

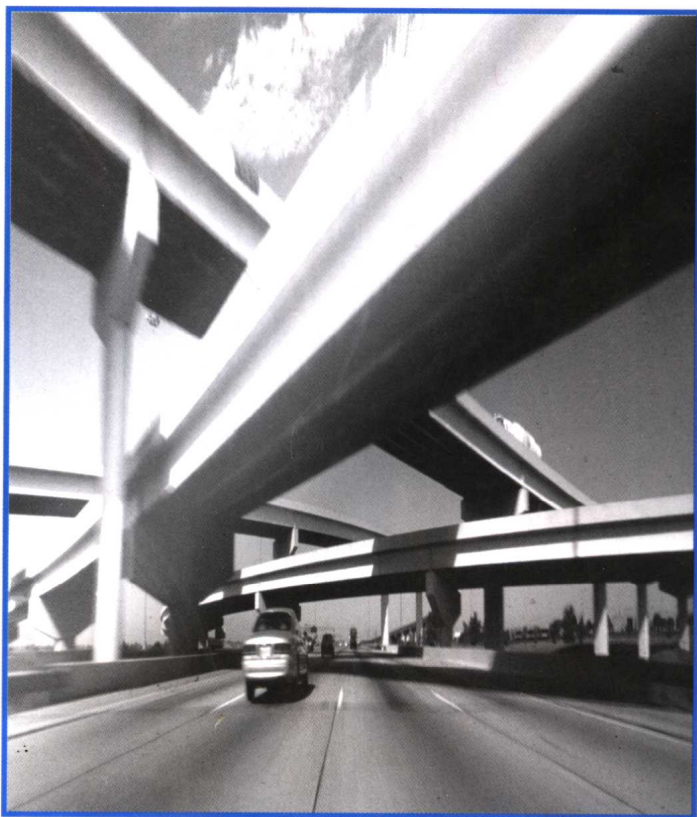
数据库系统原理与 应用技术基础(Oracle)

王 林 编著
徐汀荣 审

突出应用主题

具备理论先进性与后续性

贯彻可行性与熟练技巧培养

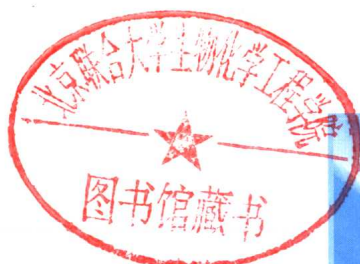


北京希望电子出版社
Beijing Hope Electronic Press
www.bhp.com.cn

TP311.13
/w00

数据库系统原理与 应用技术基础(Oracle)

王 林 编著
徐汀荣 审



突出应用主题

具备理论先进性与后续性

贯彻可行性与熟练技巧培养



北京希望电子出版社
Beijing Hope Electronic Press
www.bhp.com.cn

内 容 简 介

本书是教育部世界银行贷款 21 世纪初高等理工科教育教学改革项目中《以培养创新人才为目标的地方院校计算机科学与技术专业教学改革的研究与实践》项目成果之一,是“21 世纪高等院校计算机技术教学”丛书的一本。

本书比较系统、完整地讲述了当前常见数据库系统技术的基本原理,数据库系统的应用技术和 Oracle 数据库系统应用技术基础。全书共分 12 章,第 1 章主要介绍了数据库的基本概念,第 2~6 章讲述了数据模型、关系数据库系统及 SQL 语言、关系数据库的优化与实现技术、数据库设计、数据库系统成熟技术等内容,第 7~12 章主要讲述了常见关系数据库及其应用开发系统,Oracle 数据库各方面的内容和数据库技术的热点与前沿问题。此外,每章后均附有习题,以供课后练习。全书内容详实,结构清晰,语言流畅,力求做到理论与实践紧密结合。

本书既可作为高等院校计算机专业相关课程的教材,也可供数据库应用技术开发人员参考,本书还适用于对数据库感兴趣的各类自学人员使用。

本版 CD 为配套电子书。

盘书系列名 : 21 世纪高等院校计算机技术教学丛书 (8)

盘 书 名 : 数据库系统原理与应用技术基础 (Oracle)

总 策 划 : 北京希望电子出版社

文本著作者 : 王林 编著 徐汀荣 审

CD 制 作 者 : 希望多媒体开发中心

CD 测 试 者 : 希望多媒体测试部

责 任 编 辑 : 杨敏

出版、发行者 : 北京希望电子出版社

地 址 : 北京市海淀区知春路甲 63 号卫星大厦三层 100080

网址: www.bhp.com.cn

E-mail: lxr@bhp.com.cn

电 话 : 010-62520290,62521724,62528991,62630301,62524940,62521921,82610344

(发行) 010-82675588-202 (门市) 010-82675588-501,82675588-201 (编辑部)

