



教育部财政部职业院校教师素质提高计划职教师资培养资源开发项目
《软件工程》专业职教师资培养资源开发 (VTNE034)

软件工程专业职教师资培养系列教材

数据库设计与开发

钱 进 常玉慧 叶飞跃 主编



科学出版社

软件工程专业职教师资培养系列教材

数据库设计与开发

钱 进 常玉慧 叶飞跃 主编

科学出版社

北 京

内 容 简 介

本书是在作者 2006 年出版的《数据库原理与应用》教材基础上编写而成, 选用最新的数据库管理系统 SQL Server 2012 和系统开发平台 Visual Studio 2012, 以“网上选课系统”案例贯穿全书, 全面介绍了数据库系统、数据库设计、数据库管理、数据库应用系统开发以及数据库教学设计等方面的基本原理、项目开发过程和教学案例。

本书共 8 章, 内容包括数据库系统概述、数据库设计、关系数据库标准语言 SQL、数据库应用系统简单开发、数据库应用系统高级开发、数据库管理、数据库新技术以及数据库课程教学设计。本书理论通俗易懂, 实践与时俱进。

本书可作为高等院校软件工程、计算机科学与技术、信息管理与信息系统等相关专业的本科数据库原理与应用课程教材, 也可作为高职高专相关专业数据库技术与应用类课程教材, 同时适用于数据库领域的初学者或作为数据库应用系统开发的专业技术人员的参考书。

图书在版编目(CIP)数据

数据库设计与开发 / 钱进, 常玉慧, 叶飞跃主编. —北京: 科学出版社, 2016.10
软件工程专业职教师资培养系列教材
ISBN 978-7-03-049740-6

I. ①常… II. ①钱… ②常… ③叶… III. ①数据库系统—师资培养—教材 IV. ①TP311.13

中国版本图书馆 CIP 数据核字 (2016) 第 200308 号

责任编辑: 邹 杰 / 责任校对: 郭瑞芝

责任印制: 张 伟 / 封面设计: 迷底书装

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京京华虎彩印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2016 年 10 月第 一 版 开本: 787×1092 1/16

2016 年 10 月第一次印刷 印张: 16

字数: 400 000

定价: 49.00 元

(如有印装质量问题, 我社负责调换)

《教育部财政部职业院校教师素质提高计划成果系列丛书》

《软件工程专业职教师资培养系列教材》

项目牵头单位：江苏理工学院

项目负责人：叶飞跃

项目专家指导委员会

主 任：刘来泉

副主任：王宪成 郭春鸣

成 员：（按姓氏笔画排列）

刁哲军	王继平	王乐夫	邓泽民	石伟平	卢双盈	汤生玲
米 靖	刘正安	刘君义	孟庆国	沈 希	李仲阳	李栋学
李梦卿	吴全全	张元利	张建荣	周泽扬	姜大源	郭杰忠
夏金星	徐 流	徐 朔	曹 晔	崔世钢	韩亚兰	

丛 书 序

《国家中长期教育改革和发展规划纲要（2010—2020年）》颁布实施以来，我国职业教育进入到加快构建现代职业教育体系、全面提高技能型人才培养质量的新阶段。加快发展现代职业教育，实现职业教育改革发展新跨越，对职业学校“双师型”教师队伍建设提出了更高的要求。为此，教育部明确提出，要以推动教师专业化为引领，以加强“双师型”教师队伍建设为重点，以创新制度和机制为动力，以完善培养培训体系为保障，以实施素质提高计划为抓手，统筹规划，突出重点，改革创新，狠抓落实，切实提升职业院校教师队伍整体素质和建设水平，加快建成一支师德高尚、素质优良、技艺精湛、结构合理、专兼结合的高素质专业化的“双师型”教师队伍，为建设具有中国特色、世界水平的现代职业教育体系提供强有力的师资保障。

目前，我国共有60余所高校正在开展职教师资培养，但由于教师培养标准的缺失和培养课程资源的匮乏，制约了“双师型”教师培养质量的提高。为完善教师培养标准和课程体系，教育部、财政部在“职业院校教师素质提高计划”框架内专门设置了职教师资培养资源开发项目，中央财政划拨1.5亿元，系统开发用于本科专业职教师资培养标准、培养方案、核心课程和特色教材等系列资源。其中，包括88个专业项目、12个资格考试制度开发等公共项目。该项目由42家开设职业技术师范专业的高等学校牵头，组织近千家科研院所、职业学校、行业企业共同研发，一大批专家学者、优秀校长、一线教师、企业工程技术人员参与其中。

