



普通高等教育“十一五”国家级规划教材

高等学校计算机科学与技术系列教材

数据库原理与应用(第2版)

李 红 编著



高等教育出版社
Higher Education Press

TP311.13/317

2007

普通高等教育“十一五”国家级规划教材

高等学校计算机科学与技术系列教材

数据库原理与应用

(第2版)

李红 编著

高等教育出版社

内容提要

本书以关系数据库系统为核心,系统全面地阐述了数据库系统的基本概念、基本原理和应用技术,主要内容
包括数据库技术概述、关系数据库、关系数据库的标准语言 SQL、关系数据库设计、数据库保护、网络数据库、网
络数据库管理系统 SQL Server 2000、分布式数据库系统、XML 数据库等。

本书概念清楚、重点突出、章节安排合理,理论与实践结合紧密。每章配以丰富的习题、案例及实验,有助于
读者加深对内容的理解、掌握并巩固概念;案例为读者提供了真实的数据库应用场景,有助于读者从实际应用的
角度出发,联系所学理论,掌握所学内容;而实验为读者提供了将理论与实践相结合的具体上机操作途径,最终
巩固所学内容。习题、案例与实验的设计也是本书一个比较突出的特点。

本书既可作为高等学校本科计算机专业(信息技术方向)、信息管理与信息系统专业及相关专业数据库课
程的教学用书,也可作为从事信息领域工作的科技人员的自学参考书。

图书在版编目(CIP)数据

数据库原理与应用/李红编著. —2 版. —北京:高等教育
出版社,2007.12

ISBN 978-7-04-022511-2

I. 数... II. 李... III. 数据库系统-高等学校-教材
IV. TP311.13

中国版本图书馆 CIP 数据核字(2007)第 166678 号

出版发行	高等教育出版社	购书热线	010-58581118
社 址	北京市西城区德外大街 4 号	免费咨询	800-810-0598
邮政编码	100011	网 址	http://www.hep.edu.cn
总 机	010-58581000		http://www.hep.com.cn
经 销	蓝色畅想图书发行有限公司	网上订购	http://www.landracom.com
印 刷	北京未来科学技术研究所 有限责任公司印刷厂	畅想教育	http://www.widedu.com
开 本	787×1092 1/16	版 次	2003 年 3 月第 1 版
印 张	28.75	印 次	2007 年 12 月第 2 版
字 数	650 000	定 价	2007 年 12 月第 1 次印刷 32.00 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 22511-00

高等学校计算机科学与技术系列教材 编审委员会

主任:李 未

副主任:傅育熙 王志英 齐治昌 陈 平 蒋宗礼 马殿富

委员:(按姓氏笔画为序)

王 戟(国防科学技术大学)

宁 洪(国防科学技术大学)

刘 强(清华大学)

孙吉贵(吉林大学)

庄越挺(浙江大学)

何炎祥(武汉大学)

何钦铭(浙江大学)

张晨曦(同济大学)

李宣东(南京大学)

李晓明(北京大学)

陈 钟(北京大学)

陈道蓄(南京大学)

周立柱(清华大学)

周傲英(华东师范大学)

孟祥旭(山东大学)

岳丽华(中国科学技术大学)

罗军舟(东南大学)

姚淑珍(北京航空航天大学)

胡事民(清华大学)

骆 斌(南京大学)

徐宝文(东南大学)

黄虎杰(哈尔滨工业大学)

蒋建伟(上海交通大学)

廖明宏(哈尔滨工业大学)

熊 璋(北京航空航天大学)

樊晓桢(西北工业大学)

序

计算机和通信技术的迅猛发展,不仅形成了融合度最高、潜力最大、增长最快的信息产业,而且成为推动全球经济快速增长和全面变革的关键因素。进入 21 世纪,我国的信息产业虽然已取得了长足的发展,但与发达国家相比,还有不小的差距。国家信息化的发展和信息产业国际竞争能力的提高,迫切需要高素质、创新型的计算机专业人才。

