章新友 盖立平 韩 滨 喻 焰

焦 纯 焦 瑞 隋 虹 黎小沛

潘志方 魏本征

21世纪是信息技术高度发展且得到广泛应用的时代，信息技术从各个方面改变着人们的学习、工作和生活方式。现代计算机技术和医学技术的融合给生物医药科技带来了新的发展契机和巨大的推动力，深刻地影响并改变着传统的医药科学，给当今的医学工作者带来了难得的机遇和挑战。因此，信息社会对医药人才信息技术的培养要求不断提高，同时也促进了我国医药高校计算机基础教学的迅速发展。

在教育部教育改革、提倡教材多元化的精神指导下，我们秉承“教师为主导、学生为主体”的现代教育理念，以教育部高等学校计算机基础课程教学指导委员会颁发的《高等学校计算机基础教学发展战略研究报告暨计算机基础课程教学基本要求》为指导，在与全国数十位同行就医药类计算机公共课程的“教学目的、教学内容、教学方法和教学手段”等内容进行深入探讨的基础上，在充分调研计算机技术最新发展状况和学生计算机应用水平现状的情况下，由全国30余所医药高校具有丰富教学和教改经验的一线教师编写而成，其中主编和副主编多数是曾经著书立说的教育精英和资深教授。

本丛书由《大学计算机基础》、《Visual Basic 程序设计》、《数据库技术与应用》、《医学成像及处理技术》、《医学多媒体及其在医学中应用》和《医学信息分析与决策》共6本教材组成，涵盖了全国高等医药院校本科、专科各专业的计算机与信息技术应用基础公共课程的教学内容，以培养能够掌握医学信息学基础理论和基本技能、满足社会应用的医学人才。

在组织编写本丛书的过程中，我们始终贯彻“精英出精品，精品育精英”的理念，在内容上强调面向应用、任务驱动、注重实例分析、培养应用能力；在风格上力求文字精练、脉络清晰、图表丰富、版式明快。此外，我们结合课程内容，在适当的位置配以“文史哲经”方面的小贴士以拓宽学生的视野。

本丛书充分体现了系统完整、科学性强、重点突出、文字简练、图文并茂、易教易学、新颖实用和理论联系实际等特点，达到教材编写“三基”（基础理论、基本知识、基本技能），“五性”（思想性、科学性、先进性、启发性、适用性）的要求。

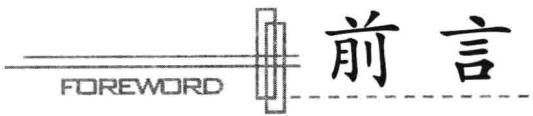
本丛书不仅适合作为医药类高校包括7年制和5年制在内的各类本科专业的教学用书，也是其他高校全日制本、专科学生或成人教育各类专业本、专科学生值得使用的好教材或教学参考书，也可作为计算机等级考试培训教材和参考书。

本丛书的出版得到了谭浩强教授的指导和帮助，他经验丰富、洞察非凡，为丛书质量把关作出了极大贡献。全国医药高校的教师对本丛书的编写提出了许多宝贵的意见和建议。中国铁道出版社对本丛书的策划、论证、组织、出版和发行等做了大量认真而卓有成效的工作。在此对曾为本丛书顺利出版作出贡献的所有人员表示衷心感谢！

全国高等院校计算机基础教育研究会

医学专业委员会

2011年1月于北京



# 前言

随着信息技术和互联网技术的迅猛发展，数据库及数据库技术在各个领域中得到了广泛的应用。在医药领域中常用于信息检索的医学生物数据库、医药信息检索数据库，以及基于医院管理的医院信息系统、实验室信息系统、药品管理系统等都是数据库技术应用的典型案例。教育部高等学校计算机基础课程教学指导委员会《高等学校计算机基础教学发展战略研究报告暨计算机基础课程教学基本要求》和《实施方案》中，也将“数据库技术与应用”列为我国高等学校理工、农林和医学三大类专业本科生计算机基础教学的核心课程和必修课程。

