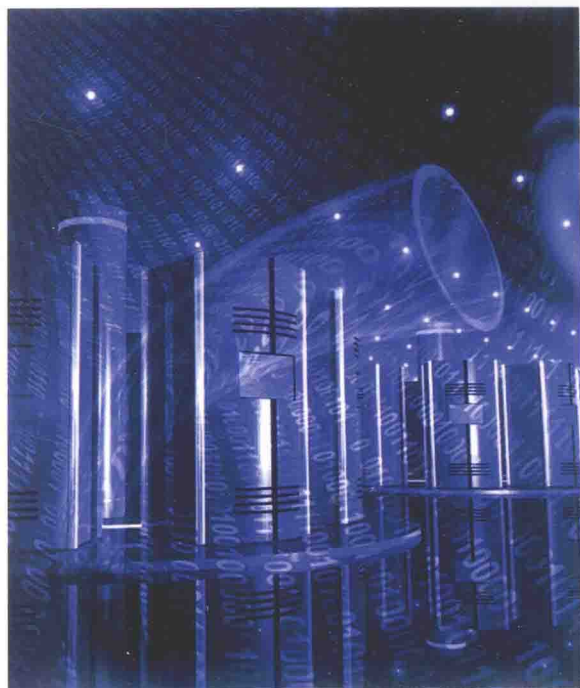


数据库原理及应用

- ◆ 数据库系统概论
- ◆ 关系数据库
- ◆ 关系数据库标准语言SQL
- ◆ 关系数据库规范化理论
- ◆ 数据库设计
- ◆ 数据库设计典型案例
- ◆ 数据库安全与维护

SQL Server数据库



倪春迪 殷晓伟 主 编

刘国成 曲丽娜 副主编



清华大学出版社

高等学校计算机应用规划教材

数据库原理及应用

倪春迪 殷晓伟 主 编
刘国成 曲丽娜 副主编



清华大学出版社
北 京

内 容 简 介

本书全面介绍了数据库系统基本原理以及数据库管理系统应用技术。全书共 8 章, 主要内容包括数据库系统概述、关系数据库系统理论基础、关系代数和关系的规范化理论、SQL Server 2012 概述与安装、数据库的创建与管理、数据表的创建与管理、视图、索引以及备份与恢复等内容。本书主要是为了满足高等院校培养应用技术型人才的需要而编写的, 具有较强的实用性。

本书内容循序渐进、深入浅出、概念清晰、结构合理, 将数据库基本原理与应用实践相结合并配有适量的例题和习题, 帮助读者从不同角度理解和掌握所学的知识, 构建完整的知识体系。

本书可作为高等院校计算机专业或数学、通信、电子类等相关专业的本科生或高职高专院校专科生的教材, 也可作为计算机爱好者的自学用书。

本书对应的电子教案和习题答案可以到 <http://www.tupwk.com.cn> 网站下载。

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。

版权所有, 侵权必究。侵权举报电话: 010-62782989 13701121933

图书在版编目(CIP)数据

数据库原理及应用 / 倪春迪, 殷晓伟 主编. —北京: 清华大学出版社, 2015

(高等学校计算机应用规划教材)

ISBN 978-7-302-40431-6

I. ①数… II. ①倪… ②殷… III. ①数据库系统—高等学校—教材 IV. ①TP311.13

中国版本图书馆 CIP 数据核字(2015)第 122535 号

责任编辑: 胡辰浩 袁建华

装帧设计: 孔祥峰

责任校对: 曹 阳

责任印制: 刘海龙

出版发行: 清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦 A 座 邮 编: 100084

社 总 机: 010-62770175 邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载: <http://www.tup.com.cn>, 010-62794504

印 刷 者: 北京富博印刷有限公司

装 订 者: 北京市密云县京文制本装订厂

经 销: 全国新华书店

开 本: 185mm×260mm 印 张: 15.5 字 数: 303 千字

版 次: 2015 年 6 月第 1 版 印 次: 2015 年 6 月第 1 次印刷

印 数: 1~3000

定 价: 33.00 元

前 言

数据库是现代化数据管理的最重要、最广泛、最先进的技术,是计算机科学的重要分支。《数据库原理及应用》为计算机及相关的众多学科提供了利用计算机技术进行数据管理的基本理论知识,是计算机科学与技术、软件工程及其相关专业学科的专业必修课。

