



数据库分析与设计

DATABASE ANALYSIS AND DESIGN

周佩德 译 姚玉洁 校

南京工学院出版社

数据库分析与设计

〔美〕 I.T.霍列基威茨 著

周 佩 德 译
姚 玉 洁 校

南京工学院出版社

内 容 提 要

本书介绍数据库应用系统的开发中数据库分析及设计的技术和方法。全书共分十七章,主要内容沿数据库生命期而展开:首先介绍关系模型及其它几种语义模型,并以这几种模型作工具阐述数据的分析和结构的过程;然后介绍三种主要数据模型及若干典型DBMS,及在这些模型上实现逻辑设计和物理设计的步骤;同时叙述对设计进行评价的技术;最后讨论如何将各阶段的设计技术集成为一个整体的设计方法。本书贯彻理论和实践并重的原则,内容深入浅出,包含大量的实例和习题。

本书可作为高等院校计算机有关专业数据库设计课程的教材或教学参考书,也可供从事计算机工作的科研人员、工程技术人员及其他有关人员参考。

数 据 库 分 析 与 设 计

(美) I. T. 霍列基威茨 著

周佩德 译

姚玉洁 校

南京工学院出版社出版

江苏南京市四牌楼2号

江苏省新华书店发行 江苏高淳印刷厂印刷

开本787×1092毫米 1/16 26.75印张 字数 610千字

1987年11月第1版 1987年11月第1次印刷

印数1—4000册

ISBN 7-81023-013-(1)/T·13

统一书号: 15409·021 定价: 4.50元

责任编辑: 孙文治

前 言

当前、数据库管理已得到了广泛的实际应用，並已成为计算机科学的一个重要研究领域。近几年来，无论是在数据库的实际应用方面，还是在其理论基础方面，都出现了很大的进展。这些进展是由于隶属于数据库方向的许多研究领域的发展而取得的，在这些领域中有些是面向实际的，而有些是面向理论的。这些领域如关系数据库理论、数据语义、数据库管理系统以及物理文件设计等等。

一位数据库的初学者在涉足这些领域时，很容易把这些领域视为互不相关的一些有趣课题。诚然，在关系数据库理论和物理文件设计之间或者在数据语义和数据库管理系统之间是很难看出直接联系的。于是，可能会有人对所有这些各别的领域都有所了解，但是却不能将这些领域有机地结合在一起。

本书的目标之一是试图把数据库诸多领域的研究成果围绕数据库设计的全过程统一为一个有机的整体。当然，有人可能会提出为什么要利用数据库设计来作这种结合领域呢？我们的回答是：数据库设计在某种意义上可以认为是数据库研究在应用方面的主要目标，亦即我们的最终目标是设计和使用数据库。如果在数据库某一方面所进行的研究期望最终得到应用，则此种研究必然要和数据库设计过程的某个阶段相联系。

除作为结合数据库各研究领域的框架外，数据库设计中还要为数据库的优劣提供评价标准，因此，对数据库各方面所进行的研究还必须为数据库设计人员提供根据这些准则评价数据库的能力。在实际设计工作中必须满足下述设计准则：

1. 所选用的数据库结构应能很容易地导出用户需求的所有联系。
2. 当用户需求发生变化时，所设计的数据库只需通过维护就可满足变化后的需求，而无须重新设计其结构和重新开发全部程序。
3. 所设计的数据库在完成用户需求中，性能应是合理的。

因此，在数据库各个领域中所开展的研究，必须有助于实现上述这些准则。应该指出，就数据库设计本身而言，目前尚不是一个完全结构化的过程。早期的设计人员在工作中几乎不能借助于什么设计技术和设计工具，一般说来，他们当时是利用文件设计技术来进行数据库设计的，其中的许多技术注重的是物理设计的准则，强调的是性能而不是数据的结构或灵活性。