高素质计算机专业人才的培养离不开高质量的计算机教育。我们的专业虽然机会多,处于非常有利的条件,但是我们同样面临着一件事,就是从规模发展向质量提高的转变。怎么提高质量?专业素质的教育和应用素质的训练非常重要。尤其是我国高等教育进入大众化发展阶段,社会对计算机专业人才呈现出了多样化的需求。而与此同时,计算机学科的发展已极大地突破了原有的学科体系框架,形成了在“计算机科学与技术”之下向多个专业方向发展的新格局。在这种背景下,教育部高等学校计算机科学与技术教学指导委员会编制了《高等学校计算机科学与技术专业发展战略研究报告暨专业规范(试行)》(以下简称“专业规范”)。专业规范按照“培养规格分类”的指导思想,提出了三种类型、四个方向,即科学型(计算机科学方向),工程型(计算机工程方向、软件工程方向),应用型(信息技术方向)的计算机专业发展建议,体现了社会对不同人才类型的需求,对于指导我国计算机教学改革与建设,规范计算机教学工作,促进计算机教学质量的提高都具有重要的意义。

高水平的教材是一流教育质量的重要保证。为了配合专业规范的试行,便于广大高校教师按照新的专业规范组织实施教学,高等教育出版社在大力支持专业规范研究与起草工作的同时,还邀请规范起草小组的有关专家成立“高等学校计算机科学与技术系列教材编审委员会”,组织规划了结合计算机专业规范、面向全国高等学校计算机专业本科生的“高等学校计算机科学与技术系列教材”。令人高兴的是,一批有创新、改革精神,

且有丰富教学经验的高等学校教师投身到新体系计算机专业教材的编写中来,他们用自己创造性的思维、辛勤的汗水诠释专业规范的思想,把新的课程体系和教学内容生动地传达给师生,并进行着有意义的教学实践。

“高等学校计算机科学与技术系列教材”以专业规范和 CC2001 - CC2005 有关教程为依据,以强化基础、突出实践、注重创新为原则,体现了学科课程体系和教学内容改革的新成果。此外,这一系列教材还配有丰富的教学辅助资源,并与现代教育技术手段相结合,充分发挥网络平台的作用,使教材更有利于广大教师和学生使用。目前,这一系列教材有不少选题已列入普通高等教育“十一五”国家级规划教材,希望这些教材的出版能够对新形势下我国高等学校计算机专业课程改革与建设起到积极的推动作用,使我国高校的计算机专业教学质量再上一个台阶。



中国科学院院士

2006—2010 年教育部高等学校计算机科学与技术教学指导委员会主任

2007 年 11 月

第2版前言

本书自2003年2月第1版出版以来,受到广大读者的欢迎,获得了2006年度大连市科学著作二等奖。本书第2版入选普通高等教育“十一五”国家级规划教材。

由于近年数据库技术的快速发展,第1版中的许多内容已无法满足客观形势发展和读者实际的需要。特别是近来网络数据库技术和分布式数据库系统的推广和普及以及XML数据库的应用,使得这些技术逐渐成为新时代数据库技术应用的主流技术。为了适应数据库技术的发展,进一步深化数据库课程的教学改革,提高人才培养的质量,作者对第1版的内容进行了修订,结构及内容上做了如下的改动:

- 第1章数据库技术概述中增加了1.5节数据库技术的发展,主要包括数据库技术发展趋势。

- 第4章关系数据库设计的内容,按照数据库设计的阶段进行了结构上的调整,删去了原来4.3节网络环境下的数据库应用规划的内容,并增加了用UML进行数据库应用设计方法的内容。

第2章关系数据库中将2.1.3节关系的完整性约束和2.4节关系数据库规范化理论的内容进行了调整。

- 第5章数据库保护中将5.4节数据库恢复的内容进行了调整。

删去的内容有:

- 第6章面向对象数据库及对象-关系数据库。
- 第7章C/S数据库和Web数据库。
- 第8章开放的客户端开发及主流数据库产品。
- 第9章数据库技术的发展。

增加了4章新的内容:

- 第6章网络数据库。
- 第7章网络数据库管理系统SQL Server 2000。

- 第8章分布式数据库系统。
- 第9章XML数据库。

此次修订的主要目的是使读者既掌握数据库技术的基本原理、方法和应用技术,也能有效地使用现用的数据库管理系统,掌握数据库应用系统的设计与开发方式,同时也适时地掌握网络数据库和分布式数据库以及XML数据库的相关技术。

