

新世纪高职高专实用规划教材

● 计算机系列

数据库系统 及应用教程

SHUJUKU XITONG JI YINGYONG JIAOCHENG

汪洋 喻梅 于健 薛辉 编著



清华大学出版社

新世纪高职高专实用规划教材 计算机系列

数据库系统及应用教程

汪洋 喻梅 于健 薛辉 编著

清华大学出版社

北 京

内 容 简 介

本书首先介绍了数据库系统的基本概念和关系数据库的理论；接着简单介绍了 Visual FoxPro 6.0，包括其性能、安装、配置和使用等基本知识；然后重点介绍了有关数据库操作的内容，包括建立数据库、数据表和索引等内容；最后介绍了 Visual FoxPro 6.0 中程序设计的方法，包括视图、表单、菜单和报表的设计，使用户对使用 Visual FoxPro 6.0 开发一套数据库应用系统有更深入的了解。

本书可以作为高职高专院校相关专业数据库应用基础课程的教材，同时也可供从事数据库研究和应用 Visual FoxPro 6.0 进行数据库系统开发的计算机专业人员参考。

版权所有，翻印必究。举报电话：010-62782989 13501256678 13801310933

本书封面贴有清华大学出版社防伪标签，无标签者不得销售。

本书防伪标签采用清华大学核研院专有核径迹膜防伪技术，用户可通过在图案表面涂抹清水，图案消失，水干后图案复现；或将表面膜揭下，放在白纸上用彩笔涂抹，图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

数据库系统及应用教程/汪洋，喻梅，于健，薛辉编著. —北京：清华大学出版社，2005.4
(新世纪高职高专实用规划教材 计算机系列)

ISBN 7-302-10496-4

I.计… II.①汪…②喻…③于…④薛… III.数据库系统—高等学校：技术学校—教材 IV.TP311.13

中国版本图书馆CIP数据核字(2005)第011989号

出版者：清华大学出版社

<http://www.tup.com.cn>

社总机：010-62770175

地 址：北京清华大学学研大厦

邮 编：100084

客户服务：010-62776969

组稿编辑：王景先

文稿编辑：许瑛琪

封面设计：陈刘源

排 版 者：王 婷

印 刷 者：北京市通州大中印刷厂

装 订 者：三河市金元装订厂

发 行 者：新华书店总店北京发行所

开 本：185×260 印张：20 字数：478千字

版 次：2005年4月第1版 2005年4月第1次印刷

书 号：ISBN 7-302-10496-4/TP·7123

印 数：1~4000

定 价：26.00元

前 言

计算机技术的进步,使得数据库管理系统软件得到了快速发展,从 dBASE、FoxBASE、FoxPro 发展到现在的 Visual FoxPro,并且随着版本的不断升级,功能也越来越强大。可视化数据库设计与管理是当前的技术热点,本书是针对目前流行的 Visual FoxPro 6.0 而编写的数据库应用基础教程。

Visual FoxPro 6.0 是由 Microsoft 公司推出的中、小型可视化数据库管理系统,属于面向对象的编程语言,是运行于 Windows 98 和 Windows NT 平台的 32 位数据库开发系统。它不但兼容早期 FoxBASE 的各种版本,同时还提供了许多基于 Windows 的崭新功能。它具有优越的性能、丰富的工具、友好的图形界面、简单的数据存取方式、优秀的跨平台特性和真正的可编译性,使组织数据、定义数据库规则和建立应用程序等工作变得简单透明。它提供了可视化的设计工具和向导,可以快捷地创建表单、查询和报表。它提供了集成化的开发环境,拥有面向对象的程序设计工具,为快速开发功能全面的应用程序创造了良好的条件。Visual FoxPro 6.0 不仅可以简化中、小型数据库的管理,也可以作为大型数据库的前端开发工具。

Visual FoxPro 6.0 包含的内容很多,本书编写时,在具体内容的安排上,从应用的角度出发,以实用性为重点,侧重于操作和应用所必需的基础知识,使读者着重理解面向对象的程序设计的思路和方法。在取材的深度和广度方面,进行了精心的优化筛选,注意内容简练,精心设计实例,用通俗易懂的语言进行叙述。为了便于学习掌握,在例题的安排上注意了连续性,各章相对独立,又互为补充,使学生可以在有限的学时内全面掌握实用技术。

本书取材注重实用,内容由浅入深,语言简练,逻辑性强,适用于教学。本书可以作为高职高专院校相关专业数据库应用基础课程的教材,同时也可供从事数据库研究和使用的 Visual FoxPro 6.0 进行数据库系统开发的计算机专业人员参考。

