

► 高等学校应用型本科“十三五”规划教材
浙江省“十一五”重点教材建设项目



- 附有模拟试题(含参考答案)和教学大纲
- 提供教学课件和相关资料

数据库技术与设计



主 编 范剑波
副主编 刘良旭



西安电子科技大学出版社
<http://www.xduph.com>

高等学校应用型本科“十三五”规划教材



浙江省“十一五”重点教材建设项目

数据库技术与设计

主 编 范剑波

副主编 刘良旭

西安电子科技大学出版社

内 容 简 介

本书详细描述了数据库系统的基础知识,阐述了关系数据库系统的设计技术,介绍了 SQL Server 中的 T-SQL 语言,最后还给出了数据库技术的新进展。全书共 10 章,内容包括数据库系统概述、关系数据模型、数据库建模、关系数据库模式设计、关系数据库设计、SQL Server 2008 综述、SQL Server 2008 的 T-SQL I 和 T-SQL II、数据库系统新进展。建议教学时数为 64 学时。

本书每章配有三个典型案例、本章小结和练习题,书末附有课程模拟试题(含参考答案)和课程教学大纲,供读者参考。本书配套教材《数据库技术与设计》学习和实验指导在近期出版,相关的教学课件和资料可以向出版社索取或在相关网站上下载。

本书可作为计算机类专业及相关专业本科教材,也可供 IT 行业的科技人员参考。

图书在版编目(CIP)数据

数据库技术与设计/范剑波主编. —西安:西安电子科技大学出版社,2016.2

高等学校应用型本科“十三五”规划教材

ISBN 978-7-5606-3995-6

I. ① 数… II. ① 范… III. ① 关系数据库系统—高等学校—教材 IV. ① TP311.138

中国版本图书馆 CIP 数据核字(2016)第 011653 号

策 划 李惠萍 胡华霖

责任编辑 任倍萱 李惠萍

出版发行 西安电子科技大学出版社(西安市太白南路 2 号)

电 话 (029)88242885 88201467 邮 编 710071

网 址 www.xduph.com 电子邮箱 xdupfb001@163.com

经 销 新华书店

印刷单位 陕西华沐印刷科技有限责任公司

版 次 2016 年 2 月第 1 版 2016 年 2 月第 1 次印刷

开 本 787 毫米×1092 毫米 1/16 印 张 17

字 数 398 千字

印 数 1~3000 册

定 价 32.00 元

ISBN 978-7-5606-3995-6/TP

XDUP 4287001-1

如有印装问题可调换

西安电子科技大学出版社
高等学校应用型本科“十三五”规划教材

编审专家委员会名单

主任：鲍吉龙（宁波工程学院副院长、教授）

副主任：彭 军（重庆科技学院电气与信息工程学院院长、教授）

张国云（湖南理工学院信息与通信工程学院院长、教授）

刘黎明（南阳理工学院软件学院院长、教授）

庞兴华（南阳理工学院机械与汽车工程学院副院长、教授）

电子与通信组

组长：彭 军（兼）

张国云（兼）

成员：（成员按姓氏笔画排列）

王天宝（成都信息工程学院通信学院院长、教授）

安 鹏（宁波工程学院电子与信息工程学院副院长、副教授）

朱清慧（南阳理工学院电子与电气工程学院副院长、教授）

沈汉鑫（厦门理工学院光电与通信工程学院副院长、副教授）

苏世栋（运城学院物理与电子工程系副主任、副教授）

杨光松（集美大学信息工程学院副院长、教授）

钮王杰（运城学院机电工程系副主任、副教授）

唐德东（重庆科技学院电气与信息工程学院副院长、教授）

谢 东（重庆科技学院电气与信息工程学院自动化系主任、教授）

楼建明（宁波工程学院电子与信息工程学院副院长、副教授）

湛腾西（湖南理工学院信息与通信工程学院教授）

计算机大组

组长：刘黎明（兼）

成员：（成员按姓氏笔画排列）

刘克成（南阳理工学院计算机学院院长、教授）

毕如田（山西农业大学资源环境学院副院长、教授）

向 毅（重庆科技学院电气与信息工程学院院长助理、教授）

李富忠（山西农业大学软件学院院长、教授）
张晓民（南阳理工学院软件学院副院长、副教授）
何明星（西华大学数学与计算机学院院长、教授）
范剑波（宁波工程学院电子与信息工程学院党委书记兼副院长、教授）
赵润林（山西运城学院计算机科学与技术系副主任、副教授）
黑新宏（西安理工大学计算机学院副院长、教授）
雷亮（重庆科技学院电气与信息工程学院计算机系主任、副教授）