然而，近几年在数据库设计方面已经开发了许多新的技术，从而逐步形成了数据库设计方法，即从用户数据结构的形式化分析开始，经过一系列结构化的设计步骤，最终由用户需求演变到数据库的物理设计。因此，我们必须首先研究有待存贮的数据，进而再选择这些数据在计算机中的表现方法。由于数据库的关系理论已为选择数据库的结构提供了实现形式化的基础，而且这种理论很快在实际工作中得到了应用，所以数据库设计方法的许

多成果是随着数据库关系理论的发展而出现的。

近年来各种各样的数据库分析和设计技术已经得到了开发, 这些技术有些是以关系理论作为基础, 其余的则是以诸如实体联系法、作用模型或聚合及归纳这样的数据抽象技术作为构造数据库的指南。

因此, 数据库系统的设计周期为各种数据库领域的集成提供了一个基础。在数据分析和结构的过程中使用关系理论和语义模型, 所产生的结构转化为基于某种数据模型上的逻辑结构, 这些模型进而由市场所提供的数据库管理系统所实现。因此在语义模型和商业软件之间产生了联系。而且最终可通过考察存取需求以及可供设计人员使用的物理结构, 对所实现的数据库进行调整, 所以构成了从逻辑结构到物理设计的联系。本书就是根据这个设计周期概括地阐述了数据库的各个研究领域以及在设计过程中各个阶段的数据库理论。本书内容包括了作者在麻省理工学院(MIT)攻读博士学位时的心得与经验, 当时作者在该校首次聆听了J.丹尼斯(Jack Dennis)教授对数据库关系理论的介绍, 后来作者曾在澳大利亚和美国两地, 从事过数据库设计的实际工作, 开设和讲授过大学本科生的数据库课程, 并致力于数据库的研究工作, 在美国作者有幸在马里兰大学与E.西勃莱(Edgar Sibley)教授一起合作共事。本书内容还包括作者多年对数据库实际工作人员讲授高级数据库课程所取得的丰富教学经验, 因此, 本书各章节中都列举了作者在此期间所积累的许多设计实例。

在本书的编写过程中承蒙许多同行提出不少建设性的意见, 使作者受益匪浅。作者在此向IBM公司的M.保尔(Marilyn Bohl)、南美以美会大学的C.M.依斯特曼(Caroline M. Eastman)、旧金山州立大学的K.M.赫特尔(Kenneth M. Hunter)、乔治亚州工学院的N.D.格列弗斯(Nancy D. Griffith)诸位同行谨表谢意。最后还得指出, 本书得以问世是与SRA出版社的大力支持分不开的, 特别要感谢本书初稿形成时A.罗厄(Alan Lowe)所给予的帮助, 以及后期阶段中E.A.伍德(E. Ann Wood)和SRA出版社编辑部工作人员给予的支持。

译者序

数据库技术是计算机数据管理的最新技术。由于数据库技术的出现和发展,促使计算机应用更加广泛地渗透到工农业、商业、行政管理、科学研究、工程技术和国防军事等各个领域。在这些需要数据库技术的领域中,当前的問題是如何更有效地应用数据库技术的问题,而设计这些数据库应用系统的中心问题就是数据库设计。

数据库设计是个复杂的过程,在这个过程中需要将现实世界中的事物最终转化为由机器世界所存贮和管理,所以要求数据库设计师必须对实际应用对象和数据库系统这两方面都有充分的了解。很长时期以来,数据库设计人员由于找不到好的设计方法和工具而只能凭经验和直觉来设计数据库,所以数据库设计往往被认为是一门技艺(art)而並不是一门科学(science)。

随着实际应用对象的日益复杂,数据库设计工程已成为一项软件工程,通过对软件工程的研究,数据库研究工作者将数据库设计分解为概念设计、逻辑设计、物理设计和系统维护等几个阶段,並逐渐对各个阶段提出了一些设计的理论和方法。但是由于数据库设计横跨了现实世界和机器世界,所以一直没有找到一种综合的方法能将数据库设计各个阶段的设计方法和工具集成为一个整体的设计方法,特别是由于现实世界的丰富语义特性,更增加了这项工作的复杂性。