经过三年的努力，培养资源开发项目取得了丰硕成果：一是开发了中等职业学校88个专业（类）职教师资本科培养资源项目，内容包括专业教师标准、专业教师培养标准、评价方案，以及一系列专业课程大纲、主干课程教材及数字化资源；二是取得了6项公共基础研究成果，内容包括职教师资培养模式、国际职教师资培养、教育理论课程、质量保障体系、教学资源中心建设和学习平台开发等；三是完成了18个专业大类职教师资资格标准及认证考试标准开发。上述成果，共计800多本正式出版物。总体来说，培养资源开发项目实现了高效益：形成了一大批资源，填补了相关标准和资源的空白；凝聚了一支研发队伍，强化了教师培养的“校-企-校”协同；引领了一批高校的教学改革，带动了“双师型”教师的专业化培养。职教师资培养资源开发项目是支撑专业化培养的一项系统化、基础性工程，是加强职教师资培养培训一体化建设的关键环节，也是对职教师资培养培训基地教师专业化培养实践、教师教育研究能力的系统检阅。

自2013年项目立项开题以来，各项目承担单位、项目负责人及全体开发人员做了大量深入细致的工作，结合职教师资培养实践，研发出很多填补空白、体现科学性和前瞻性的成果，有力推进了“双师型”教师专门化培养向更深层次发展。同时，专家指导委员会的各位专家以及项目管理办公室的各位同志，克服了许多困难，按照两部对项目开发工作的总体要求，为实施项目管理、研发、检查等投入了大量时间和心血，也为各个项目提供了专业的咨询和指导，有力地保障了项目实施和质量成果。在此，我们一并表示衷心的感谢。

编写委员会

2016年3月

前 言

数据库技术是数据管理的最新技术，是计算机科学的一个重要分支，与计算机网络、人工智能一起被称为计算机技术界三大热门技术，是现代化管理的重要工具。随着 IT 技术的发展，数据库技术应用已从事务处理扩大到计算机辅助设计、人工智能、决策支持系统和网络应用等新的领域。

本书作为教育部“软件工程”本科专业职教师资培养资源开发项目的特色教材，在编写过程中将理论与实践紧密结合。作者在 2006 年编著《数据库原理与应用》教材的基础上，结合多年从事数据库课程教学经验，并汲取了其他同类教材的精华，又增加了实用性的数据库应用系统，力求使本教材体现“理论通俗易懂，实践与时俱进”，真正满足培养计算机应用型人才和软件工程职教师资的需要。

本书共 8 章，打破传统教材章节编写次序，从网上选课系统需求出发，先简单介绍数据库基本知识，然后针对网上选课系统需求，开始设计数据库，接着利用 SQL 语言创建数据库和表，进行数据库应用系统简单开发。为了提升数据库应用系统性能和安全性，介绍存储过程、视图等，进行数据库应用高级开发。针对数据库应用系统使用过程中出现的状况，讲解数据库管理技术。最后，介绍数据库最新技术以及如何进行数据库课程教学。