机电组

组 长：庞兴华（兼）

成 员：（成员按姓氏笔画排列）

丁又青（重庆科技学院机械与动力工程学院副院长、教授）
王志奎（南阳理工学院机械与汽车工程学院系主任、教授）
刘振全（天津科技大学电子信息与自动化学院副院长、副教授）
何高法（重庆科技学院机械与动力工程学院院长助理、教授）
胡文金（重庆科技学院电气与信息工程学院系主任、教授）

前 言

数据库技术是计算机科学技术中发展最快的领域之一，也是应用最广泛的技术之一，它已成为计算机信息系统与应用系统的核心技术和重要基础。

应用型本科教材的编写应该注重与工程应用相结合，注重与能力培养相联系，注重与培养目标相一致。在这样的思想指导下，我们对十多年数据库课程教学与科研工作的实践进行了总结，对教育部计算机示范性专业教学改革的工作进行了探索和实践。从 2004 年起，我们构建了数据库课程体系，该体系属于与 IT 职业岗位群直接相关的专业课程体系，是培养软件行业卓越工程师的核心内容，是培养学生核心能力的关键环节。数据库课程体系由程序设计类(面向对象程序设计、Java 程序设计*/C#程序设计*)、数据库设计类(数据库技术与设计、数据库技术课程设计)、数据库应用开发实践类(J2ME 移动应用开发*/ASP.NET Web 应用开发*、Java EE 应用开发*/.NET 应用开发*)和毕业设计提高类(开发网络数据库应用系统)构成(注：课程名后标注*的课程为计算机科学与技术专业培养计划中校企合作置换课程，课程 1/课程 2 表示学生选课时为二选一)。该课程体系在教学计划中处于关键的地位，在计算机类专业集中实践环节中占有相当大的比重，并对该专业学生顺利完成毕业设计提供了技术支持，已成为教学计划得以顺利实施的关键环节。

本书的特点主要体现在下述四个方面：

一是集理论、应用和系统于一体。理论的目标是最终能进行关系数据库的设计；应用的目标是能按软件工程的规范和数据库设计的步骤来进行数据库应用系统的设计；系统的目标是通过使用 SQL Server 来进行数据库应用系统的开发。

二是将 SQL Server 的内容结合到数据库原理中去讲解，避免了部分内容的重复，也避免了在原理部分讲解时无实际系统作为参照时的空对空现象。

三是在每章增加了三个典型案例分析的内容，使学生能对所学的知识有一个很好的复习。

四是增加了面向对象数据库系统、分布式数据库系统、UML 语言和数据挖掘等内容，展示了数据库系统的新进展和广阔的发展前景与应用趋势。

本书详细介绍了数据库的基本知识，阐述了关系数据库系统的理论、方法和设计技术，介绍了 SQL Server 中的 T-SQL 语言，给出了数据库应用系统设计的实例，最后还介绍了数据库技术的新进展。本书共 10 章，内容包括：数据库系统概述、关系数据模型、数据库建模、关系数据库模式设计、关系数据库设计、SQL Server 2008 综述、SQL Server 2008 的 T-SQL I 和 T-SQL II、数据库系统新进展。

本书每章配有三个典型案例，供读者复习之用；每章后附有本章小结和习题，供读者总结、练习之用；本书最后附有课程模拟试题(含参考答案)和课程教学大纲，供读者参考。

本书可作为普通高等学校计算机类专业及相关专业本科生教材,也可供 IT 行业的科技人员参考。

本书作为教材使用时,建议总学时至少为 64 学时,其中理论要求 48 学时,实验要求 16 学时,部分章节内容根据需要可以有选择地进行讲解。

本书是在作者之前出版的教材和教学项目的基础上完成的,其相关获奖情况如下:

- 《数据库理论与技术》(范剑波等编著)教材被评为 2010 年浙江省“十一五”重点教材建设项目(浙教高教〔2011〕10 号);

