



数据库技术与应用

主 编 瞿有甜 副主编 何小卫
主 审 胡维华

数据库技术与应用

主 编 瞿有甜

副主编 何小卫

编著者 郭艳华 柯海丰 方 宁



ZHEJIANG UNIVERSITY PRESS

浙江大学出版社

内容简介

本书详细介绍了数据库技术与应用的基本原理、方法及应用实践。内容包括数据库技术基础和新趋势、关系代数、关系数据库语言 SQL、范式及规范化方法、关系数据库的设计、数据库安全和保护、网络数据库技术与 XML 数据库等。

本书理论知识介绍简练清晰,应用实践方法介绍具体详实,较好地做到了理论与实践并重,突出能力培养的目的。

本书既可作为计算机相关专业的本、专科生教材,也可作为其他工科类非计算机专业本科生的教材;可作为国家和浙江省计算机等级考试数据库技术的教材或参考用书,还可作为信息系统开发与维护的工程技术人员、管理人员的参考用书或培训教材。

图书在版编目(CIP)数据

数据库技术与应用/瞿有甜主编. —杭州:浙江大学出版社,2010.3
ISBN 978-7-308-07438-4

I. ①数… II. ①瞿… III. ①数据库系统 IV.
①TP311.13

中国版本图书馆 CIP 数据核字(2010)第 045579 号

数据库技术与应用

主编 瞿有甜 副主编 何小卫

丛书策划 希言 黄娟琴

责任编辑 许佳颖

出版发行 浙江大学出版社

(杭州市天目山路 148 号 邮政编码 310007)

网址: <http://www.zjupress.com>

排 版 杭州大漠照排印刷有限公司

印 刷 德清县第二印刷厂

开 本 787mm×1092mm 1/16

印 张 17.25

字 数 430 千

版 印 次 2010 年 3 月第 1 版 2010 年 3 月第 1 次印刷

书 号 ISBN 978-7-308-07438-4

定 价 29.00 元

版权所有 翻印必究 印装差错 负责调换

浙江大学出版社发行部邮购电话(0571)88925591

前 言

数据库技术一直是计算机科学技术中发展最快的领域之一。随着 Internet 的进一步发展,数据库和网络将构成计算机应用的两个重要平台。近年来,数据库发展的技术趋势有三个主要特征:支持 XML 数据格式、支持商业智能、支持 SOA 架构。为适应这种技术发展的需要,数据库专家们就 XML 数据库、Web 数据库、智能数据库、特种数据库、数据仓库与数据挖掘、数据库应用开发方法与工具等诸多方面开展了卓有成效的研究,并取得了丰硕的研究成果。考虑到数据库技术课程的内容十分丰富,涉及的相关知识领域又十分广泛,我们本着了解最新发展动态,掌握成熟应用开发技术的指导思想来规划和设计教材的内容和结构。期望读者通过本教材的学习,达到拓展视野、学会并掌握应用开发技术的基本目的。

数据库课程是一门实践性非常强的课程。本教材在介绍数据库基本概念的基础上,重点突出读者实践动手能力的培养和提高,以成熟的数据库应用开发技术、方法和模型为基础,结合 SQL Server 2000、网络应用展开了讨论,同时也兼顾了读者对了解和掌握数据库新技术的需求。教材第 1 章介绍了数据库基本概念,讨论了当前流行实用的数据库管理系统、数据库技术的研究领域、发展现状和应用情况;第 2 章介绍了关系数据库;第 3 章介绍了 SQL 语言的语法、应用方法,还讨论了存储过程的创建和应用;第 4 章讨论关系模式设计的基础理论以及对关系模式进行规范化的方法;第 5 章详细介绍基于 E-R 的数据库设计的 7 个步骤,在详细介绍各个设计阶段的任务和设计方法的同时,还给出了若干典型应用的设计案例;第 6 章围绕流行 DBMS Microsoft SQL Server 的安装、运行环境、数据管理及相关工具等展开介绍;第 7 章结合 SQL Server 介绍数据保护功能,保证数据的安全可靠和正确有效,包括数据库恢复、数据库完整性、数据库安全性和并发控制等措施;第 8 章讨论数据库技术与 Web 技术的融合应用;第 9 章对 XML 相关标准以及 XML 数据的操纵作入门介绍,并简单介绍有关 XML 数据管理的一些知识。