本书理论部分由钱进、常玉慧、叶飞跃编写，实践部分由钱进和习海旭共同编写。

本书在编写时参考了中国期刊网相关教改论文和书籍资料，谨向有关作者表示感谢。

由于编者水平有限，书中难免存在不足之处，敬请广大读者批评指正。编者 E-mail: qjqlqyf@163.com。

编 者

2015 年春

目 录

丛书序

前言

第 1 章 数据库系统概述.....1

1.1 教学案例——网上选课系统.....1

1.1.1 问题的引出.....1

1.1.2 网上选课系统简介.....1

1.2 数据库系统基本概念.....2

1.2.1 数据与数据管理.....2

1.2.2 数据管理技术.....3

1.2.3 数据库系统.....7

1.3 数据模型.....11

1.3.1 数据模型及数据建模.....11

1.3.2 信息的三种世界.....12

1.3.3 概念模型与 E-R 图.....13

1.3.4 常见的逻辑数据模型.....15

1.4 关系模型.....17

1.4.1 关系模型的数据结构.....17

1.4.2 关系模型的数据操作.....19

1.4.3 关系模型的数据约束.....23

1.5 数据抽象与数据库的三级模式.....24

1.5.1 数据抽象.....24

1.5.2 数据库的三级模式结构.....25

1.5.3 数据库的二级映像功能与 数据独立性.....26

1.6 案例实践——网上选课系统.....27

1.7 项目实践——图书馆管理系统.....27

1.8 本章小结.....27

1.9 习题.....28

第 2 章 数据库设计.....29

2.1 教学案例——网上选课系统 数据库设计.....29

2.1.1 问题的引出.....29

2.1.2 网上选课系统数据库设计.....29

2.2 数据库设计概述.....30

2.2.1 数据库设计的基本步骤.....30

2.2.2 数据库设计的主要内容.....31

2.3 需求分析.....32

2.3.1 需求分析调查的内容.....32

2.3.2 需求分析调查的方法.....33

2.3.3 需求分析的具体步骤.....33

2.3.4 数据流图和数据字典.....34

2.4 概念结构设计.....38

2.4.1 概念结构概述.....39

2.4.2 采用 E-R 模型方法的概念 结构设计.....40

2.4.3 局部概念结构设计.....45

2.4.4 全局概念结构设计.....46

2.5 逻辑结构设计.....50

2.5.1 E-R 图向关系模型的转换.....50

2.5.2 函数依赖.....53

2.5.3 范式.....55

2.5.4 关系模式的规范化.....59

2.5.5 模式评价与优化.....63

2.5.6 设计用户子模式.....64

2.6 数据库的物理设计.....65

2.7 数据库的实施.....66

2.8 数据库运行与维护.....68

2.9 案例实践——网上选课系统数据 库设计.....69

2.9.1 网上选课系统数据库概念 结构设计.....69

2.9.2 网上选课系统数据库逻辑 结构设计.....72

2.10 项目实践——图书馆管理系统 数据库设计.....75

2.11 本章小结.....75

2.12 习题.....75

第3章 关系数据库标准语言 SQL	76
3.1 教学案例——网上选课系统 SQL 操作	76
3.1.1 问题的引出	76
3.1.2 网上选课系统成绩查询 SQL 操作	76
3.2 关系数据库标准语言 SQL 概述	77
3.2.1 SQL 的发展历程	77
3.2.2 SQL 语言基本知识	77
3.2.3 SQL 数据库的体系结构	79
3.2.4 SQL 的组成	79
3.3 SQL 的数据定义	80
3.3.1 数据库的定义、修改和删除	80
3.3.2 基本表的创建、修改和删除	83
3.3.3 SQL 数据更新	87
3.4 SQL 的数据操纵	90
3.4.1 SELECT 语句	91
3.4.2 单表查询	92
3.4.3 连接查询	95
3.4.4 嵌套查询	97
3.4.5 集合查询	101
3.5 视图	102
3.5.1 视图的作用	102
3.5.2 创建视图	103
3.5.3 管理视图	105
3.5.4 查询视图	106
3.5.5 更新视图	106
3.6 索引	107
3.6.1 索引的分类	107
3.6.2 创建索引	107
3.6.3 管理索引	108
3.7 案例实践——网上选课系统 SQL 操作	109
3.7.1 网上选课系统数据表的创建	109
3.7.2 网上选课系统数据维护	110
3.7.3 网上选课系统查询操作	112
3.7.4 网上选课系统视图操作	112
3.8 项目实践——图书馆管理系统数据表的创建与查询	113
3.9 本章小结	113
3.10 习题	113
第4章 数据库应用系统简单开发	115
4.1 教学案例——网上选课系统简单开发	115
4.1.1 问题的引出	115
4.1.2 网上选课系统简单开发	115
4.2 数据库应用系统简单开发概述	116
4.3 ADO.NET 访问 SQL SERVER 技术	117
4.3.1 ADO.NET 概述	118
4.3.2 ADO.NET 对象模型	118
4.3.3 ADO.NET 访问数据库方式	120
4.3.4 ADO.NET 对象及其编程	121
4.3.5 ADO.NET 访问数据库步骤	125
4.4 案例实践——网上选课系统主要功能实现	127
4.4.1 网上选课系统开发环境	127
4.4.2 网上选课系统主框架布局设计	127
4.4.3 网上选课系统数据库连接信息设置	127
4.4.4 网上选课系统公共类	128
4.4.5 网上选课系统部分视图	134
4.4.6 网上选课系统部分界面	134
4.5 项目实践——图书馆管理系统简单开发	147
4.6 本章小结	147
4.7 习题	147
第5章 数据库应用系统高级开发	148
5.1 教学案例——网上选课系统学分自动更新	148
5.1.1 问题的引出	148