- “数据库课程体系的改革研究与实践”项目在 2009 年被评为浙江省高等学校教学成果二等奖(浙教高教〔2009〕153 号);

- “数据库课程体系教学团队”被评为浙江省高等学校省级教学团队(浙教高教〔2009〕212 号)。

本书由宁波工程学院范剑波教授任主编,刘良旭副教授任副主编。刘良旭老师编写了第 5.7、5.8、9.3、9.4 节以及附录 1 和附录 2 等内容,其余章节由范剑波老师编写,全书由范剑波老师统稿。

本书的出版要感谢宁波工程学院的大力支持,更要感谢出版社同志们的辛勤付出。

限于作者水平,书中难免存在错误和不妥之处,殷切期望广大读者给予指正。

作 者

2015 年 11 月于宁波

目 录

第一篇 数据库系统基础篇

第 1 章 数据库系统概述	2	练习题	26
1.1 数据库系统的应用和研究	2		
1.1.1 数据库系统的应用	2		
1.1.2 数据库系统的研究	3		
1.2 文件系统与数据库系统	4		
1.2.1 数据、信息和数据处理	4		
1.2.2 文件系统的特点与局限性	4		
1.2.3 数据库系统的发展及其特点	6		
1.3 数据库系统的组成和模式结构	9		
1.3.1 数据库系统的组成	9		
1.3.2 数据库系统的模式结构	11		
1.3.3 数据与程序的独立性	12		
1.4 数据描述和数据模型	13		
1.4.1 数据描述的领域	13		
1.4.2 数据联系的描述	14		
1.4.3 数据模型的概念	15		
1.4.4 数据模型及实例	17		
1.5 典型案例分析	21		
典型案例 1 高校组织结构概念 数据模型的设计(1)	21		
典型案例 2 学生在不同部门数据的 结构化	22		
典型案例 3 数据库管理系统的选择	23		
本章小结	26		
		第 2 章 关系数据模型	29
		2.1 关系模型的基本概念	29
		2.1.1 关系的通俗定义	29
		2.1.2 关系的数学定义	30
		2.1.3 关系模型	32
		2.2 关系代数	33
		2.2.1 传统的集合运算	34
		2.2.2 专门的关系运算	35
		2.2.3 关系代数表达式及实例	36
		2.3 关系数据库查询的优化	38
		2.3.1 查询优化问题的提出	38
		2.3.2 关系代数的等价变换	41
		2.3.3 查询优化的一般策略	41
		2.3.4 查询优化的步骤	42
		典型案例 1 关系代数表达式的查询	44
		典型案例 2 关系代数表达式的优化	45
		典型案例 3 关系代数表达式查询 时间的计算	46
		2.4 典型案例分析	44
		本章小结	47
		练习题	47

第二篇 数据库系统设计篇

第 3 章 数据库建模	50		
3.1 数据库建模概述	50		
3.2 E-R 图的设计	51		
3.2.1 E-R 图的设计方法	51		
3.2.2 E-R 图的设计原则	53		
3.3 E-R 图的子类和继承	54		
		3.3.1 E-R 图的子类	54
		3.3.2 E-R 图的继承	55
		3.4 E-R 图的约束建模	55
		3.4.1 实体集的键码	56
		3.4.2 单值约束	56
		3.4.3 完整性约束	57

3.4.4 其他类型约束	57
3.5 典型案例分析	58
典型案例 1 高校组织结构 E-R 图的设计(2)	58
典型案例 2 在线考试系统 E-R 图的设计	59
典型案例 3 图书网上销售系统 E-R 图的设计	60
本章小结	61
练习题	61
第 4 章 关系数据库模式设计	62
4.1 关系模式的存储异常和数据依赖	62
4.2 函数依赖的概念	63
4.2.1 函数依赖的定义	64
4.2.2 完全函数依赖和部分函数依赖	64
4.2.3 传递函数依赖	65
4.2.4 关系模式的键码	65
4.3 函数依赖的规则	66
4.3.1 三个推理规则	66
4.3.2 闭包的计算	67
4.4 关系的规范化	68
4.4.1 第一范式	69
4.4.2 第二范式	69
4.4.3 第三范式	70
4.4.4 BCNF 范式	71
4.4.5 多值依赖和第四范式	72
4.5 模式分解的优劣	74
4.5.1 模式分解的等价性	74
4.5.2 模式分解的规则和方法	76
4.6 案例分析	79
典型案例 1 产品订货系统关系数据库模式的设计	79
典型案例 2 在线考试系统关系数据库模式的设计	80