全书共分9章:第1章数据库技术概述,包括数据库的作用及相关概念、数据库管理技术的发展、数据库系统结构、数据模型以及数据库技术的发展。第2章全面介绍关系数据库,包括关系模型、关系模式、关系代数以及关系数据库规范化理论。第3章介绍了关系数据库标准语言SQL的组成、功能、特点,重点介绍了数据定义、数据更新、数据查询、数据控制语言以及嵌入式SQL。第4章关系数据库设计,介绍了数据库设计的方法和基本步骤以及数据库设计各阶段的目标和方法。着重介绍了概念模型的基本概念、实体-联系方法以及如何设计数据库的E-R模型并将E-R模型转换为关系模型;使用UML进行数据库设计的方法以及UML模型向关系模型映射的方法。第5章数据库保护,包括数据库安全性控制、并发性控制、完整性控制以及数据库恢复与备份。第6章网络数据库,包括网络数据库的产生与发展、网络数据库的概念与特点、网络数据库系统的工作模式、网络数据库的访问技术、网络数据库的数据交换技术、网络数据库的Web应用以及目前主流的网络数据库产品简介。第7章网络数据库管理系统SQL Server 2000,包括SQL Server 2000的体系结构及特点、SQL Server 2000的管理工具和实用程序、SQL Server服务器管理、SQL Server 2000数据库操作、SQL Server 2000数据库对象操作以及有关SQL Server 2000编程结构的基础知识。第8章分布式数据库系统,包括分布式数据库系统的含义、特点、优缺点、分类及环境;分布式数据存储、分布式数据库系统的体系结构及其分布式处理。第9章XML数据库,包括XML文档及类型、XML数据模型、纯XML数据库、XML数据库的数据存储以及常用的XML查询(XQuery及SQL/XML)。书中加“*”号的部分可作为扩展内容,教师可根据具体情况进行选择。

本书由浅入深、循序渐进、理论与实践并重,力求使读者通过学习本书,能对数据库技术有比较全面的了解,掌握数据库理论与数据库应用的基本知识,了解数据库应用系统的开发模式,并具有初步的数据库应用开发能力。

本书以培养21世纪计算机专业及信息处理专业人才为目的,同时满足社会对数据库管理人员的需要,让更多的人可以更快地学到实用的数据库理论、技术与方法。本书以简明实用、便于学习、反映数据库技术的最新发展和应用为特色,在每章的前面分别列出了本章的学习目标,指出了本章的主要学习内容以及应掌握的知识点;每章都配以丰富的习题、案例及实验。习题帮助读者加深对内容的理解、掌握并巩固概念;案例为读者提供真实的数据库应用场景,帮助读者从实际应用的角度出发联系所学理论,掌握所学内容;而实验则为读者

提供理论与实践相结合的具体上机操作途径,帮助读者巩固所学内容。

本书的教学(包括上机实践)拟采用72个课时,其中应重点讲授第1~5章的内容,对于第6、7、8、9章的内容,教师可以有针对性地选择其中某些内容进行讲授。此外,也可以根据教学需要安排章节顺序,如讲授完第1、2章之后,为便于学生掌握SQL语言,在讲授第3章之前先讲授第7章。

本书既可作为高等学校本科计算机专业(信息技术方向)、信息管理与信息系统专业及相关专业数据库课程的教学用书,也可作为从事信息领域工作的科技人员的自学参考书。

本书的修订得益于大量的国内外相关的参考资料。在此,谨向书后参考文献中列出的作者及参考网站表示感谢!同时,感谢本书的责任编辑彭立辉周密、细致、耐心的工作,避免了许多错误的发生。最后,我要感谢我远在澳大利亚新南威尔士大学计算机科学与工程系学习的儿子,在我对本书的修订过程中所给予的支持和鼓励。

在本次修订中,虽对书中内容反复多次进行修改,以尽量减少错误,但由于本人的水平有限,难免会有各种错误和疏漏,因此恳切希望能够得到广大读者的批评指正,请将意见发送至邮箱 dbms-2@163.com 中,谢谢。