本书由浙江传媒学院瞿有甜老师担任主编,负责全书的统稿和审稿工作。浙江师范大学何小卫老师担任副主编,参与了全书的审稿工作。第 1 章由杭州电子科技大学郭艳

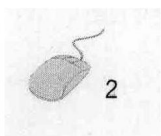
华老师编写,第2章由浙江大学城市学院柯海丰老师编写,第3章由浙江大学方宁老师编写,第4章由温州大学王咏老师编写,第5章由浙江传媒学院瞿有甜老师编写,第6章由杭州电子科技大学张海平老师编写,第7章由浙江师范大学何小卫老师编写,第8章由浙江理工大学的陶灵兵老师编写,第9章由浙江工业大学的杨良怀老师编写。另外,参与本书编写工作的还有王超男、钟丽丽等。宁波职业技术学校曾棕根老师为本教材提供了部分范例。

浙江省高校计算机教学研究会的多位专家,特别是杭州电子科技大学胡维华教授,浙江大学陈天洲教授等对本书定位、结构安排及内容组织等提出了许多宝贵意见和建议,并给予了多方面的指导和关注,在此,谨表衷心感谢!

限于编著者的水平,书中不妥之处在所难免,恳请读者批评指正。

编著者

2010.1.16



目 录

第 1 章 数据库技术基础	(1)
1.1 数据库技术概述	(2)
1.1.1 数据与信息	(2)
1.1.2 数据管理技术的发展	(2)
1.1.3 数据库系统	(6)
1.2 数据库系统的模式结构	(9)
1.2.1 三层模式结构	(9)
1.2.2 两层模式映射	(10)
1.3 数据库系统的数据模型	(12)
1.3.1 数据模型	(12)
1.3.2 概念模型	(13)
1.3.3 常用逻辑数据模型	(15)
1.4 数据库管理系统	(19)
1.4.1 数据库管理系统的主要功能	(19)
1.4.2 数据库管理系统的组成	(20)
1.5 主流网络数据库产品简介	(21)
1.6 数据库技术与应用新趋势*	(24)
1.6.1 数据库技术的研究领域	(24)
1.6.2 数据库技术的发展	(25)
1.6.3 数据库技术的新应用	(28)
1.6.4 数据库技术与应用新趋势	(30)
1.7 小 结	(32)
习题 1	(32)
第 2 章 关系数据库基础	(35)
2.1 关系模型的基本概念	(35)
2.1.1 关系模型形式定义	(35)
2.1.2 关系模型的基本术语	(36)



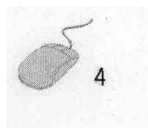
2.1.3	关系的三类完整性规则	(37)
2.1.4	关系查询语言和关系运算	(40)
2.2	关系代数	(40)
2.2.1	传统的集合运算	(40)
2.2.2	专门的关系运算	(42)
2.2.3	关系代数运算的应用举例	(48)
2.3	关系演算	(52)
2.3.1	元组关系演算	(52)
2.3.2	域关系演算	(53)
2.4	小 结	(53)
	习题 2	(53)
第 3 章	SQL 语言	(55)
3.1	SQL 概 述	(55)
3.1.1	SQL 的发展简史	(55)
3.1.2	SQL 的特点	(55)
3.1.3	SQL 的功能	(56)
3.1.4	网上书店数据库	(56)
3.2	SQL 表达式	(61)
3.2.1	常量与变量	(61)
3.2.2	函 数	(64)
3.2.3	表达式	(66)
3.2.4	新增数据类型	(68)
3.3	SQL 数据定义	(68)
3.3.1	创建、修改与删除基表	(68)
3.3.2	创建与删除索引	(70)
3.3.3	创建与删除视图	(71)
3.4	SQL 数据查询	(71)
3.4.1	SELECT 语句基本格式	(71)
3.4.2	单表查询	(72)
3.4.3	连接查询	(78)
3.4.4	嵌套查询	(81)
3.4.5	集合查询	(86)
3.5	SQL 数据操作	(87)
3.5.1	数据插入	(87)
3.5.2	数据修改	(87)