本书在编写中遵循“加强基础，培养能力，提高素质，注重应用”的原则，注重培养学生的科学观念与思维方式，融合医学信息学知识，以案例为主线贯穿整个教材。本书共 8 章，内容包括数据库技术概述、数据存储、数据查询、SQL 查询、窗体和报表设计、宏、VBA 程序设计。前 4 章以介绍数据库基本知识和概念为主，后 4 章以 Access 数据库管理系统为应用平台，讲解面向对象的概念以及程序设计的相关知识。

本书由郭永青任主编，焦纯任副主编，其中郭永青编写第 1 章，王路漫编写第 2 章，彭琳编写第 3 章，孙静编写第 4 章，赵晓英编写第 5 章，焦纯编写第 6 章，孙志勇编写第 7 章和第 8 章。在此对所有付出艰辛劳动的作者表示感谢。

本书编写联合了多所医学院校的一线老师，分享了他们多年教学经验，在编写过程中得到中国铁道出版社的大力支持，在此表示感谢。在此还要感谢北京大学医学部公共教学部计算机教研室给予的支持。本书从酝酿到完成，经过多次商讨和修改，但仍难免存在疏漏和不足，诚请广大读者批评指正。

编 者

2012 年 3 月

# 目录

<b>第1章 数据库技术概述</b>	1
1.1 数据库相关的基础知识	1
1.1.1 基本概念	1
1.1.2 数据模型	4
1.2 数据库技术的产生与发展	5
1.2.1 计算机数据管理	5
1.2.2 数据库技术的发展	6
1.3 数据库系统	7
1.3.1 数据库系统的组成	7
1.3.2 关系数据模型的主要概念	8
1.3.3 关系数据库设计	9
1.4 数据库应用系统开发过程简介	10
思考与练习	11
<b>第2章 数据存储</b>	12
2.1 创建数据库	12
2.1.1 自定义创建空数据库	12
2.1.2 利用模板创建空数据库	13
2.2 数据库的简单操作	14
2.3 构建数据表	15
2.3.1 Access 数据类型	16
2.3.2 字段的属性设置	16
2.3.3 创建表	18
2.3.4 表维护	22
2.3.5 创建索引与主键	29
2.4 表之间的关系	30
2.4.1 建立表之间的关系	31
2.4.2 实施参照完整性	32
2.5 数据库的备份及安全性	33
2.5.1 数据库备份与还原	33
2.5.2 数据库安全性设置	34
思考与练习	37
<b>第3章 数据查询</b>	42
3.1 查询概述	42
3.1.1 查询的种类	42

3.1.2 查询环境 .....	43
3.2 选择查询 .....	45
3.2.1 使用查询向导创建选择查询 .....	45
3.2.2 使用查询设计创建选择查询 .....	48
3.2.3 在查询中设置条件 .....	51
3.3 参数查询 .....	59
3.3.1 单参数查询 .....	59
3.3.2 多参数查询 .....	60
3.4 交叉表查询 .....	62
3.4.1 使用查询向导创建交叉表查询 .....	62
3.4.2 使用“查询设计”方式创建交叉表查询 .....	64
3.5 操作查询 .....	66
3.5.1 生成表查询 .....	66
3.5.2 删除查询 .....	67
3.5.3 更新查询 .....	68
3.5.4 追加查询 .....	69
思考与练习 .....	71
<b>第 4 章 SQL 查询 .....</b>	<b>73</b>
4.1 SQL 概述 .....	73
4.1.1 查看 SQL 语句 .....	74
4.1.2 SQL 的 Select 语句 .....	74
4.2 基本查询 .....	75
4.2.1 简单的条件查询 .....	75
4.2.2 查询全部列 .....	77
4.2.3 设定排序条件 .....	77
4.2.4 限定重复记录 .....	78
4.2.5 显示前面的若干条记录 .....	78
4.3 使用列表达式 .....	79
4.3.1 计算列值 .....	79
4.3.2 修改查询结果的列标题 .....	80
4.3.3 Where 从句的进一步使用 .....	80
4.3.4 数据汇总 .....	82
4.4 多表查询 .....	84
4.4.1 连接查询 .....	84
4.4.2 嵌套查询 .....	85
4.5 创建 SQL 的特定查询 .....	87
4.5.1 创建联合查询 .....	87
4.5.2 创建传递查询 .....	88
4.5.3 创建数据定义查询 .....	89