本书第 1~3 章由喻梅编写,第 4~6 章由汪洋编写,第 7~9 章由于健编写,第 10~11 章由薛辉编写。

由于时间仓促,作者水平有限,书中难免有不当之处,敬请读者批评指正。

编者

2005 年 1 月

目 录

第 1 章 数据库系统的基本概念 1	
1.1 信息、数据与数据处理..... 1	
1.1.1 信息与数据..... 1	
1.1.2 数据处理..... 2	
1.2 数据管理技术的发展..... 2	
1.2.1 人工管理阶段..... 2	
1.2.2 文件系统阶段..... 3	
1.2.3 数据库系统阶段..... 4	
1.3 数据模型..... 6	
1.3.1 数据描述的三个领域..... 6	
1.3.2 数据模型..... 8	
1.3.3 概念数据模型..... 9	
1.3.4 结构数据模型..... 13	
1.4 数据库的体系结构..... 17	
1.4.1 数据库系统的模式结构..... 17	
1.4.2 三级模式结构..... 19	
1.4.3 两级模式映像及数据 独立性..... 20	
1.5 数据库系统..... 20	
1.5.1 数据库系统的组成..... 20	
1.5.2 数据库管理系统(DBMS)..... 22	
1.5.3 数据库管理员..... 24	
1.6 习题..... 25	
第 2 章 关系数据库 26	
2.1 关系模型的基本概念..... 26	
2.1.1 关系模型的基本术语..... 26	
2.1.2 关系的定义和性质..... 28	
2.1.3 关系模型的三要素..... 29	
2.2 关系代数..... 30	
2.2.1 传统的集合运算..... 31	
2.2.2 专门的关系运算..... 32	
2.2.3 关系代数表达式及其 应用实例..... 35	
2.3 关系规范化..... 36	
2.3.1 关系模式的设计问题..... 37	
2.3.2 函数依赖..... 39	
2.3.3 关系模式的范式与规范化..... 39	
2.4 习题..... 42	
第 3 章 数据库设计 44	
3.1 数据库设计概述..... 44	
3.1.1 数据库设计的内容..... 44	
3.1.2 数据库设计的方法..... 45	
3.1.3 数据库设计的步骤..... 46	
3.2 需求分析..... 48	
3.2.1 需求分析的任务..... 48	
3.2.2 需求分析的基本步骤..... 49	
3.3 概念设计..... 50	
3.3.1 概念设计的目标和策略..... 50	
3.3.2 采用 E-R 方法的数据库 概念设计..... 51	
3.4 逻辑设计..... 54	
3.4.1 逻辑设计的步骤..... 54	
3.4.2 E-R 模型向关系数据 模型的转换..... 55	
3.4.3 关系数据库的逻辑设计..... 56	
3.5 物理设计..... 58	
3.5.1 物理设计的内容..... 58	
3.5.2 物理设计的性能..... 59	
3.6 实现与维护..... 60	
3.6.1 数据库的实现..... 60	
3.6.2 数据库的其他设计..... 61	
3.6.3 数据库的运行与维护..... 61	
3.7 习题..... 62	

第 4 章 Visual FoxPro 6.0 概述	64	5.3.4 逻辑表达式.....	87
4.1 Visual FoxPro 6.0 特性.....	64	5.3.5 时间日期表达式.....	87
4.1.1 FoxPro 发展简史.....	64	5.3.6 操作符的优先级.....	88
4.1.2 Visual FoxPro 的特性.....	65	5.3.7 空值的处理.....	88
4.2 Visual FoxPro 6.0 的安装、 启动与退出.....	66	5.4 常用函数.....	90
4.2.1 Visual FoxPro 6.0 的安装.....	66	5.4.1 数值函数.....	90
4.2.2 Visual FoxPro 6.0 的启动.....	67	5.4.2 字符函数.....	92
4.2.3 Visual FoxPro 6.0 的退出.....	67	5.4.3 日期和时间函数.....	94
4.3 Visual FoxPro 6.0 系统界面.....	68	5.4.4 类型转换函数.....	95
4.3.1 Visual FoxPro 6.0 的 主窗口.....	68	5.4.5 测试函数.....	97
4.3.2 Visual FoxPro 6.0 的向导.....	70	5.5 习题.....	98
4.3.3 Visual FoxPro 6.0 的 生成器.....	71	第 6 章 数据表的创建与基本操作	99
4.3.4 Visual FoxPro 6.0 的 设计器.....	72	6.1 数据表的创建.....	99
4.4 Visual FoxPro 6.0 系统环境的 配置.....	72	6.1.1 数据表结构.....	99
4.4.1 通过【选项】对话框设置 系统环境.....	72	6.1.2 创建数据表.....	102
4.4.2 通过配置文件设置 系统环境.....	74	6.1.3 显示和修改表结构.....	108
4.5 Visual FoxPro 6.0 的常用 文件类型.....	74	6.2 表数据的输入.....	110
4.6 Visual FoxPro 6.0 性能指标.....	75	6.2.1 创建表时立即输入.....	111
4.7 Visual FoxPro 6.0 的命令格式.....	76	6.2.2 以追加方式输入数据.....	112
4.8 习题.....	76	6.2.3 在表中插入数据.....	113
第 5 章 数据与数据运算	78	6.3 记录的显示与修改.....	114
5.1 Visual FoxPro 6.0 数据类型.....	78	6.3.1 表的打开与关闭.....	114
5.2 常量与变量.....	81	6.3.2 使用“编辑”或“浏览” 方式显示和修改数据.....	115
5.2.1 常量.....	81	6.3.3 使用浏览命令显示和 修改数据.....	116
5.2.2 变量.....	82	6.3.4 列表显示表中数据.....	117
5.2.3 数组.....	84	6.4 记录的定位.....	119
5.3 操作符与表达式.....	85	6.4.1 使用菜单移动记录.....	119
5.3.1 数值表达式.....	85	6.4.2 移动记录命令.....	120
5.3.2 字符表达式.....	86	6.4.3 与数据库相关的函数.....	121
5.3.3 关系表达式.....	86	6.5 记录的删除与恢复.....	123
		6.5.1 逻辑删除记录.....	123
		6.5.2 删除记录的恢复.....	125
		6.5.3 物理删除.....	126
		6.5.4 一次性删除表中 全部记录.....	127
		6.6 索引与排序.....	127