经 销 : 各地新华书店、软件连锁店

排 版 : 希望图书输出中心 马君

CD 生 产 者 : 北京中新联光盘有限责任公司

文本印刷者 : 北京双青印刷厂

开本 / 规格 : 787 毫米×1092 毫米 16 开本 18.75 印张 427 千字

版次 / 印次 : 2003 年 3 月第 1 版 2003 年 3 月第 1 次印刷

印 数 : 0001~5000 册

本 版 号 : ISBN 7-89498-097-8

定 价 : 28.00 元 (本版 CD)

说明 : 凡我社产品如有残缺,可执相关凭证与本社调换。

21世纪高等院校计算机技术教学丛书

编委会成员名单

顾问委员：（以姓氏笔划为序）

刘 璟 教授 南开大学

何炎祥 教授 武汉大学

施伯乐 教授 复旦大学

钱培德 教授 苏州大学

主任委员：左孝凌 教授 上海交通大学

副主任委员：徐汀荣 教授 苏州大学

委员：（以姓氏笔划为序）

左孝凌 教授 上海交通大学

朱巧明 教授 苏州大学

宋方敏 教授 南京大学

余雪丽 教授 太原理工大学

徐汀荣 教授 苏州大学

陶树平 教授 同济大学

崔志明 教授 苏州大学

曾 明 教授 西安交通大学

谢康林 教授 上海交通大学

序

本丛书是苏州大学主持的高等院校计算机技术教学丛书，其编委会聚集了国内很多院校的专家，编委会为本丛书的组织、审定起到集思广益、推波助澜的作用。

苏州大学建校百年，蜚声海内外，学校治学严谨、理论底蕴深厚，理科各系，久负盛名。近年来学校多次调整院系，全校各系蓬勃发展，尤其是电子信息各科发展迅速。当前信息时代，信息技术的应用已经深入各行各业，而苏州大学更以信息技术的应用作为各个专业的发展之本。

发展信息技术，苏州大学具有以下三个方面的独特优势：

一是苏州大学计算机工程系成立于1987年10月，至今已15年，这正是信息科学突飞猛进的时期。当前计算机技术的应用已经遍及生产、生活各个方面，苏州大学计算机工程系顺时应势，使全系的教学科研迅速崛起。目前系内除有211工程重点建设的计算机应用学科外，还有江苏省重点实验室（计算机信息处理技术实验室）以及纵横汉字信息技术研究所、多媒体应用技术研究所等，在国内外极具影响的研究机构。苏州大学计算机工程系现有博士点一个、硕士点两个，这些都显示了该系具有强劲的得天独厚的科研教育环境。

二是苏州大学地处“天堂”苏州，南邻上海，北依南京，近年新加坡和中国在苏州合建工业园区，而世界著名信息公司Motorola也移师毗邻，使苏州成为其主要研发基地。现在苏州已经成为世界信息技术产业中心之一，信息技术产业产值已达一千亿元以上。苏州大学身临其境，近水楼台，耳濡目染，凭借着培养人才的基础以及软件开发的优势，为新加坡工业园区的建设以及Motorola的芯片研发等项目进行了广泛的优势互补的合作。

三是十几年来信息技术一直是苏州大学发展重点，校长钱培德教授是国内著名信息技术学科的学术带头人，强将手下无弱兵。在苏州大学有一批奋发有为、敬业创新在教学科研上崭露头角的中青年教师队伍，他们把长期从事科学研究的经验与信息教学的实践紧密结合起来编写了一套以理论底蕴为基础，面向应用为目标，充分反映当前计算机进展的实用教材。本丛书即是重要组成部分。

本丛书包括现代软件工程、单片机技术、数据库系统及SYBASE技术、计算机组网技术、计算机图形图像技术、Delphi程序设计、Java程序设计、计算机网页设计等。这些教材有的是计算机专业核心教材，有的是电子类相关专业的实用性应用教材。这套丛书的立足点是以计算机教学应用为主体，也能反映当前的发展动向，在理论上具有先进性和后续性，而在教学实践中努力贯彻可行性和熟练技巧培养。为了改革精简教学内容，突出应用主题，我们把网络原理等大课，分解为组网技术和网页设计等不同主题，以便适应有关专业的不同需要。