4.6 其他 SQL 应用 .....	92
4.6.1 使用 Select...Into 语句实现生成表查询.....	92
4.6.2 使用 Delete 语句实现删除查询 .....	93
4.6.3 使用 Insert Into 语句实现追加查询 .....	94
思考与练习 .....	95
<b>第 5 章 窗体和报表设计 .....</b>	<b>96</b>
5.1 窗体 .....	96
5.1.1 创建窗体 .....	96
5.1.2 窗体的编辑操作 .....	101
5.1.3 特殊窗体 .....	118
5.2 报表 .....	128
5.2.1 使用报表向导 .....	128
5.2.2 报表中的计数 .....	138
思考与练习 .....	143
<b>第 6 章 宏 .....</b>	<b>146</b>
6.1 宏的基础知识 .....	146
6.2 宏的创建 .....	150
6.3 宏的调试 .....	156
6.4 宏的运行 .....	157
思考与练习 .....	158
<b>第 7 章 VBA 程序设计 .....</b>	<b>161</b>
7.1 模块的基本概念 .....	161
7.1.1 模块的定义 .....	161
7.1.2 类模块 .....	162
7.1.3 标准模块 .....	163
7.2 创建模块 .....	163
7.2.1 模块的创建 .....	163
7.2.2 模块的转换 .....	164
7.3 面向对象程序设计 .....	165
7.3.1 面向对象程序设计的基本概念 .....	165
7.3.2 对象的模型 .....	166
7.4 VBA 编程环境 .....	170
7.5 数据类型 .....	173
7.5.1 基本数据类型 .....	173
7.5.2 数据常量 .....	173
7.5.3 数据变量 .....	174
7.5.4 有效使用数据类型 .....	176
7.6 常用函数 .....	176

7.6.1 数学函数.....	177
7.6.2 字符串函数.....	177
7.6.3 日期与时间函数.....	178
7.6.4 类型转换函数.....	179
7.6.5 其他函数.....	180
7.7 过程 .....	180
7.7.1 过程的创建.....	181
7.7.2 过程调用中的参数传递.....	182
7.8 流程控制语句 .....	183
7.8.1 条件判断结构语句.....	183
7.8.2 循环结构语句.....	187
7.9 数组 .....	189
7.9.1 数组的声明.....	190
7.9.2 数组的应用.....	190
7.10 调试 .....	191
7.10.1 错误类型.....	192
7.10.2 错误的检测.....	193
7.10.3 排错.....	194
7.10.4 设置错误陷阱.....	194
7.11 VBA 的数据库编程.....	195
7.11.1 数据库访问接口 .....	195
7.11.2 数据访问对象 .....	196
7.11.3 ActiveX 数据对象 .....	198
思考与练习 .....	201
<b>第 8 章 数据库系统开发应用举例——MiniHIS 系统.....</b>	<b>202</b>
8.1 系统功能描述 .....	202
8.2 MiniHIS 系统的实现过程 .....	203
思考与练习 .....	223
<b>参考文献.....</b>	<b>224</b>

# 第1章 数据库技术概述

## 教学目标：

通过本章的学习，初步认识数据、数据库和数据库技术，并了解数据库技术在医学领域中的应用。

## 教学重点和难点：

- 数据
- 数据模型
- 数据库
- 数据库系统

在当今社会，信息资源日益成为人们关注的重点，尤其是随着信息技术和互联网技术的迅猛发展，数据库及数据库技术在各个领域中得到了广泛的应用。例如，在医药领域中常用于信息检索的医学生物数据库、医药信息检索数据库、基于医院管理的医院信息系统、实验室信息系统、药品管理系统等都是数据库及数据库技术应用的典型案例。对于一个医学院校的学生来说，学习和了解数据库知识是非常必要的。

## 1.1 数据库相关的基础知识

计算机技术的实现，孕育了数据管理技术。将现实世界中事物和信息转换为计算机能识别、处理的数据进行存储，从而加以处理，存储的数据是有结构的，这种结构反映了事物之间的相互关系。数据库是存储数据的一种组织管理方式，数据库中的数据是按照一定的结构关系存放的。本节对数据及数据库相关的常用术语进行介绍。