3.5.3 数据删除	(88)
3.6 SQL 数据控制	(88)
3.6.1 授 权	(88)
3.6.2 权限收回	(89)
3.7 视 图	(89)
3.7.1 视图的概念	(89)
3.7.2 定义视图	(89)
3.7.3 删除视图	(90)
3.7.4 视图的作用	(90)
3.8 存储过程	(91)
3.8.1 存储过程的概念	(91)
3.8.2 创建和执行存储过程	(91)
3.9 用户自定义函数*	(93)
3.9.1 函数的概念	(93)
3.9.2 创建和调用标量函数	(93)
3.9.3 创建和调用内嵌表值函数	(94)
3.9.4 创建和调用多语句表值函数	(94)
3.9.5 更改和删除函数	(95)
3.10 嵌入式 SQL*	(96)
3.10.1 嵌入式 SQL 的实现方式	(96)
3.10.2 嵌入式 SQL 的使用规定	(97)
3.10.3 嵌入式 SQL 的使用技术	(98)
3.10.4 动态 SQL 语句	(99)
3.11 小 结	(100)
习题 3	(101)
第 4 章 关系模式设计基础	(102)
4.1 关系模式的设计原则	(102)
4.1.1 关系模式的数据冗余和异常问题	(102)
4.1.2 关系模式的非形式化设计准则	(103)
4.2 函数依赖	(104)
4.2.1 函数依赖的定义	(104)
4.2.2 部分函数依赖和完全函数依赖	(106)
4.2.3 传递函数依赖	(106)
4.3 范式和规范化方法	(107)
4.3.1 第一范式(1NF)	(107)

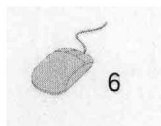


4.3.2	第二范式(2NF)	(109)
4.3.3	第三范式(3NF)	(110)
4.3.4	BCNF	(110)
4.3.5	多值依赖与 4NF	(111)
4.3.6	模式设计方法的原则	(113)
4.3.7	数据依赖的公理系统*	(114)
4.4	小结	(115)
	习题 4	(115)
第 5 章	基于 E-R 的数据库设计	(117)
5.1	数据库设计的基本步骤	(117)
5.1.1	数据库设计概述	(117)
5.1.2	规划阶段	(120)
5.1.3	需求分析阶段	(120)
5.1.4	概念结构设计阶段	(124)
5.1.5	逻辑结构设计阶段	(127)
5.1.6	数据库物理设计阶段	(129)
5.1.7	数据库实施阶段	(129)
5.1.8	数据库的运行与维护	(130)
5.2	E-R 模型	(131)
5.2.1	E-R 模型的基本元素	(131)
5.2.2	属性的分类	(134)
5.2.3	E-R 模型的操作	(136)
5.2.4	基于 E-R 模型的数据库概念设计步骤	(137)
5.3	E-R 模型到关系模型的转换	(142)
5.3.1	E-R 图转换成关系模式集的算法	(142)
5.3.2	基于 E-R 模型的逻辑结构设计步骤	(143)
5.4	E-R 模型设计举例	(144)
5.4.1	学籍管理 E-R 模型	(144)
5.4.2	图书借阅系统的 E-R 模型	(145)
5.4.3	网上书店的 E-R 模型	(146)
5.4.4	旅客登机 E-R 模型	(147)
5.4.5	银行借贷系统 E-R 模型	(148)
5.5	其他数据库设计方法简介*	(149)
5.6	小结	(150)
	习题 5	(150)



第 6 章 SQL Server 基础	(152)
6.1 SQL Server 概述	(152)
6.2 安装与卸载	(152)
6.2.1 准备工作	(152)
6.2.2 安装过程与选项	(153)
6.2.3 安装完成后的测试	(155)
6.2.4 卸 载	(155)
6.3 常用工具简介	(156)
6.3.1 服务管理器	(156)
6.3.2 企业管理器	(156)
6.3.3 查询分析器	(157)
6.3.4 数据导入导出工具	(158)
6.4 数据库的创建与管理	(159)
6.4.1 SQL Server 数据库的文件组成	(159)
6.4.2 数据库的创建	(159)
6.4.3 数据库的修改	(162)
6.4.4 数据库的删除	(164)
6.5 数据库对象基本操作	(164)
6.5.1 SQL Server 表操作	(164)
6.5.2 SQL Server 编程结构	(167)
6.6 小 结	(168)
习题 6	(169)
第 7 章 数据库安全与保护技术	(172)
7.1 数据库备份和恢复	(172)
7.1.1 事 务	(172)
7.1.2 数据库故障类型	(173)
7.1.3 SQL Server 数据库备份	(174)
7.1.4 SQL Server 数据库恢复	(177)
7.2 数据库完整性	(178)
7.2.1 数据库完整性概述	(178)
7.2.2 SQL Server 基本数据完整性约束的实现	(179)
7.3 数据库安全性	(182)
7.3.1 数据库安全性控制	(183)
7.3.2 SQL Server 的安全控制和认证模式	(184)
7.3.3 SQL Server 登录账户和角色管理	(185)

