计 算 机 科 学 丛 书

数据库系统实现

Database System Implementation

(美) Hector Garcia-Molina
Jeffrey D. Ullman
Jennifer Widom
著
(斯坦福大学)
杨冬青 唐世渭 徐其钧 等译

机械工业出版社
China Machine Press
Prentice Hall

计算机科学丛书

数据库系统实现

Hector Garcia-Molina
(美) Jeffrey D. Ullman 著
Jennifer Widom
(斯坦福大学)
杨冬青 唐世渭 徐其钧 等译

机械工业出版社
China Machine Press

本书是斯坦福大学计算机科学专业数据库系列课程第二门课的教科书。书中对数据库系统实现原理进行了深入阐述，并具体讨论了数据库管理系统的三个主要成分——存储管理器、查询处理器和事务管理器的实现技术。书中还对信息集成的最新技术，例如数据仓库、OLAP、数据挖掘、Mediator、数据立方体系统等进行了介绍。

本书适合于作为高等院校计算机专业研究生的教材或本科生的教学参考书，也适合作为从事相关研究或开发工作的专业技术人员的高级参考资料。

本书版权登记号：图字：01-2000-0304

图书在版编目（CIP）数据

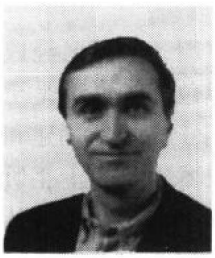
数据库系统实现/（美）加西亚（Garcia-Molina, H.），（美）沃尔曼（Ullman, J. D.），（美）威德姆（Widom, J.）著；杨冬青等译. - 北京：机械工业出版社，2001.3
（计算机科学丛书）
书名原文：Database System Implementation

ISBN 7-111-07887-X

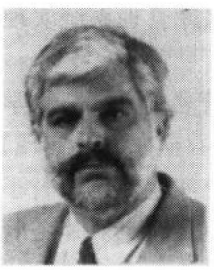
Ⅰ.数… Ⅱ.①莫…②沃…③威…④杨… Ⅲ.数据库系统-理论研究 Ⅳ.TP311.13
中国版本图书馆CIP数据核字（2000）第80012号

机械工业出版社（北京市西城区百万庄大街22号 邮政编码 100037）
责任编辑：陈贤舜
北京忠信诚胶印厂印刷·新华书店北京发行所发行
2001年3月第1版第1次印刷
787mm×1092mm 1/16·29.75印张
印数：0 001-5 000册
定价：45.00元

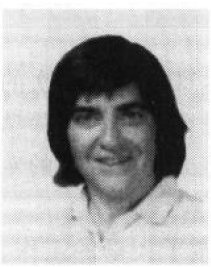
作者简介



Hector Garcia-Molina是斯坦福大学计算机科学与电气工程系的Leonard Bosack和Sandra Lerner教授。他在数据库系统、分布式系统和数字图书馆领域中发表了大量论文。他的研究兴趣包括分布式计算系统、数据库系统和数字图书馆。



Jeffrey D. Ullman是斯坦福大学的Stanford W. Ascherman计算机科学教授。他作为作者或合作者出版了15本著作，发表了170篇技术论文，其中包括《A First Course in Database Systems》（Prentice Hall 出版社，1997）和《Elements of ML Programming》（Prentice Hall 出版社，1998）。他的研究兴趣包括数据库理论、数据库集成、数据挖掘和利用信息基础设施进行教育。他获得了Guggenheim Fellowship等多种奖励，并被推选进入国家工程院。他还被授予1996年Sigmod贡献奖和1998年Karl V. Karstrom杰出教育家奖。



Jennifer Widom是斯坦福大学计算机科学与电气工程系的副教授。她是多个编辑委员会和程序委员会的成员，在计算机科学会议和杂志上发表了许多文章。她还是《A First Course in Database Systems》的作者之一。她的研究兴趣包括半结构化数据的数据库系统和XML、数据仓库以及主动数据库系统。