5.1.2 网上选课系统学分自动更新	148	6.4.2 完整性约束分类	178
5.2 数据库应用系统高级开发概述	148	6.4.3 完整性的定义与验证	181
5.3 存储过程	149	6.4.4 SQL Server 数据库完整性	182
5.3.1 存储过程概述	149	6.5 数据库的并发控制	182
5.3.2 存储过程的相关操作	149	6.5.1 事务及并发控制概述	182
5.4 触发器	153	6.5.2 封锁协议	186
5.4.1 触发器概述	153	6.5.3 封锁带来的问题	189
5.4.2 触发器的相关操作	154	6.5.4 并发调度的可串行性	190
5.5 SQL 程序设计	156	6.5.5 SQL Server 的并发控制	192
5.5.1 SQL 程序基本成分	156	6.6 数据库的恢复	196
5.5.2 SQL 程序流程控制语句	159	6.6.1 数据库恢复概述	196
5.5.3 游标	162	6.6.2 恢复的实现技术	196
5.5.4 SQL 程序实例	163	6.6.3 故障的种类及恢复策略	198
5.6 案例实践——网上选课系统学分自动更新代码实现	164	6.6.4 数据库镜像	200
5.6.1 网上选课系统存储过程设计	164	6.6.5 SQL Server 数据库的恢复	200
5.6.2 网上选课系统学分自动更新代码实现	165	6.7 案例实践——网上选课系统数据库安全性实现	202
5.7 项目实践——图书馆管理系统高级开发	167	6.7.1 用户权限实现	202
5.8 本章小结	168	6.7.2 角色权限实现	203
5.9 习题	168	6.8 项目实践——图书馆管理系统数据库安全性实现	204
第 6 章 数据库管理	169	6.9 本章小结	204
6.1 教学案例——网上选课系统用户登录	169	6.10 习题	205
6.1.1 问题的引出	169	第 7 章 数据库新技术	206
6.1.2 网上选课系统用户登录	169	7.1 教学案例——网上选课系统课程类关联规则	206
6.2 数据库管理概述	169	7.1.1 问题的引出	206
6.3 数据库的安全性管理	170	7.1.2 网上选课系统课程类关联规则	206
6.3.1 安全性概述	170	7.2 数据库新技术概述	207
6.3.2 自主存取控制	171	7.3 数据仓库与数据挖掘	208
6.3.3 强制存取控制	174	7.3.1 数据仓库的概念与结构	208
6.3.4 安全性的其他技术	175	7.3.2 数据仓库的设计与实现	213
6.3.5 SQL Server 安全认证模式	177	7.3.3 联机分析处理	214
6.4 数据库的完整性管理	178	7.3.4 数据挖掘	216
6.4.1 完整性概述	178	7.4 XML 数据库	221
		7.4.1 XML 概述	221
		7.4.2 XML 表示	222
		7.4.3 XML 数据模型	224

7.4.4 XML 数据库	229	8.2 数据库课程教学设计概述	233
7.5 案例实践——网上选课系统 数据挖掘	230	8.3 数据库课程教学标准	237
7.6 项目实践——图书馆管理系统 数据挖掘	231	8.3.1 数据库课程概述	237
7.7 本章小结	231	8.3.2 数据库课程内容和要求	237
7.8 习题	231	8.4 案例实践——网上选课系统 数据库教学设计	238
第 8 章 数据库课程教学设计	232	8.4.1 项目式数据库教学设计	238
8.1 教学案例——网上选课系统 SQL 查询教学设计	232	8.4.2 案例情景式数据库教学 设计	240
8.1.1 问题的引出	232	8.5 项目实践	243
8.1.2 网上选课系统 SQL 查询 教学设计	232	8.6 本章小结	243
		8.7 习题	243
		参考文献	244

第 1 章 数据库系统概述

数据库技术是计算机和软件工程学科中的一个重要分支，已经广泛应用到各行各业中。要想掌握好数据库系统技术，必须弄清数据、数据管理、数据库、数据模型、关系模型等专业术语，了解数据库的发展过程和数据库应用系统的特点，分清数据库、数据库管理系统以及数据库应用系统三者之间的关系，以便今后更好地进行数据库设计和开发工作。

1.1 教学案例——网上选课系统

1.1.1 问题的引出

随着学分制的普及，网上选课系统已成为高校信息管理系统中的重要组成部分。一般来说，每个同学根据自己的兴趣爱好选修不同的课程，从而获得相应的选修学分。开学某一天，班主任叫逸凡同学统计上学期全班同学选课情况，要求能够查阅每个学生的学号、姓名、所选修的课程和成绩，并计算平均分。逸凡冥思苦想，决定通过下面的步骤完成任务。

第一步，根据要收集的数据设计一个空白的表格，其格式可能如表 1-1 所示。

表 1-1 学生选课基本信息表

学号	姓名	课程 1	成绩 1	课程 2	成绩 2	课程 3	成绩 3	课程 4	成绩 4	平均分

第二步，按顺序在表格中添加信息，同时还要对里面的信息进行修改和删除。在填写过程中，需要不断查阅已填的学生选课信息，不能重复填写，不能遗漏。

第三步，在填写完所有信息后，需要对表格信息进行校对，若发现错误，则重复第二步。

完成上述任务时，逸凡无形中发现或多或少已用到了一些数据库技术。所设计的二维表类似于关系模型的数据结构，增删改信息和不断查阅则对应于关系模型的数据操纵，发现错误并校正则对应了关系模型的数据约束。采用何种方式处理和保存该信息，则对应于数据管理技术。同时，他还发现表 1-1 存在如下问题：①若选修课程门数不限，一部分同学选修课程门数很多时，表格将会太长；若少数同学只选修一两门课程时，表格中将出现大量空白；②若某一列不是存放同一门课程，可能无法统计某门课程选修人数；若某一列存放的是同一门课程，则又冗余存储；③需要统计一行中课程门数才能计算出选修课程平均分，比较麻烦。