编 者

2007年8月

目 录

第 1 章 数据库技术概述 1

学习目标	1
1.1 数据库的作用及相关概念	2
1.1.1 数据与信息	2
1.1.2 数据管理与数据库	3
1.1.3 数据库管理系统与信息系统	5
1.2 数据管理技术的发展	6
1.2.1 手工管理阶段	6
1.2.2 文件系统阶段	6
1.2.3 数据库系统阶段	7
1.3 数据库系统结构	9
1.3.1 数据库系统的组成	9
1.3.2 数据库体系结构	14
1.3.3 数据库应用系统体系结构	19
1.4 数据模型	22
1.4.1 概述	22
1.4.2 层次数据模型	24
1.4.3 网状数据模型	26
1.4.4 关系数据模型	28
1.4.5 面向对象数据模型	29
1.5 数据库技术的发展	30
1.5.1 数据库系统发展的 3 个阶段	30
1.5.2 现代应用对数据库系统的新要求	32
1.5.3 数据库技术的发展趋势	34
本章小结	39
习题 1	39
案例 1	40

第 2 章 关系数据库 41

学习目标	41
2.1 关系模型	42
2.1.1 关系数据结构	42
2.1.2 关系操作	43
2.1.3 关系的完整性约束	44
2.2 关系模式	47
2.2.1 关系概念模式	47
2.2.2 关系内模式	48
2.2.3 关系外模式	48
2.3 关系代数	49
2.3.1 传统的集合运算	50
2.3.2 特殊的关系运算	52
2.4 关系数据库规范化理论	55
2.4.1 关系模式规范化的必要性	56
2.4.2 函数依赖	58
2.4.3 关系的范式及规范化	60
2.4.4 关系分解原则	65
本章小结	68
* 习题 2	68
案例 2	71
第 3 章 关系数据库的标准语言 SQL	72

学习目标	72
3.1 SQL 概述	73

3.1.1 SQL 的发展历程	73
3.1.2 SQL 的组成与功能	73
3.1.3 SQL 的特点	74
3.1.4 SQL 基本知识	75
3.2 数据定义语言	79
3.2.1 数据库	79
3.2.2 表	81
3.2.3 索引	88
3.2.4 视图	92
3.2.5 存储过程	95
3.3 数据更新语言	97
3.3.1 插入数据	98
3.3.2 修改数据	99
3.3.3 删除数据	100
3.4 数据查询语言	101
3.4.1 SELECT 语句	101
3.4.2 单表查询	102
3.4.3 联接查询	103
3.4.4 嵌套查询	107
3.4.5 集合查询	110
3.4.6 函数查询	111
3.5 数据控制语言	113
3.5.1 数据控制方法	114
3.5.2 数据控制语句	115
3.5.3 事务控制语句	116
3.6 嵌入式 SQL	118
3.6.1 嵌入式 SQL 的形式及处理过程	118
3.6.2 SQL 与宿主语言间的通信	119
3.6.3 嵌入式 SQL 的实现	120
3.6.4 游标	121
3.6.5 嵌入式 SQL 的使用技术	127
3.6.6 动态 SQL	131
本章小结	133
习题 3	133
案例 3	135
实验	136

第 4 章 关系数据库设计 138

学习目标

4.1 数据库设计概述

4.1.1 数据库的设计特点

4.1.2 数据库的设计方法

4.1.3 数据库设计的基本步骤

4.1.4 数据库设计过程中的各级模式

4.2 需求分析

4.2.1 需求分析的任务及其工作

4.2.2 高校管理系统的信息需求

4.3 概念结构设计

4.3.1 重要性、设计策略及步骤

4.3.2 概念模型

4.3.3 实体-联系方法

*4.3.4 UML 方法

4.4 逻辑结构设计

4.4.1 E-R 模型向关系模型的转换

*4.4.2 UML 模型向关系模型的映射

4.4.3 数据模型的优化

4.4.4 设计外模式

4.5 物理结构设计

4.5.1 物理结构设计的内容和方法

4.5.2 关系模式存取方法选择

4.5.3 确定数据库的存储结构

4.5.4 评价物理结构

4.6 数据库实施

4.6.1 数据装载和应用程序的 编制调试

4.6.2 数据库的试运行

4.7 数据库运行和维护

本章小结

习题 4

案例 4

实验

第 5 章 数据库保护 211

学习目标	211
5.1 数据库完整性	212
5.1.1 完整性约束的定义	212
5.1.2 完整性约束的检查和处理	215
5.2 数据库安全性	217
5.2.1 用户验证	218
5.2.2 存取控制	218
5.3 数据库并发控制	220
5.3.1 事务	220
5.3.2 事务的调度	222
5.3.3 事务并发执行所带来的问题	224
5.3.4 封锁	225
5.4 数据库恢复	229
5.4.1 故障的种类与恢复原理	229
5.4.2 故障恢复技术	231
5.4.3 故障恢复策略	237
本章小结	240
习题 5	240
案例 5	240
第 6 章 网络数据库	242

