



普通高等教育“十一五”国家级规划教材

21世纪高等学校计算机规划教材

21st Century University Planned Textbook of Computer Science

# 数据库原理及 应用教程 (第二版)

## Database Principles and Applications

陈志泊 主编 陈志泊 王春玲 编著

- 强调关系数据库理论基础
- 突出SQL Server技术实践
- 一个应用系统贯穿全书



精品系列

 人民邮电出版社  
POSTS & TELECOM PRESS



普通高等教育“十一五”国家级规划教材

21世纪高等学校计算机规划教材

21st Century University Planned Textbook of Computer Science

TP311.13/176=2

2008

# 数据库原理及 应用教程 (第二版)

Database Principles and Applications

陈志泊 主编 陈志泊 王春玲 编著



精品系列

人民邮电出版社

北京

## 图书在版编目 (CIP) 数据

数据库原理及应用教程 / 陈志泊主编; 王春玲编著. —2版. —北京: 人民邮电出版社, 2008.3

21 世纪高等学校计算机规划教材. 普通高等教育“十一·五”国家级规划教材

ISBN 978-7-115-16416-2

I. 数… II. ①陈…②王… 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 144684 号

## 内 容 提 要

全书共七章, 全面系统地讲述了数据库技术的基本原理和应用。主要内容包括: 数据库系统概述、关系数据库、关系数据库标准语言 (SQL)、关系数据库理论、数据库安全保护、数据库设计和 SQL Server 2000 高级应用。本书还以 SQL Server 2000 为背景介绍了数据库技术的实现, 包括数据库和数据表的维护、查询与统计、视图管理、存储过程和触发器的管理、用户管理、约束和默认管理、数据库的备份和还原、Transact-SQL 程序设计等内容, 使读者可以充分利用 SQL Server 2000 平台深刻理解数据库技术的原理, 达到理论和实践紧密结合的目的。

本书内容循序渐进, 深入浅出, 概念清晰, 条理性强, 每一章节都给出了大量的实例, 并进行解释说明。同时, 每章后面附有大量的各种类型的习题, 从各种不同的侧面帮助读者了解和掌握所学知识点, 书后所附的习题答案可以帮助读者检验学习和练习效果。

本书可作为大学计算机及相关专业的本、专科教材, 也可供从事计算机软件工作的科技人员、工程技术人员以及其他相关人员参阅。

普通高等教育“十一·五”国家级规划教材

21 世纪高等学校计算机规划教材

数据库原理及应用教程 (第二版)

- 
- ◆ 主 编 陈志泊
  - 编 著 陈志泊 王春玲
  - 责任编辑 滑 玉
  - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号  
邮编 100061 电子函件 315@ptpress.com.cn  
网址 <http://www.ptpress.com.cn>  
三河市海波印务有限公司印刷  
新华书店总店北京发行所经销
  - ◆ 开本: 787×1092 1/16  
印张: 20  
字数: 477 千字 2008 年 3 月第 2 版  
印数: 68 001 - 71 000 册 2008 年 3 月河北第 1 次印刷

---

ISBN 978-7-115-16416-2/TP

定价: 29.00 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223

反盗版热线: (010)67171154

## 出版者的话

计算机应用能力已经成为社会各行业最重要的工作要求之一，而计算机教材质量的好坏会直接影响人才素质的培养。目前，计算机教材出版市场百花争艳，品种急剧增多，要从林林总总的教材中挑选一本适合课程设置要求、满足教学实际需要的教材，难度越来越大。

人民邮电出版社作为一家以计算机、通信、电子信息类图书与教材出版为主的科技教育类出版社，在计算机教材领域已经出版了多套计算机系列教材。在各套系列教材中涌现出了一批被广大一线授课教师选用、深受广大师生好评的优秀教材。老师们希望我社能有更多的优秀教材集中地呈现在老师和读者面前，为此我社组织了这套“21世纪高等学校计算机规划教材-精品系列”。

“21世纪高等学校计算机规划教材-精品系列”具有下列特点。

(1) 前期调研充分，适合实际教学需要。本套教材主要面向普通本科院校的学生编写，在内容深度、系统结构、案例选择、编写方法等方面进行了深入细致的调研，目的是在教材编写之前充分了解实际教学的需要。

