



普通高等教育“十一五”规划教材

数据库



原理与应用

SHUJUKU YUANLI YU YINGYONG

蔡延光 主编



机械工业出版社
CHINA MACHINE PRESS



普通高等教育“十一五”规划教材

数据库原理与应用

主编 蔡延光

参编 黄永慧 邢延 张钢



机械工业出版社

本书以数据库应用系统的生存期模型为主线,从数据库需求分析、数据库设计、数据库标准语言 SQL、数据库实现、数据库测试、数据库维护、数据库应用系统开发实践、关系数据库理论等方面阐述数据库课程的理论与实践教学内容,并附有实验方案和课程设计指导。

本书体系结构新颖、内容完整、素材新、工程背景强、概念清楚、重点突出、叙述流畅、通俗易懂,可作为高等学校电子信息类相关专业(如网络工程、自动化、电气工程及其自动化、机械设计及其自动化、信息工程)本科生教材,也可作为从事计算机软件研究与开发工作的科研人员和工程技术人员的参考书。

本书配有免费电子课件,欢迎选用本书作教材的老师登录 www.cmpedu.com 下载或发邮件到 EdmondYan@hotmail.com 索取。

图书在版编目 (CIP) 数据

数据库原理与应用/蔡延光主编. —北京:机械工业出版社, 2009. 1

普通高等教育“十一五”规划教材

ISBN 978 - 7 - 111 - 25669 - 4

I. 数… II. 蔡… 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 (2008) 第 186480 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

责任编辑: 贡克勤 版式设计: 霍永明 责任校对: 姚培新

封面设计: 张 静 责任印制: 邓 博

北京京丰印刷厂印刷

2009 年 1 月第 1 版 · 第 1 次印刷

184mm × 260mm · 16.25 印张 · 399 千字

标准书号: ISBN 978 - 7 - 111 - 25669 - 4

定价: 30.00 元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

销售服务热线电话: (010) 68326294

购书热线电话: (010) 88379639 88379641 88379643

编辑热线电话: (010) 88379711

封面无防伪标均为盗版

前言

数据库技术是计算机科学技术中发展最快、应用最广泛的领域之一。数据库应用系统在我们日常生活和工作中随处可见,如生产管理系统、销售管理系统、医院门诊管理系统、手机短信管理系统、图书管理系统、学生成绩管理系统等。目前数据库技术课程已成为大学本科电子信息类专业的专业基础课或专业选修课。本书就是针对高等学校电子信息类相关专业(如网络工程、自动化、电气工程及其自动化、机械设计及其自动化、信息工程)本科教学要求编写的。

本书以数据库应用系统的生存期模型为主线,从数据库需求分析、数据库设计、数据库标准语言 SQL、数据库实现、数据库测试、数据库维护、数据库应用系统开发实践、关系数据库理论等方面阐述数据库课程的理论与实践教学内容,并附有实验方案和课程设计指导。本书主要特色有两个。一是体系结构新颖、内容完整:数据库需求分析部分内容新颖且完整,数据库测试、数据库应用系统开发部分分别独立成章;学生掌握本书内容后,不需学习管理信息系统或软件工程等后续课程就会运用数据库技术解决实际问题。二是素材新、工程背景强:除经典案例外,本书的不少例题、习题的素材来自于编者们的科研实践;本书特别注意选取一些现实生活中与数据库相关的热点问题素材,同时注意选取嵌入式系统应用背景素材。

全书共 11 章:第 1 章介绍了数据库系统基础知识,第 2 章介绍数据库需求分析,第 3 章介绍数据库设计,第 4 章介绍关系数据库的 SQL 语言,第 5 章介绍数据库实现,第 6 章介绍数据库测试,第 7 章介绍数据库维护,第 8 章介绍数据库应用系统开发,第 9 章介绍关系数据库理论,第 10 章及第 11 章分别为本书配套实验及课程设计的及要求及实施内容。

本书第 1 章、第 2 章 2.1~2.10 节与第 6 章由蔡延光编写,第 2 章 2.11 节、第 3 章、第 4 章、第 10 章与第 11 章由黄永慧编写,第 7 章与第 9 章由邢延编写,第 5 章与第 8 章由张钢编写。全书由蔡延光统稿。