我们感谢北京希望电子出版社领导对本套丛书的立题、审稿、编辑、出版等做的大量工作。编委会希望广大读者能够指出丛书中的各种疏漏，期望这套丛书能在信息技术日新月异飞速发展的年代起到一点卵石铺路的作用。

高等院校计算机技术教学丛书编委会 左孝凌

2002年8月

前 言

本书为 21 世纪高等院校计算机技术教学丛书之一。

数据库是普通高校计算机专业和信息管理等专业的专业基础课。其主要任务是研究如何存储、使用和管理数据。目的是使学生掌握数据库的基本原理、方法和应用技术，能有效使用现有的数据库管理系统和软件开发工具，掌握数据库结构的设计和数据库应用系统的开发方式。

数据库技术是计算机软件领域的一个重要分支，随着数据库系统的推广，计算机应用已深入到社会各个角落。当今的管理信息系统（MIS）、办公信息系统（OIS）、计算机辅助设计与制造（CAD/CAM）、计算机集成制造系统（CIMS）、地理信息系统（GIS）等领域都以数据库技术为基础。在 20 世纪 90 年代初，我国已在国民经济各领域装备了以数据库为基础的大型计算机系统，从那时起，数据库技术亦日渐普及，数据库应用也日渐拓宽，据不完全统计，在计算机应用中涉及到数据库应用的百分比已占 60% 以上，数据库不仅在传统的商业领域中、管理领域中发挥着主要作用而且在非传统应用中也起到越来越大的作用，如在工程领域中，在数据集成领域中，在科学统计领域以及在图形、图像、CAD、多媒体等领域中起着关键作用。另一方面，人类进入了 21 世纪知识经济时代，信息已成为经济发展的战略资源，信息技术已成为社会生产力中重要的组成部分。人们充分认识到，数据库是信息化社会中信息资源管理与开发利用的基础。与此同时，随着技术的逐渐成熟，数据库自身也进行着显著的变化，关系数据库技术业已成为当前数据库发展的主流，层次与网络数据库已经退出历史舞台，另外，分布式数据库系统及面向对象数据库系统正在兴起。对于一个国家或地区，数据库的建设规模、使用水平已成为衡量其信息化程度的重要标志。因此，数据库是国内外计算机专业的一门重要的课程。

从另一方面来看，数据库是一门实践性极强的课程，我们编写本书的宗旨就是集理论、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信息管理系统的开发；系统的目标是通过使用某种商品型数据库，在选择数据库原理的实验对象时，我们注意到美国 ORACLE 公司的数据库产品自进入我国后，在我国不断地拓宽市场，引进最新技术和研究手段。随着 ORACLE 数据库版本的不断升级，它在中国软件产业的已经成为大型关系数据库系统中具有最大用户群的产品。

我们通过多年的数据库课程教学与科研工作的实践，形成了计算机及相关专业《数据库系统原理与应用》课程的教学模式：由实践（实用数据库语言程序设计）→理论（数据库系统原理与应用）→再实践（数据库课程设计）→再提高（毕业设计、科研），取得了比较明显的效果，学生的数据库设计能力及实际编程动手能力均有了一定程度的提高。本书就是在此基础上，并在作者历经使用国内外多本数据库教程后，参照其多方面的优点改写而成的。本书内容包括：数据库技术概论、关系数据库、结构化查询语言、关系数据库设计理论、数据库查询的实现与优化、数据库保护、数据库设计、数据库应用系统开发与维护和 ORACLE 应用技术基础。

鉴于以上有关理论与技术的考虑，我们既结合了数据库新技术有关材料以适应当前

数据库应用的发展,又涉及数据库理论自身的发展,正基于此目的,本书具有如下几个特点:

- 突出以关系数据库技术为主要内容。
- 加强应用技能的比重,尤其是 SQL 的应用比重。
- 增加数据库开发工具的内容。
- 介绍了分布式数据库及面向对象数据模型等内容。
- 强化了数据库设计的内容。
- 加入商品型数据库 ORACLE 的应用简介。
- 减少并弱化层次、网络模型的内容。

此外,本书增加了数据库应用系统的开发和维护的简介内容,使学生能运用关系数据库的设计理论和软件工程的规范并参照相关的文献来开发一个实际的数据库应用系统。

本书系教育部世界银行贷款 21 世纪初理工科教育教学改革项目——《以培养创新人才为目标的地方院校计算机科学技术专业教学改革的研究与实践》的成果之一。本书可适用于作为高等学校数据库课程教材,也可用于数据库应用开发人员参考之用,同时也适用于各类自学人员使用。