本书作者在这方面做了大量的工作,书中首先介绍了关系模型及其它几种语义模型,並以这几种模型作工具阐述数据的分析和结构的过程;接着介绍了三种主要数据模型及若干典型的DBMS,以及在这些模型上实现逻辑设计和物理设计的步骤,然后叙述对设计进行评价的技术,最后讨论如何将各阶段的设计技术集成为一个整体的设计方法。这样,本书作者提出了如何以一种统一的语义模型作基础,而构造一个集成的设计方法的道路。本书作者多年从事数据库设计的理论工作和实践工作,积累了丰富的实践经验,书中的例子都有很强的实践背景,而且收集了大量的设计实例作为习题,所以可以说本书以理论和实践並重、深入浅出作为其第二个特点。同时,书末还列出了丰富的参考书目,为希望对数据库设计进行深入研究的人员提供了方便。译者在从事数据库的数学和数据库设计的研究中,深切地感到数据库设计的专著实在太少,特别是这种将数据库设计作全面而统一论述的专著更为稀缺,所以翻译了这本书,愿为有关院校提供一本教学用书,同时希望能对数据库技术在我国推广有所裨益。

在本书第十一章至第十三章及设计课题的翻译中,曾参阅了中国科学院应用数学研究所乐庆同志的译文;高等教育出版社姚玉洁同志审阅了全书,对原著的某些笔误及译者的译文作了精心校订;同时,本书在翻译出版的过程中得到南京工学院王能斌教授的大力支持帮助,译者在此致以谢忱。

译文改正了原著的若干笔误和印刷错误,限于篇幅在译文中並未逐一注明。译者虽力求翻译准确,但限于水平,误译之处在所难免,谨望广大读者不吝指正。

周佩德 于南京工学院 1987.4

目 录

第一章 引论	1
§ 1.1 计算机和数据库	1
§ 1.2 数据库设计过程的历史沿革	2
§ 1.3 数据库的直接设计过程	3
§ 1.4 分阶段设计过程	3
§ 1.5 分析和设计的工具	4
§ 1.6 数据库需求的说明和设计	6
§ 1.7 数据库的技术设计	6
§ 1.8 数据库设计的技术	7
§ 1.9 本书的结构	8
1.9.1 数据分析	8
1.9.2 实现模型	8
1.9.3 习题	9
第二章 关系模型	10
§ 2.1 引言	10
§ 2.2 关系模型的历史回顾	10
§ 2.3 关系模型的基本结构	10
2.3.1 术语	12
2.3.2 一致性	14
§ 2.4 函数依赖	17
§ 2.5 范式	18
2.5.1 举例	19
2.5.2 BC范式	20
§ 2.6 函数依赖和关系设计	21
2.6.1 函数依赖的特性	21
2.6.2 极小化算法	23
§ 2.7 关系语言	24
2.7.1 关系演算	24
2.7.2 QUEL	26
2.7.3 SQL	26
2.7.4 关系代数	30
§ 2.8 本章小结	32
习 题	33

第三章 关系的设计	39
§ 3.1 引言	39
§ 3.2 泛关系假设	39
§ 3.3 关系设计准则	42
3.3.1 满足可表述准则	42
3.3.2 无损分解	44
3.3.3 无损分解的条件	44
3.3.4 冗余性准则	44
§ 3.4 关系设计过程	45
§ 3.5 分解	45
3.5.1 简单分解算法	45
3.5.2 简单分解算法的局限性	47
3.5.3 多值依赖	49
3.5.4 多值依赖的关系属性相关性	51
3.5.5 第四范式	51
3.5.6 分解算法2	51
3.5.7 多值依赖和函数依赖	53
3.5.8 分解算法的一些复杂问题	54
§ 3.6 合成	55
3.6.1 函数依赖的语义	55
3.6.2 合成算法	56
§ 3.7 关系理论的最新进展	60
§ 3.8 本章小结	61
习 题	62
第四章 数据分析中的语法设计及语义设计问题	64
§ 4.1 引言	64
§ 4.2 系统分析中的关系模型	64
§ 4.3 基于记录的数据模型	67
§ 4.4 关系模型的扩充	67
§ 4.5 语义模型	68
§ 4.6 语义模型对组织的模型化	69
4.6.1 语义模型	69
4.6.2 语义模型的表示	70
§ 4.7 本章小结	71
习 题	72
第五章 语义模型I——实体及联系	73
§ 5.1 引言	73
§ 5.2 实体—联系模型	73