从本书的撰写到出版过程中,章云教授、王钦若教授、程良伦教授、唐平教授、何小敏副教授、刘治副教授、鲍芳副教授、黄英副教授对本书体系结构和内容取舍等方面提出了许多建设性意见;研究生丁志勇、李永生、林灼强、程明君、魏明、张敏捷、宋康等也做了许多有益的工作;机械工业出版社贡克勤老师为本书的编辑出版付出了很多心血。作者对他们支持和帮助表示诚挚的谢意。

由于作者水平有限,本书肯定存在不少的缺点和不足,欢迎专家和读者批评指正。

编者

目录

前言

第1章 数据库系统概述	1
1.1 数据库系统的基本概念	1
1.1.1 信息的定义与特征	1
1.1.2 数据的定义与特征	2
1.1.3 数据类型	3
1.1.4 数据库及其特点	4
1.1.5 数据库管理系统	5
1.1.6 数据库系统	6
1.2 数据库技术的发展历程	10
1.2.1 人工管理阶段	10
1.2.2 文件系统阶段	11
1.2.3 数据库系统阶段	13
1.3 数据库应用系统的生存期模型	14
1.3.1 数据库应用系统的生存期	14
1.3.2 瀑布模型	15
1.4 数据库技术的发展趋势	16
习题	19
第2章 数据库需求分析	21
2.1 概述	21
2.1.1 数据库需求分析的定义	21
2.1.2 数据库需求分析的内容	22
2.1.3 数据库需求分析的原则	22
2.1.4 数据库需求分析的步骤	24
2.1.5 数据库系统需求分析的方法和工具	25
2.2 数据库需求调查	27
2.2.1 数据库需求调查的内容	27
2.2.2 数据库需求调查的原则	28
2.2.3 数据库需求调查的方法	28
2.3 数据字典	29
2.3.1 数据字典的基本概念	29
2.3.2 数据项	30
2.3.3 数据结构	31
2.3.4 外部实体	32
2.3.5 数据流	32

2.3.6 数据存储	33
2.3.7 处理逻辑	34
2.3.8 数据字典编制的结构化方法	35
2.4 信息分类与编码	36
2.4.1 信息分类	37
2.4.2 信息编码	38
2.5 数据定义分析	41
2.6 数据操纵分析	41
2.6.1 数据更新分析	41
2.6.2 数据查询分析	42
2.6.3 数据统计和排序分析	42
2.7 数据完整性分析	42
2.7.1 数据项取值范围	43
2.7.2 关键字	43
2.7.3 数据关联	43
2.8 数据安全性分析	44
2.8.1 数据安全的概念	44
2.8.2 数据访问权限	44
2.8.3 数据加密	47
2.9 并发处理分析	47
2.10 数据库性能分析	48
2.11 E-R 图	49
2.11.1 E-R 图的基本概念	49
2.11.2 E-R 图设计的原则	52
2.11.3 E-R 图设计的方法	52
2.11.4 E-R 图的设计步骤	53
习题	56
第3章 数据库设计	58
3.1 概述	58
3.1.1 数据库设计的定义	58
3.1.2 数据库设计的任务	59
3.1.3 数据库设计的原则	59
3.1.4 数据库设计的方法	60
3.1.5 数据库设计的步骤	60
3.2 关系数据库的基本概念	61
3.2.1 数据模型概述	61
3.2.2 关系模型	62
3.3 关系操作	65
3.3.1 关系操作概述	65
3.3.2 关系代数	65
3.3.3 存储过程	70
3.4 完整性设计	71
3.4.1 完整性约束	72