本教材主要介绍数据库的基本理论和数据库管理系统的基本应用技术。全书共 8 章。第 1 章数据库概述,主要介绍数据库的产生与发展,包括数据库、数据库管理系统和数据库系统的基本概念,数据模型与数据库体系结构等内容。第 2 章关系数据库系统模型,主要介绍关系数据模型、关系模型的完整性规则、关系代数的基本运算等。第 3 章结构化查询语言 SQL,主要介绍标准 SQL 概述,SQL 的数据定义、数据查询、数据更新等。第 4 章关系数据库的设计理论,主要介绍实体类型的属性关系、数据的函数依赖、关系数据库模式的规范化理论、关系模式的分解算法等。第 5 章数据库设计,包括数据库设计概述、需求分析、概念结构设计、逻辑结构设计、物理结构设计、数据库的实施和维护。第 6 章给出了一个数据库设计的实例。第 7 章数据库的安全与维护,主要介绍数据库的安全性、完整性、并发控制、备份与恢复等。第 8 章介绍 SQL Server 2012 数据库管理系统的基本操作。

本书可以作为各高等学校计算机及相关专业“数据库原理及应用”课程的教材,也可作为广大计算机爱好者的自学用书。

全书每章后面均配有适量的习题,以加强读者对本章所学知识的理解和掌握。本书对应的电子教案和习题答案可以到 <http://www.tupwk.com.cn> 网站下载。

本书由黑龙江工程学院倪春迪编写第 1、2、3、7 章;东北林业大学殷晓伟编写第 5、8 章;刘国成编写第 4 章;曲丽娜编写第 6 章。此外,参加本书编写的人员还有孙琪、张玉琪、宋仁涛、何东昌、鲁月新、范博文、黄修尧、李罚、吴兴、宋泽辉、金敬杰、张雪峰、钱程、李飞龙、程淇、张可心。全书由刘国成统稿。

本书的出版得到了清华大学出版社相关同志的热情关心和大力支持。许多老师和读者也对本书的编写提出了许多宝贵建议和修改意见,我们对此一并表示由衷的感谢。由于时间仓促,加之编者水平有限,书中错误和不当之处难免,恳请读者批评指正。我们的电话是 010-62796045,信箱是 huchenhao@263.net。

编 者

2015 年 3 月

目 录

第 1 章 数据库系统概论	1
1.1 数据管理技术的发展	1
1.1.1 人工管理阶段	2
1.1.2 文件管理阶段	2
1.1.3 数据库系统阶段	4
1.2 数据描述	7
1.2.1 数据描述的 3 个领域	7
1.2.2 数据联系的描述	8
1.3 数据模型	10
1.3.1 数据模型的概念	10
1.3.2 实体联系模型	10
1.3.3 结构数据模型	11
1.4 数据库系统的体系结构	16
1.4.1 内模式	16
1.4.2 模式	17
1.4.3 外模式	17
1.4.4 数据独立性	18
1.4.5 映射	19
1.5 数据库管理系统	20
1.5.1 数据库管理系统的主要功能	20
1.5.2 数据库管理系统的组成	21
1.6 数据库系统	22
1.7 小结	23
1.8 习题	23
第 2 章 关系数据库	25
2.1 关系数据模型	25
2.1.1 关系的数学定义	26