在本书的编写过程中,参考和引用了国内外杰出 IT 科学工作者和教育家的大量成果和著述,并得到出版社和本套丛书的多位专家顾问的指导和帮助,尤其是本书主审徐汀荣教授的悉心关怀,在此表示衷心的感谢。

作者通过编写本书学到了很多,也使自己在 ORACLE 数据库等方面的知识得到充实和进一步系统化。编者虽然竭尽全力,但由于水平有限,错误仍然在所难免。欢迎广大读者批评,提出宝贵的修改意见。

编者

2002 年 10 月于江南

目 录

第 1 章 数据库系统的基本概念	6
1.1 数据库、数据库管理系统与数据库系统	6
1.1.1 数据库	6
1.1.2 数据库管理系统	6
1.1.3 数据库应用	7
1.1.4 数据库管理员	8
1.1.5 系统硬件及相关软件	8
1.2 数据库系统的基本特点	8
1.3 数据库系统的工作过程	10
1.4 数据库技术的研究要点	11
习题 1	12
第 2 章 数据模型	14
2.1 数据模型的基本概念	14
2.1.1 三个抽象层次的数据模型	14
2.1.2 数据模型的三个要素	15
2.2 数据模型所对应的不同抽象层次	15
2.3 概念世界与概念模型	16
2.3.1 E-R 模型	16
2.3.2 EE-R 模型	21
2.3.3 面向对象模型	23
2.4 信息世界与逻辑模型	24
2.4.1 概述	24
2.4.2 层次模型与层次模型数据库系统	25
2.4.3 网络模型与网络模型数据库系统	26
2.4.4 关系模型与关系模型数据库系统	28
2.5 计算机世界与物理模型	30
2.5.1 概述	30
2.5.2 文件系统	31
2.5.3 关系模型的存储结构	32
2.6 描述数据库系统结构的重要概念——数据模式	32
习题 2	34
第 3 章 关系数据库系统及其标准语言 SQL	36
3.1 关系数据库系统特点及其标准	36
3.1.1 关系数据库系统概述	36
3.1.2 对关系数据库系统的衡量和评价准则	38
3.2 关系模型的数学理论基础	39
3.2.1 关系的表示	39
3.2.2 数据操纵的表示	40
3.2.3 关系模型与关系代数	42

3.2.4	关系代数中的扩充运算	42
3.2.5	关系代数运算实例	44
3.3	SQL 及其特点	46
3.3.1	SQL 概述	46
3.3.2	SQL 的数据查询	47
3.3.3	SQL 的数据更新	58
3.3.4	SQL 的数据定义	60
3.3.5	定义视图和操作视图	64
3.3.6	嵌入式 SQL	66
3.4	SQL 总结与应用举例	71
3.4.1	SQL 的数据定义小结	71
3.4.2	SQL 的数据操纵举例	72
3.4.3	SQL 的更新功能举例	78
3.4.4	有关视图的举例	79
3.4.5	嵌入式 SQL 小结	81
3.4.6	SQL 的特点小结	82
习题 3		83
第 4 章	关系数据库的优化理论与系统实现技术	87
4.1	关系数据库的规范化理论	87
4.1.1	问题的引出	87
4.1.2	规范化理论	88
4.1.3	规范化所引起的相关问题	98
4.2	数据库的并发控制	98
4.2.1	数据的不一致性与并发控制	98
4.2.2	封锁	99
4.2.3	事务	100
4.3	数据库的完整性	101
4.3.1	完整性约束	101
4.3.2	触发器	102
4.4	数据库的安全性	102
4.4.1	使用权的鉴别	103
4.4.2	使用范围限制	103
4.4.2	存取控制权鉴别	103
4.5	数据库的恢复技术	105
4.5.1	数据转储	106
4.5.2	日志	106
4.6	数据库的物理存储问题	106
4.6.1	数据存储介质	107
4.6.2	文件组织	108
4.6.3	文件中记录的组织	110
4.6.4	索引技术	111
4.6.5	散列技术	115
习题 4		117