6.6.1 排序	128	7.3.4 定义触发器	173
6.6.2 索引	128	7.3.5 定义索引	174
6.6.3 创建索引	131	7.4 建立表间的关系	175
6.6.4 使用索引	135	7.4.1 建立表间的永久关系	175
6.6.5 表的查询	139	7.4.2 设置参照完整性	178
6.7 表的计算与汇总	142	7.5 习题	180
6.7.1 计数命令 COUNT	142	第 8 章 查询与视图	182
6.7.2 求和命令 SUM	143	8.1 创建查询	182
6.7.3 求平均值命令 AVERAGE	143	8.1.1 使用查询向导创建 查询	182
6.7.4 分类汇总命令 TOTAL	143	8.1.2 使用查询设计器 创建查询	185
6.8 多表的关联、连接与更新	144	8.1.3 定义结果	188
6.8.1 不同工作区中表的操作	145	8.1.4 保存查询	193
6.8.2 不同工作区中表的 逻辑关联	147	8.1.5 运行查询	193
6.8.3 数据工作期窗口与 视图文件	148	8.2 创建视图	194
6.8.4 不同工作区中表的 物理连接	151	8.2.1 创建视图	194
6.8.5 不同工作区中表的更新	152	8.2.2 设置视图的更新条件	199
6.9 习题	153	8.2.3 使用视图	201
第 7 章 数据库的创建与管理	155	8.3 习题	201
7.1 数据库的创建	155	第 9 章 程序设计基础	203
7.1.1 数据库的设计	155	9.1 程序文件的建立与执行	203
7.1.2 创建数据库	156	9.1.1 Visual FoxPro 6.0 的 工作方式	203
7.1.3 显示数据库结构	161	9.1.2 程序与程序文件	204
7.2 使用数据库	162	9.1.3 程序文件的建立与编辑	205
7.2.1 打开、关闭和删除 数据库	162	9.1.4 程序文件的执行	206
7.2.2 向数据库中添加与 删除表	163	9.2 程序的交互式命令	206
7.3 设置数据库表的字段属性	166	9.2.1 人机交互式输入命令	206
7.3.1 长表名和表中的 长字段名	166	9.2.2 中止与挂起程序运行 命令	208
7.3.2 字段的标题和注释、 默认值	169	9.2.3 屏幕格式命令	208
7.3.3 设置字段级及记录级 规则	171	9.3 结构化程序设计	211
		9.3.1 程序设计的概念和步骤	211
		9.3.2 程序流程图	212
		9.3.3 程序控制结构	213
		9.4 顺序结构程序设计	215