(2) 编写目标明确，读者对象针对性强。每一本教材在编写之前都明确了该教材的读者对象和适用范围，即明确面向的读者是计算机专业、非计算机理工类专业还是文科专业的学生，尽量符合目前普通高等教学计算机课程的教学计划、教学大纲以及发展趋势。

(3) 精选作者，保证质量。本套教材的作者，既有来自院校的一线授课老师，也有来自IT企业、科研机构等单位的资深技术人员。通过他们的合作使老师丰富的实际教学经验与技术人员丰富的实践工程经验相融合，为广大师生编写出适合目前教学实际需求、满足学校新时期人才培养模式的高质量教材。

(4) 一纲多本，适应面宽。在本套教材中，我们根据目前教学的实际情况，做到“一纲多本”，即根据院校已学课程和后续课程的不同开设情况，为同一科目提供不同类型的教材。

(5) 突出能力培养，适应人才市场要求。本套教材贴近市场对于计算机人才的能力要求，注重理论与实际应用的结合，注重实际操作和实践动手能力的培养，为学生快速适应企业实际需求做好准备。

(6) 配套服务完善，共促提高。对于每一本教材，我们在教材出版的同时，都将提供完备的PPT课件，并根据需要提供书中的源程序代码、习题答案、教学大纲等内容，部分教材还将在作者的配合下，提供疑难解答、教学交流等服务。

在本套教材的策划组织过程中，我们获得了来自清华大学、北京大学、人民大学、浙江大学、吉林大学、武汉大学、哈尔滨工业大学、东南大学、四川大学、上海交通大学、西安交通大学、电子科技大学、西安电子科技大学、北京邮电大学、中国林业大学等院校老师的大力支持和帮助，同时获得了来自信息产业部电信研究院、联想、华为、中兴、同方、爱立信、摩托罗拉等企业和科研单位的领导和技术人员的积极配合。在此，人民邮电出版社向他们表示衷心的感谢。

我们相信，“21世纪高等学校计算机规划教材-精品系列”一定能够为我国高等院校计算机课程教学做出应有的贡献。同时，对于工作欠缺和不妥之处，欢迎老师和读者提出宝贵的意见和建议。

人民邮电出版社

## 编者的话

本书自第一版出版以来，受到了广大师生的普遍欢迎，并于 2005 年被评为北京市高等教育精品教材，2006 年被评为普通高等教育“十一五”国家级规划教材。但是，随着数据库技术的飞速发展，新技术、新知识层出不穷，一本教材的内容不可能一成不变，也不可能包罗万象。同时，在第一版的使用过程中，很多读者也给我们提出了很多有益的建议和意见，在这种情况下，作者深入调查了目前许多高校讲授数据库课程的详细情况，同时参阅和借鉴了国内外许多优秀教材的内容和优点，并进一步吸取了第一版教材的优点和教学过程中的体会和经验。将第一版中的部分内容进行了合理的取舍，进行了大量的修改、补充和完善。

与第一版相比，本书体现了以下几方面的特点。

(1) 更加符合初学者学习数据库课程的认识规律，进一步体现了概念讲解的条理性、逻辑性，使内容讲解循序渐进，深入浅出，易于读者学习和掌握。

(2) 进一步完善和扩充了关系数据库系统的模式结构、数据模型的有关概念和理论，使概念更加清晰，条理性更强。

(3) 进一步完善和扩充了关系数据库的基本理论知识，主要包括函数依赖、函数依赖的逻辑蕴涵、Armstrong 公理、关键字的求解理论与算法、关系模式分解及其有关算法等。

(4) 进一步丰富了例题，并将所学内容和相关知识点组织到相应例题中，使读者能很容易地加深对知识点的理解和掌握。

(5) 进一步丰富了习题的类型和数量，除了简答题之外，还增加了选择题、填空题等，从各种不同的侧面进一步帮助读者理解和掌握所学知识点；另外，书后还附有习题答案，帮助读者检验学习和练习的效果。

(6) 进一步完善了有关 SQL Server 2000 的高级应用知识，使读者很容易学会利用 SQL Server 2000 环境进行数据库的管理工作，真正做到学以致用。

为方便读者学习和教师讲课，本书还提供了 PowerPoint 电子教案，读者可到人民邮电出版社的网站下载。

本书内容全面，深入浅出，概念清晰，条理清楚，不仅适合于教学，也适合读者自学。如果作为教材，建议总学时为 60 学时，其中主讲学时 40 学时。由于课程学时的限制，实验学时各学校可适当调整，一般为 20 学时左右；另外，除实验学时外，最好安排学生自由上机的时间，加强学生的实际动手能力。