第 5 章 数据库设计	120
5.1 数据库设计方法与软件工程学理念	120
5.1.1 数据库工程	120
5.1.2 数据库系统生存期	121
5.1.3 数据库设计的具体步骤	121
5.2 规划阶段	122
5.3 需求分析阶段	122
5.4 概念设计阶段	124
5.4.1 概念设计的重要性	124
5.4.2 数据抽象	124
5.4.3 基本 ER 模型的扩充	124
5.4.4 采用 E-R 方法的概念设计步骤	125
5.5 逻辑设计阶段	125
5.6 物理设计阶段	127
5.7 编码测试与数据库的实现	127
5.7.1 编码与编译	127
5.7.2 测试与调试	129
5.7.3 数据库的实现	131
5.8 数据库的运行和维护	131
习题 5	132
第 6 章 数据库系统的成熟新技术	134
6.1 分布式数据库系统	134
6.1.1 分布式数据库概述	134
6.1.2 分布式数据库系统的分类和结构	137
6.1.3 数据分布策略	141
6.1.4 查询的分解和优化	143
6.1.5 分布式数据库系统中的并发控制和恢复技术	149
6.2 面向对象数据库技术	153
6.2.1 新型数据库应用	153
6.2.2 面向对象数据模型	154
6.2.3 面向对象语言	159
6.2.4 持久化程序设计语言	159
6.2.5 持久化 C++ 系统	162
6.3 对象关系数据库技术	164
6.3.1 嵌套关系	164
6.3.2 复杂类型和面向对象	165
6.3.3 复杂类型查询	169
6.3.4 面向对象和对象关系数据库的比较	173
习题 6	174
第 7 章 常见关系数据库及应用开发系统简介	176
7.1 流行数据库应用开发系统简介	176
7.1.1 常见的桌面数据库产品	176
7.1.2 数据库应用程序开发工具	177

7.1.3	数据库应用程序的开发步骤.....	178
7.2	主流关系数据库简介.....	179
7.2.1	ORACLE.....	180
7.2.2	SYBASE.....	185
7.2.3	INFORMIX.....	190
7.2.4	DB2.....	194
7.2.5	INGRES.....	199
7.3	关系数据库管理系统产品特性.....	202
习题 7		204
第 8 章	Oracle 数据库简介.....	205
8.1	Oracle 数据库的结构.....	205
8.1.1	数据库结构概述.....	205
8.1.2	Oracle 数据库的物理结构.....	205
8.1.3	数据库的逻辑结构.....	206
8.1.4	数据库的存储对象.....	206
8.1.5	数据字典.....	208
8.1.6	ORACLE 多线索服务器结构.....	208
8.2	应用 Oracle 数据库应该了解的预备知识.....	209
8.2.1	ORACLE 数据库的技术特点.....	209
8.2.2	分布式数据库管理功能简介.....	210
8.2.3	ORACLE 的互联网功能.....	211
8.3	ORACLE 的安装.....	211
8.4	验证安装的结果.....	220
8.5	ORACLE 的启用.....	220
8.6	ORACLE 的卸载.....	221
习题 8		222
第 9 章	Oracle 数据库的 SQL.....	223
9.1	ORACLE 中 SQL 的分类与 SQL*PLUS 的应用.....	223
9.1.1	ORACLE 中 SQL 的分类.....	223
9.1.2	SQL*PLUS 的构成.....	223
9.1.3	程序块结构语言 PL/SQL.....	224
9.1.4	用 PL/SQL 说明变量.....	225
9.1.5	一些常见的 PL/SQL 程序控制结构.....	227
9.2	ORACLE 中 SQL 的使用.....	227
9.2.1	SQL DML 语句的类型.....	227
9.2.2	创建若干 Oracle 表.....	227
9.2.3	使用 INSERT 语句.....	229
9.2.4	SELECT 语句.....	230
9.2.5	使用 UPDATE 语句和 DELETE 语句.....	231
9.2.6	PL/SQL 对异常的处理.....	232
9.2.7	使用 LOCK TABLE 语句.....	232
9.2.8	事务控制语句.....	233
9.2.9	记录的使用.....	233