典型案例 3 图书网上销售系统的关系数据库模式的设计	80
本章小结	81
练习题	82
第 5 章 关系数据库设计	84
5.1 数据库设计概述	84
5.1.1 数据库设计的方法	84
5.1.2 数据库设计的特点	85
5.1.3 数据库设计的步骤	85
5.2 需求分析	86
5.2.1 需求分析的任务	86
5.2.2 需求分析的结构化分析方法	87
5.3 概念结构设计	89
5.3.1 概念结构设计任务	89
5.3.2 概念结构设计的方法与步骤	90
5.4 逻辑结构设计	99
5.4.1 逻辑结构设计任务	99
5.4.2 逻辑结构设计的方法与步骤	100
5.5 数据库的物理设计	103
5.6 数据库应用系统的实施与调优	103
5.7 案例分析	104
典型案例 1 某仓储超市 POS 系统关系数据库的设计	104
典型案例 2 某宾馆信息管理系统关系数据库的设计	108
典型案例 3 某公司活动信息采集系统关系数据库的设计	110
5.8 JDBC 的数据库连接方法	113
5.8.1 Windows 下 Java 开发环境的配置	113
5.8.2 Java 使用 JDBC 连接数据库的方法	114
本章小结	115
练习题	115

第三篇 数据库系统实践篇

第 6 章 SQL Server 2008 综述	118
6.1 SQL Server 2008 概述	118

6.2 SQL Server 2008 的安装	120
6.2.1 SQL Server 2008 的安装环境	120

6.2.2 SQL Server 2008 的安装过程.....	120	7.5.1 视图的建立	161
6.3 SQL Server 2008 管理工具	131	7.5.2 视图的查询和删除	162
6.3.1 SQL Server 2008 服务器的配置.....	132	7.6 SQL Server 的函数	163
6.3.2 SQL Server 2008 服务器的注册和 连接	133	7.6.1 内置函数的分类	163
6.3.3 SQL Server 2008 服务器的关闭和 启动	135	7.6.2 用户定义函数	164
6.3.4 SQL Server 2008 的常用工具.....	136	7.7 典型案例分析	165
6.4 典型案例分析	138	典型案例 1 SQL Server 2008 基本表 查询的应用	165
典型案例 1 SQL Server 2008 联机丛书的 查询	138	典型案例 2 SQL Server 2008 视图 查询的应用	168
典型案例 2 数据库系统管理员 sa 密码的 设定	139	典型案例 3 SQL Server 2008 用户定义 函数的应用	168
典型案例 3 SQL Server 2008 数据库的 附加和分离	140	本章小结	169
本章小结	142	练习题	169
练习题	142		
第 7 章 SQL Server 2008 的 T-SQL I	143	第 8 章 SQL Server 2008 的 T-SQL II	171
7.1 SQL Server 的数据库	143	8.1 SQL Server 的流程控制语言	171
7.1.1 数据库及其组成	143	8.1.1 批处理、脚本和变量	171
7.1.2 创建用户数据库	144	8.1.2 顺序结构语句	174
7.1.3 管理用户数据库	146	8.1.3 分支结构语句	175
7.2 SQL Server 的数据表	147	8.1.4 循环结构语句	178
7.2.1 数据表结构和内容	147	8.1.5 其他语句	179
7.2.2 创建用户数据表	149	8.2 SQL Server 的游标	183
7.2.3 管理用户数据表	151	8.2.1 游标的概念	183
7.3 SQL Server 的数据更新	153	8.2.2 游标的使用	185
7.3.1 数据插入	153	8.3 SQL Server 的存储过程和触发器	187
7.3.2 数据修改	154	8.3.1 存储过程的建立、执行和删除	187
7.3.3 数据删除	155	8.3.2 触发器的建立和删除	188
7.4 SQL Server 的数据查询	156	8.4 SQL Server 的数据库保护	190
7.4.1 数据基本查询	156	8.4.1 数据库的安全性	190
7.4.2 数据分组查询	157	8.4.2 数据库的完整性	192
7.4.3 多表连接查询	157	8.4.3 数据库的恢复技术	197
7.4.4 数据子查询	159	8.4.4 数据库的并发控制	202
7.4.5 附加子句	160	8.5 典型案例分析	207
7.5 SQL Server 的视图	161	典型案例 1 SQL Server 2008 游标的 应用	207
		典型案例 2 SQL Server 2008 触发器的 应用	208