7.3.4	SQL Server 数据库用户和角色的管理	(187)
7.3.5	SQL Server 权限管理	(189)
7.4	数据库并发控制	(191)
7.4.1	并发控制概述	(191)
7.4.2	封锁和封锁协议	(193)
7.4.3	封锁出现的问题和解决方法	(196)
7.4.4	SQL Server 封锁机制	(198)
7.5	小 结	(199)
	习题 7	(199)
第 8 章	网络数据库技术	(201)
8.1	Web 数据库概述	(201)
8.1.1	Web 简介	(201)
8.1.2	静态网页与动态网页	(202)
8.1.3	Web 数据库应用的三层体系结构	(203)
8.2	网络数据库系统的工作模式	(204)
8.2.1	C/S 模式	(204)
8.2.2	B/S 模式	(205)
8.2.3	混合模式及 N 层模式	(206)
8.3	网络数据库的访问技术	(207)
8.3.1	ODBC 技术	(207)
8.3.2	OLE DB 技术	(210)
8.3.3	JDBC 技术	(211)
8.4	ASP+ADO 模式	(212)
8.4.1	ASP 概述	(212)
8.4.2	ASP 的运行环境	(212)
8.4.3	ASP 的工作模式	(214)
8.4.4	ASP 的优点	(215)
8.4.5	数据操作对象 ADO	(215)
8.5	ASP.Net+ADO.Net 模式	(223)
8.5.1	.Net 概述	(223)
8.5.2	ASP.Net 概述	(224)
8.5.3	ASP.Net 的运行环境	(224)
8.5.4	ASP.Net 的工作模式	(226)
8.5.5	ASP.Net 的优势	(226)
8.5.6	ADO.Net	(227)



8.6 JSP+JDBC 模式	(235)
8.6.1 J2EE 概述	(235)
8.6.2 JSP 概述	(236)
8.6.3 JSP 的运行环境	(237)
8.6.4 JSP 的实现方式	(239)
8.6.6 JSP 的优点	(239)
8.6.7 JDBC	(240)
8.7 小 结	(244)
习题 8	(244)
第 9 章 XML 数据库*	(246)
9.1 XML 概述	(246)
9.1.1 XML 的诞生	(246)
9.1.2 XML 文档	(247)
9.1.3 XML 文档类型定义	(250)
9.1.4 XML 模式	(253)
9.1.5 XML 数据库	(254)
9.2 XML 编程接口	(255)
9.2.1 文档对象模型	(255)
9.2.2 简单的应用程序设计接口	(256)
9.3 常用的 XML 查询	(257)
9.3.1 XQuery 查询	(257)
9.3.2 SQL/XML	(258)
9.3.3 MS SQL Server 对 XML 数据的支持	(259)
9.4 小 结	(261)
习题 9	(261)
参考文献	(262)



第 1 章

数据库技术基础

随着计算机技术、通信技术和网络技术的发展,人类社会已经进入了信息化时代。数据库技术是计算机技术的重要分支,是数据库管理的实用技术。数据库技术于 20 世纪 60 年代中期产生,是目前最活跃、发展最快的 IT 技术之一。随着网络和多媒体技术的迅猛发展,数据库技术的应用也得到了更广泛的拓展,并进入了一个崭新的时期。现代数据库技术融合了多种新技术,并正在不断地以超出我们想象的速度飞速发展。

如今,信息资源成为最重要和最宝贵的资源之一,数据库技术已经成为信息社会中对大量数据进行组织与管理的信息系统核心技术和网络信息化管理系统的重要基础。从某种意义上讲,建立一个行之有效的信息系统是企业组织生存和发展的重要前提条件,而数据库的建设规模、数据库信息量的大小和使用频率已经成为衡量一个企业甚至国家信息化程度的重要标志。