9.5 分支结构的程序设计.....215	10.3.1 控件对象..... 250
9.5.1 单向分支结构.....216	10.3.2 标签(Label)控件..... 251
9.5.2 双向分支结构.....216	10.3.3 文本框(Text)控件..... 251
9.5.3 多向分支结构.....217	10.3.4 编辑框(Edit Box)控件..... 252
9.6 循环结构的程序设计.....219	10.3.5 命令按钮(Command Button) 控件..... 253
9.6.1 DO WHILE...ENDDO 循环语句.....219	10.3.6 命令按钮组 (Command Group)控件..... 254
9.6.2 FOR...ENDFOR 循环 语句.....222	10.3.7 选项按钮组(Option Group) 控件..... 254
9.6.3 SCAN...ENDSCAN 循环语句.....224	10.3.8 复选框(Check Box)控件..... 255
9.6.4 循环语句的嵌套结构.....225	10.3.9 列表框(List Box)控件..... 255
9.7 子程序.....225	10.3.10 组合框(Combo Box) 控件..... 256
9.7.1 子程序的调用与返回.....226	10.3.11 微调(Spinner)控件..... 257
9.7.2 子程序的嵌套调用.....227	10.3.12 表格(Grid)控件..... 257
9.7.3 子程序调用中的参数传递.....228	10.3.13 图像(Image)控件..... 258
9.8 过程与过程文件.....229	10.3.14 计时器(Timer)控件..... 259
9.8.1 过程文件的建立.....230	10.3.15 综合实例..... 260
9.8.2 过程文件的打开.....231	10.4 保存和运行表单..... 262
9.8.3 过程文件的关闭.....231	10.4.1 保存表单..... 262
9.9 用户自定义函数.....232	10.4.2 运行表单..... 263
9.9.1 用户自定义函数的定义.....232	10.5 习题..... 263
9.9.2 用户自定义函数的调用.....232	第 11 章 菜单和报表设计..... 265
9.10 习题.....233	11.1 菜单设计..... 265
第 10 章 表单设计.....236	11.1.1 菜单的组成..... 265
10.1 创建表单.....236	11.1.2 使用菜单设计器设计 菜单..... 266
10.1.1 表单的设计过程.....236	11.1.3 设置菜单的选项..... 272
10.1.2 使用表单向导创建表单.....237	11.1.4 使用命令方式设计菜单..... 274
10.1.3 使用表单设计器 创建表单.....240	11.2 报表设计..... 274
10.1.4 表单生成器.....244	11.2.1 设计报表布局..... 275
10.1.5 使用命令方式创建表单.....245	11.2.2 创建报表..... 275
10.2 表单的属性、事件和方法.....245	11.2.3 报表的相关命令..... 285
10.2.1 定义表单的属性.....245	11.3 习题..... 285
10.2.2 设置数据环境.....246	第 12 章 项目管理器与应用程序..... 287
10.2.3 编写事件代码.....248	12.1 项目管理器..... 287
10.2.4 编写方法程序.....249	
10.3 添加控件.....250	

12.1.1 创建项目	287	12.2.2 设置初始化环境.....	297
12.1.2 项目管理器中的选项.....	289	12.2.3 显示初始界面.....	298
12.1.3 项目管理器的基本操作.....	291	12.2.4 控制事件循环.....	298
12.1.4 项目管理器中的命令 按钮	294	12.2.5 恢复初始环境.....	298
12.2 应用程序的开发.....	296	12.2.6 包含和排除文件.....	298
12.2.1 设置主文件.....	297	12.3 习题.....	299
		参考答案	301

第 1 章 数据库系统的基本概念

教学提示：本章主要对数据库系统的基本概念做介绍，包括数据与信息概念；数据管理技术的发展阶段；数据模型的两个层次；数据库系统的体系结构及数据库系统的组成。

教学目标：了解数据、信息及其关联与区别；数据管理技术发展的三个阶段及其特点；掌握概念数据模型与结构数据模型的基本概念；掌握数据库系统的三级模式、两级映像及两级独立性。

1.1 信息、数据与数据处理

在科学、技术、经济、文化和军事等各个领域里，我们会遇到大量的数据，这些数据是复杂而且数据量大，在科学、文化、经济、军事等一切领域的发展中都离不开数据，因此如何科学地管理数据是一个极为重要的课题。

数据库技术是使用计算机来管理数据的一门最新科学技术。经过多年的研究和实践，数据库技术已发展成为一门完整的学科，人们已开发出多种数据库管理系统，使用这些数据库系统能够科学有效地管理大量数据，它们正在为各领域的发展发挥着重要的作用。

1.1.1 信息与数据

计算机的出现，开辟了数据处理的新纪元。数据处理的基本要素是数据的组织、存储、检索、维护和加工利用，这些正是数据库系统所要解决的问题。

数据是数据库系统研究和处理的对象。数据与信息是分不开的，它们既有关联又有区别。

1. 信息

随着社会的发展和科学技术的进步，人们对信息这个名词已经不陌生了，然而对于信息的定义，从不同角度又有着不同的解释。一般认为，信息是人们进行各种活动所需要的知识，是现实世界各种状态的反映。合理利用信息可以增加人们的知识，提高人们对事物的认识能力。现代社会已进入信息化的时代，不论是生产、科学研究和社会活动，还是个人的生活都离不开信息。

2. 数据

数据是描述信息的符号，符号的形式多种多样，如数值、文本、图形、图像、声音等类型的数据，用来反映不同类型的信息。利用计算机进行信息处理，就得把信息转换为计算机能够识别的符号，即用 0 和 1 两个编码符号来表示各种各样的信息。从这个意义上说，数据是信息的载体。