5.2.1	实体集和联系集	74
5.2.2	E-R模型的图形表示	74
5.2.3	多值属性	75
5.2.4	非函数联系	75
5.2.5	标识符	76
5.2.6	使用E-R模型进行模型化	78
§ 5.3	实体、联系及关系	79
§ 5.4	实体和联系的抉择	81
§ 5.5	实体间的依赖	82
5.5.1	使用组合码的问题	82
5.5.2	组合码的错误使用	84
§ 5.6	多值联系	84
§ 5.7	二元联系和n元联系	87
§ 5.8	实体模型	91
§ 5.9	联系的不可表达性	91
§ 5.10	关系的最小化	92
§ 5.11	本章小结	92
	习 题	92
第六章	语义模型Ⅱ——作用及类型	95
§ 6.1	引言	95
§ 6.2	递归联系	96
§ 6.3	作用模型	98
6.3.1	模型化过程中作用的特性	99
6.3.2	作用概念的其它优点	101
6.3.3	非同构性	101
6.3.4	非同构及函数依赖	103
6.3.5	模型化过程中作用的结构	103
6.3.6	作用的标识符和作用的结构	105
§ 6.4	作用、源及关系	107
6.4.1	作用结构与函数依赖	107
6.4.2	同质作用集	107
6.4.3	异质作用集	107
6.4.4	简化方法	111
§ 6.5	实体类型	113
6.5.1	实体类型的模型化	114
6.5.2	关系和异质类型集	115
6.5.3	简化方法	118
§ 6.6	聚合客体	118

§ 6.7	实体类型和聚合客体的组合	121
§ 6.8	本章小结	122
	习 题	123
第七章	语义模型化过程的统一及问题	129
§ 7.1	引言	129
§ 7.2	聚合和归纳	130
§ 7.3	二元模型	134
§ 7.4	函数模型	137
§ 7.5	本章小结	138
	习 题	138
第八章	数据库说明	139
§ 8.1	引言	139
§ 8.2	转化方法	140
§ 8.3	语义模型的规范化	141
§ 8.4	“规范化”的语义模型到记录说明	143
8.4.1	同质作用的逻辑结构	144
8.4.2	异质强制作用	146
8.4.3	异质选择作用	148
8.4.4	关于作用转化和类型转化的说明	148
§ 8.5	关系模型转化为逻辑记录结构	149
§ 8.6	逻辑记录结构的最小化(MLRS)	150
§ 8.7	存取需求说明	150
8.7.1	存取路径	150
8.7.2	伪码的使用	152
§ 8.8	量化数据说明	155
§ 8.9	本章小结	156
	习 题	156
第九章	实现模型I——文件结构	157
§ 9.1	引言	157
§ 9.2	实现模型	159
9.2.1	层次总体结构	160
9.2.2	层次间映射的产生	163
9.2.3	分层结构的作用	164
§ 9.3	物理记录界面	164
9.3.1	磁盘传送的存取方法软件	165
9.3.2	其它一些物理界面	166
9.3.3	页面	166
9.3.4	可变长记录	166