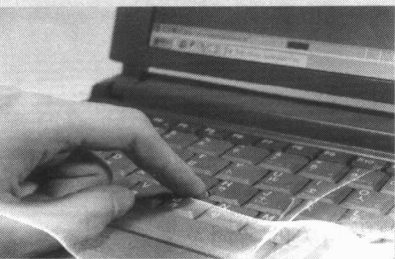
学习目标	242
6.1 网络数据库的产生与发展	243
6.1.1 Web 的产生与发展	243
6.1.2 网络数据库的发展历程	246
6.2 网络数据库的概念与特点	249
6.2.1 网络数据库的概念	249
6.2.2 网络数据库的特点	250
6.3 网络数据库系统的工作模式	251
6.3.1 C/S 模式	251
6.3.2 B/S 模式	254
6.3.3 混合模式及 N 层模式	256
6.4 网络数据库的访问技术	258
6.4.1 ODBC 技术	258
6.4.2 OLE DB 技术	270
6.4.3 JDBC 技术	272
6.4.4 对象/关系映射	274

6.5 网络数据库的数据交换技术	275
6.5.1 XML 技术概述	276
6.5.2 主要的数据库交换技术	280
6.6 网络数据库的 Web 应用	281
6.6.1 Web 应用体系结构	281
6.6.2 常用的 Web 应用开发方式	284
6.7 目前主流的网络数据库产品 简介	288
6.7.1 微软公司的 SQL Server	288
6.7.2 Oracle 产品	289
6.7.3 IBM DB2	290
6.7.4 MySQL 数据库	290
6.7.5 Sybase	290
6.7.6 Informix Dynamic Server 产品	291
本章小结	291
习题 6	292
实验	292

第 7 章 网络数据库管理系统 SQL Server 2000 295

学习目标	295
7.1 SQL Server 2000 简介	296
7.1.1 SQL Server 2000 的体系结构	296
7.1.2 SQL Server 2000 的特点	298
7.1.3 SQL Server 2000 的版本及安装 环境	299
7.2 SQL Server 2000 的管理工具和 实用程序	302
7.2.1 服务管理器	302
7.2.2 企业管理器	302
7.2.3 查询分析器	304
7.2.4 其他管理工具程序介绍	305
7.3 SQL Server 服务器管理	308
7.3.1 创建服务器组	308
7.3.2 服务器注册	309
7.3.3 运行服务器	314
7.3.4 服务器配置选项设置	316

7.3.5 系统数据库	320	8.4.1 查询处理	372
7.3.6 系统表	321	8.4.2 事务处理	376
7.4 SQL Server 2000 数据库操作	322	8.4.3 并发控制	377
7.4.1 SQL Server 2000 数据库基本 结构	322	8.4.4 分布恢复	378
7.4.2 SQL Server 2000 数据库的建立	324	本章小结	380
7.4.3 显示数据库信息	327	习题 8	380
7.4.4 修改数据库	328	讨论题	381
7.4.5 删除数据库	330	第 9 章 XML 数据库	382
7.5 SQL Server 2000 数据库对象 操作	331	学习目标	382
7.5.1 SQL Server 2000 表操作	331	9.1 概述	383
7.5.2 SQL Server 2000 视图操作	336	9.1.1 XML 文档	383
7.5.3 SQL Server 2000 索引操作	339	9.1.2 XML 数据库	385
7.5.4 SQL Server 2000 存储过程	342	9.2 XML 数据模型	386
7.5.5 SQL Server 2000 的编程结构	345	9.2.1 XML DTD	387
本章小结	349	9.2.2 XML Schema	389
习题 7	349	9.3 纯 XML 数据库	391
实验	349	9.3.1 基本概念	391
第 8 章 分布式数据库系统	352	9.3.2 纯 XML 数据库管理系统	393
学习目标	352	9.4 XML 数据库的数据存储	396
8.1 概述	353	9.4.1 纯 XML 数据库	396
8.1.1 分布式数据库系统的含义	353	9.4.2 基于关系的 XML 数据库	398
8.1.2 分布式数据库系统的特点	355	9.5 XML 查询	403
8.1.3 分布式数据库系统的优缺点	357	9.5.1 XQuery 查询	403
8.1.4 分布式数据库系统的分类及 环境	358	9.5.2 SQL/XML	414
8.2 分布式数据存储	359	本章小结	426
8.2.1 数据分片	359	习题 9	427
8.2.2 数据分布	361	实验	429
8.3 分布式数据库系统的体系结构	362	附录	430
8.3.1 分布式数据库系统的模式结构	362	参考文献	439
8.3.2 分布透明性	364	相关网站	441
8.3.3 分布式数据库管理系统	368		
8.4 分布式处理	371		