本书由陈志泊担任主编，陈志泊、王春玲编写。第 1, 2, 4 章由陈志泊编写；第 3, 5, 6, 7 章主要由王春玲编写，全书内容由陈志泊修改定稿。另外，黄金燕、张惠萍、王鹏等参加了文字编辑、Powerpoint 讲稿制作工作。

由于时间仓促，加之作者水平有限，书中不当之处在所难免，敬请广大读者批评指正。

编者

# 目 录

第 1 章 数据库系统概述 .....	1
1.1 信息、数据、数据处理与数据管理 .....	1
1.1.1 数据与信息 .....	1
1.1.2 数据处理与数据管理 .....	2
1.2 数据库技术的产生、发展 .....	3
1.3 数据库系统的组成 .....	8
1.4 数据库系统的模式结构 .....	10
1.4.1 数据库系统的三级模式结构 .....	10
1.4.2 数据库系统的二级映像与数据独立性 .....	12
1.4.3 数据库系统的三级模式与二级映像的优点 .....	13
1.5 数据库系统的外部体系结构 .....	13
1.6 数据库管理系统 .....	15
1.6.1 DBMS 的主要功能 .....	15
1.6.2 DBMS 的组成 .....	16
1.6.3 DBMS 的数据存取的过程 .....	17
1.7 数据模型 .....	18
1.7.1 数据模型的概念 .....	18
1.7.2 三个世界的划分及其有关概念 .....	19
1.7.3 数据模型的分类 .....	21
1.7.4 实体联系模型及 E-R 图 .....	22
1.7.5 数据模型的组成要素 .....	23
1.8 四种数据模型 .....	24
1.8.1 层次模型 .....	25
1.8.2 网状模型 .....	27
1.8.3 关系模型 .....	28
1.8.4 面向对象模型 .....	30
1.9 数据库系统的发展 .....	31
1.9.1 第一代数据库系统 .....	31
1.9.2 第二代数据库系统 .....	32
1.9.3 第三代数据库系统 .....	33
1.10 数据库技术与其他相关技术的结合 .....	34
1.10.1 分布式数据库 .....	35
1.10.2 主动数据库 .....	37

1.10.3	多媒体数据库	39
1.10.4	数据库技术的研究领域	39
1.11	小结	40
	习题	41
<b>第2章</b>	<b>关系数据库</b>	<b>44</b>
2.1	关系模型的数据结构及其形式化定义	44
2.1.1	关系的形式化定义及其有关概念	44
2.1.2	关系的性质	47
2.1.3	关系模式	48
2.1.4	关系数据库与关系数据库模式	49
2.2	关系的键与关系的完整性	49
2.2.1	候选键与主关系键	49
2.2.2	外部关系键	50
2.2.3	关系的完整性	51
2.3	关系代数	52
2.3.1	关系代数的分类及其运算符	53
2.3.2	传统的集合运算	53
2.3.3	专门的关系运算	55
2.4	关系演算*	59
2.4.1	元组关系演算语言	60
2.4.2	域关系演算语言 QBE	65
2.5	小结	68
	习题	69
<b>第3章</b>	<b>关系数据库标准语言——SQL</b>	<b>72</b>
3.1	SQL 的基本概念与特点	72
3.1.1	SQL 的发展及标准化	72
3.1.2	SQL 的基本概念	73
3.1.3	SQL 的主要特点	74
3.2	了解 SQL Server 2000	74
3.2.1	SQL Server 2000 的主要组件	74
3.2.2	企业管理器	76
3.2.3	查询分析器	77
3.3	创建与使用数据库	78
3.3.1	SQL Server 的系统数据库	79
3.3.2	SQL Server 的实例数据库	80
3.3.3	创建用户数据库	80