1.1.2 网上选课系统简介

网上选课系统管理员通过输入自己的账号和密码，进入到网上选课系统主界面，如图 1-1 所示。网上选课系统主要涉及学院信息、教师信息、班级信息、学生信息、课程信息、开课

班信息、上课时间信息、选课信息、任课信息、先修课程、个人信息等。如此庞大复杂的信息，手工处理显然不合适了。尽管用户界面隐藏了访问数据库的细节，大多数人甚至还没有意识到他们正在和一个数据库打交道，然而访问数据库已经成为当今生活中不可缺少的一部分。网上选课系统需要使用数据库技术进行管理和存储这些信息。如何更好地处理和解决这些问题，首先必须理解数据库系统基本概念、数据模型和关系模型等知识点。

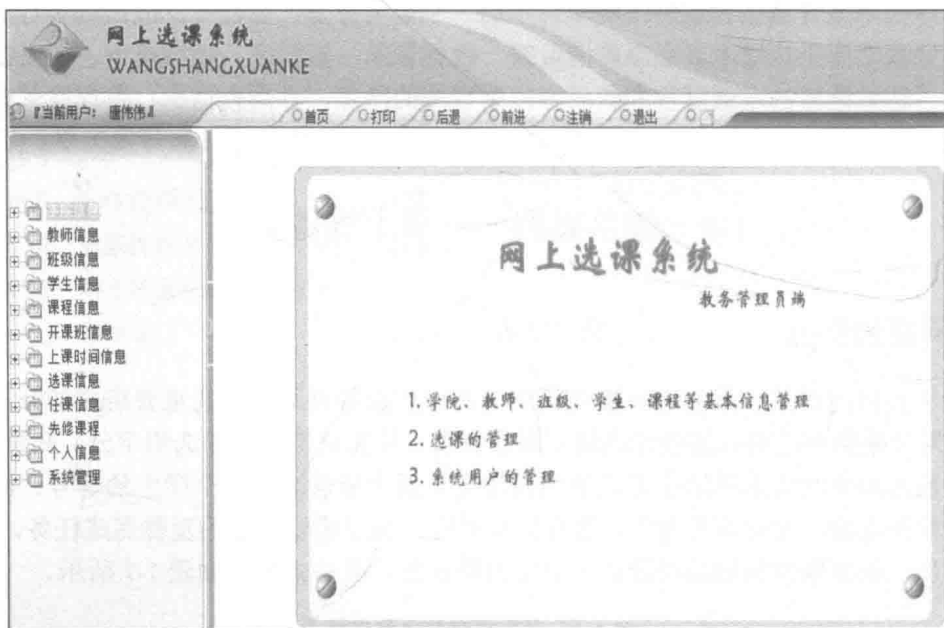


图 1-1 网上选课系统

1.2 数据库系统基本概念

1.2.1 数据与数据管理

1. 数据

描述事物的符号记录称为数据，如数值数据、文本数据、音频数据、视频数据等，能够被计算机所识别、存储和处理。

数据具有四个基本特征：

(1) 数据有“类型”和“值”之分。

例如，学生姓名的类型为字符型，其值为“张三”等。

(2) 数据受数据类型和取值范围的约束。

例如，学生年龄为整数类型，正常取值范围为 18~50。

(3) 数据有定性数据和定量数据之分。

例如，学生身高可以用“很高”来定性表示，也可以用 1.75 米来定量表示。

(4) 数据具有多种表示形式。

例如，学生性别可以用“男”或“女”来表示，也可以用 0 或 1 来表示。

数据是数据库中存储的对象,通过计算机对其进行适当的管理和操作。用数据描述客观事物,即用数据的表现形式表示事物的特征。例如,一个学生数据如学号,姓名,性别,年龄和班级等,描述了一个学生的有关特征和特性。这些特征也就是相应数据的语义。数据是需要解释的,也就是要对数据的含义或语义进行解释才有意义,因为数据的形式即符号本身并不能完全表达其内涵。对于以表格形式描述的对象,表头栏目名就是对表中数据的语义解释。