第1章

数据库技术概述



信息资源已成为社会各行各业的重要资源和财富,作为实施有效信息处理的信息系统已成为一个企业或组织生存和发展的重要基础条件。由于数据库技术是信息系统的核心和基础,因而得到快速的发展和越来越广泛的应用。数据库技术主要研究如何科学地组织和存储数据、高效地获取和处理数据,是数据管理的最新技术,是计算机科学与技术的重要分支。数据库技术可以为各种用户提供及时的、准确的、相关的信息,满足用户各种不同的需要。

学习目标

- 数据库相关的概念及数据库的作用。
- 数据管理技术的发展阶段及各阶段的特点。
- 数据库系统的组成、各部分的功能及其相互之间的关系。
- 数据库体系的三级模式结构、两级映像及数据独立性概念及其作用。
- 数据库应用系统的体系结构中数据的存储、管理和应用的关系与层次。
- 数据库应用系统结构的不同结构类型及各自的特点。
- 数据模型的概念、分类及其组成。
- 层次、网状、关系和面向对象数据模型的特点及各自的优缺点。
- 数据库系统发展的特点及主流数据库技术的发展趋势。

1.1 数据库的作用及相关概念

为了便于理解什么是数据库以及数据库的作用,首先了解一下与数据库相关的名词及概念,主要有数据与信息、数据管理与数据库、数据库管理系统与信息管理等。

1.1.1 数据与信息

1. 信息

(1) 概念

信息泛指通过各种方式传播的,可被感受的数字、文字、图像和声音等符号所表征的某一事物的新的消息、情报和知识。它是观念性的东西,是人们头脑对现实事物的抽象反映,与载体无关。

(2) 特征

① 信息的内容是关于客观事物或思想方面的知识,即信息的内容能反映已存在的客观事实,能预测未发生事物的状态和能用于指挥、控制事物发展的决策。

② 信息是有用的,它是人们活动的必需知识,利用信息能够克服工作中的盲目性,增加主动性和科学性,可以把事情办得更好。

③ 信息能够在空间和时间上传递,在空间上传递信息称为信息通信,在时间上传递信息称为信息存储。

④ 信息需要一定的形式表示,信息与其表现符号不可分离。

(3) 作用

信息对于人类社会的发展有着重要的意义,它可以提高人们对事物的认识,减少人们活动的盲目性;信息是社会机体进行活动的纽带,社会的各个组织通过信息网相互了解并协同工作,使整个社会协调发展;社会越发展,信息的作用就越突出;信息又是管理活动的核心,要想把事务管理好,就需要掌握更多的信息,并利用信息进行工作。

2. 数据

(1) 概念

数据是记录信息的物理符号,是表达和传递信息的工具。尽管信息有多种表现形式,可以通过手势、眼神、声音或图形等方式表达,但数据是信息的最佳表现形式。由于数据能够书写,因而它能够被记录、存储和处理,从中挖掘出更深层的信息。必须指出的是,在许多不严格的情况下,对数据和信息两个概念不进行区分而是混为一谈。

(2) 特征

① 数据有“型”和“值”之分。

数据的型是指数据的结构,而数据的值是指数据的具体取值。数据的结构指数据的内部构成和对外联系。

② 数据受数据类型和取值范围的约束。

数据类型是针对不同的应用场合设计的。数据类型不同,数据的表示形式、存储方式及数据能进行的操作和运算也各不相同。在使用计算机处理数据时,应当特别重视数据类型,为数据选择合适的类型。数据的取值范围亦称数据的值域,为数据设置值域是保证数据的有效性避免数据输入或修改时出现错误的重要措施。