3.3.4	修改用户数据库	83
3.3.5	删除用户数据库	86
3.3.6	查看数据库信息	86
3.4	创建与使用数据表	87
3.4.1	数据类型	88
3.4.2	创建数据表	89
3.4.3	定义数据表的约束	91
3.4.4	修改数据表	95
3.4.5	删除基本表	97
3.4.6	查看数据表	97
3.5	创建与使用索引	98
3.5.1	索引的作用	98
3.5.2	索引的分类	98
3.5.3	创建索引	99
3.5.4	查看与修改索引	103
3.5.5	删除索引	105
3.6	数据查询	105
3.6.1	SELECT 命令的格式与基本使用	105
3.6.2	条件查询	107
3.6.3	常用库函数及统计汇总查询	110
3.6.4	分组查询	111
3.6.5	查询的排序	112
3.6.6	数据表连接及连接查询	113
3.6.7	子查询	116
3.6.8	合并查询	120
3.6.9	存储查询结果到表中	120
3.7	数据操纵	121
3.7.1	添加数据	121
3.7.2	修改数据	122
3.7.3	删除数据	123
3.8	视图	124
3.8.1	创建视图	124
3.8.2	修改视图	126
3.8.3	删除视图	127
3.8.4	查询视图	128
3.8.5	更新视图	128
3.9	数据控制	129
3.9.1	权限与角色	129
3.9.2	系统权限与角色的授予与收回	130



3.9.3 对象权限与角色的授予与收回	130
3.10 小结	132
习题	132
<b>第4章 关系数据库理论</b>	<b>137</b>
4.1 规范化问题的提出	137
4.1.1 规范化理论的主要内容	137
4.1.2 不合理的关系模式存在的存储异常问题	137
4.2 函数依赖	140
4.2.1 函数依赖的定义	140
4.2.2 函数依赖的逻辑蕴涵定义	141
4.2.3 函数依赖的推理规则	142
4.2.4 完全函数依赖与部分函数依赖	144
4.2.5 传递函数依赖	144
4.2.6 属性集的闭包及其算法	144
4.2.7 候选键的求解理论和算法	145
4.2.8 函数依赖推理规则的完备性	147
4.2.9 函数依赖集的等价、覆盖和最小函数依赖集	148
4.3 关系模式的分解*	150
4.3.1 模式分解问题	150
4.3.2 无损连接分解	151
4.3.3 无损分解的测试算法	152
4.3.4 保持函数依赖的分解	154
4.4 关系模式的范式	155
4.4.1 第一范式	155
4.4.2 第二范式	156
4.4.3 第三范式	159
4.4.4 BC 范式	162
4.4.5 多值依赖与第四范式	165
4.5 关系模式的规范化	169
4.5.1 关系模式规范化的目的和原则	169
4.5.2 关系模式规范化的步骤	170
4.5.3 关系模式规范化的要求	170
4.6 小结	172
习题	173
<b>第5章 数据库安全保护</b>	<b>177</b>
5.1 数据库的安全性	177

5.1.1	数据库安全性的含义	177
5.1.2	安全性控制的一般方法	177
5.1.3	SQL Server 2000 的数据安全性机制	182
5.1.4	SQL Server 2000 的身份验证模式	182
5.1.5	SQL Server 的登录账号和服务器角色	184
5.1.6	SQL Server 的数据库用户账号和数据库角色	188
5.2	完整性控制	194
5.2.1	数据库完整性的含义	194
5.2.2	完整性规则的组成	194
5.2.3	完整性约束条件的分类	195
5.2.4	数据完整性的实施	196
5.2.5	规则	197
5.2.6	默认	200
5.3	并发控制与封锁	203
5.3.1	数据库并发性的含义	203
5.3.2	事务 (Transaction)	203
5.3.3	并发操作与数据的不一致性	205
5.3.4	封锁	206
5.4	数据库的恢复	212
5.4.1	数据库恢复的含义	212
5.4.2	数据库恢复的原理及其实现技术	212
5.4.3	数据库的故障和恢复的策略	213
5.5	小结	216
	习题	216
<b>第 6 章</b>	<b>数据库设计</b>	<b>219</b>
6.1	数据库设计概述	219
6.1.1	数据库设计的任务、内容和特点	219
6.1.2	数据库设计方法简述	220
6.1.3	数据库设计的步骤	222
6.2	系统需求分析	225
6.2.1	需求分析的任务	225
6.2.2	需求分析的方法	226
6.3	概念结构设计	228
6.3.1	概念结构设计的必要性	228
6.3.2	概念模型的特点	229
6.3.3	概念结构设计的方法与步骤	230
6.4	逻辑结构设计	238