说到数据库,或许你并未真正意识到,其实有关你的很多信息早就成为数据库中信息的一部分,或者说你每天或多或少地都在与数据库直接或间接地打着交道。

- 身份证或户籍信息存储在居民信息管理中心数据库中。
- 手机或电话信息存储在电信公司信息管理中心数据库中。
- 银行(信用)卡信息存储在银行信息管理中心数据库中。
- 通过电话或网络订购机票或旅馆时,这些信息有相应的数据库存储和处理。
- 去图书馆查阅书籍时,已经不需要翻阅书籍目录卡片了,而是通过相应的计算机软件直接搜索,搜索的信息也有相应的数据库支持。
- 网络聊天 QQ 号或网络邮箱账号信息存储在网站信息管理中心数据库中。
- 上网浏览或购物交易的信息存储在各个网站的后台信息管理中心数据库中。
- 学籍信息存储在学校学籍信息管理中心数据库中。
- 职工信息存储在单位信息管理中心数据库中。

因此,数据库应用已经与我们的工作或生活不可分割了。

因特网(Internet)的出现和发展,还有随之而来的企业内部网(Intranet)和企业外部网(Extranet)以及虚拟私有网(Virtual Private Network,VPN)的产生和应用,已经将整个世界连成一个小小的地球村,人们可以跨越时空在网上交换信息和协同工作。这样,展现在人们面前的已不是局限于本部门、本单位和本行业的庞大数据库,而是浩瀚无垠的信息海洋。所以说,现在的数据库系统通常是借助于计算机网络进行接收、传递和反馈信息的。

那么,什么是数据库技术,什么是数据库系统,什么是数据库?数据库中的信息是如何组织、存储、提取和处理的?本章首先对数据库技术基础进行概括性叙述。

1.1 数据库技术概述

数据库技术是随着数据管理技术的需要和发展应运而生的,数据管理技术是指对数据的分类、组织、编码、存储、检索和维护的技术,其发展是和计算机技术及其应用的发展密不可分的。简言之,数据库技术就是运用计算机进行数据管理的新技术。

数据库,顾名思义,是用来存储数据的仓库,但是数据应该以怎样的形式、以何种关系、以什么样的结构进行存储,可以使得原本看似无意义的、离散的原始数据变成有关联、有价值和有寓意的信息,并便于信息的访问、查询、统计和输出?这应该是数据库技术的关键所在,所以在介绍数据库的相关知识之前,首先了解一下数据库的基本组成元素——数据和信息,以及数据与信息的关系。

1.1.1 数据与信息

1. 数据与信息

任何事物的属性都是通过数据来表示的。数据是信息的物理表示和载体,数据经过处理、组织并赋予一定关联和意义后即可成为信息。

一个简单的例子可以用来理解数据与信息的关系。“070211008”,“010001”,“95”,这是三个数据,两个字符数据,一个数值数据,如果不赋予其含义和关联,它们没有任何意义,只是一些离散的数据而已。若对上述数据作这样的解释:学号为“070211008”的学生,其课程编号为“010001”的课程的期末成绩为95分,于是它便成为一条有意义的信息。

在数据库中,如果将相关数据关联起来,还可以得出更详尽的信息:学号为“070211008”的“计算机应用”专业的“男”学生“张无忌”,其课程编号为“010001”,“第一学期”学分为4的“必修课”课程“高等数学”的期末成绩为95分。

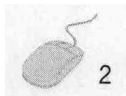
数据是描述事物的符号记录。描述事物的符号可以通过特定设备输入到计算机中,并可以进行储存、处理和传输的各种数字、字母、文字、声音、图片和视频等。数据需要经过解释才能产生有意义的内容,数据的解释是指对数据的含义的说明,数据的含义称为数据的语义,数据和数据的语义是不可分的。

人们通过解释、推理、归纳、分析和综合等方法,从数据中获得的有意义的内容称为信息。数据是信息存在的一种表现形式,且只有通过解释或处理的数据才能真正成为有用的信息。

1.1.2 数据管理技术的发展

数据处理的一个重要方面就是数据管理,计算机对数据的管理是指对数据的组织、分类、编码、存储、检索和维护提供操作手段和途径。

数据管理经历了由低级到高级的发展过程,随着计算机硬件和软件技术的发展而不断提高,数据管理技术的发展大体分为三个阶段:人工管理阶段、文件系统阶段和数据库系统阶段。



1. 人工管理阶段

在 20 世纪 50 年代中期以前,由于计算机软、硬件技术发展水平的限制,计算机系统中没有对数据进行管理的软件。数据管理任务,包括存储结构、存取方法、输入输出方式等,都是针对每个具体应用,由编程人员单独设计解决的。这一时期数据和程序之间的关系可以用图 1-1 来表示。