9.2.10 伪列	238
习题 9	239
第 10 章 Oracle 的 DBA	243
10.1 Oracle 的 DBA 的分类	243
10.2 可供 DBA 使用的工具	243
10.2.1 DBA Studio	243
10.2.2 SQL*Plus Worksheet	244
10.3 例程管理	245
10.4 方案管理	246
10.5 安全性管理	249
10.6 存储管理	250
10.7 利用 ORACLE 提供的数据库管理向导	250
10.7.1 分析向导	251
10.7.2 导出向导	251
10.7.3 导入向导	252
10.7.4 装载向导	252
10.7.5 备份向导	253
10.7.6 其他向导	253
10.8 数据库人工管理简介	255
10.8.1 数据库的存取管理	255
10.8.2 数据库程序结构管理	259
10.8.3 数据库维护 and 安全性管理	259
习题 10	259
第 11 章 数据库系统的基本概念	261
11.1 Oracle 的数据库的安全性	261
11.1.1 数据库的存取控制	262
11.1.2 特权和角色	264
11.1.3 审计	265
11.2 Oracle 的数据完整性	266
11.2.1 完整性约束	266
11.2.2 数据库触发器	267
11.3 Oracle 的并发控制技术	268
11.3.1 数据库不一致的类型	269
11.3.2 封锁	269
11.3.3 Oracle 多种一致性模型	269
11.3.4 封锁机制	270
11.3.5 手工的数据封锁	271
11.4 Oracle 的数据库备份和恢复	272
11.4.1 数据库恢复所使用的结构	272
11.4.2 在线日志	273
11.4.3 归档日志	274
11.4.4 数据库备份	275
11.4.5 数据库恢复	276

习题 11	277
第 12 章 数据库技术的热点与前沿问题.....	278
12.1 数据库系统的崛起及其快速的发展道路	278
12.2 数据库技术的发展现状.....	279
12.3 新型数据库的应用.....	280
12.4 对数据库技术的展望.....	284
习题 12	286
参考文献	287

引 言

1 我们所面对的世界

显然,我们所面对的是一个纷繁复杂的世界。在人类诞生以来,人们就以不同的方式与手段在改造着这个世界,以使我们的周围——即我们每个人所接触的世界的一部分更加适合于我们人类的意愿,这种趋势在刚刚过去的几百年,尤其是在约近 100 年中表现尤甚。

而人类改造的对象,总的来说表现在对于资源的占有与利用上,从刀耕火种时代的对于土地、领地等以空间为主的资源的占有,到目前计算机的时代的对于速度、性能等以时间为主的强烈需求,已经越来越清晰地表现出对于科学与技术的越来越强烈的依赖意识和不争的事实。姑且不论这种改变的社会意义如何,仅就技术层面而言,可以说是人们改变的欲求一直推动着技术的出现、发展和进步。

人们千百年所追求的资源,到了今天,总的说来,无非主要表现在了如下的几个方面:能源、材料、信息和人才与社会因素等。当前人类社会正处于信息社会,人类知识以惊人的速度增长。如何组织和利用这样庞大的知识成为急需解决的问题。由此,信息科学与信息管理技术的重要性可见一斑。

作为涉足信息科学与技术的一员,我们应当充分了解数据信息对于现代社会的极其重要的作用。

数据信息的社会属性,主要表现在:

- 数据信息是有用的,有时是有极大作用的。
- 数据信息是具有时效性的,有时是表现在时、分、秒,甚至是更快的数量级别上的。
- 从人类社会的发展来看,人们积累的信息是不断增长的。

由此可见,信息已经成为社会上各行各业不可缺少的重要资源之一。另外,信息还有其固有的一些自然属性,诸如:

- 信息来源于物质和能量;
- 信息是可以感知的;
- 信息是可以存储的;
- 信息是可以加工、传递和再生的等等。

以上两点,对于现代化程度越高的社会,对于数据与信息的依赖程度和信息时效性越强,数据与信息的增长速率越大。

因此,我们这门课程就是围绕以上列举的人们各种需求,研究如何管理日益增长的数据信息,并在此基础上能够尽量快速与高效地为使用者提供有用的信息的一门学问。

2 信息与数据的关系

从本质上讲,数据是对客观对象特征的抽象的、符号化的表示,即用一定的符号表示那些从诸如观察、测量等活动中收集到的基本事实。当然,采用什么符号完全是一种人为的规定,例如从原始人类的结绳计日,到古代的各种计数制的运用,直到现在计算机内部用 0 和 1 的高速运算,数据无时无刻不在伴随着人类向前迈出的每一步。而在数据的基础

上，人类总结和归纳出了许许多多的经验和规律，如历法、天文甚至整个人类的文化，对于这些在数据的基础上产生出的具有思想性的推演结果，我们称之为信息。可见，信息实际上是指经过加工、处理后的数据，是经过“消化”了的数据。