9.3.5 可变长物理记录	168
9.3.6 多缓冲区	169
9.3.7 物理界面对设计的影响	169
§ 9.4 逻辑记录存取	169
§ 9.5 顺序存取	170
§ 9.6 直接存取	172
9.6.1 散列存取方法	173
9.6.2 索引实现	173
9.6.3 B树	177
9.6.4 各种直接存取方法的比较	180
§ 9.7 多索引存取	180
§ 9.8 索引顺序存取方法	181
§ 9.9 链结文件	182
9.9.1 符号指引元	182
9.9.2 逻辑记录指引元	183
§ 9.10 COBOL语言的存取方法	184
9.10.1 顺序文件组织	185
9.10.2 索引顺序文件组织	185
9.10.3 相对文件组织	187
§ 9.11 本章小结	188
习 题	189
第十章 实现模型 II —— 数据库管理系统	191
§ 10.1 引言	191
§ 10.2 自然的用户界面	191
§ 10.3 用户视图	193
§ 10.4 数据独立性	193
§ 10.5 DBMS 的用户界面软件	195
§ 10.6 DBMS 的结构	198
10.6.1 三层次结构	198
10.6.2 模式—子模式结构	200
§ 10.7 数据库的操作环境	200
10.7.1 操作环境的类型	201
10.7.2 处理方式	201
§ 10.8 支持多用户环境的数据库功能	204
10.8.1 数据库的完整性	204
10.8.2 恢复	206
10.8.3 数据库的安全性	206
10.8.4 数据库的分布	207

10.8.5	数据字典	209
10.8.6	DBMS系统软件的典型结构	210
§ 10.9	本章小结	211
	习 题	211
第十一章	关系数据库管理系统	212
§ 11.1	引言	212
§ 11.2	System R	213
11.2.1	SQL语言	213
11.2.2	数据库定义	214
11.2.3	数据库的赋值与更新	214
11.2.4	修改数据库定义	215
11.2.5	定义用户视图	215
11.2.6	嵌入式SQL	216
11.2.7	实现	218
§11.3	QBE (Query—By—Example)	221
11.3.1	条件检索	221
11.3.2	多个关系的检索	223
11.3.3	逆条件	224
11.3.4	函数	225
11.3.5	集合比较	225
11.3.6	插入、删除、更新	226
11.3.7	数据库定义的建立和修改	227
§ 11.4	个人计算机系统	228
11.4.1	用户界面	228
11.4.2	个人计算机数据库的使用	228
11.4.3	某些更高级的功能	230
§ 11.5	本章小结	231
	习 题	231
第十二章	网状数据模型	233
§ 12.1	引言	233
§ 12.2	网状模型	233
§ 12.3	网状模型的实现	237
§ 12.4	DBTG实现方法	242
12.4.1	DBTG模式	244
12.4.2	域	244
12.4.3	记录条目	245
12.4.4	系条目	247
12.4.5	DBTG网状数据库的存取	253

12.4.6 子模式	259
§ 12.5 本章小结	261
习 题	261
第十三章 层次数据库管理系统	267
§ 13.1 引言	267
§ 13.2 层次模型	267
13.2.1 层次数据结构	267
13.2.2 层次数据结构的数据存取	268
§ 13.3 IMS	270
13.3.1 IMS的结构	270
13.3.2 IMS的物理结构	273
13.3.3 IMS数据库的定义	277
13.3.4 辅助索引	282
13.3.5 IMS数据库的存取	283
13.3.6 其它操作特性	285
§ 13.4 通用选择法	286
§ 13.5 System 2000	286
13.5.1 System 2000的数据结构	286
13.5.2 System 2000数据库的定义	287
13.5.3 System 2000数据库的存取	288
13.5.4 数据库的物理结构	293
13.5.5 System 2000的一些操作特性	294
§ 13.6 本章小结	294
习 题	296
第十四章 设计过程	300
§ 14.1 引言	300
§ 14.2 设计目标	301
§ 14.3 设计技术	301
§ 14.4 初始设计	302
§ 14.5 设计的迭代过程	305
14.5.1 性能问题	306
14.5.2 设计问题	306
14.5.3 设计策略	306
§ 14.6 逻辑设计的策略	306
14.6.1 减少存取步的数目	306
14.6.2 派生联系	307
14.6.3 数据项的重复	308
14.6.4 文件的组合	308