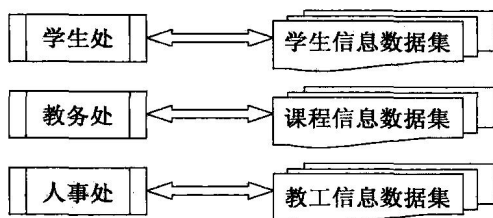


图 1-1 人工管理阶段的数据和程序之间的关系

学习过计算机高级语言(C 或 VB 等)的读者并不难理解人工管理阶段的特点,举例如下。

要编写一个程序(学生处),求一个班 N 个学生每个人 M 门课程的平均成绩并排序。那么这 $M \times N$ 个成绩就必须在编写程序时考虑其输入方式和存储形式(数组),输入随程序的运行而完成,输出结果无法长期保存。

要再编写一个程序(教务处),求某一门课程需要参加补考的学生人数。那么这门课程所有学生的成绩必须在编写这个程序时考虑其输入方式和存储形式(简单变量),输入随程序的运行而完成(尽管这个数据与前面程序中的数据有重复,但无法共享),输出结果同样也无法长期保存。

人工管理阶段有如下几个特点。

(1) 数据不保存。计算机系统不提供对用户数据的管理功能。应用程序包含自己要使用的全部数据,用户编写程序时,必须全面考虑好相关的数据,包括数据的定义、存储结构及存取方法等。程序和数据是一个不可分割的整体。数据脱离了程序就无任何存在的价值,数据无独立性。

(2) 数据不能共享。不同的程序均有各自的数据,这些数据对不同的程序通常是不相同的,不可共享;即使不同的程序使用了相同的一组数据,这些数据也不能共享,程序中仍然需要各自加入这组数据,谁也不能省略。数据的不可共享性必然导致程序与程序之间存在大量的重复数据,浪费存储空间。

(3) 数据不具有独立性。基于数据与程序是一个整体,数据只为本程序所使用,数据只有与相应的程序一起保存才有价值,即数据面向应用,否则就毫无用处。所以,所有程序的数据均不单独保存。

2. 文件系统阶段

20 世纪 50 年代后期至 60 年代中后期,数据管理工作开始借助计算机完成,大量的数据存储、检索和维护工作提上议事日程。此时,在硬件方面可直接存取的磁鼓、磁带、磁盘逐渐变成主要外存。软件方面出现了高级语言和操作系统。操作系统中的文件管理模

块(即输入输出控制模块)的重要功能之一是管理外存储器中的数据。

在文件系统的支持下,数据开始从程序中逐步地独立出来,数据文件可以独立地、长期地存储,数据的逻辑结构和物理结构有了一定的区别;数据的逻辑结构是指呈现在用户面前的数据结构形式;数据的物理结构是指数据在计算机存储设备上的实际存储结构。

例如,用户看到的记录是按照记录号顺序排列的,实际上这些记录可能是分散存储在磁盘的不同扇区里,用链接的方式组织起来的。逻辑结构与物理结构之间的转换由文件系统的存取方式来实现。用户访问记录时只需给出文件名、逻辑记录号,而不必关心记录在存储器上的地址和内、外存交换数据的过程。这一时期数据和程序之间的关系可以用图 1-2 来表示。

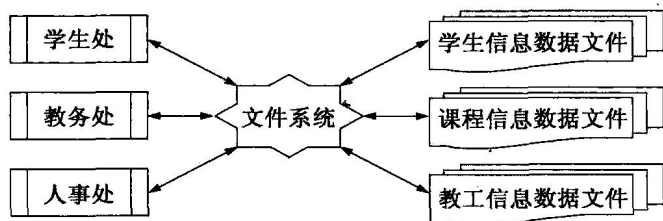


图 1-2 文件系统阶段的数据和程序之间的关系

理解这个阶段的特点,同样可以利用计算机高级语言(C 或 VB 等)的(文件操作)编程经验,举例如下。

要编写一个程序(学生处),求一个班 N 个学生每个人 M 门课程的平均成绩并排序。那么这 $M \times N$ 个成绩就必须事先通过某种方式先写到一个数据文件(顺序文件)中,而数据文件中的数据存储形式必须要与对应程序中数据读取方式契合,虽然这个文件可以长期存储并多次读取,但其具有极强的专一性和依赖性,共享性差。