3.4.2	完整性设计中应注意的问题	73
3.4.3	完整性设计的原则	74
3.4.4	触发器	74
3.5	安全性设计	75
3.5.1	访问权限设计	75
3.5.2	加密算法设计	76
3.6	索引设计	77
3.6.1	索引的概念	77
3.6.2	索引的作用	78
3.6.3	索引的设计原则	79
3.7	事务与并发控制	79
3.7.1	事务	79
3.7.2	并发控制	80
3.8	E-R 图转换	85
3.8.1	实体集向关系模式的转换	85
3.8.2	联系集向关系模式的转换	85
3.8.3	弱实体集向关系模式的转换	89
3.9	层次数据库与网状数据库	89
3.9.1	层次数据库	89
3.9.2	网状数据库	91
	习题	92
第4章	数据库标准语言 SQL	94
4.1	SQL 概述	94
4.1.1	SQL 简介	94
4.1.2	SQL 的主要功能	94
4.1.3	SQL 语言的主要特点	95
4.1.4	SQL 语法	96
4.2	数据定义	97
4.2.1	数据库	98
4.2.2	基本表	99
4.2.3	索引	103
4.3	数据查询	104
4.3.1	概述	105
4.3.2	简单查询	105
4.3.3	连接查询	109
4.3.4	使用 SQL 函数查询	112
4.3.5	分组查询	114
4.3.6	嵌套查询	115
4.3.7	集合查询	118
4.4	数据更新	119
4.4.1	插入数据	119
4.4.2	修改数据	120
4.4.3	删除数据	121

4.5 视图	123
4.5.1 定义视图	123
4.5.2 删除视图	123
4.5.3 查询视图	124
4.5.4 更新视图	124
4.6 数据控制	125
4.6.1 授权	125
4.6.2 收权	126
4.6.3 拒绝访问	126
4.7 嵌入式 SQL	127
4.7.1 嵌入式 SQL 概述	127
4.7.2 不用游标的 SQL	129
4.7.3 使用游标的 SQL	131
4.8 存储过程的创建与维护	133
4.9 触发器的创建与维护	134
4.9.1 定义触发器	134
4.9.2 修改触发器	135
4.9.3 删除触发器	136
习题	136
第5章 数据库实现	138
5.1 概述	138
5.2 主流 DBMS	139
5.2.1 Microsoft SQL Server	139
5.2.2 Access	139
5.2.3 MySQL	140
5.2.4 Oracle	141
5.2.5 Sybase	141
5.2.6 Informix	142
5.3 Microsoft SQL Server 2000	142
5.3.1 Microsoft SQL Server 2000 的安装与使用	142
5.3.2 查询分析器	145
5.3.3 企业管理器	147
习题	148
第6章 数据库测试	149
6.1 概述	149
6.1.1 数据库测试的定义	149
6.1.2 数据库测试的原则	150
6.1.3 数据库测试的方法	150
6.1.4 数据库测试的步骤	151
6.1.5 数据库测试输出结果的展现	152
6.2 测试用例设计	152
6.2.1 白盒测试的测试用例设计	153
6.2.2 黑盒测试的测试用例设计	156

6.3 数据库模式测试	159
6.3.1 基本表数据结构测试	160
6.3.2 数据库完整性测试	160
6.3.3 视图测试	160
6.4 数据库功能测试	160
6.4.1 数据定义功能测试	160
6.4.2 数据操纵功能测试	161
6.4.3 数据库安全性测试	162
6.4.4 并发处理测试	162
6.5 数据库性能测试	163
6.5.1 数据库性能测试的概念	163
6.5.2 影响数据库性能的因素	163
6.6 数据库调试	164
6.6.1 数据库调试的原则	164
6.6.2 数据库调试的步骤	165
6.6.3 数据库调试的方法	166
习题	166
第7章 数据库维护	168
7.1 概述	168
7.1.1 数据库维护的定义	168
7.1.2 数据库维护的任务	168
7.1.3 数据库维护的原则	169
7.1.4 数据库维护的方法	169
7.1.5 数据库维护的步骤	169
7.2 运行日志	170
7.2.1 日志文件的格式和内容	170
7.2.2 日志文件的作用	170
7.2.3 日志文件的使用	170
7.2.4 日志文件的维护	171
7.3 数据库故障及其排除	171
7.3.1 故障的种类	171
7.3.2 故障的排除方法	172
7.4 数据库备份与恢复	172
7.4.1 数据库备份的原则和方式	172
7.4.2 数据库恢复策略	173
7.4.3 具有检查点的数据库恢复技术	175
7.4.4 数据库的镜像功能	176
习题	176
第8章 数据库应用系统开发	178
8.1 概述	178
8.1.1 数据库应用系统的基本结构	178
8.1.2 数据库应用系统开发的要求	179
8.2 数据库接口技术	180