经过进一步加工后具有新的事实知识的数据称为信息。具体来说,信息是对数据进行加工、处理的结果,这种数据形式对接受者来说是有意义的,而且对当前和将来的决策具有明显的或实际的价值。在现实生活中,数据与信息两个术语常常被混淆,但是数据与信息在概念上还是有区别的。不是所有数据都能成为信息,只有抽取加工之后,具有新的事实知识的数据才能成为信息。例如,某经销部门看到的日销售额或月销售额,是经过累加一天或一个月的销售记录而得到的。这个数据,对于销售总管来说是一个原始数据,他需要根据各地区各商店的销售额来规划产品的市场销售计划和策略,从而提高其产品的市场占有率和利润。同样是这个数据,对于本地或本商店的销售人员也可以看做是信息,可用来规划下一步的促销行动。由此可以看出,数据和信息之间的关系类似于原料和成品之间的关系,同时又具有相对性。

2. 数据管理

数据处理是指将数据转换成信息的过程,即对数据进行收集、组织、整理、加工、存储和传播等一系列活动过程。数据处理的过程主要包括数据管理——收集信息、表示数据、组织和保存数据;数据加工——对数据进行变换、抽取和运算;数据传播——在空间或时间上以各种形式传播信息(不改变数据的结构、性质和内容)。通过数据管理获得数据;通过数据加工得到更有用的信息(以指导或控制人的决策行为或事物的变化趋势);通过数据传播,使更多的人共享信息。

数据管理具体就是指人们对数据进行收集、分类、组织、存储、检索和维护的一系列活动,它是数据处理的中心问题。数据组织管理技术的发展及其应用的广度和深度,极大地影响着人类社会发展的进程。现在都利用计算机进行数据管理,数据管理软件则成为不可缺少的一部分。

1.2.2 数据管理技术

伴随着计算机软硬件技术的发展,数据管理技术经历了人工管理、文件系统、数据库系统三个阶段。每一阶段的发展以数据存储冗余不断减小、数据独立性不断增强、数据操作更加简单方便为标志,各有特点。

1. 人工管理阶段

人工管理阶段主要是指20世纪50年代中期以前的这段时间。这一阶段计算机没有完整的操作系统,更没有管理数据的软件,计算机主要用于数值计算,数据处理的方式是批处理。因此,处理的数据量很小,数据无结构,数据依赖于特定的应用,而且数据间缺乏逻辑组织,缺乏独立性。

这一阶段的数据管理主要有如下特点。

1) 数据不能保存

当时计算机的硬件存储设备十分昂贵,计算机主要用于计算,计算完后,数据也不需要长期保存。在需要计算时,利用卡片、纸带等将数据输入,经过运算后输出结果,数据处理的过程就结束了,数据空间随程序空间一起被释放。

2) 数据不能独立

数据是程序的输入、输出，数据和程序作为一个不可分割的整体同时提供给计算机来计算。程序员不仅要知道数据的逻辑结构，也要规定数据的物理结构，程序员对存储结构，存取方法及输入、输出的格式有绝对的控制权，要修改数据必须修改程序。要对 100 组数据进行同样的运算，就要给计算机输入 100 个独立的程序，因为数据无法独立存在。

3) 数据是面向程序的

一组数据对应一个程序。不同程序的数据之间是相互独立、彼此无关的，即使两个不同程序涉及相同的数据，也必须各自定义，无法互相利用，互相参照。数据不但高度冗余，而且不能共享。

这一阶段应用程序与数据之间的关系如图 1-2 所示。

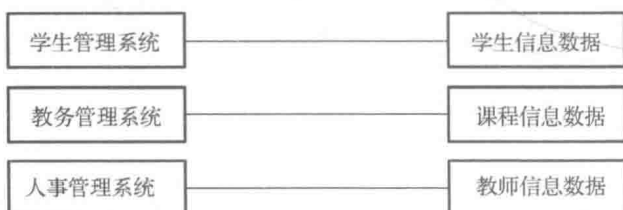


图 1-2 人工管理阶段程序与数据的关系

2. 文件管理系统阶段

文件管理系统阶段主要是指 20 世纪 50 年代后期到 20 世纪 60 年中期的这段时间。这一阶段的数据管理是把计算机中的数据组织成相互独立的数据文件，并可按文件的名字进行访问，实现对数据的存取。数据处理方式不仅有批处理，还有联机实时处理。

这一阶段数据管理主要有如下特点。

1) 数据长期保留

数据可以长期保留在外存上反复处理，即可以经常进行查询、修改和删除等操作。所以计算机大量用于数据处理。

2) 数据的独立性

由于有了操作系统，利用文件系统进行专门的数据管理，使得程序员在保存数据时，只需编写保存指令，读取数据时，只要给出文件名，而不必考虑计算机如何物理地保存数据，不必知道文件的具体存放地址。文件的逻辑结构和物理存储结构由系统进行自动转换，程序与数据之间有了一定的独立性，数据在存储上的改变不一定要引起程序的变化。例如，数据文件中有 100 条记录，使用某一个查询程序，而当文件中有 1000 条记录时，仍可使用同一个查询程序。