典型案例 3 SQL Server 2008 完整性	本章小结	210
机制的应用	练习题	210

第四篇 数据库系统新进展篇

第 9 章 数据库系统新进展.....	9.4 数据挖掘.....	214	239
9.1 面向对象数据库系统	9.4.1 数据挖掘的定义	214	239
9.1.1 面向对象数据模型的基本概念	9.4.2 数据挖掘的任务	214	241
9.1.2 对象定义语言 ODL 概述	9.4.3 数据挖掘的流程和方法	218	241
9.1.3 面向对象数据库系统的发展前景	本章小结	224	244
9.2 分布式数据库系统	练习题	224	244
9.2.1 分布式数据库系统概述	附录	224	246
9.2.2 分布式数据库系统的查询处理和 优化	附录 1 “数据库技术与设计”课程模拟 试题及参考答案	226	246
9.2.3 分布式数据库系统的发展前景	附录 2 “数据库技术与设计”课程 教学大纲(仅供参考)	229	258
9.3 XML 语言	参考文献	230	262
9.3.1 XML 概述.....		230	
9.3.2 XML 语法.....		232	
9.3.3 XML 文档类型定义.....		235	

第一篇

数据库系统基础篇

第 1 章 数据库系统概述

本章导言

计算机诞生早期,其主要用于解决科学研究和工程设计中的数值计算问题,这个时期计算机的技术与应用只被少数科学家所拥有。随着时代的进步和科学的发展,计算机的应用迅速发展,涉及领域非常广泛。目前科学计算(或数值计算)、数据处理(或信息处理)、过程控制(或实时控制)、计算机辅助、人工智能(或智能模拟)和网络应用等技术已渗透到了社会的各行各业,正改变着传统的工作、学习和生活方式,推动着社会的发展。据统计,80%以上的计算机主要用于数据处理,这类工作决定了计算机应用的主导方向。数据库系统技术就是作为数据处理的一门新技术而迅速发展起来的。

本章重点

本章主要介绍数据库系统的应用和研究、文件系统和数据库系统、数据库系统的组成和模式结构、数据描述和数据模型,这些内容是学习后面章节的基础。本章学习的重点是掌握数据库系统的特点、数据库系统的模式结构、数据模型的概念等内容。

1.1 数据库系统的应用和研究

1.1.1 数据库系统的应用

数据库系统是在计算机系统中引入数据库技术以后所形成的系统,其应用范围非常广泛,以下是一些具有代表性的应用。

- (1) 银行业:用于存储客户的信息、账户、贷款以及银行的交易记录。
- (2) 航空业:用于存储订票和航班的信息。航空业是最先以地理分布的方式使用数据库的行业之一。
- (3) 大学:用于存储学生的信息、课程注册和成绩的信息。
- (4) 信用卡交易:用于记录信用卡消费的情况和产生每月清单。
- (5) 电信业:用于存储通话记录,产生每月账单,维护预付电话卡的余额和存储通信网络的信息。
- (6) 金融业:用于存储股票、债券等金融票据的持有、出售和买入的信息;也可用于存储实时的市场数据以便客户能够进行联机交易,公司能够进行自动交易。

(7) 销售业：用于存储客户、产品及购买信息。

(8) 联机的零售商：用于存储以上所述的销售数据，以及实时的订单跟踪，推荐品清单的生成，还有实时的产品评估的维护。

(9) 制造业：用于管理供应链，跟踪工厂中产品的生产情况、仓库和商店中产品的详细清单以及产品的订单。

(10) 人力资源：用于存储雇员、工资、所得税和津贴等信息。

正如以上所列举的，数据库系统的应用已经成为当今几乎所有企业、学校、行政和事业单位不可缺少的组成部分。

在 20 世纪最后的 40 年中，数据库的使用在所有的企业中都有所增长。在早期，很少有人直接和数据库系统打交道，尽管没有意识到这一点，但他们还是与数据库间接地打着交道。例如，通过打印的报表(如信用卡的对账单)或者通过代理(如银行的出纳员和机票预订代理等)间接与数据库发生关系；自动取款机的出现，使用户可以直接和数据库进行交互。