2.1.2	关系模式及关系实例	27
2.1.3	关系中的基本术语	28
2.1.4	关系的性质	29
2.1.5	关系模型的完整性规则	30
2.2	关系代数	31
2.2.1	关系操作	31
2.2.2	传统的集合运算	33
2.2.3	专门的关系运算	35
2.2.4	关系代数小结	38
2.2.5	关系代数操作举例	39
2.3	小结	41
2.4	习题	41
第3章	关系数据库标准语言 SQL	45
3.1	SQL 概述	45
3.2	SQL 语法	47
3.3	CREATE TABLE 语句	48
3.4	INSERT 语句	50
3.5	SELECT 语句	51
3.5.1	默认查询	51
3.5.2	查询指定的字段	52
3.5.3	按条件查询	53
3.5.4	排序	55
3.5.5	多表连接查询	56
3.5.6	子查询	56
3.5.7	GROUP BY 语句	58
3.6	UPDATE 语句	59
3.7	DELETE 语句	59
3.8	DROP TABLE 语句	60
3.9	SQL 查询综合应用	60
3.10	小结	64
3.11	习题	64

第 4 章 关系数据库规范化理论	67
4.1 为什么要规范化	67
4.1.1 问题的提出	68
4.1.2 解决方法	69
4.1.3 关系模式规范化理论概述	70
4.2 函数依赖	71
4.2.1 函数依赖的定义	71
4.2.2 Armstrong 公理系统	72
4.2.3 函数依赖的分类	72
4.2.4 属性间联系类型与函数依赖	73
4.2.5 属性集闭包	73
4.2.6 最小函数依赖集及其算法	75
4.2.7 函数依赖与候选关键字	77
4.3 关系模式的规范化	81
4.3.1 概述	81
4.3.2 关系模式的级别	81
4.3.3 关系模式分解	86
4.4 小结	93
4.5 习题	94
第 5 章 数据库设计	95
5.1 数据库设计概述	95
5.1.1 数据库设计的任务、内容和特点	95
5.1.2 数据库设计的基本方法	96
5.1.3 数据库设计的步骤	97
5.2 需求分析	98
5.2.1 需求分析的任务	99
5.2.2 需求分析的方法	99
5.2.3 数据字典	100
5.2.4 数据流图	102
5.3 概念结构设计	102
5.3.1 E-R 模型基本概念	103
5.3.2 E-R 模型图的绘制	105

5.3.3	概念结构设计的方法和步骤	107
5.3.4	数据抽象和局部 E-R 模型设计	108
5.3.5	全局 E-R 模型设计	110
5.4	逻辑结构设计	112
5.4.1	逻辑结构设计任务和步骤	112
5.4.2	E-R 模型向关系模型的转换	112
5.4.3	关系模型的优化	113
5.4.4	设计用户子模式	114
5.5	物理结构设计	114
5.5.1	确定数据库的物理结构	115
5.5.2	评价物理结构	117
5.6	数据库的实施和维护	117
5.6.1	数据的载入	117
5.6.2	数据库试运行	118
5.6.3	数据库的运行与维护	119
5.7	小结	120
5.8	习题	120
第 6 章	数据库设计案例	123
6.1	需求分析	123
6.1.1	功能需求分析	123
6.1.2	数据流图	124
6.1.3	数据字典	126
6.2	概念结构设计	129
6.2.1	实体	129
6.2.2	实体之间的联系	129
6.2.3	局部 E-R 图	130
6.2.4	全局 E-R 图	130
6.3	逻辑结构设计	131
6.3.1	E-R 图到关系模式的转换	132
6.3.2	关系模式的优化	132
6.3.3	表结构设计	134
6.4	物理结构设计	136

6.4.1 确定存储介质	136
6.4.2 创建数据库	137
6.4.3 建立表	139
6.4.4 设计索引	141
6.4.5 数据库服务器性能优化	141
6.6 小结	142
6.7 习题	142
第7章 数据库安全与维护	145
7.1 数据库的安全性控制	145
7.1.1 数据库系统安全主要风险	145
7.1.2 安全性控制	146
7.2 数据库系统的完整性	149
7.2.1 完整性约束条件	150
7.2.2 完整性控制	152
7.3 并发控制	152
7.3.1 事务	153
7.3.2 并发操作引起的问题	154
7.3.3 封锁	155
7.4 数据库备份	159
7.4.1 基本概念	159
7.4.2 数据备份系统的分类	159
7.4.3 数据备份系统的工作原理	160
7.4.4 数据备份系统的主流备份技术	161
7.5 数据库恢复技术	163
7.5.1 数据库恢复概述	163
7.5.2 恢复的基本原理	163
7.5.3 故障恢复的方法	165
7.6 Oracle 备份与恢复的基本策略	167
7.7 备份与恢复系统总体设计	168
7.8 小结	169
7.9 习题	170