6.4.1	逻辑结构设计的任务和步骤	238
6.4.2	初始关系模式设计	238
6.4.3	关系模式规范化	239
6.4.4	模式评价与改进	240
6.5	物理结构设计	242
6.5.1	确定物理结构	242
6.5.2	评价物理结构	244
6.6	数据库实施	244
6.6.1	建立实际数据库结构	244
6.6.2	装入数据	244
6.6.3	应用程序编码与调试	245
6.6.4	数据库试运行	245
6.6.5	整理文档	245
6.7	数据库运行和维护	246
6.7.1	维护数据库的安全性与完整性	246
6.7.2	监测并改善数据库性能	246
6.7.3	重新组织和构造数据库	246
6.8	小结	247
	习题	247
<b>第7章</b>	<b>SQL Server 2000 高级应用</b>	<b>251</b>
7.1	Transact-SQL 程序设计	251
7.1.1	Transact-SQL 程序的结构与批处理	251
7.1.2	变量	252
7.1.3	流程控制命令	253
7.1.4	常用命令	258
7.1.5	常用函数	261
7.2	存储过程	269
7.2.1	存储过程的概念、优点及分类	269
7.2.2	创建存储过程	271
7.2.3	查看存储过程	274
7.2.4	重新命名存储过程	274
7.2.5	删除存储过程	274
7.2.6	执行存储过程	275
7.2.7	修改存储过程	275
7.3	触发器	276
7.3.1	触发器的概念、分类与作用	276
7.3.2	触发器的工作原理	277

---

7.3.3	创建触发器 .....	278
7.3.4	查看触发器 .....	281
7.3.5	修改触发器 .....	282
7.3.6	删除触发器 .....	283
7.4	备份和还原 .....	283
7.4.1	备份和还原概述 .....	283
7.4.2	创建备份设备 .....	287
7.4.3	备份数据库 .....	289
7.4.4	还原数据库 .....	290
7.4.5	备份和还原系统数据库 .....	291
7.5	小结 .....	292
	习题 .....	292
	SQL Server 2000 综合练习 .....	294
	<b>习题参考答案</b> .....	<b>299</b>
	<b>参考文献</b> .....	<b>304</b>

# 第 1 章 数据库系统概述

数据库技术是数据管理的最新技术，是计算机科学技术中发展最快的领域之一，也是应用最广的技术之一。它已成为各行各业存储数据、管理信息、共享资源的最先进最常用的技术。因此，数据库课程不仅是计算机科学与技术专业、信息管理与信息系统专业的必修课程，也是许多非计算机专业的选修课程。

数据库 (database)，简单地说就是数据的仓库，即数据存放的地方。我们周围有许多数据库的例子，如通讯录是一个小数据库，图书馆则是一个典型的大型数据库。小数据库尚可用手工管理，而大型数据库必须由计算机进行管理。在计算机三大主要应用领域（科学计算、过程控制和数据处理）中，数据处理所占比例约为 70%。20 世纪 60 年代末，数据库技术作为数据处理的最新技术应运而生。

本章首先回顾数据管理技术的三个发展阶段，然后介绍数据库中有关的主要概念。学习本章后，读者应了解数据库的三个发展阶段及各阶段的主要特点，掌握数据库中的有关的基本概念、数据库系统的组成及各部分的主要功能，重点掌握实体、属性定义和实体型之间的联系类型及特点，了解三种数据模型的特点及区别，为后续各章的学习打下基础。

## 1.1 信息、数据、数据处理与数据管理

### 1.1.1 数据与信息

在数据处理中，我们最常用到的基本概念就是数据和信息，二者之间既有区别又有联系。

#### 1. 信息 (information)

(1) 信息的定义。在信息社会，信息是一种资源，它与能源、材料一起构成客观世界的三大要素。信息是关于现实世界事物的存在方式或运动状态反映的综合。信息是客观存在的，人类有意识地对信息进行采集并加工、传递，从而形成了各种消息、情报、指令、数据及信号等。例如，对于学生基本情况来说，某同学的学号是 S1，姓名是赵亦，性别是女，年龄是 17 岁，所在系别是计算机等，这些都是关于某个同学的具体信息，是该同学当前存在状态的反映。