20 世纪 90 年代末的互联网革命急剧地增加了用户对数据库的直接访问。例如，当你访问一家在线书店，浏览一本书或一个音乐集时，其实你正在访问存储在某个数据库中的数据，并且当你确认了一次网上订购，你的订单也就保存在了某个数据库中；当你访问一个银行网站，检索你的账户余额和交易信息时，这些信息也是从银行的数据库系统中取出来的。

因此，尽管用户界面隐藏了访问数据库的细节，大多数人甚至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几乎每个人生活中不可缺少的组成部分。也可以从另一个角度来评判数据库系统的重要性。如今，像 Oracle 这样的数据库系统厂商是世界上最大的软件公司之一，并且在微软和 IBM 等这些有多样化产品的公司中，数据库系统的产品也是其产品线的一个重要组成部分。

1.1.2 数据库系统的研究

数据库系统的研究包括以下三个方面。

1) 数据库管理系统的研究

这方面的研究包括数据库管理系统(Data Base Management System, DBMS)应具有什么样的功能和应如何实现等问题的研究。当前，数据库管理系统的研究已从集中式数据库管理系统向分布式数据库管理系统(DDBMS)、知识库管理系统(KBMS)等方面延伸，以至延伸到适应各种应用领域。

2) 数据库理论的研究

这方面的研究主要围绕关系数据库理论、事务理论、逻辑数据库、面向对象数据库、知识库等方面的研究，探索新思想的表达、提炼、简化，最后使其为人们所理解；研究新算法以提高数据库效率。

3) 数据库设计方法及工具的研究

数据库设计的主要含义是在数据库管理系统的支持下，按照应用要求为某一部门或组织设计一个结构良好、使用方便、效率较高的数据库及其应用系统。目前，这一领域正在进行数据库设计方法、设计工具和理论的研究，数据模型和数据建模的研究，计算机辅助

数据库设计方法及其软件系统的研究, 数据库设计规范和标准的研究, 等等。

1.2 文件系统与数据库系统

1.2.1 数据、信息和数据处理

数据是数据库系统研究和处理的对象, 本质上是对信息的一种符号化表示。数据分为狭义和广义数据, 其中狭义数据可以是字符(含数字等)、文字等数据, 广义数据可以是声音、图像和视频等数据。总体上说, 凡是能输入到计算机中并能被计算机处理的东西都是数据。信息则是现实世界中各种事物的存在特征、运动形态以及不同事物间的相互联系等因素在人脑中的抽象反映, 进而形成概念。注意: 数据与信息是分不开的, 二者既有联系又有区别, 数据是信息的载体, 信息是数据的内涵。

数据处理是指对各种数据进行收集、存储、整理、分类、统计、加工、利用、传播等一系列活动的统称, 这些是数据处理的基本环节, 是任何数据处理任务必有的共性部分。数据处理的目的是从大量的、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; 目的之二是为了借助计算机科学地保存和管理复杂的大量数据, 以便人们能方便而充分地利用这些宝贵的信息资源。

数据库技术所研究的问题就是如何科学地组织和储存数据, 如何高效地获取和处理数据。数据库技术作为数据管理的最新技术, 目前已广泛应用于各个领域, 数据库系统已成为当今计算机系统的重要组成部分。

1.2.2 文件系统的特点与局限性

在 20 世纪 50 年代末到 60 年代末之间, 计算机不但用于科学计算, 而且已经大量用于处理数据, 而此时, 计算机外存储器有了磁鼓、磁盘等直接存取的存储设备, 这为计算机进行事务管理奠定了硬件基础; 与此同时, 随着数据结构设计和数据管理技术研究的软件技术的迅速发展, 出现了专门的管理数据的软件, 这就是所谓的“文件系统”阶段。

在文件系统阶段, 数据管理的主要特点可概括为如下几点:

(1) 数据以文件的形式保存在计算机的外存中, 用户可以随时通过程序对文件进行查询、修改和增删等处理。

(2) 文件系统有专用的数据管理软件, 它能对驻留在外存上的数据文件实施统一的管理, 这种专用数据管理软件构成了操作系统的一个重要组成部分。由于应用程序不再需要了解数据在存储介质上的实际地址, 因而大大减少了程序设计的工作量。

(3) 文件组织已日益多样化, 文件组织形式不再局限于顺序文件, 还出现了索引文件和链表文件等。因此, 对文件的访问形式既可以是顺序存取, 也可以是直接存取。

(4) 数据的存取基本上以记录为单位。

文件技术的上述特点使得数据管理在 20 世纪 60 年代中得到了充分的发展, 将计算机应用推向了一个新的高潮。但由于文件数据结构的设计仍然是基于特定用途的, 因此数据结构与程序之间的依赖关系并未根本改变, 从而限制了它的进一步发展。

在 20 世纪 60 年代中期以后, 计算机在数据处理领域的应用迅速发展, 由个别部门的应用逐步发展成多个部门的普遍应用, 由简单孤立的单项应用发展为彼此相关的复杂应用, 从而使管理的规模更加庞大, 数据量急剧增长, 共享性也不断增强。这就带来了数据管理上的一些新问题, 现举一例加以说明。

例 1.1 某学校的学生处、教务处和图书馆均要使用计算机对学生的有关信息进行管理, 但其各自处理的内容不同, 如用文件系统实现, 可按如下方式进行组织:

学生处要处理的信息包括学号、姓名、系名、年级、专业、年龄、性别、籍贯、政治面貌、家庭住址、个人简历、社会关系等。为此, 学生处的应用程序员必须定义一个文件 F1, 该文件结构中的记录应包括上述几个数据项。

教务处要处理的信息包括学号、姓名、系名、年级、专业、课名、成绩、学分等。显然, 教务处的应用程序员需定义一个文件 F2, 该文件结构中的每一记录包括以上几个数据项。

类似地, 当图书馆要记录和处理学生的有关图书借阅信息时, 其创建的文件 F3 应包括学号、姓名、系名、年级、专业、图书编号、图书名称、借阅日期、归还日期、滞纳金等。

当上述三个部门都使用计算机对学生的有关信息进行管理时, 就要在计算机的外存中分别保存 F1、F2 和 F3 三种文件, 可这三种文件中均有学生的学号、姓名、系名、年级和专业等信息, 因此重复的数据项达到了 1/3 以上, 数据冗余将会产生以下问题:

(1) 数据冗余不仅浪费存储空间, 更严重的是存在潜在的不一致性。由于数据存在多个副本, 所以当发生数据更新时, 很可能发生某些副本被修改而另一些副本被遗漏的情况, 从而使数据发生不一致现象, 影响数据的正确性和可靠性。比如, 某学生因故需从计算机科学与技术专业转到网络工程专业, 当学生处得到该信息后, 将该生所属的专业名改为网络工程, 因而 F1 文件中保存了正确的信息。但若教务处和图书馆没有得到此信息, 或者没有及时更改 F2 和 F3 文件, 这就造成了数据的不一致性。由于数据的使用价值很大程度上依赖其可靠性, 所以, 这种不一致的后果是不可忽视的。可以想象, 当这种情况发生在军事、航天、金融等行业部门时, 其后果是非常严重的。

(2) 文件是为某一特定应用服务的, 每个文件对应于一个应用程序(见图 1.1), 这就造成了应用程序与数据结构过分地互相依赖, 而且系统很难扩充。一旦逻辑结构发生改变, 就必须修改应用程序中文件结构的定义; 反之, 应用程序的改变, 也会影响文件的数据结构的改变。例如, 学生处在管理学生信息时发现只登记学生年龄是不够的, 还应记录学生的出生日期, 这样必须在 F1 的文件结构中增加一个数据项, 同时还要修改

涉及这个数据项的应用程序; 另一方面, 若用户原先用 VB 语言编制访问 F1 文件的应用程序, 现在想改用 C 语言访问, 那么, F1 文件的数据结构也必须随之改变。