第 8 章 SQL Server 2012 数据库	171
8.1 SQL Server 2012 简介	171
8.1.1 SQL Server 2012 的新功能与优势	171
8.1.2 SQL Server 2012 的组成	172
8.1.3 SQL Server 2012 的版本	173
8.2 安装 SQL Server 2012	174
8.2.1 SQL Server 2012 Express 简介	174
8.2.2 下载 SQL Server 2012 Express	175
8.2.3 安装 SQL Server 2012 Express	176
8.3 SQL Server 管理工具	180
8.3.1 SQL Server 配置管理器	180
8.3.2 SQL Server Management Studio	184
8.4 创建和修改数据库	186
8.4.1 SQL Server 数据库概述	186
8.4.2 创建数据库	189
8.4.3 修改数据库	197
8.4.4 删除数据库	198
8.5 数据表的操作	199
8.5.1 表的概念	199
8.5.2 数据类型	199
8.5.3 表结构的创建、修改	203
8.5.4 表数据的操作	209
8.5.5 删除表	213
8.5.6 创建约束	214
8.5.7 索引	223
8.5.8 视图	226
8.6 数据库的备份与恢复	229
8.7 小结	233
8.8 习题	233
参考文献	235

第1章 数据库系统概论

数据库技术是现代信息科学与技术的重要组成部分，是计算机数据处理与信息管理的核心。数据库技术拟研究和解决在计算机信息处理过程中如何有效地组织和存储大量数据的问题，以及在数据库系统中减少数据冗余、实现数据共享、保障数据安全，高效地检索数据和处理数据的问题。

本章介绍与数据库有关的一些基本概念、数据管理技术的发展阶段、数据模型、数据库系统结构及其组成等。要求掌握数据库的基本概念；了解数据管理的3个发展阶段以及各阶段的主要特点；了解数据库系统的组成及各组成部分的主要功能。重点掌握实体、属性的定义，实体之间的联系类型及其特点，E-R图、三级模式与二级映像等。

学习本章的目的在于了解本课程相关的基础知识，掌握基本概念，为以后的学习打下坚实的基础。有些概念一开始接触会感觉比较抽象，但随着学习进度的逐渐推进，在后续章节中，这些抽象的概念会逐渐变得清晰和具体起来。

1.1 数据管理技术的发展

数据库技术产生于60年代末，是数据管理的最新技术，是计算机科学的重要分支。数据库技术是信息系统的核心和基础，它的出现极大地促进了计算机应用向各行各业的渗透。数据库的建设规模、数据库信息量的大小和使用频度已经成为衡量一个国家信息化程度的重要标志。

数据管理(Data Management)是指数据的收集、整理、组织、存储、维护、检索、传送等操作，这些操作是数据处理业务的基本环节。数据管理技术的优劣，将直接影响数据处理的效率。用计算机进行数据管理是将大量的数据有组织地存储在计算机的存储介质中，以便于维护、检索、重组和处理。

计算机的软、硬件发展水平，特别是存储技术的发展水平与数据管理技术的发展密切相关。数据管理技术的发展大致经历了3个阶段：人工管理阶段、文件系统阶段和数据库系统阶段。

1.1.1 人工管理阶段

人工管理阶段大约在 20 世纪 50 年代中期以前,当时的计算机主要用于科学计算,外存只有磁带、纸带和卡片等,软件只有汇编语言,尚无数据管理方面的软件,数据处理方式以批处理为主。这个时期的数据不保存在机器中,没有专用的软件对数据进行管理,数据与程序在一起,数据程序对应。在人工管理阶段,应用程序与数据之间的对应关系可用图 1-1 来表示。人工管理数据具有以下特点。