3) 实时处理

由于有了直接存取设备，也有了索引文件、链接存取文件、直接存取文件等存储机制，故可以采用顺序批处理或实时处理方式。数据的存取以记录为基本单位。

文件系统数据管理如图 1-3 所示。文件系统实现了记录内的结构化，但从文件的整体来看却是无结构的，其数据面向特定的应用程序。因此，数据共享性和独立性差，冗余度较大，管理和维护的代价也很大。

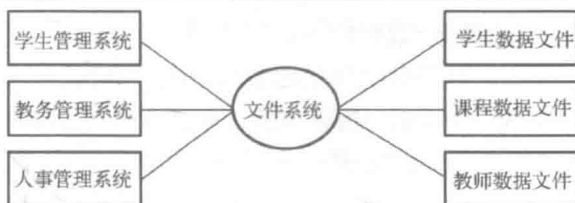


图 1-3 文件管理系统阶段程序与数据的关系

3. 数据库管理系统阶段

从 20 世纪 60 年代后期开始, 数据管理进入了数据库管理系统阶段。这一时期, 计算机管理的规模日益庞大, 应用越来越广泛, 数据量急剧增长, 要求共享数据的呼声越来越强。此时的计算机有了大容量磁盘, 硬件价格下降, 编制软件和维护软件的费用相对在增加。联机实时处理的要求更多, 并开始提出和考虑并行处理。在这样的背景下, 数据库技术应运而生, 出现了统一管理数据的专门软件——数据库管理系统(Database Management System, DBMS)。

从文件管理系统到数据库管理系统, 数据管理技术进行了质的飞跃。20 世纪 80 年代不仅在大、中型计算机上实现并应用了数据库管理软件, 如 Oracle、Sybase、Informix 等, 在微型计算机上也可使用数据库管理软件, 如常见的 Access、FoxPro 等软件, 使数据库技术得到广泛应用和普及。数据库不再只针对某一个特定的应用, 而是面向全组织, 具有整体的结构性, 共享性高, 冗余度小, 具有一定的程序与数据之间的独立性, 并且对数据进行统一的控制, 保证数据的完整性和安全性, 提高了数据管理效率。在应用程序和数据库之间, 由数据库管理系统(DBMS)把所有应用程序中使用的相关数据汇集起来, 按统一的数据模型, 以记录为单位存储在数据库中, 方便各个应用程序快捷地查询和使用数据。这一阶段管理数据的方法如图 1-4 所示。



图 1-4 数据库管理系统阶段程序与数据的关系

这一阶段数据库管理系统主要有如下特点。

1) 数据结构化

整体数据结构化是数据库的主要特征之一, 也是数据库管理系统与文件系统的本质区别。所谓整体结构化, 一是指数据不再仅仅针对某一个具体应用, 而是面向全组织的所有应用; 二是指数据不仅仅是数据内部结构化, 其整体也是结构化的, 数据与数据之间具有联系。数据的结构用数据模型描述, 无需程序定义和解释; 数据可以变长, 数据的最小存取单位是数据项, 从而数据库中实现的是数据的真正结构化。

2) 数据的共享性高, 冗余度低, 易扩充

数据库管理系统从整体角度看待和描述数据, 数据面向整个系统, 可以被多个用户、多个应用系统共享使用。数据共享一方面可以减少数据冗余, 节约存储空间, 另一方面避免了数据之间的一致性。

所谓数据的不一致性,是指同一数据的不同副本的值不一样。采用人工管理或文件系统管理时,不同的应用在修改数据时没有把其他副本进行同样的修改,这样造成了数据的不一致性。在数据库管理阶段,数据仅存储一份,减少了数据冗余造成的一致现象。

由于数据面向整个系统,不仅可以被多个应用共享使用,而且容易增加新的应用,这使得数据库系统易于扩充,可以适应各种用户需求。例如,可以选取整体数据的各种子集用于不同的应用,也可以选取不同的子集再加上一部分数据,满足新的应用需求。

3) 数据独立性高

数据独立是指数据的使用(即应用程序)与数据的说明(即数据的组织结构与存储方式)分离,使应用程序只考虑如何使用数据,而无须关心它们是如何构造和存储的,因而各方(在一定范围内)的变更互不影响。

数据独立性是指数据库中的数据与应用程序之间相互独立,是数据库领域中一个常用术语,包括数据的物理独立性和数据的逻辑独立性,是通过数据库系统的两层映像功能来实现的。

物理独立性是指用户的应用程序与存储在磁盘上的数据库中数据是相互独立的。当数据的物理存储改变了,通过数据库管理系统的相应改变以保持数据的逻辑结构不变,从而应用程序不用改变,因为应用程序要处理的只是数据的逻辑结构。