因此，数据与信息二者的关系密不可分，它们既有联系，又有区别。数据既为信息提供了源泉，又反过来表示了信息，而信息只有通过数据形式表示出来才能够被人们理解和接受。尽管数据和信息在概念上不尽相同，但通常人们并不严格地区分它们。例如，一般人们并不认为数据处理和信息处理有什么区别。

3 如何管理数据与信息

数据的基本含义是从某些已知的数据出发，推导出一些新的数据，这些新的数据又表示了新的信息。具体操作中，涉及到数据收集、管理、加工乃至信息输出的演变与推导的全过程。

在数据处理过程中，通常计算比较简单，一般不涉及复杂的数值算法和超大的数值计算量，但表现的特点是：

- 参与算法的数据量十分巨大，但一般在数据处理过程中其数值变化不大，甚至决大多数的数据的值是不变的。
- 处理过程涉及的数据结构比较复杂，同时大量的数据之间又具有复杂的逻辑联系。

因此，数据处理本身的矛盾焦点不是数值计算，而是数据管理——即数据的收集、整理、组织、存储、维护、检索、传送等操作，这些操作是数据处理的基本环节，而且也是任何数据处理系统中必须具备的共性部分。

现在，让我们举一个实际的例子。例如管理学生的一些数据和信息。

每位学生具有如下信息：学号（S#）、姓名（SNAME）、性别（SEX）、年龄（AGE）。

设想一下：如果只是管理几十位或者上百位甚至千位数量级的学生的信息，那么采用传统的以记事本为基础的人工管理应该是能够胜任的。如果是百位级别的，我们写成如下的形式就可以了。

（2002123001，赵一，男，18），（2002123002，钱二，男，19）

但如果是千位级别的呢？那么如下的表格形式可能更适合问题的要求：

学号（S#）	姓名（SNAME）	性别（SEX）	年龄（AGE）
2002123001	赵一	男	18
2002123002	钱二	男	19
2002123003	孙三	男	18
.....			

可是如果我们要管理的对象的数量级很大，如达到数万甚至更大，那会出现什么问题呢？则可供我们选择的对策有：

- 用多个人来管理，其中每个人分担一定数量级（如千级别）。
- 采用新的管理技术与管理设备，减少人员开支，提高信息管理效率（如采用计算机等技术设备）。

两者相较，显然后者为首选。

那么就让我们看一看采用计算机进行信息处理的技术背景吧。

4 信息管理的历史背景与技术背景

从 20 世纪 50 年代中期开始, 计算机的应用由科学研究部门扩展到企业、行政部门, 数据处理迅速上升为计算机应用的主要方面。在计算机的三大应用(科学计算、数据处理与过程控制)中, 数据处理所占比重随着社会的发展和时间的推移变得越来越大, 现在已达到近 85% 左右的使用水平。

那么人类在利用计算机来管理信息时走过哪些道路呢?

- 最早采用计算机进行数据管理的我们称为自由管理阶段

最早的计算机, 没有必要的软件支持, 因此用户使用计算机进行数据处理不得不自行管理数据, 程序员在程序中既要考虑数据处理方法, 又要管理数据的组织与存储。这一阶段, 存取数据是根据设备的物理地址进行的, 这种方式迫使程序员直接与物理设备打交道。管理数据的存放地址, 导致程序与数据紧密相关, 程序高度依赖于数据。因此数据稍有变动, 整个程序就必须进行全面修改。这一阶段的特点是, 编程效率低, 程序不灵活而且容易出错, 但当时计算机性能较低且为数寥寥, 主要用来进行数值计算, 极少用于数据处理。

- 接着当推文件系统阶段

文件系统是计算机系统的既有组成部分, 这个阶段出现于上个世纪 50 年代至 60 年代中, 它提供了简单的数据共享与数据管理功能, 但它无法提供完整统一管理 with 数据共享能力, 由于它的功能较为简单, 一般它均附属于操作系统而并不成为独立的软件。因此, 这一阶段的数据处理能力取决于计算机系统, 尤其是操作系统的文件处理能力。

- 其次是初级数据库系统阶段

自 60 年代起各种数据库系统相继问世, 首推层次模型与网络模型的数据库系统, 它们为统一管理与共享数据提供了有力支撑。在这一时期, 由于数据库系统的蓬勃发展, 从而形成了历史上著名的“数据库时代”。当然, 这两种类型的数据库系统也有其不足地方, 最主要是它们均脱胎于文件系统比较简单的数据结构, 因此, 既受文件中的物理结构影响较大, 用户在使用数据库时需要数据的物理结构有详细的了解, 这对使用数据库带来了诸多的麻烦, 同时, 数据库中表示数据模式的结构方式又过于烦琐, 直接影响了数据库应用中越来越多的对数据结构的复杂要求的实现。