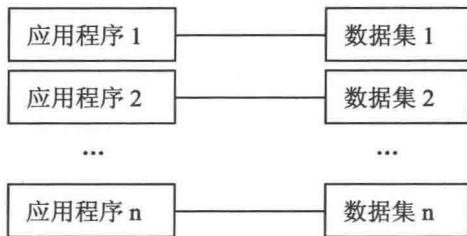


图 1-1 人工管理阶段应用程序与数据之间的对应关系

(1) 数据不保存

此阶段的计算机主要用于科学计算,并不对数据进行其他操作,一般不需要将数据长期保存,只是在计算某一课题时将数据批量输入,数据处理完以后不保存原始程序和数据。计算机断电之后计算结果也会随之消失。

(2) 应用程序管理数据

数据需要应用程序自己管理,没有相应的软件系统负责数据的管理工作。程序员不仅要规定数据的逻辑结构,而且要设计数据的物理结构,包括存储结构、存取方法和输入输出方式等,使得程序员负担很重。

(3) 数据不共享

一组数据只能对应一个程序,数据是面向应用的。各个应用程序的数据各自组织,无法互相利用和互相参照,因此,程序与程序之间有大量的冗余数据。

(4) 数据不具有独立性

数据的逻辑结构和物理结构都不具有独立性。当数据的逻辑结构或物理结构发生变化后,必须对应用程序做相应的修改,从而给程序员设计和维护带来繁重的工作。

1.1.2 文件管理阶段

文件(File System)管理阶段大约在 20 世纪 50 年代后期至 60 年代中期,当时的计算机不仅用于科学计算,还用于信息管理。在软件方面,数据结构和数据管理软件迅速发展,出现了高级语言和操作系统;硬件方面,已经有了磁盘、磁鼓等存储设备。

在这一阶段,数据可长期保存在外存的磁盘上,数据的逻辑结构与物理结构有了区别,文件组织呈现多样化,有索引文件、链接文件和散列文件等。但文件之间相互独立,缺乏联系。数据不再属于某个特定的程序,可以重复使用。

随着数据管理规模的扩大和数据量急剧增加,文件系统显露出了一些缺陷,可能会造成数据冗余以及数据不一致等现象。这些问题促使人们研究新的数据管理技术,因此,在20世纪60年代末产生了数据库技术。

文件系统阶段应用程序与数据之间的关系如图1-2所示。

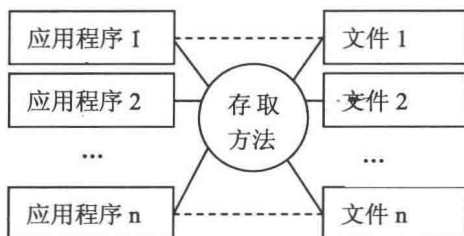


图 1-2 文件系统阶段应用程序与数据之间的对应关系

1. 文件系统的优点

此阶段数据管理具有如下优点:

(1) 数据可以长期保存

由于计算机大量用于数据处理,数据需要长期保留以便在外存上反复进行查询、修改、插入和删除等操作。

(2) 文件系统管理数据

由专门的软件即文件系统进行数据管理,文件系统把数据组织成相互独立的数据文件,利用“按文件名访问,按记录进行存取”的管理技术,可以对文件中的数据进行修改、插入和删除等操作。文件系统实现了记录内的结构化,但就文件整体而言是无结构的。程序和数据之间由文件系统提供的存取方法进行转换,使应用程序与数据之间有了一定的独立性。程序员可以不必过多地考虑物理细节,将精力集中于算法。而且数据在存储上的改变不一定反映在程序上,这就大大节省了程序的维护工作。

2. 文件系统的缺点

尽管文件系统有上述优点,但它仍存在一些缺点,主要表现在以下几个方面:

(1) 数据的共享性差,冗余度高

在文件系统中,数据的建立、存取都依赖于应用程序,基本是一个(或一组)数据文件对应一个应用程序,即数据仍然是面向应用的。当不同的应用程序具有部分相同的数据时,也必须建立各自的文件,而不能共享相同的数据,因此,数据的冗余度大,

浪费存储空间。同时,由于相同数据的重复存储和各自管理,容易造成数据的不一致,给数据的修改和维护带来困难。

(2) 数据的独立性不足

文件系统中的数据虽然有了一定的独立性,但是由于数据文件只存储数据,由应用程序来确定数据的逻辑结构,设计数据的物理结构,一旦数据的逻辑结构或物理结构需要改变,则必须修改应用程序;或者由于语言环境的改变需要修改应用程序时,也将引起文件数据结构的改变。因此,数据与应用程序之间的逻辑独立性不强。另外,要想对现有的数据再增加一些新的应用会很困难,系统不容易扩充。

(3) 并发访问容易产生异常

文件系统缺少对并发操作进行控制的机制,虽然系统允许多个用户同时访问数据,但是由于并发的更新操作相互影响,容易导致数据的不一致。

(4) 数据的安全控制难以实现

数据不是集中管理。在数据的结构、编码、表示格式、命名以及输出格式等方面不容易做到规范化、标准化,所以其安全性、完整性得不到可靠保证,而且文件系统难以实现对不同用户的不同访问权限的安全性约束。

1.1.3 数据库系统阶段

在 20 世纪 60 年代末,磁盘技术取得了重要进展,具有数百兆容量和快速存取的磁盘陆续进入市场,为数据库技术的产生提供了良好的物质条件。

数据管理技术进入数据库阶段的标志,是 20 世纪 60 年代末发生的 3 件大事:

(1) 1968 年,美国 IBM 公司推出层次模型的 IMS 系统。

(2) 1969 年 10 月,美国数据系统语言协会(CODASYL)的数据库任务组(DBTG)发表关于网状模型的 DBTG 报告(1971 年通过)。

(3) 1970 年, E.F.Codd 连续发表论文,提出关系模型,奠定了关系数据库的理论基础。

20 世纪 70 年代以来,数据库技术迅速发展,不断有新的产品投入运行。数据库系统克服了文件系统的缺陷,使数据管理变得更有效、更安全。

数据库系统阶段的数据管理方式具有以下特点:

(1) 数据结构化

数据结构化是数据库与文件系统的本质区别。数据库系统阶段采用复杂的数据模型来表示数据结构。

(2) 数据共享性高、冗余度低、易扩充

数据库系统从整体角度看待和描述数据,数据不再是面向某个应用而是面向整个系统,因此,数据可以被多个用户、多个应用共享使用。数据共享可以大大减少数据冗余,节约存储空间,还能够避免数据之间的不相容性与不一致性。所谓数据的不一致性是指同一数据不同拷贝的值不一样。采用人工管理或文件系统管理时,由于数据被重复存储,当不同的应用使用和修改不同的拷贝时,就很容易造成数据的不一致。在数据库中数据共享减少了由于数据冗余造成的一致性问题。由于数据面向整个系统,是有结构的数据,不仅可以被多个应用共享使用,而且容易增加新的应用。这就使得数据库系统弹性大,易于扩充,可以适应各种用户的要求。可以取整体数据的各种子集用于不同的应用系统,当应用需求发生改变或增加时,只要重新选取不同的子集便可满足新的需求。

(3) 数据独立性高

数据独立性是数据库领域的一个常用术语,包括数据的物理独立性和数据的逻辑独立性。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。也就是说,数据在磁盘上的数据库中如何存储是由 DBMS 管理的,用户程序不需要了解,应用程序要处理的只是数据的逻辑结构,这样,当数据的物理存储改变了,应用程序也不用改变。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。也就是说,数据的逻辑结构改变了,用户程序也可以不变。(数据独立性是由 DBMS 的二级映像功能来保证的,将在后面讨论。)

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

(4) 数据由数据库管理系统(DBMS)统一管理和控制

数据由数据库管理系统(DBMS)统一管理和控制,DBMS提供了以下几方面的数据控制功能。

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,使每个用户只能按规定对某些数据以某些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。完整性检查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。