14.6.5 组合为非同构文件	308
§ 14.7 物理设计的策略	309
14.7.1 问题4——指引元的管理	309
14.7.2 问题5——需要的存贮量太大	310
14.7.3 问题6——系统开销	310
§ 14.8 本章小结	310
习 题	311
第十五章 初始设计	312
§ 15.1 引言	312
§ 15.2 关系数据库的设计	312
§ 15.3 网状数据库的设计	312
15.3.1 转化规则的特性	313
15.3.2 作用和类型的转化	316
15.3.3 两种转化方法的选择	317
15.3.4 网状数据库的初始物理设计	318
§ 15.4 层次数据库的设计	321
15.4.1 层次设计方法的选择	322
15.4.2 第1类设计方法	323
15.4.3 分部设计法	326
15.4.4 第2类设计方法	326
15.4.5 第3类设计方法	329
15.4.6 作用和类型的转化	332
15.4.7 初始物理设计	334
§ 15.5 COBOL语言的文件设计	334
15.5.1 码冲突	334
15.5.2 作用和类型的转化	336
§ 15.6 本章小结	338
习 题	339
第十六章 设计的评价	340
§ 16.1 引言	340
§ 16.2 DBMS的选择	340
§ 16.3 数据库的监测	341
§ 16.4 性能估价	341
16.4.1 逻辑设计分析与物理设计分析	342
16.4.2 分析过程	343
16.4.3 存贮需要量	344
16.4.4 逻辑分析	347
16.4.5 物理分析	350

16.4.6 DBMS环境中的性能估计	352
§ 16.5 模型化的分析方法	353
16.5.1 概率估计	354
16.5.2 响应时间估计	356
16.5.3 事务按常数分布	359
16.5.4 分析模型的推广	359
§16.6 本章小结	359
习 题	360
第十七章 设计方法的选择	366
§ 17.1 引言	366
§ 17.2 设计方法的构造	366
17.2.1 设计技术的选择	366
17.2.2 文本化方法的选择	367
§ 17.3 选择1: 数据的收集方法	367
17.3.1 信息来源	367
17.3.2 数据收集的目的	368
§ 17.4 选择2: 分析组织数据的方法	374
17.4.1 数据流分析	374
17.4.2 面向数据的分析方法	374
§ 17.5 选择3: 数据库说明的构成	375
§ 17.6 选择4: 向逻辑设计转化	375
§ 17.7 选择5: 选择设计结果的分析模型	376
§ 17.8 本章小结	376
数据库设计课题	377
设计课题A: 合同——供应系统	378
设计课题B: 供求系统	380
设计课题C: 车间统计系统	383
设计课题D: 课程系统	387
附录 样本数据库	391
参考文献	399

第一章 引 论

数据库是一个组织的信息系统的基本部分。在组织的工作过程中，信息系统管理工作过程所需的数据，辅助用户用这些数据进行决策，所以，信息系统是该组织工作过程的支柱。数据库是信息系统中数据的仓库，所以在一个组织的工作过程中，数据库发挥着核心作用。

用户在决策过程中，通过访问数据库获得数据，决策完成后再将决策结果存贮在数据库中。数据库的布局和存取功能对信息系统的工作效率影响极大，如果在组织的所有工作地点都能很容易存取各种数据，则能使信息系统快速响应组织内决策人员的要求；反之，存取不易当然会妨碍快速响应。如果获取数据很困难，则不可能进行及时的决策，或者只能根据不完全的数据进行决策，其结果都可能导致系统以后不能正常工作。

同时，数据库必须具有灵活的结构。这样，即使组织对数据的需求发生了变化，数据库也能满足该组织的需要。当一个组织要完成一项新的任务时，就要进行相应的新决策，这时数据库必须存贮新的数据以及数据间新的联系，以支持新的决策。所以，数据库必须具备这种能很容易进行相应变更的功能。

本书的主题是：为满足信息系统的需求如何进行数据库设计。为此，书中讨论了数据库设计的一系列重要问题：诸如数据模型、数据库中数据记录的结构及数据库的物理组织方式对系统性能的影响，等等。