8.2.1	ODBC	180
8.2.2	ADO	181
8.2.3	JDBC	182
8.3	Java 数据库应用系统开发	183
8.3.1	Java 语言	183
8.3.2	Java 数据库应用系统开发实例	183
8.4	C++ 数据库应用系统开发	191
8.4.1	C++ 语言	191
8.4.2	C++ 数据库应用系统开发实例	192
	习题	202
第9章	关系数据库理论	203
9.1	关系模式的规范化理论概述	203
9.1.1	关系模式规范化的必要性	203
9.1.2	关系模式规范化的概念	205
9.2	函数依赖及范式	205
9.2.1	属性间的联系	205
9.2.2	函数依赖	206
9.2.3	第一范式	207
9.2.4	第二范式	207
9.2.5	第三范式	208
9.2.6	Boyce Codd 范式	208
9.3	多值依赖及范式	209
9.3.1	多值依赖的定义和性质	209
9.3.2	第四范式	210
9.4	连接依赖及范式	210
9.4.1	连接依赖的定义	210
9.4.2	第五范式	211
9.4.3	小结	211
9.5	模式分解	212
9.5.1	模式分解的概念	212
9.5.2	模式分解的算法	215
9.6	查询优化	217
9.6.1	查询优化的必要性	218
9.6.2	查询优化的一般准则	219
9.6.3	关系代数等价变换规则	220
9.6.4	关系代数表达式的优化算法	221
9.6.5	优化的一般步骤	223
	习题	226
第10章	实验方案	228
10.1	实验目的与实验要求	228
10.2	实验条件	228
10.3	实验内容	229
10.3.1	实验1 数据库需求分析	229

10.3.2	实验2 关系数据库设计	230
10.3.3	实验3 数据定义	230
10.3.4	实验4 数据更新	232
10.3.5	实验5 简单查询和连接查询	232
10.3.6	实验6 嵌套查询	233
10.3.7	实验7 集合查询、分组查询、函数查询	234
10.3.8	实验8 存储过程	234
10.3.9	实验9 触发器	235
第11章	课程设计指导	236
11.1	概述	236
11.1.1	课程设计的目标	236
11.1.2	课程设计的要求	237
11.1.3	课程设计的方法与步骤	237
11.2	成绩管理系统	237
11.2.1	背景简介	237
11.2.2	系统目标	238
11.2.3	功能要求	238
11.2.4	运行环境、开发环境与开发工具	239
11.3	图书管理系统	239
11.3.1	背景简介	239
11.3.2	系统目标	239
11.3.3	功能要求	240
11.3.4	运行环境、开发环境与开发工具	240
11.4	采购管理系统	240
11.4.1	背景简介	240
11.4.2	系统目标	241
11.4.3	功能要求	241
11.4.4	运行环境、开发环境与开发工具	241
11.5	库存管理系统	242
11.5.1	背景简介	242
11.5.2	系统目标	242
11.5.3	功能要求	242
11.5.4	运行环境、开发环境与开发工具	243
11.6	银行储蓄卡管理系统	243
11.6.1	背景简介	243
11.6.2	系统目标	243
11.6.3	功能要求	243
11.6.4	运行环境、开发环境与开发工具	244
11.7	房地产中介管理系统	244
11.7.1	背景简介	244
11.7.2	系统目标	244
11.7.3	功能要求	244
11.7.4	运行环境、开发环境与开发工具	245

11.8 手机联系人管理系统	245
11.8.1 背景简介	245
11.8.2 系统目标	245
11.8.3 功能要求	245
11.8.4 运行环境、开发环境与开发工具	246
参考文献	247

第 1 章

数据库系统概述

数据库技术是研究数据的组织、存储、管理和使用的一门计算机科学分支，其应用非常广泛。要掌握好数据库技术，首先必须掌握数据库系统的基本概念，了解数据库应用系统的开发过程。本章介绍信息、数据、数据类型、数据库、数据库系统、数据库管理系统、数据库应用系统等基本概念，介绍数据库应用系统的生存期及瀑布模型、数据库技术的发展历程和发展趋势。