(2) 信息的特征。信息具有如下特征。

① 信息源于物质和能量。信息不可能脱离物质而存在，信息的传递需要物质载体，信息的获取和传递要消耗能量。如信息可以通过报纸、电台、电视、计算机网络进行传递。

② 信息是可以感知的。人类对客观事物的感知，可以通过感觉器官，也可以通过各种

仪器仪表和传感器等,不同的信息源有不同的感知形式。如报纸上刊登的信息通过视觉器官感知,电台中广播的信息通过听觉器官感知。

③ 信息是可存储、加工、传递和再生的。人们用大脑存储信息,叫做记忆。计算机存储器、录音、录像等技术的发展,进一步扩大了信息存储的范围。借助计算机,还可对收集到的信息进行整理。

## 2. 数据(data)

(1) 数据的定义。数据是用来记录信息的可识别的符号,是信息的具体表现形式。例如,上面提到的描述某个同学的信息,可用一组数据“S1、赵亦、女、17、计算机”表示。由于这些符号在此已被赋予了特定的语义,因此,它们就具有传递信息的功能。

(2) 数据的表现形式。可用多种不同的数据形式表示同一信息,而信息不随数据形式的不同而改变。如“2000年硕士研究生将扩招30%”,其中的数据可改为汉字形式“两千年”、“百分之三十”,而表达的信息是一致的。

由于早期的计算机系统主要用于科学计算,因此计算机中处理的数据主要是整数、浮点数等传统数学中的数字。但是,在现代计算机系统中,数据的概念是广义的,数据的概念内涵已被大大地拓宽了,其表现形式不仅包括数字,还包括文字、图形、图像和声音等,它们都可以经过数字化后存储到计算机中。

## 3. 数据与信息的联系

通过前面的分析可以看出,信息与数据之间存在着固有的联系:数据是信息的符号表示或载体,信息则是数据的内涵,是对数据的语义解释。如上例中的数据“2000”、“30%”被赋予了特定的语义,此处的2000表示的是“2000年”,30%表示的是“研究生将扩招30%”,因此,它们就具有了传递信息的功能。

数据表示了信息,而信息只有通过数据形式表示出来才能被人们理解和接受。尽管两者在概念上不尽相同,但通常人们并不严格去区分它们。

### 1.1.2 数据处理与数据管理

数据处理是将数据转换成信息的过程,包括对数据的收集、管理、加工利用乃至信息输出的演变与推导等一系列活动。其目的是从大量的原始数据中抽取和推导出有价值的信息,作为决策的依据。

我们可用下式简单地表示出信息、数据与数据处理的关系:

$$\text{信息} = \text{数据} + \text{处理}$$

数据是原料,是输入,而信息是产出,是输出结果。“信息处理”的真正含义应该是为了产生信息而处理数据。

在数据处理中,通常计算比较简单,而数据的管理比较复杂。数据管理是指数据的收集、整理、组织、编码、存储、维护、检索、传输等操作,这些操作是数据处理业务的必不可少的基本环节。对数据管理部分,应研制出一个通用、高效而又使用方便的管理软件,把数据有效地管理起来,以便最大限度地减轻程序员的负担;至于处理业务中的加工计算,因不同业务各不相同,要靠程序员根据业务情况编写应用程序加以解决。所以,数据处理是与

数据管理相联系的，数据管理技术的优劣，将直接影响数据处理的效率。数据库技术正是瞄准这一目标研究、发展并完善起来的专门技术。

## 1.2 数据库技术的产生、发展

通过前面的学习可知，数据处理的中心问题是数据管理。随着计算机硬件和软件的发展，数据管理经历了人工管理阶段、文件系统阶段和数据库系统阶段三个发展阶段。数据库技术正是应数据管理任务的需要而产生、发展的。

### 1. 人工管理阶段

在 20 世纪 50 年代中期以前，计算机主要用于科学计算，当时硬件中的外存只有卡片、纸带、磁带，没有磁盘等直接存取设备。软件只有汇编语言，没有操作系统和管理数据的软件。数据处理的方式基本上是批处理。因此，称这样的数据管理方式为人工管理数据。

人工管理数据有如下几个特点。

(1) 数据不保存。因为当时计算机主要用于科学计算，对于数据保存的需求尚不迫切。在计算某一课题时将原始数据随程序一起输入内存，运算处理结束后将结果数据输出，随着计算任务的完成，数据和程序一起从内存中被释放；若再计算同一课题时，还需要再次输入原始数据和程序。