数据是信息的具体表现形式，信息是有一定意义的数据的集合，它们既关联又有一定的区别，如果我们把客观世界的某种现象或观念所反映的知识用一定的方法描述出来，那么前者是信息而后者是数据。因为信息和数据都是现象和概念所反映的知识。这是它们的共同点，因此当不需要严格区分时，我们把这两者不加区分地使用，如“数据处理”与“信息处理”是一样的。

信息以数据的形式处理，而处理的结果又可能产生新的信息。

1.1.2 数据处理

数据处理是指对各种形式的数据进行收集、存储、加工和传播的一系列活动的总和。其目的是从大量的、原始的数据中抽取、推导出对人们有价值的信息以作为行动和决策的依据；数据处理从根本上来说是为了借助计算机科学地保存和管理复杂的大量的数据，以便人们能方便而充分地利用这些宝贵的信息资源。

在数据处理的一系列活动中，数据的收集、组织、存储、传播、检索、分类等活动是基本环节，这些基本环节统称为数据管理或信息管理。数据处理中，对数据的加工、计算、打印报表等操作对不同的业务部门可以有不同的内容。数据库技术所研究的问题就是如何科学地组织和存储数据，如何高效地获取和处理数据。数据库技术是数据管理的最新技术。数据库系统是当代计算机系统的重要组成部分。

1.2 数据管理技术的发展

数据管理技术的发展与硬件(主要是外存)、软件、计算机应用的范围有密切的关联。数据管理技术的发展经过三个阶段：人工管理阶段、文件系统阶段和数据库系统阶段。

1.2.1 人工管理阶段

人工管理阶段是从 20 世纪 40 年代中期电子计算机问世到 50 年代中期，这一阶段计算机主要用于科学计算。在该阶段从硬件看，外存只有磁带、卡片、纸带，速度低、内存小，没有磁盘等直接存取的存储设备。从软件看，没有操作系统，没有管理数据的软件。数据处理方式是批处理。

在人工管理阶段，数据管理的特点是：

- (1) 数据不保存在机器中。因为计算机主要用于科学计算，一般不需要将数据长期保存。在计算时将数据输入，计算完毕将数据输出。
- (2) 没有软件系统对数据进行管理。程序员不仅要规定数据的逻辑结构，而且还要在程序中设计物理结构，包括存储结构、存取方法、输入输出方式等。因此程序中存取数据的子程序随着存储结构的改变而改变，使得数据与程序不具有独立性，这样不仅程序员必须花费许多精力在数据的物理布置上，而且一旦数据在存储结构上有一些改变、就必须修改程序。
- (3) 只有程序的概念，没有文件的概念，数据的组织方式必须由程序员自行设计。
- (4) 数据是面向应用的。一组数据对应一个程序，即使两个应用程序涉及某些相同的数据，也必须各自定义，所以程序与程序之间有大量重复数据，如图 1.1 所示。

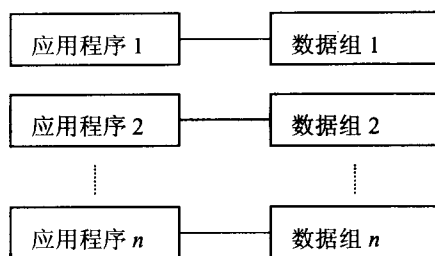


图 1.1 人工管理阶段程序与数据的关系

1.2.2 文件系统阶段

从 20 世纪 50 年代中期到 60 年代中期是文件系统阶段。这一阶段计算机不仅用于科学计算，还大量用于管理。计算机硬件比过去有了较大的发展，外存储器有了磁盘、磁鼓等直接存取的存储设备。在软件方面，操作系统中已经有了专门的数据管理软件，一般称为文件系统。处理方式上不仅能够进行文件批处理，而且能够实现联机实时处理。

在这一阶段数据管理有以下几个特点：

- (1) 数据可以长期保存在外存储设备上。由于计算机大量用于数据处理，数据需要长期保留在外存上进行反复处理，即进行查询、修改、插入和删除等操作。
- (2) 数据的逻辑结构与物理结构有了区别。由于有了数据管理软件，程序和数据之间由软件提供存取方法进行转换，有共同的用于数据查询、修改的管理模块，文件的逻辑结构与存储结构由系统进行转换，使程序与数据有了一定的独立性。这样程序员可以集中精力于算法，而不必过多地考虑物理细节。
- (3) 文件组织呈现多样化。由于已有了直接存取存储设备，也就有了索引文件、链接文件和直接存取文件等。
- (4) 数据不再属于某个特定的程序，可以重复使用。但文件结构的设计仍然是基于特定的用途，程序基于特定的存储结构和存取方法，因此程序与数据结构之间的依赖关系并未根本改变。