1.1 数据库系统的基本概念

1.1.1 信息的定义与特征

1. 信息的定义

“信息”是日常生活中的常见名词。简单地讲，信息就是新的、有用的事实和知识。它具有时效性、有用性和知识性，是客观世界的反映。

信息的确切定义呈现出多定义而又无定论的局面，不同学科对信息的定义差异度较大。

信息论的创始人香农（C. E. Shannon）1948年在《通信的数学理论》一书中指出：“凡是在一种情况下能减少不确定性的任何事物都叫做信息。”这一定义指出了信息的一个价值——减少不确定性。

控制论的创始人维纳（N. Wiener）在《人有人的用处——控制论和社会》中说：“信息这个术语的内容就是我们对外界进行调节并使我们的调节为外界所了解时而与外界交换来的东西。”这一定义深入到了人与世界的交换关系，涉及交换内容，因此比前一个定义更适合用于研究探索人类的信息传播。

我国国家标准《情报与文献工作词汇基本术语》中，关于“信息”的解释是：“信息是物质存在的一种方式、形态或运动状态，也是事物的一种普遍属性，一般指数据、消息中所包含的意义，可以使消息中所描述事件的不定性减少。”这个定义首先明确了信息的本质是

物质的属性，而不是物质实体本身。客观存在的一切事物，包括自然界、人体本身和人类社会，都是在不断运动着的。运动的物质必然会产生相互作用和影响，从而引起物质结构、数量等多方面的变化。事物的这些变化便成为信息产生的物质基础。因此信息不是事物本身，而是由事物发出的数据中所包含的意义。

2. 信息的特征

信息的特征是指信息区别于其他事物的本质属性。信息的基本特征有以下几个方面。

(1) 信息的普遍性和客观性 世界是客观存在的物质世界，物质是运动的，运动的物质既产生也携带信息。因而信息是客观的、普遍存在的。信息的客观性表现为它是以物质客观存在为前提的。即使是主观信息，如决策、判断、指令等，也有它的客观实际背景，并受客观实践的检验。因此信息必须真实、准确地反映客观实际。

(2) 信息是有用的 信息的有用性是指利用信息可以创造价值，如提高决策的科学性、提高生产效率、提高产品销售量。

(3) 信息的依附性 信息的依附性是指信息总是依附于某种载体表现的，信息不能脱离载体而独立存在。信息的载体（或信息的表现形式）包括语言、文字、数据、符号、图像、声音、情景、表情、动作、状态等。

(4) 信息的共享性 信息的共享性是指同一信息可被许多人分享，且信息不会因为分享而减少或丧失。信息的共享性有两层含义：一是信息交换的双方，即传播者和接受者都可以享有被交换的同一信息；二是信息在交换或交流中，可以同时为众多的接受者接收和利用。

信息的共享性是相对而言的，信息的共享范围往往是受限制的。例如，某超市的销售信息对其管理人员来说是共享的，而对其竞争对手来说是不共享的。

值得注意的是，信息的共享性特征也引起了信息的安全性和保密性等一系列问题。

(5) 信息的可传输性 把信息从时间或空间的某一点向其他点移动的过程称为信息传输。信息的可传输性是指信息可以通过一定的传输工具和载体进行传输。信息可以通过多种渠道、采用多种方式进行传输。语言、表情、动作、报刊、书籍、广播、电视、电话、电报、通信卫星等是常用的信息传输方式。

(6) 信息的可识别性 信息的可识别性是指人们不仅可以通过眼、鼻、耳等感觉器官去感知信息，而且可以通过各种仪表器械去检测、识别和处理。

信息的识别方式分为直接识别和间接识别。直接识别是指通过感官的识别，间接识别是指通过各种测试手段的识别。不同的信息源有不同的识别方法。

(7) 信息的可压缩性 信息的可压缩性是指人们通过对信息进行加工、整理、概括、归纳等手段而使之精练、浓缩，用较少的信息量描述一事物主要特征。信息压缩在实际中很有必要，因为我们可能没有能力也没有必要收集、存储一个事物的全部信息。

(8) 信息的可转换性 信息的可转换性是指信息可以从一种表现形式转换为另一种表现形式。如自然信息可转换为语言、文字和图像等形式，也可转换为电磁波信号和计算机代码。

1.1.2 数据的定义与特征

1. 数据的定义

数据是信息的最佳表现形式之一，它通过可书写的信息编码表示信息。

2. 数据的特征

数据除了具有信息的所有特征外，还具有以下一些特征。

(1) 数据有“型”和“值” 数据的型指数据的结构，即数据的形式架构。数据的值是指数据的具体取值。如“学生”数据的型由“学号”、“姓名”、“性别”、“所在系”等构成，其中“学生”为数据名，而“2006010210、王燕、女、自动化系”是“学生”数据的一个值。