(2) 系统没有专用的软件对数据进行管理。数据需要由应用程序自己管理，没有相应的软件系统负责数据的管理工作。每个应用程序不仅要规定数据的逻辑结构，而且要设计物理结构，包括数据的存储结构、存取方法和输入方式等，因此，程序员的负担也很重。

(3) 数据不共享。数据是面向程序的，一组数据只能对应一个程序。多个应用程序涉及某些相同的数据时，也必须各自定义，无法互相利用、互相参照，因此，程序之间有大量的冗余数据。

(4) 数据不具有独立性。程序依赖于数据，如果数据的类型、格式或输入输出方式等逻辑结构或物理结构发生变化，必须对应用程序作出相应的修改，因而，数据与程序不具有独立性，这也进一步加重了程序员的负担。

在人工管理阶段，程序与数据之间是一一对应的关系，如图 1-1 所示。

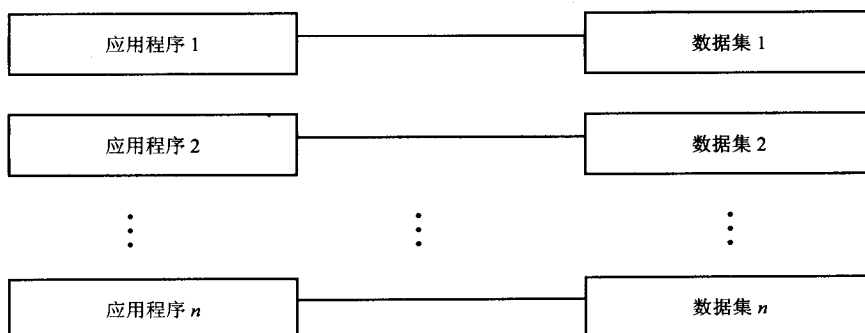


图 1-1 人工管理阶段应用程序与数据之间的对应关系

## 2. 文件系统阶段

在 20 世纪 50 年代后期至 60 年代中期, 计算机应用范围逐步扩大, 不仅用于科学计算, 还大量用于信息管理。随着数据量的增加, 数据的存储、检索和维护成为紧迫的需要。

此时, 在硬件方面, 已有了磁盘、磁鼓等直接存取设备; 在软件方面, 出现了高级语言 and 操作系统, 操作系统中有了专门管理数据的软件, 一般称为文件系统, 处理方式上不仅有文件批处理, 而且能够联机实时处理。

文件系统阶段数据管理有如下特点。

① 数据以文件形式长期保存。数据以文件的组织方式, 保存在计算机的存储设备上, 可以被多次反复使用。应用程序可对文件进行查询、修改和增删等处理。

② 由文件系统管理数据。文件系统把数据组织成内部有一定结构的记录并以文件的形式存储在存储设备上, 这样, 程序只与存储设备上的文件名打交道, 不必关心数据的物理存储 (存储位置、存储结构等), 而由文件系统提供的存取方法实现数据的存取, 从而实现了“按文件名访问, 按记录进行存取”的数据管理技术。

③ 程序与数据间有一定独立性。由于文件系统在程序与数据文件之间的存取转换作用, 使得程序和数据之间具有“设备独立性”, 即当改变存储设备时, 不必改变应用程序。程序员也不必过多地考虑数据存储的物理细节, 而将精力集中于算法设计上, 从而大大减少了维护程序的工作量。