§1.1 计算机和数据库

显然，数据必须存贮在物理介质上；另外，在数据库及其用户之间传送数据也必须借助物理的手段。

信息系统的设计者必须选择存贮数据的物理介质。过去，存贮数据长期用的是文件柜、纸张、分类卡片、个人文件等等；而现在，存贮数据已越来越广泛地使用了计算机，同时，计算机还辅助用户完成大量数据处理工作。当然，这并不意味着组织中所有的数据都要由计算机进行管理，考虑到当前的经济性，某些数据不必由计算机存贮，某些数据要经过一段时间才能逐步存贮到计算机系统中去。

计算机系统除存贮数据外，还提供了存取数据的功能，使位于不同工作场所的人都可以存取这些数据。用户使用存取语言来存贮数据和检索数据，远距离的终端通过计算机通讯网络访问计算机来传送数据。信息系统的目标是使用户方便地存取数据，计算机系统的上述灵活性满足了信息系统的需要。

§1.2 数据库设计过程的历史沿革

随着计算机数据库应用范围的扩大,如何设计数据库的问题越来越受到人们的重视;随着数据库新应用领域的开拓,数据库设计过程也不断地发展变化。早期,数据库通常只支持一个组织的某个单方面的应用,例如,财会系统已有了一个数据库,仓库管理系统可能有另一个数据库,等等,而且这些大部分为批处理系统,只是提供定期的规格化报表。现在,许多数据库系统和这些早期的系统已有很大的不同,用户的功能已逐步集成化,许多数据库能完成组织的若干方面的功能,同时,数据库已和用户环境更紧密地结合起来,数据库除提供定期的规格化报表以外,还允许用户进行联机访问,不仅能产生规格化报表而且可完成用户即席设计的报表。

随着数据库系统的发展,数据库日益复杂化,数据库和用户功能日益集成化,从而出现了许多问题,迫切需要在数据库设计过程中能解决这些问题。和早期的设计过程相比,数据库设计过程已发生了很大变革。许多早期的设计过程强调的是数据库的物理实现,主要注重于数据记录的存取方法,设计人员往往只需考虑系统某单项功能的实现,从而也只考虑一个文件至多少数几个文件,然后选择适当的索引技术,以满足实现这个功能时的性能要求;而现在、对数据库设计已提出许多准则,其中包括:

- 减少数据的冗余量;
- 提供稳定的数据结构,在用户需求发生改变时,该数据结构能随之迅速变化;
- 允许用户对数据进行即席访问;
- 在数据元素间维持复杂的联系;
- 支持多种多样的决策需要。

在这种环境中进行数据库设计往往是一件相当复杂的任务,为有效地完成这一任务特别需要一些合适的技术,同时还要求将这些设计技术正确组织起来,构成一个有序的设计过程。

我们特地将设计技术和设计过程区别开来。所谓设计技术是指数据库设计人员所使用的设计工具,其中包括各种算法、文本化方法、用户组织的图形表示法、各种转化规则、数据库定义的方法及编程技术;而设计过程则确定了这些技术的使用顺序。例如,在一个规范的设计过程中,可能要求设计人员首先用图形表示用户数据,再使用转化规则生成数据库结构,下一步则用某确定的算法优化这一结构,这些工作完成后,就可进行数据库的定义工作和程序开发工作。

一般说来,设计技术分为下列两类:

- 数据分析技术,数据分析技术用以分析用户数据的语义;
- 技术设计技术,技术设计技术用以将数据分析的结果转化为数据库的技术实现。

上述两类技术处理的是两种不同的问题。第一类问题考虑的是正确地结构数据,这类问题通过使用诸如消除数据冗余技术、保证数据库稳定性技术、结构数据技术来解决,其目的是使用户易于存取数据,从而满足用户对数据的各种需求;第二类问题是保证所实现的数据库能有效地使用数据资源,解决这类问题要用到一些技术设计技术,例如选择合适的存贮结构以及采用有效的存取方法等等。