在文件系统阶段，由于具有设备独立性，因此改变存储设备，不必改变应用程序。但这只是初级的数据管理，还未能彻底体现数据的逻辑结构独立于数据的物理存储结构的要求。在数据的物理结构需要修改时，仍然需要用户修改应用程序。

随着数据管理规模的扩大，数据量急剧增加，文件系统显露出以下缺陷。

1. 数据冗余度大

当相同的数据存在多份时，称为数据冗余。文件系统文件基本上是对应于某个应用程序的，也就是说，数据还是面向应用的。即使不同的应用程序所需要的数据有部分相同时，也必须建立各自的文件，而不能共享相同的数据，因此数据冗余度大，浪费存储空间，并且由于相同数据的重复存储、各自管理，给数据的修改和维护带来了困难，容易造成数据的不一致性。

2. 数据和程序缺乏独立性

文件系统中的文件是为某一特定应用服务的，文件的逻辑结构对该应用程序来说是优化的。一旦数据的逻辑结构改变，则必须修改应用程序，修改文件结构的定义。而应用程序的改变，也将影响文件数据结构的改变，使得数据和程序缺乏独立性。这个时期程序与数据的关系如图 1.2 所示。

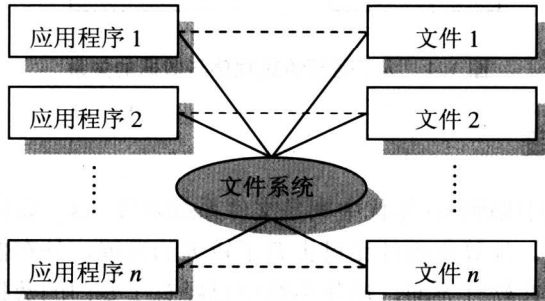


图 1.2 文件系统阶段程序与数据的关系

3. 数据间关联弱

文件与文件之间是独立的，文件之间的关联必须通过程序来完成。文件系统是一个无弹性的、无结构的数据集合，不能反映现实世界事物之间的关联。

随着人们对数据处理需求的增加，也随着计算机科学的不断发展，如何能对数据进行有效、科学、正确、方便的管理就成为人们的迫切需求。针对文件系统的缺陷，人们逐步发展了以统一管理和共享数据为主要特征的数据库管理系统。

1.2.3 数据库系统阶段

到 20 世纪 60 年代后期，计算机的软硬件进一步得到发展，已配备了速度高、容量大的磁盘，各种软件系统进一步完善，而且需要管理的数据量急剧增加。人们在数据管理方面已积累了丰富经验，数据管理技术研究取得了很大进展，为数据库系统的研究提供了良好的技术基础。

1968 年美国 IBM 公司成功开发了世界上第一个数据库管理系统 IMS(Information Management System); 1969 年美国 CODASYL 委员会(Conference On Data System Language)的 DBTG(DataBase Task Group)小组公布了它的研究成果 DBTG 报告; 1970 年 IBM 公司的研究员 E. F. Codd 发表了题为《大型共享数据库数据的关系模型》等一系列关系数据库论文。这三大事件标志着数据处理已进入了数据库技术的新时代。

20 世纪 70 年代以来，数据库技术得到迅速发展，开发了许多有效的产品并投入运行。数据库系统克服了文件系统的缺陷，提供了对数据更高级更有效的管理。

数据库系统阶段的管理方式具有以下特点。

1. 面向全组织的复杂的数据结构

这就要求在描述数据时不仅要描述数据本身，还要描述数据之间的关联。文件系统中

尽管记录内部已经有了某些结构,但记录之间是没有关联的,是孤立的。因此数据的结构化是数据库的主要特征之一,也是数据库与文件系统的根本区别。

2. 数据冗余度小容易扩充

由于数据库是从整体观点来看待和描述数据的,数据不再是面向某一应用,而是面向整个系统,这就大大减少了数据的冗余度,既节约存储空间,减少存取时间,又可避免数据之间的不相容性和不一致性。

3. 具有较高的数据和程序的独立性

所谓数据的独立性,就是应用程序不必因为数据存储结构的变化而修改,即应用程序与数据的结构之间不存在依赖关系,这是数据库系统所努力追求的一个目标。数据库系统结构之所以复杂,这是一个重要的原因。数据库系统的数据独立性分为两级:

- (1) 物理独立性。数据库物理结构的变化(如物理设备的更换、物理位置的变化、存取方法的改变等),不影响数据库的逻辑结构,从而也就影响不到应用程序,不会导致应用程序的修改。
- (2) 逻辑独立性。数据库逻辑结构的变化(如数据定义的修改,新数据类型的增加,数据间关联的变更等),不会影响到用户原有应用程序的修改。