- 再次是关系数据库系统阶段

关系型数据库系统的理论出现于 70 年代初, 其系统形成于 70 年代中并在 80 年代得到了充分的发展, 它具有简单的结构方式与较少的物理表示, 使用与操作又更加符合人们日常的处理方式并且极为方便, 因此, 在 80 年代它逐步取代层次与网络型数据库成为占主导地位的数据库, 到目前为止, 关系型数据库系统仍占领数据库应用的主要地位。

- 目前即将进入高级数据库系统阶段

由于计算机应用的发展, 计算机已从传统的科学计算、事务处理等领域逐步扩展到工程设计统计、人工智能、多媒体、分布式等领域, 这些新的领域需要有新的数据库支撑, 而传统关系数据库系统是以商业应用、事务处理为背景而发展起来的, 它并不完全适用于新领域的应用, 因此, 需要有新的数据库系统以满足不同领域的要求, 自 80 年代中开始, 各种适应不同领域的新型数据库系统不断涌现, 如工程数据库、多媒体数据库、CAD 数据库、图形数据库、图像数据库、智能数据库、分布式数据库以及面向对象数据库等, 其中

特别是面向对象数据库系统由于其通用性强,适应面广而受到青睐,在90年代后期形成了多种数据库系统共同支撑应用的局面。但关系数据库系统仍将是其中一支强有力的主力军,而面向对象数据库将会起主导作用。

在我国,数据库的发展也经历了上述的类似的几个阶段,但是略有不同。由于我国计算机应用发展较技术先进的国家略迟一步,因此,文件系统阶段较为短暂,而早期、初级数据库系统阶段基本没有经历,自80年代中起即直接进入关系数据库系统阶段。特别是以dBase-III与Foxbase为代表的初级关系型数据库系统阶段和90年代以来以ORACLE、Sybase等为代表的成熟的大型关系数据库系统阶段,基本弥补了我国在高级数据库系统的应用上的不足。

5 现在的技术前提和技术手段

当前人类社会正处于信息社会,人类知识与信息以惊人的速度增长,引发了如何组织和利用这样庞大的知识和信息,成为了急需解决的问题。在比较了不同领域和种类的数据处理系统之后,人们发现了数据处理系统的一些共性,诸如数据共享、减少冗余、程序(应用)独立于数据、数据完整、数据安全等等。基于这些共性,人们陆续制作出一些在大多数领域“普适”的软件部件,并随着这一部件的性能优化和功能明确,将其命名为数据库管理系统(DBMS),同时围绕其功能和性能完善了许多外围工作,形成一个个的功能完善且独立运作的数据库系统,终于在20世纪60年代的“软件危机”中,数据库技术作为一门软件学科应运而生了。

据查,数据库(DataBase,记作DB)这个名词起源于20世纪50年代,当时美国为了战争的需要,把各种情报集中在一起,存储在计算机里,称为Information Base或DataBase。

- 1963年,美国Honeywell公司的IDS(Integrated Data Store)系统投入运行,揭开了数据库技术的序幕。
- 1965年,美国一家火箭公司利用数据存储系统帮助设计了阿波罗登月火箭,推动了数据库技术的产生。当时美国社会上产生了许多形形色色的DataBase或Databank,但是,它们基本上都是文件系统的扩充或是倒排文件系统。
- 1968年,美国IBM公司推出层次模型的IMS数据库系统(1969年形成产品);1969年,美国CODASYL(Conference On Data System Language,数据系统语言协会)组织的数据库任务组(DBTG)发表关于网状模型的DBTG报告(1971年正式通过)。
- 1970年,IBM公司的E.F.Codd发表论文提出关系模型。

以上几个具有历史意义的事件奠定了现代数据库技术发展的基础。

20世纪70年代是数据库蓬勃发展的年代。网状系统和层次系统占据了整个商用市场,而关系系统仅处于实验阶段。许多计算机厂商研制了各种数据库管理系统,许多学者发表了大量的探讨性论文,使数据库技术在实践中和理论上得到飞速发展。数据库技术也日趋成熟。

到80年代,关系系统由于使用简便,逐步代替网状系统和层次系统占领了市场。同时关系数据库的理论日趋完善,走向更高级的阶段。数据库的理论研究以及新的数据库系统的研究仍在继续,不断深化。