(2) 数据受数据类型和取值范围的约束 数据的类型包括字符型、数值型、日期型等简单数据类型，也包括由若干数据类型构造出的复杂数据类型。例如，“姓名”的数据类型为字符型，“年龄”的数据类型为数值型。数据类型不同，其表示形式、存储方式、操作运算一般也不同。例如，数值型数据可以进行加、减、乘、除等算术运算，而字符型数据则可以进行连接、取子串、查子串等运算。

数据是有取值范围约束的。例如，“年龄”不可能取值负数，也不可能太大，其取值范围一般为0~150。

(3) 数据有定性表示和定量表示 例如，职工的年龄既可以定性地表示为“老”、“中”、“青”，也可以定量地用具体岁数表示。

1.1.3 数据类型

常见的数据类型有字符型、数值型、日期型、逻辑型、备注型。

1. 字符型

字符型数据由中文字符、英文字符、数字字符、空格和其他专用符号组成。字符型数据一般有长度限制，最大长度不超过254B（字节）。例如，姓名、公司名称属于字符型数据。

2. 数值型

数值型数据只能由数字、小数点和正负号组成。数值型数据介于 -2^{63} 和 2^{64} 之间，数值的最大精度为16位有效数字，小数最多15位。数值型数据可以分为整数型、实数型和浮点型等。例如，人数、商品单价属于数值型数据。

3. 日期型

日期型数据是用来表示日期的特殊数据，有固定的格式。日期型数据长度固定为8位。一般日期格式“yyyy-mm-dd”表示“年-月-日”，美国日期格式“mm/dd/yyyy”表示“月/日/年”。例如，2008年10月1日的一般日期格式表示为“2008-10-1”、美国日期格式表示为“10/01/2008”。

4. 逻辑型

逻辑型数据是表示逻辑判断结果的值，只能取真（T, Y）或假（F, N）。逻辑型数据的长度固定为1bit（位）。

5. 备注型

备注型也称为记忆型、文本型。

当需要存储大量的字符时，应该使用备注型数据。注意，备注型数据一般长度较大，而字符型数据长度较小。

1.1.4 数据库及其特点

数据库已经与我们的日常工作、学习和生活密切相关，如学生成绩数据库、员工档案数据库、库存商品数据库、图书数据库、银行数据库、手机短信数据库等。

1. 数据库的定义

数据库 (Database, DB) 是一个按一定数据结构来组织、存储和管理的数据集合。

2. 数据库的特点

(1) 数据结构化 数据结构化是指数据库按照一定的数据模型来组织和存放数据。这种结构化了的数据能反映数据之间的自然联系。而传统的文件是无模型的，结构非常简单，因此不能反映数据之间的自然联系。数据结构化是数据库和文件之间的一个重要的本质差别，是实现对数据的集中控制和减少数据冗余的前提和保证。这里，数据冗余是指同一组数据在多个文件中同时出现所引起的数据重复现象。

(2) 数据独立性 数据独立性是指数据库的数据与应用程序之间不存在相互依赖关系，也就是数据的逻辑结构、存储结构和存取方法等不因应用程序的修改而修改，反之亦然。这是数据库与文件之间的另一个重要区别。数据独立性分为物理独立性和逻辑独立性两级。

1) 数据物理独立性是指数据的物理结构（或存储结构）与应用程序相互之间有较强的独立性。当数据的物理结构改变（如物理存储设备的更换、物理存储位置的变更、存取方法的改变）时，应用程序不需要修改仍能正常工作，反之亦然。

2) 数据逻辑独立性是指当数据的逻辑结构改变时，如修改数据结构、增加新的数据类型、改变数据间的联系等，无需修改原来的应用程序。

一个具有数据独立性的系统称之为以数据为中心的系统，或面向数据的系统。

(3) 实现数据共享 数据共享是指多个用户、多种程序语言、多个应用程序使用一些共同的数据。数据共享是促成发展数据库技术的重要原因之一，也是数据库技术先进性的一个重要体现。数据库中的数据可供多个应用程序、多个用户共同使用，这些应用程序可以用不同的程序设计语言编写，用户可以处于不同的业务部门、不同的物理位置。