这两种独立性统称为数据独立性。数据独立性的目的,就是使应用程序尽可能不受数据的影响。这两种数据独立性是靠数据库管理系统来实现的,从而大大减轻了程序员的负担。

4. 统一的数据控制功能

数据库是系统中各用户的共享资源。计算机的共享一般是并发的,即许多用户同时使用数据库。因此,系统必须提供以下三个方面的数据控制功能。

- (1) 数据的安全性控制。数据的安全性是指保护数据以防止不合法的使用造成数据的泄密和破解。这就要求采取一定的安全保密措施。例如,系统用检查口令或其他手段来检查用户身份,合法用户才能进入数据库系统。同时,系统要求提供用户访问级别和数据存取权限,当用户对数据库执行操作时,系统自动检查用户能否执行这些操作,检查通过后才执行允许的操作。
- (2) 数据的完整性控制。数据的完整性是指数据的正确性、有效性与相容性。系统提供必要的功能,保证数据库中的数据在输入、修改过程中始终符合原来的定义和规定。例如:学生性别只能是男或女;学号是惟一的等。

此外,当计算机软、硬件发生故障而破坏了数据或当对数据库数据的操作发生错误时,系统能进行应急处理,把数据库恢复到正确状态。

- (3) 并发控制。当多个用户的并发进程同时存取、修改数据库时,可能会发生互相干扰而得到错误的结果并使数据库的完整性遭到破坏,因此必须对多用户的并发操作加以控制、协调。

5. 数据的最小存取单位是数据项

既可以存取数据库中某一个数据项或一组数据项,也可以存取一个记录或一组记录。

这个阶段程序和数据的关系如图 1.3 所示。

数据库是一个通用化的综合性数据集合，它可以供各种用户共享且具有最小的数据冗余度和较高的数据与程序的独立性。由于多种程序并发地使用数据库，为了能有效及时地处理数据，并保证数据库的安全性和完整性，必须有一个软件系统——数据库管理系统 DBMS(DataBase Management System)在建立、运用和维护时对数据库进行统一控制。

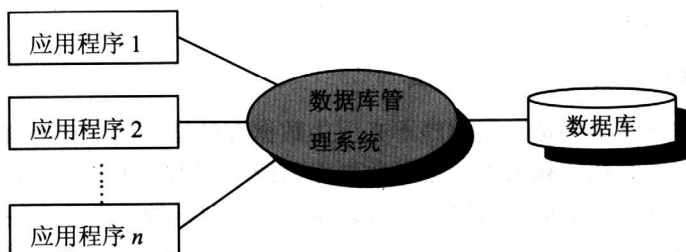


图 1.3 数据库系统阶段程序与数据的关系

1.3 数据模型

1.3.1 数据描述的三个领域

对用户来说，数据库是模拟现实世界中某些企业活动的信息集合。数据库中所存储的数据来源于现实世界的信息流，是用来描述现实世界中一些事物的某些方面的特征及其相互关联的。在处理这些信息之前，必须先分析它，选择一种方法描述这些待处理对象，并将这种描述转换成计算机所能接收的数据形式。在数据处理中，数据描述将涉及不同的范畴。从事物的特性到计算机中的数据表示，涵盖了三个领域：现实世界、信息世界和机器世界。

1. 现实世界

现实世界是指存在于人脑之外的客观世界，泛指客观存在的事物及其相互间的关联。一个实际存在并且可以识别的事物称为客观事物。客观事物可以是一个具体的事物，如一个学生、一台计算机，一本书等，也可以是一个抽象的事物，如一次比赛、一次借书等。

每个客观事物都有自己的特征，用以区别于其他客观事物，如学生有姓名、性别、年龄、身高、体重等许多特征来标识自己，但是在研究客观事物时，往往只选择其中对研究有意义的特征。

把具有相同特征的客观事物称为同类客观事物，所有同类事物的集合称为总体。例如所有的“学生”、所有的“课程”都是一个总体。

所有这些客观事物是信息的源泉，是设计数据库的出发点。这些事物是数据库技术接触到的最原始的数据，数据库设计是对这些原始数据进行综合处理，抽取出数据库技术所需要的数据。

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印象和概念，产生认识，这就是信息，即进入信息世界。在信息世界中，每一个被认识的客观事物称为实体，这是具体事物在人们头脑中产生的概念，是信息世界的基本单位。另外，客观事物的特征在头脑形成的知识称为属性。所以属性是事物某一方面的特征，即属性是反映实体的某一特征的。一个实体是由它所有的属性表示的。例如，一本书是一个实体，可以由书号、书名、作者、出版社、单价 5 个属性来表示。

在信息世界里，主要研究的不是个别的实体，而是它们的共性。把具有相同属性的实体称为同类实体，同类实体的集合为实体集，例如，所有的男学生组成了男学生实体集。能惟一标识每个实体的属性或属性集称为实体标识符，例如，书的书号可以作为书的实体标识符。