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,必须对多用户的并发操作加以控制和协调。

④ 数据库恢复

计算机系统的硬件故障、软件故障、操作员的失误以及故意的破坏都会影响数据库中数据的正确性,甚至造成数据库部分或全部数据的丢失。DBMS 必须具有将数据库从错误状态恢复到某一已知的正确状态(亦称为完整状态或一致状态)的功能,这就是数据库的恢复功能。

数据库阶段的程序和数据之间的联系可以用图 1-3 表示。

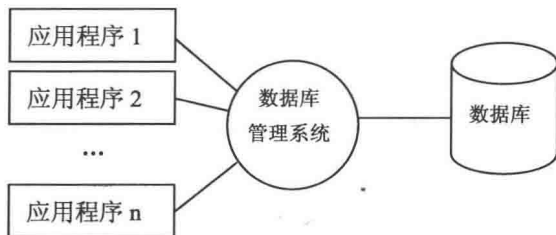


图 1-3 程序和数据库间的联系

在数据库技术中经常会涉及到以下 5 个名词:

数据(Data): 描述事物的符号记录。记录形式可以是文字、图形、图像、声音等。

数据库(Database, DB): 长期存储在计算机内、有组织、可共享的数据集合。具有较小的数据冗余,较高的数据独立性和易扩展性。

数据库管理系统(Database Management System, DBMS): 位于用户与操作系统之间的一层数据管理软件,为用户或应用程序提供访问数据库的方法,包括数据库的建立、查询、更新以及各种数据控制,能够确保数据的完整性和安全性。

数据库系统(Database System, DBS): 实现有组织地、动态地存储大量关联数据,方便多用户访问的计算机软件、硬件和数据资源组成的系统,即采用了数据库技术的计算机系统。一般由数据库、数据库管理系统、应用系统、数据库管理员(Database Administrator, DBA)和用户(User)构成。

数据库技术(Database Technology): 一门研究数据库的结构、存储、管理和使用的软件学科。数据库技术是在操作系统的文件系统基础上发展起来的,要用到数据结构、集合论、数理逻辑等理论知识。因此,数据库技术是一门综合性较强的学科。

1.2 数据描述

1.2.1 数据描述的3个领域

从事物的特征到计算机中数据的表示,数据描述要经历3个不同的领域: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们头脑之外的客观世界,它是原始数据(记录)的来源。

2. 信息世界

信息世界是指现实世界在人们头脑中的反映。人们用文字、符号、图形、图像、声音等方式记载下现实世界的信息,我们称之为信息世界。

在信息世界中,常用的术语有以下4个:

实体(Entity):客观存在、可以相互区别的事物称为实体。实体可以是具体的对象,如:一个员工、一本教材等。也可以是抽象的事件,如:一次旅游、一场球赛等。

实体集(Entity Set):性质相同的同类实体的集合称为实体集。如:学生等。

属性(Attribute):实体有很多特性,每个特性称为一个属性。每个属性有一个数据类型和值域。例如:学生有学号、姓名、年龄、性别等属性。学号、姓名、性别的数据类型是字符串,而年龄的数据类型是整数,其值域在13~30之间。

实体标识符(Identifier):能唯一标识每个实体的属性或属性集称为实体标识符,或简称键。例如,学生的学号可以作为学生实体标识符。

3. 机器世界

信息世界的信息在机器中以数据形式存储,成为机器世界。在机器世界中,常用的术语有以下4个:

字段(Field):标记实体属性的命名单位称为字段或数据项。它是可以命名的最小信息单位,所以又称为数据元素或初等项。字段的命名往往与属性名相同。例如:一个学生记录中有学号、姓名、年龄、性别等字段。

记录(Record):字段的有序集合称为记录。一般用一个记录描述一个实体,所以记录又可以定义为能完整地描述一个实体的字段集。例如:一个学生记录由有序的字段集组成:(学号,姓名,年龄,性别)。