(4) 数据的冗余度小 在数据库产生以前，每个应用都需要建立自己的数据文件，从而造成了数据冗余。数据冗余带来的弊病有：

1) 浪费存储空间。

2) 为了更新某些冗余副本，保持数据的一致性，必须执行多次更新操作，从而增加不必要的机器时间。

3) 由于数据的不同副本可能处于不同的更新阶段，从而可能给出不一致的信息。

数据库则是以整体观点来组织和存储数据的，数据是集成化、结构化的。它是为多种应用所共享的，从而大大减少了数据冗余。

(5) 避免了数据的不一致性 当本应相同的数据在不同的应用中出现不同的值时，我们就说数据出现了不一致。例如，同一职工的工资，如果在工资单上和人事档案中具有不同的值就出现了不一致。数据的不一致性主要是由数据冗余引起的。数据库在理论上消除了数据冗余，因而可以避免数据的不一致性。即使存在某些冗余，数据库也提供了对数据的控制和检查机制，保证在更新数据的同时更新其所有副本，从而保证了数据的一致性。

(6) 有利于数据的安全性 数据的安全性是指保护数据,防止不合法使用数据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访问和处理。只有 DBA (Database Administrator, 数据库管理员) 对数据库中的数据拥有完全的操作权限, DBA 可以规定各用户的权限。每当用户企图存取敏感数据时,数据库就进行安全性检查。对于数据库中数据的各种操作(查询、修改、删除等),数据库都可以实施不同级别的安全性检查。

(7) 有利于保证数据的完整性 数据的完整性是指数据的正确性、有效性和相容性,即将数据控制在有效的范围内,或要求数据之间满足一定的关系。数据的不一致性是失去完整性的例子。当然,数据冗余可能引起数据的不完整性,但是没有数据冗余同样可能出现不正确的数据而使数据库失去数据完整性。例如,在一个数值型数据中出现了字母、特殊符号等,或一个工人一周的工作时间超过 200h 等都是失去完整性的例子,它们与是否存在数据冗余无关。数据库的集中控制可以避免这些情况的出现。它通过 DBA 定义的完整性规则,对每一次更新操作实施完整性检查,保证数据的完整性。数据的完整性检查对于多用户系统尤为重要。例如,当一个用户正在修改某一数据时,别的用户也去修改这个数据就可能产生不完整的数据。

(8) 可以发现故障和恢复正常状态 实际工作中,人们很难保证数据库在运行过程中不受到破坏。实际上,硬件、软件故障及用户操作的失误,随时有可能使数据库遭到局部性的或全局性的破坏。数据库有一套及时发现故障,并迅速地把数据库恢复到故障以前的正确状态的措施,如转储、日志、检查点等方法。

1.1.5 数据库管理系统

1. 数据库管理系统的定义

数据库管理系统(Database Management System, DBMS)是对数据库进行管理的软件系统,它负责数据库中的数据组织、数据操纵、数据维护、数据控制及保护和数据服务等。

DBMS 建立在操作系统之上,实施对数据库的统一管理和控制。基于数据库的实际应用系统通过 DBMS 操作数据库中的数据。

2. 数据库管理系统的功能

DBMS 的主要功能包括以下几个方面:

(1) 数据定义 DBMS 通过其数据定义语言(Data Definition Language, DDL)定义数据库及其组成元素的结构。

(2) 数据操纵 DBMS 通过其数据操纵语言(Data Manipulation Language, DML)操纵数据库中的数据,即对数据库中的数据进行插入、修改、删除、查询、统计和排序等操作。

(3) 数据控制 DBMS 通过其数据控制语言(Data Control Language, DCL)对数据库中的数据进行安全性控制,以保证数据的安全。

(4) 数据组织与存取 DBMS 提供数据在外围存储设备上的物理组织与存取方法。

(5) 数据库建立与维护 数据库的建立是指数据的载入、转储、重组与恢复等。数据库的维护是指数据库及其组成元素的结构修改、数据备份等。

(6) 运行管理 DBMS 提供并发控制、数据存取控制、完整性控制、运行日志,以保证