3. 机器世界

信息世界中的有些信息，可以直接用数字表示，如学生的成绩、年龄，图书的书号等；有些是由符号、文字或其他形式来表示的。在计算机中，所有信息只能用二进制数表示，一切信息进入计算机时，必须是数据化的。可以说，数据是信息的具体表现形式。在计算机世界中涉及以下术语。

- 数据项。数据项是实体属性的数据表示，它是可以命名的最小信息单位，又称数据元素或字段。如职工的职工号、姓名等。
- 记录。记录是实体的数据表示，是数据项的有序集合。如一个职工就是一个记录，它由职工号、姓名、性别、职称等数据项组成。
- 文件。文件是实体集的数据表示，是同类记录的集合，如所有职工的登记表组成一个文件。
- 关键字。能惟一标识文件中每个记录的数据项或数据项的集合，称为记录的关键码(或“键”)。如职工的职工号可以作为职工记录的关键码。

由此可见，现实世界、信息世界、机器世界这 3 个领域是由客观到认识、由认识到使用管理的 3 个不同层次，而且后一领域是前一领域的抽象描述。

3 个领域之间的术语对应关系如图 1.4 所示。

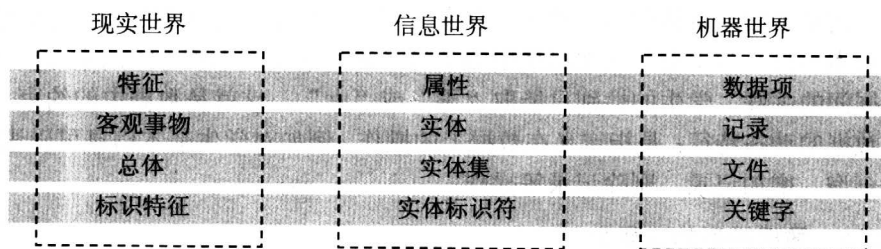


图 1.4 三个领域之间术语的对应关系

1.3.2 数据模型

1.3.2.1 数据模型的基本概念

模型是对现实世界特征的模拟和抽象。数据模型也是一种模型，它是对现实世界数据特征的抽象。数据库是某个企业、组织或部门所涉及的数据的综合，它不仅反映数据本身的内容，而且反映数据之间的关联。在数据库系统的形式化结构中如何抽象、表示、处理现实世界中的数据 and 信息呢？在数据库中是用数据模型来对现实世界进行抽象的。数据模型是数据库系统中用于提供信息表示和操作手段的形式构架。

数据模型应满足三个方面的要求：一是能比较真实地模拟现实世界；二是容易被人们理解；三是便于在计算机上实现。用一种模型来很好地满足这三方面的要求在目前是困难的。在数据库系统中针对不同的使用对象和应用目的，采用不同的数据模型。

不同的数据模型是提供了不同的模型化数据和信息的工具。根据模型应用的不同目的，可以将模型分为两类，它们属于两个不同的层次：一是概念数据模型(也称信息模型)，一是结构数据模型。

概念数据模型用于信息世界的建模，它是从数据的语义视角来抽取模型并按用户的观点对数据和信息建模。这类模型强调语义表达能力，要能够较方便、直观地表达应用中的各种语义知识。这类模型应当概念简单、清晰、易于用户理解，它是现实世界到信息世界的第一层抽象，是用户和数据库设计人员之间进行交流的语言。因此这类模型主要用于数据库的设计阶段，常用的概念数据模型是实体关联模型。

结构数据模型用于机器世界，它是从数据的组织层次来描述数据并按计算机系统的观点对数据建模。这类模型通常需要有严格的形式化定义，而且常常会加上一些限制或规定，以便于在机器上的实现。它通常有一组严格定义了语法和语义的语言，可以使用它来定义、操纵数据库中的数据。它是信息世界到机器世界的第二层抽象，主要用于数据库管理系统(DBMS)的实现，包括层次模型、网状模型、关系模型和面向对象模型等。

1.3.2.2 数据描述

数据的描述包括两个方面。

- 数据的静态描述：包括数据的基本结构、数据间的关联和数据中的约束。例如学生的基本信息包括学号、姓名、性别、年龄、专业，这些都是数据静态结构中的基本信息。学生进行选课时使用的学号必须是学生基本信息中的学号，这就是数据间的关联。学生的性别只能取“男”或“女”，这就是数据中的约束。
- 数据的动态特征：是指定义在数据上的操作。例如对学生基本信息可以进行修改、查询、增加记录、删除记录等操作。

1.3.2.3 数据模型的三要素

数据模型是严格定义的概念集合。这些概念精确地描述系统的静态特性、动态特性和完整性约束条件。因此，数据模型通常由数据结构、数据操作和完整性约束三部分组成。