③ 数据有定性表示和定量表示之分。

例如,在表示职工的年龄时,可以用“老”、“中”、“青”定性表示,也可以用具体岁数定量表示。由于数据的定性表示是带有模糊因素的粗略表示方式,而数据的定量表示是描述事物的精确表示方式,所以在计算机软件设计中,应尽可能地采用数据的定量表示方式。

④ 数据应具有载体和多种表现形式。

数据是对客观物体或概念的属性的记录,它必须有一定的物理载体。当数据记录在纸上时,纸张是数据的载体;当数据记录在计算机的外存上时,硬盘、软盘或磁带就是数据的载体。数据具有多种表现形式,它可以用报表、图形、语音及不同的语言符号来表示。

(3) 数据与信息的关系

数据与信息是两个既有联系、又有区别的概念。数据是信息的载体,信息则是对数据加工的结果,是对数据的解释。数据与信息的关系如图 1-1 所示。

计算机系统的每项操作,均是对数据进行某种处理。数据输入计算机后,经存储、传送、排序、计算、转换、检索、制表及仿真等操作,输出人们需要的结果,即产生信息。

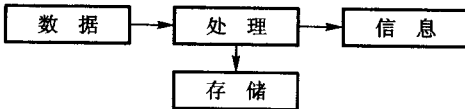


图 1-1 数据与信息的关系

1.1.2 数据管理与数据库

数据管理是数据处理的基础工作,数据库是数据管理的技术和手段。

1. 数据处理

(1) 概念

围绕着数据所做的工作均称为数据处理。数据处理是指对数据的收集、组织、整理、加工、存储和传播等工作。

(2) 分类

① 数据管理:主要任务是收集信息、将信息用数据表示并按类别组织、保存,其目的是在需要的时候,为各种应用和数据处理提供数据。

② 数据加工:主要任务是对数据进行变换、抽取和运算,通过数据加工会得到更有用的数据,以指导或控制人的行为或事物的变化趋势。

③ 数据传播:是指在空间或时间上以各种形式传播信息,而不改变数据的结构、性质和内容,数据传播会使更多的人得到并理解信息,从而使信息的作用充分发挥出来。

2. 数据管理

(1) 作用

在数据处理中,最基本的工作是数据管理工作。数据管理是其他数据处理的核心和基础。

在实际工作中数据管理的地位很重要。工作中有许多人从事各种行政管理工作,实际上这些管人、管财、管物或管事的工作就是数据管理工作。而人、财、物和事又可统称为事务。在事务管理中,事务以数据的形式被记录和保存。例如,在财务管理中,财务科通过对各种账目的记账、对账或查账等来实现对财务数据的管理。传统的数据管理方法是人工管理方式,即通过手工记账、算账和保管账的方法实现对各种事务的管理。计算机的发展为科学地进行数据管理提供了先进的技术和手段,目前许多数据管理工作都利用计算机进行,而数据管理也成了计算机应用的一个重要分支。

(2) 内容

数据管理工作应包括以下3项内容:

① 组织和保存数据,即将收集到的数据合理地分类组织,将其存储在物理载体上,使数据能够长期地被保存。

② 数据维护,即根据需要随时进行插入新数据、修改原数据和删除失效数据的操作。

③ 提供数据查询和数据统计功能,以便快速地得到需要的正确数据,满足各种使用要求。

3. 数据库

(1) 概念

数据库(DataBase, DB)是一个长期存储在计算机内的、有组织的、可共享的、统一管理的数据集合。它是一个按数据结构来存储和管理数据的计算机软件系统。数据库的概念实际上包括两层意思:

① 数据库是一个实体,它是能够合理保管数据的“仓库”,用户在该“仓库”中存放要管理的事务的数据,“数据”和“库”两个概念结合成为“数据库”。

② 数据库是数据管理的新方法和技术,它能够更合理地组织数据、更方便地维护数据、更严密地控制数据和更有效地利用数据。

(2) 数据库中数据的性质

① 数据整体性:数据库是一个单位或一个应用领域的通用数据处理系统,它存储的是属于企业和事业部门、团体和个人的有关数据的集合。数据库中的数据是从全局观点出发建立的,它按一定的数据模型进行组织、描述和存储,其结构基于数据间的自然联系,从而可提供一切必要