在文件系统阶段, 程序与数据之间的对应关系如图 1-2 所示。

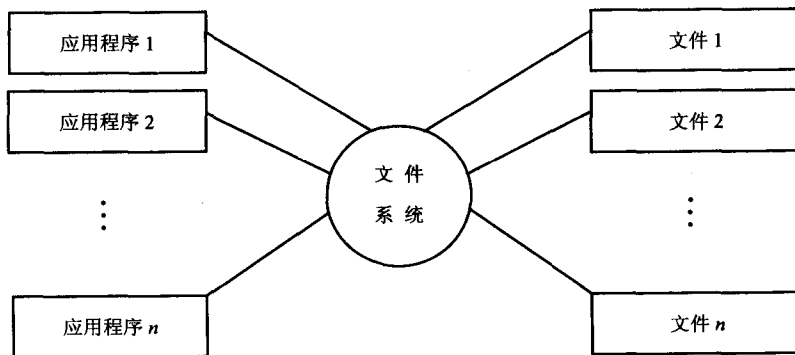


图 1-2 文件系统阶段应用程序与数据间的对应关系

与人工管理阶段相比, 文件系统阶段对数据的管理有了很大的进步, 但一些根本性问题仍没有彻底解决, 主要表现在以下几方面。

① 数据共享性差、冗余度大、易造成数据不一致。各数据文件之间没有有机的联系, 一个文件基本上对应于一个应用程序, 即文件仍然是面向应用的。当不同的应用程序所使用的数据具有共同部分时, 也必须分别建立自己的数据文件, 数据不能共享。同时, 由于相同数据的重复存储、各自管理, 在对数据进行更新操作时, 不但浪费磁盘空间, 同时也容易造成数据的不一致性。

② 数据独立性差。在文件系统阶段, 尽管程序与数据之间有一定的独立性, 但是这种独立性主要是指设备独立性, 还未能彻底体现用户观点下的数据逻辑结构独立于数据在外部



存储器的物理结构要求。因此，在文件系统中，一旦改变数据的逻辑结构，必须修改相应的应用程序，修改文件结构的定义。而应用程序发生变化，如改用另一种程序设计语言来编写程序，也将引起文件的数据结构的改变。

可见，文件系统阶段，文件之间相互独立，缺乏联系，因此，不能反映现实世界事物之间的内在联系。

### 3. 数据库系统阶段

从20世纪60年代后期开始，计算机用于管理的规模更加庞大，应用越来越广泛，数据量急剧增加，同时，多种应用、多种语言互相覆盖地共享数据集合的要求也越来越强烈。

在这种情况下，为了提高数据管理的效率，人们开始对文件系统进行扩充，但这样做没能解决问题。而与此同时，硬件方面出现了大容量、存取快速的磁盘，使计算机联机存取大量数据成为可能，这为数据库技术的产生和发展提供了物质条件。同时，硬件价格下降，而软件价格上升，使开发和维护系统软件的成本相对增加。因此，文件系统的数据库管理方法已无法适应各种应用的需要。于是为解决多用户、多个应用程序共享数据的需求，数据库技术应运而生，出现了统一管理数据的专门软件系统，即数据库管理系统。

20世纪60年代末期出现的对数据库管理技术有着奠基作用的三件大事，标志着以数据库系统为基本手段的数据管理新阶段的开始。

(1) 1968年，美国IBM公司推出了商品化的基于层次模型的信息管理系统(IMS)。

(2) 1969年，美国数据系统语言协会(Conference On Data System Language, CODASYL)的数据库任务组(DataBase Task Group, DBTG)发布了一系列研究数据库方法的DBTG报告，提出了网状数据模型。

(3) 1970年，美国IBM公司的E. F. Codd连续发表文章，提出了关系模型，奠定了关系数据库管理系统的理论基础，一直沿用至今。

20世纪70年代以来，数据库技术得到迅速发展，开发出了许多产品，并投入运行。数据库系统克服了文件系统的缺陷，提供了对数据更高级、更有效的管理，与人工管理和文件系统相比，数据库系统阶段管理数据的特点有如下几个方面。

(1) 数据结构化。文件系统中的文件之间不存在联系，从总体看，其数据是没有结构的；而在数据库系统中，将各种应用的数据按一定的结构形式(即数据模型)组织到一个结构化的数据库中，不仅考虑了某个应用的数据结构，而且考虑了整个组织(即多个应用)的数据结构，也就是说，数据库中的数据不再仅仅针对某个应用，而是面向全组织，不仅数据内部是结构化的，整体也是结构化的；不仅描述了数据本身，也描述了数据间的有机联系，从而较好地反映了现实世界事物间的自然联系。

例如：要建立学生成绩管理系统，系统包含学生(学号、姓名、性别、系别、年龄)、课程(课程号、课程名)、成绩(学号、课程号、成绩)等数据，分别对应三个文件。若采用文件处理方式，因为文件系统只表示记录内部的联系，而不涉及不同文件记录之间的联系，要想查找某个学生的学号、姓名、所选课程的名称和成绩，必须编写一段比较复杂的程序来实现，即不同文件记录间的联系只能写在程序中，在编写程序时考虑和实现。而采用数据库方式，由于数据库系统不仅描述数据本身，还描述数据之间的联系，记录之间的联系可以用参照完整性来表述，上述查询可以非常容易地联机查到。