(3) 文件系统缺乏对数据操作进行控制的方法, 对于数据的安全性、完整性和并发访问异常等问题的控制, 完全要用户程序自己负责, 这使得应用程序的编制相当繁琐。

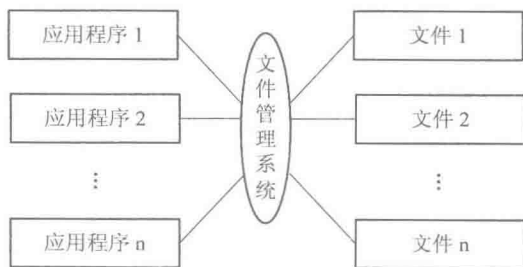


图 1.1 文件系统阶段应用程序与数据之间的关系

综上所述,传统的文件系统有许多缺点,不能满足人们的要求,因此迫切需要新的数据管理技术来实现对数据的共享,实现数据与程序的独立性,并提供安全性、完整性和并发控制的功能。也就是说,在操作系统之上必须有一个软件系统,即数据库管理系统(DBMS),它在数据库的建立、运用和维护时,对数据库进行统一的控制,这就是数据库系统。

1.2.3 数据库系统的发展及其特点

1. 数据库系统的发展

数据库系统从 20 世纪 60 年代末产生到现在只有 40 多年历史,但其发展速度很快、使用范围很广。以下几个事件标志着数据库系统日益成熟的过程。

(1) 1969 年,IBM 公司研制、开发了 DBMS 的商品化软件 IMS(Information Management System),它是层次模型数据库系统的典型代表(第一代数据库)。

(2) 20 世纪 60 年代末至 70 年代初,美国数据系统语言协商会下属的数据库任务组(Data Base Task Group, DBTG)对数据库方法进行了系统的研究、讨论,并提出了 DBTG 报告。该报告确定并建立了数据库系统的许多概念、方法和技术,它是网状模型数据库系统的典型代表(第一代数据库)。

(3) 1970 年,IBM 公司 San Jose 研究室的 E.F.Codd 在美国计算机学会会刊(Communication of the ACM)上发表了著名论文《大型共享数据库数据的关系模型》,提出了数据库的关系模型,开创了数据库关系方法和关系数据理论的研究,为关系数据库技术奠定了理论基础。1974 年,IBM 公司提出了一种基于关系方法实现对数据库存取的 SQL(Structured Query language)语言,并先后成功研制了能实现 SQL 语言的关系数据库管理系统原型 System R 和产品 DB2,表明了关系模型数据库管理系统获得了成功。目前,有影响的基于关系模型的数据库管理系统有许多,国际上较流行的有 Oracle、SQL Server、MySQL、Sybase、DB2 等,国内的有北京人大金仓信息技术股份有限公司的 Kingbase、武汉达梦数据库有限公司的 DM、天津南大通用数据技术股份有限公司的 GBase、北京神舟航天软件技术有限公司的神舟 OSCAR 等。它们都是关系模型数据库系统的典型代表(第二代数据库)。目前关系模型数据库系统已经淘汰了网状模型和层次模型数据库系统,成为当今最为流行的商用数据库系统。E.F.Codd 本人作为关系数据库的创始人和奠基人,获得了 1981 年 ACM 图灵奖。

(4) 从 20 世纪 80 年代以来,数据库系统在商业领域的巨大成功刺激了其他领域对数据库技术需求的迅速增长,这些新的领域为数据库应用开辟了新的天地。另一方面,在应用中提出的一些新的数据管理的需求也直接推动了数据库系统的研究与发展,尤其是面向对象数据库系统(Object Oriented DataBase System)的研究与发展。1990 年,美国高级 DBMS 功能委员会发表了《第三代数据库系统宣言》一文,提出了第三代 DBMS 应具有的三个基本特征:第三代数据库系统应支持数据管理、对象管理和知识管理;第三代数据库系统必须保持或继承第二代数据库系统的技术;第三代数据库系统必须对其他系统开放。

数据库技术与网络通信技术、人工智能技术、面向对象程序设计技术、并行计算技术等互相渗透、互相结合,成为当前数据库技术发展的主要特征。我国数据库系统的发展主要以 1977 年的第一次数据库系统研讨会为标志,以后几乎每年由中国计算机学会数据库专