### 1.1.1 基本概念

#### 1. 信息

信息（Information）从广义上讲有着悠久的历史，作为日常用语，“信息”经常是指“音讯、消息”的意思。从狭义上来讲，从信息论、本体论和认识论，都可以对“信息”给出具体的定义。

借助认识论层次对“信息”概念进行解释，“信息”是主体所感知或表述的事物存在的方式和运动状态，人作为主体所感知的是外部世界向主体输入的信息，主体所表述的则是主体向外部世界输出的信息。而计算机所处理的“信息”，可以认为是将人作为主体所感知的外部世界，利用计算机存储处理，再由计算机处理后呈现出来的结果的描述。例如一条李明看病的信息：李明星期一起床后感到浑身无力、怕冷，在8:10分时测量体温为38.9℃。可以从这条信息中得

到病人的姓名、体温、时间、病情等具体的信息描述。

## 2. 数据

描述事物的符号记录称为数据 ( Data ), 它是信息的具体表现形式。例如, 获得一个电话号码, 信息指所获得的使用电话联系方式的描述, 数据则指具体的数字组成的数据本身。抽象来讲, 数据是指存储在某种介质上能够识别的物理符号, 其中数据的概念包含两个方面: 一是描述事物的符号, 如姓名、年龄、性别, 测量血压、脉搏值以及测量的时间, 胎心音、X 光成像等; 另一方面是存储数据的载体, 如记录这些符号的纸张、胶片、磁盘等存储介质。也就是说, 计算机处理的数据不再是由单纯的数字组成, 它包括文本、图像、声音、影像视频等。医学数据包括在医药领域中所涉及的描述各种活动的记录, 都可以作为数据进行收集、存储。

例如, 从李明看病的信息描述可以得到以下数据:

姓名: 李明	字符型数据
检查时间: 8:10	日期时间型数据
体温: 38.9℃	数值型数据
主诉: 无力、怕冷	字符型数据

## 3. 数据处理

数据处理包括对数据的收集、存储、检索和维护等“加工”过程, 其目的是从大量的原始数据中抽取和推导出有价值的信息, 如图 1-1 所示。在计算机中, 数据处理是指将数据转换成信息的过程, 数据被加工后, 形成特定形式的数据表示, 对于数据接收者来讲是有意义的。

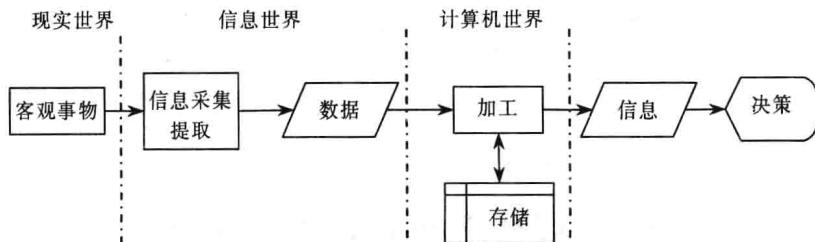


图 1-1 数据和信息

下面通过实际医疗过程, 举例说明通过对现实世界的认识、模拟和抽象到信息世界的数据采集、分类加工等, 再到计算机世界存储、处理、决策的过程:

2011 年 4 月 1 日, 门诊病历中记录上午 8 时, 患者: 张三, 男, 27 岁; 主诉: 右下腹持续疼痛 3 小时; 处理: 血常规、尿常规检测……

对上述信息进行加工, 提取得到结构化数据并输入医院信息系统 ( HIS ):

病历号	姓名	性别	年龄	门诊时间	主诉	处理记录	诊断
1	张三	男	27	2011-4-1	右下腹持续疼痛 3 小时	血常规、尿常规	阑尾炎?

对医疗仪器检验结果转换为计算机可存储的数据:

血液常规检验: WBC, RBC, HGB, HCT, MCV, MCH, MCHC, PLT, LY, MO, NE, EO, ...

尿液常规检验: YS, TMD, PRO, BLO, LEU, NIT, SG, PH, GLU, URO, BI, ...