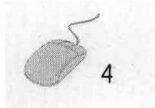
要再编写一个程序(教务处),求某一门课程需要参加补考的学生人数。那么这门课程所有学生的成绩又必须事先通过某种方式先写到一个数据文件中,而这个数据文件中的数据存储形式必须要与对应程序中数据读取方式契合,尽管这些数据与前面程序中的数据有重复,但碍于文件存储形式的差异,共享性差。当然这个数据文件同样可以长期存储并多次读取,但其同样具有极强的专一性和依赖性,它基本上为一个程序服务,共享性差,同时数据存储结构的扩展修改空间有限。

文件系统对计算机数据管理能力的提高虽然起了很大的作用,但是仍然存在许多根本性问题。文件系统管理数据有如下特点。

(1) 数据可以长期保存。数据以文件形式可长期保存在外部存储器的磁盘上。

(2) 应用程序管理数据。数据的逻辑结构与物理结构有区别,但比较简单。程序与数据之间具有设备独立性,即程序只需用文件名就可与数据打交道,不必关心数据的物理位置。由操作系统的文件系统提供存取方法(读/写),对数据的操作以记录为单位;文件中只存储数据,不存储文件记录的结构描述信息。文件的建立、存取、查询、插入、删除、修改等所有操作,都要用程序来实现。

(3) 数据依赖性强。虽然文件组织已多样化,有索引文件、链接文件和直接存取文件



等,但文件之间相互独立、缺乏联系。数据之间的联系要通过程序去构造。虽然数据不再属于某个特定的程序,可以重复使用,但是文件结构的设计仍然是基于特定的用途,程序基于特定的物理结构和存取方法,因此程序与数据结构之间的依赖关系并未根本改变,即数据依然是面向应用的。

(4) 数据共享性差,冗余度大。由于文件之间缺乏联系,造成每个应用程序都有对应的数据文件,有可能同样的数据在多个文件中重复存储。

(5) 数据不一致性。数据冗余往往造成数据不一致,在进行更新操作时,稍不谨慎,就可能使同样的数据在不同的文件中不一样。

(6) 数据之间联系弱。这是由于文件之间相互独立、缺乏联系造成的。

文件系统阶段是数据管理技术发展中的一个重要阶段。在这一阶段中,得到充分发展的数据结构和算法丰富了计算机科学,为数据管理技术的进一步发展打下了基础。

3. 数据库系统阶段

20世纪60年代后期开始,计算机用于数据管理的规模迅速扩大,对数据共享的需求日益增强,为解决数据的独立性问题,实现数据统一管理达到数据共享的目的,发展了数据库技术。同时,计算机的软、硬件技术,特别是磁盘技术的逐渐成熟也给联机存取的数据库技术的实现提供了有力的支持。

数据库技术克服了前几个阶段管理方式的缺点,试图提供一种完善的、更高级的数据管理方式,它的基本思想是解决多用户数据共享的问题,实现对数据的集中统一管理,具有较高的数据独立性,并为数据提供各种保护措施。数据库系统阶段的数据库管理软件作为用户与数据的接口,程序和数据的关系见图1-3。

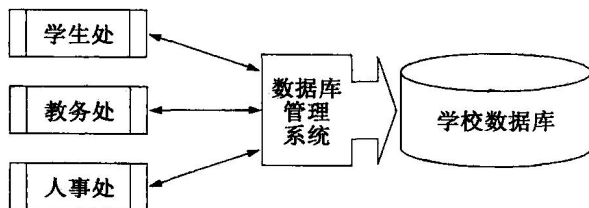


图 1-3 数据库系统阶段的数据和程序之间的关系

概括起来,数据库系统阶段的数据管理具有以下特点。

(1) 数据模型表示复杂的数据结构。数据模型不仅描述数据本身的特征,还要描述数据之间的联系,这种联系通过存取路径实现。

(2) 具有较高的数据共享性和较小的数据冗余度。通过所有存取路径表示自然的数据联系是数据库与传统文件的根本区别。这样,数据不再面向特定的某个或多个应用,而是面向整个应用系统。数据冗余明显减少,实现了数据共享。

(3) 具有较高的数据独立性。数据的逻辑结构与物理结构之间的差别可以很大。用户以简单的逻辑结构操作数据而无需考虑数据的物理结构。数据库的结构分成用户的局部逻辑结构、数据库的整体逻辑结构和物理结构三级。用户(应用程序或终端用户)的数据和外存中的数据之间的转换由数据库管理系统实现。