逻辑独立性是指用户的应用程序与数据库的逻辑结构是相互独立的,在数据的逻辑结构发生改变时,用户程序也可以不变。例如,在网上选课系统中,原来课程表中存储的字段有(课程编号,课程名称,学分),现增加(课程性质,总学时)等字段,虽然原来数据库的逻辑结构(课程编号,课程名称,学分)变为(课程编号,课程名称,学分,课程性质,总学时),但在查询学生选修课成绩时,不需要改变应用程序,整个系统仍然正常运行。

数据与程序的独立是把数据的定义从程序中分离出去,而数据的存取又由 DBMS 负责,从而简化了应用程序的编制,大大减少了应用程序的维护和修改。

4) 数据由 DBMS 统一管理和控制

数据库管理系统提供了数据共享,允许多个用户或多个应用共享数据库中的数据,还允许它们同时访问数据库中的同一数据。为了保证数据完整性和安全性,数据库管理系统还必须提供如下几个方面的数据库控制功能。

(1) 数据的安全性保护

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坏。每个用户只能在权限范围内对某些数据以某种方式进行使用和处理。

(2) 数据的完整性检查

数据的完整性是指数据的正确性、有效性和相容性,将数据控制在有效的范围内,或保证数据之间满足一定的关系。例如,对于百分制成绩,应该输入数值,而实际输入了字符,即不正确;应该输入 0~100 的数据,而实际输入了负数,即无效;在统计成绩时,优秀、良好、中等、及格、不及格的百分比之和应为 100%,而实际输入数据加起来大于 100%,即不相容。

(3) 并发控制

当多用户的并发进程同时存取、修改同一数据时,可能会发生相互干扰而得到错误的结果或使得数据库的完整性遭到破坏,因此必须控制和协调多用户的并发操作。例如,客户同时上网购买火车票,学生进行网上选课等操作都必须进行并发控制。

(4) 数据库恢复

数据库管理系统必须具有一个系统恢复子系统,能使系统从软硬件故障修复后将数据库

恢复到正确状态。例如，当数据库管理系统正在执行一个数据库的更新操作时，系统发生了硬件故障，使得更新中某些基本操作没能完成，从而使数据库处于错误状态。当系统故障发生后，系统恢复子系统必须对数据库实施修复，使其能恢复到正确状态。

综上所述，数据管理三个阶段技术特点的简单比较如表 1-2 所示。

表 1-2 数据管理三个阶段的比较

		人工管理	文件管理	数据库管理
背景	应用背景	科学计算	科学计算、管理	大规模管理
	硬件背景	无直接存取存储设备	磁盘、磁鼓	大容量磁盘
	软件背景	没有操作系统	有文件系统	有数据库管理系统
	处理方式	批处理	联机实时处理、批处理	联机实时处理、分布处理批处理
特点	数据的管理者	用户(程序员)	文件系统	数据库系统
	数据的针对者	特定应用程序	面向某一应用	面向整体应用
	数据的共享性	无共享	共享差，冗余大	共享好，冗余小
	数据的独立性	无独立性	独立性差	独立性好
	数据的结构化	无结构	记录有结构，整体无结构	整体结构化
	数据控制能力	应用程序自己控制	应用程序自己控制	由数据库管理系统提供数据安全性、完整性、并发控制和恢复

1.2.3 数据库系统

数据库(DataBase, DB)、数据库管理系统(DataBase Management System, DBMS)和数据库系统(DataBase System, DBS)是数据库技术中最基本、最常用的三个既有区别又有联系的术语。

1. 数据库

数据库是长期储存在计算机内的、有组织的、可共享的数据集合，也是现实世界中相互关联的大量数据及数据间关系的集合。数据库中的数据按一定的数据模型组织、描述和储存，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户共享。例如，在网上选课系统中可能同时有描述学生的数据(学号，姓名，性别，年龄，所属学院，院长)和教师的数据(工号，姓名，性别，出生年月，所属学院，院长)。在这两个数据中，所属学院、院长是重复的，称为冗余数据。在构造数据库时，由于数据可以共享，只要存储一份数据即可。

数据库具有如下特性：

- (1)数据库是具有逻辑关系和确定意义的数据集合。逻辑上无关的数据集合不能称为数据库。
- (2)数据库是针对明确的应用目标而设计、建立和加载的。每个数据库都具有一组用户，并为这些用户的应用程序服务。
- (3)一个数据库表示了现实世界的某些方面(称为小世界)。一个数据库应随其所表示的小世界的改变而改变。

2. 数据库管理系统

数据库管理系统是一个通用的软件系统，是位于用户和操作系统之间的数据管理软件,由