可存放在专门的数据表中。

在医院信息系统中, 管理检验所得的各种数据的系统称为 LIS 系统 ( Laboratory Information System ), 即实验室 ( 检验科 ) 信息系统。

通过对医院管理系统中的数据进行存储、检索、管理、组织、显示等，医生可以根据化验结果和体征检查，确定诊断为阑尾炎。按照临床路径对阑尾炎疾病诊疗方案，系统提示性给出遵循的顺序性、时间性，从而制订出整体诊治计划，使患者在最短的时间内获得最有效、最佳的医疗诊治、预防和康复效果，并能提高医疗单位服务质量、规范医疗行为、合理利用医疗资源、提高工作效率、降低医疗费用等。

#### 4. 信息技术

“信息技术（Information Technology）”是关于信息的产生、采集、传输、存储、加工、检索、传播等应用技术的总称。也就是利用电子计算机、现代通信、遥感、智能控制等技术获取、传递、存储、显示和应用信息的技术。由于计算机是信息管理的中心，计算机部门通常被称为“IT 部门”，但“信息技术”并非仅限计算机技术。

一般将 IT 应用技术分为 5 类：输入技术、输出技术、软件技术、存储技术、通信技术，如图 1-2 表示。

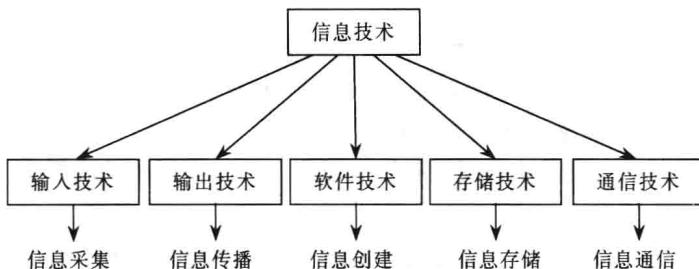


图 1-2 信息技术

#### 5. 数据管理技术

数据管理就是对数据进行组织、存储、检索和维护等工作，是数据处理的核心。数据管理技术经历了人工管理、文件系统管理、数据库系统 3 个阶段。数据库技术是 20 世纪 60 年代开始兴起的一个信息管理自动化的学科，是计算机学科中的重要分支。数据库是数据管理的产物。

数据库（ DataBase ）是按照一定方式组织起来的有联系的数据集合。数据库中的数据按一定的数据模型组织、描述和存储，具有较小的冗余度、较高的数据独立性和易扩展性，并能为各种用户所共享。数据库不仅存放了数据，而且也存放了数据之间的关系。计算机不能直接处理现实世界的具体事物，必须要先将具体事物转换为计算机能够存储和处理的数据形式，即从具体事物及关系中抽象建立出数据模型。数据库技术就是基于某种数据模型来表示和处理现实世界的信息和数据。

例如，在学校管理中，学校行政机构的层次关系为：学校下属有学院，学院下属有系，系下属有教师、学生等，按照层次关系整理出数据进行管理存储；也可以将院系学生、教师、课程、宿舍、教室等整理出各种网状关系数据进行存储；在教学管理中，经常将学生的姓名、性别、出生年月、籍贯、入学日期等，将课程编号、课程名称、学时、学分等分别组成二维关系表进行存储管理。

在医疗过程中，将实际医疗流程整理记录下来形成下列各关系表：门诊记录表（病案号，姓名，性别，年龄，门诊时间，主诉，处理记录，诊断）、血液常规记录表（样本编号，病案号，WBC, RBC, HGB, HCT, MCV, MCH, …）、尿液常规记录表（样本编号，病案号，YS, TMD, PRO, BLO, LEU, NIT, SG, PH, …）、检验记录表（样本编号，采样时间，标本种类，送检日期，科别，送检

日期), 然后将这些相关的数据表组织在一个“医疗记录数据库”中。

再如在每年例行的健康体检的“健康体检记录”表中, 每一个体检人是一条记录, 每条记录都有体检日期、病案号、姓名、出生日期、家庭住址和邮编等属性; 在“超声波”表中, 每人一条记录, 每条记录都有日期、病案号、姓名、超声检查结果、检查医生等属性; 在“血脂化验记录表”中, 每一张化验单是一条记录, 每条记录都有送检日期、病历号、姓名、甘油三酯、总胆固醇、高密度脂蛋白等属性。这些表可以组成一个“体验记录数据库”, 利用数据库管理技术将这两个数据库建立一定的联系, 进行数据的整合, 整合为一个完善的健康档案的数据库。

### 1.1.2 数据模型

数据模型 (Data Model) 是现实世界数据特征的抽象结构, 也就是说, 它是将具体事物本身或事物之间的各种关系转换成计算机能够处理的数据的一种方式。数据模型包括以下组成部分:

- ① 数据结构: 描述系统的静态特性, 即组成数据库的对象类型。
- ② 数据操作: 一般有检索、更新(插入、删除、修改)操作。
- ③ 数据的约束条件: 完整性规则的集合。

到目前为止, 数据库的逻辑数据模型有层次模型、网状模型和关系模型3种。

#### 1. 层次模型

层次模型 (Hierarchical Model) 是数据库系统中最早使用的模型, 它的数据结构类似一棵倒置的树, 只有一个根结点, 若干个枝叶结点(见图 1-3), 每个结点表示一个记录类型。层次模型的特征是: ①有且仅有一个结点没有双亲结点, 这个结点成为根结点; ②除根结点外的其他结点有且只有一个双亲结点。层次模型结构简单, 各个数据之间的联系一目了然, 对于某些特定的应用系统效率很高, 其缺点是不能表示复杂的数据关系, 如学校的行政机构设置。

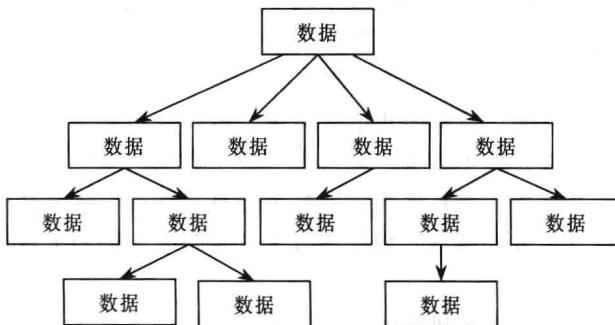


图 1-3 层次模型

#### 2. 网状模型

网状模型 (Network Model) 可以看做是层次模型的一种扩展, 它采用网状结构表示实体及其之间的联系。网状模型的特征是: ①允许一个以上的结点没有父结点; ②一个结点可以有多个父结点。网状模型(见图 1-4)与层次模型相比, 提供了更大的灵活性, 能更直接地描述现实世界, 性能和效率也比较好。网状模型的缺点是结构复杂, 用户不易掌握, 记录类型之间的联系发生变动后涉及链接指针的调整, 扩充和维护都比较复杂, 如公路交通的设置。

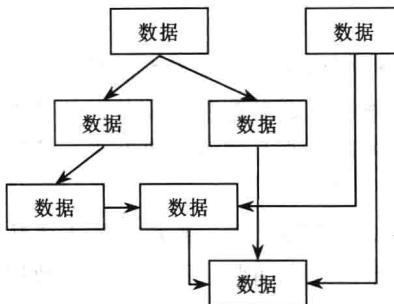


图 1-4 网状模型

### 3. 关系模型

用表格来记录数据和数据之间的关系是现实社会常用的方式。关系模型（Relation Model）是一个二维表格，它能够表示数据之间的复杂关系，且便于管理。关系模型是以集合论中的关系概念为基础发展起来的，它将数据组织成由若干行、每行又由若干列组成的表格形式，这种表格在数学上称为关系。表格中存放了表示实体本身的数据和实体之间的联系。一个数据表是具有若干相同属性的记录的集合。在表格中，行代表记录，列代表各种属性（数据项或字段），不允许有复合数据项，如表 1-1 所示的视力记录。如果用关系模型表示时就必须用表 1-2 所示的形式表示。

表 1-1 体检记录表——复合表格表示

姓名	身高	体重	视力	
			左	右
章林	1.60	48.5	4.6	6.7
刘晓	1.78	66.0	5.0	5.1
周凡	1.82	61.5	4.9	5.0
宋星	1.56	50.0	4.8	4.6

表 1-2 体检记录表——关系模型

姓名	身高	体重	左视力	右视力
章林	1.60	48.5	4.6	6.7
刘晓	1.78	66.0	5.0	5.1
周凡	1.82	61.5	4.9	5.0
宋星	1.56	50.0	4.8	4.6

由于关系模型直观简便，符合人们日常处理数据的习惯，自 20 世纪 70 年代提出后，便得到了迅速推广。

## 1.2 数据库技术的产生与发展

随着计算机的发展，特别是高效率存储设备的出现，使得数据处理进入了应用计算机的时代。数据库技术主要研究如何存储、使用和管理数据，是计算机技术中发展最快、应用最广的技术之一。

### 1.2.1 计算机数据管理

从数据处理和管理的层面上来看，计算机在数据管理方面经历了人工管理阶段、文件系统管理阶段、数据库管理阶段。

#### 1. 人工管理

早期的计算机主要用于科学计算，硬件只有纸卡、磁带等外围存储设备，对数据的管理基本是采用人工处理或随程序存放。

## 2. 文件系统管理

20世纪50年代后期计算机的应用逐渐进入人类活动的各个领域，大量的数据需要处理，当时作为计算机的存储设备，开始出现磁盘等存储设备，软件方面也出现了高级语言和操作系统，包括专门的数据管理软件，可以迅速有效地对数据进行管理，称为文件系统。

## 3. 数据库管理

20世纪60年代中期，产生了数据库技术。数据库技术将数据独立集中存放，并独立于具体的应用程序，不仅可以解决数据的冗余问题，而且做到了数据为所有应用程序所共享，保证了数据的安全和统一。

数据库管理阶段的数据管理方式有以下几个特点：

- ① 数据结构化。
- ② 数据的共享性能高，冗余度低，易扩充。
- ③ 数据独立性高。
- ④ 数据由数据库管理系统统一控制。

### 1.2.2 数据库技术的发展

数据库技术的形成和发展大概经历了第一代层次和网状数据库管理系统、第二代关系数据库管理系统，以及新一代数据库管理系统。

早在20世纪60~70年代时期，数据库系统使用的是层次和网状数据库系统。层次数据库模型的代表是IMS(Information Management System)数据库，其数据以一种树形的逻辑拓扑结构进行存储，是由IBM公司于1966—1968年为NASA(美国国家航空航天局)的阿波罗登月计划专门开发的。在阿波罗登月计划之后，IMS数据库又被广泛应用在金融、保险、制造等多个行业，至今已逾40年。网状数据模型的典型代表是DBTG系统，又称CODASYL系统，它是20世纪70年代数据系统语言协会(CODASYL)下属的数据库任务组(DataBase Task Group,DBTG)提出的一个系统方案。

在20世纪70年代，IBM公司San Jose研究所的E.F.Codd博士发表了题为“大型共享数据库的数据关系模型”论文，开创了数据库的关系方法和关系规范化的理论研究，奠定了关系数据库的理论基础。关系型数据库系统逐渐取代了层次和网状数据库系统，它的代表是IBM公司开发的System R和加州大学Berkley分校开发的INGRES数据库。在新发展的DBMS产品中，近90%是采用关系数据模型。例如，小型数据库系统有dBase、Paradox、FoxPro、Microsoft Access等，大型数据库系统有DB2、Oracle、Informix、Sybase、SQL Server、MySQL等。

20世纪80年代至今，数据库理论和应用进入成熟发展时期。随着信息技术和市场的发展，人们发现虽然关系型数据库系统技术很成熟，并能很好地处理所谓的“表格型数据”，但其局限性也是显而易见的：数据库技术一直随着计算机技术的发展而不断进步，当今信息的复杂性与关系型数据库理论产生的时代已不可同日而语，信息的结构日益复杂；随着个性化服务等需求的增长，信息的多样性和差异性都得到大幅度增加，使关系型理论在管理信息复杂性方面的不足日益明显。

20世纪90年代以后，随着计算机科学和技术的发展，数据库技术与通信技术、多媒体技术、人工智能技术、面向对象程序设计技术、并行计算技术等相互渗透、相互结合，使数据库系统得到了新的发展，并成为当代数据库技术发展的主要特征。