译者序

随着计算机硬件、软件技术的飞速发展和计算机系统在各行各业的广泛应用，数据已经成为各种机构的宝贵资源，数据库系统对于当今科研部门、政府机关、企事业单位等来说都是至关重要的。而数据库系统中的核心软件是数据库管理系统（DBMS）。DBMS用于高效地创建和存储大量的数据，并对数据进行有效的管理、处理和维护，是数据库专家和技术人员数十年研究开发的结果，是当前最复杂的系统软件之一。要深入掌握数据库系统的原理和技术，进而从事数据库管理软件和工具的开发，必须学习和研究数据库管理系统实现技术。要深入了解数据库系统的内部结构，以开发出高效的数据库应用系统，也需要学习和研究数据库管理系统实现技术。

Hector Garcia-Molina、Jeffrey D. Ullman和Jennifer Widom是斯坦福大学著名的计算机科学家，多年来他们在数据库系统领域中做了大量的开创性工作。由他们撰写的《数据库系统实现》一书是关于数据库系统实现方面内容最为全面的著述之一。书中对数据库系统实现原理进行了深入阐述，并具体讨论了数据库管理系统的三个主要成分——存储管理器、查询处理器和事务管理器的实现技术。书中还对信息集成的最新技术，例如数据仓库、OLAP、数据挖掘、Mediator(集成层软件)、数据立方体系统等进行了介绍。该书已经作为斯坦福大学计算机科学专业数据库系列课程第二门课的教科书使用。我们在北京大学计算机系研究生课程的教学中也使用了该书中的部分内容。

我们认为该书内容深入且全面，技术实用且先进，叙述深入浅出，是一本难得的高层次的教科书。我们将这本书译成中文，介绍给国内广大读者。我们认为这本书既适合于作为高等学校计算机专业研究生教材或本科生课程参考书，又适合于作为从事相关的研究或开发工作的专业技术人员的高级参考资料。

杨冬青全面组织了本书的翻译，唐世渭和徐其钧在本书的翻译和审校中做了大量的工作。参加翻译的还有杨良怀、王爱华、王腾蛟、叶茂盛、赵绍军、赵畅。另外，高桂英协助进行了译稿的整理、录入等工作。

在本书的翻译过程中，译者参照该书的WWW主页中的勘误表，对书中的疏漏之处进行了更正。此外，对于未包括在勘误表中的明显的笔误和排版错误，我们也做了订正。

限于译者水平，译文中疏漏和错误难免，欢迎批评指正。

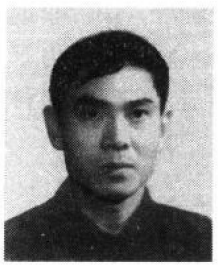
译者

2000年10月于北京大学

译者简介

杨冬青，北京大学计算机系教授，博士生导师，数据库与信息系统领域负责人。中国计算机学会数据库专委会委员，中国计算机学会普及工作委员会副主任。1969年毕业于北京大学数学力学系数学专业，从事数据库与信息系统领域研究、开发与教学二十余年，曾获国家科技进步二等奖等多项奖励。

唐世渭，北京大学计算机系教授，博士生导师。北京大学信息科学中心主任，视觉与听觉信息处理国家重点实验室主任。中国计算机学会数据库专委会副主任。1964年毕业于北京大学数学力学系计算专业，从事数据库与信息系统领域研究、开发与教学二十余年，曾获国家科技进步二等奖等多项奖励。



徐其钧，北京大学计算机系高级工程师。从事数据库与信息系统领域的研究、开发十余年。参加过国家科技攻关项目数个，主持过大型数据库应用系统项目，并编著过数据库领域的书籍数部。

前　言

本书是为斯坦福大学数据库系列课程的第二门课CS245设计的。第一门课程CS145的内容包括数据库设计和数据库编程，Jeff Ullman和Jennifer Widom为该课程写的教科书《数据库系统入门教程》（A First Course in Database Systems）于1997年由Prentice-Hall出版社出版。CS245的内容包括DBMS实现技术，特别是存储结构、查询处理和事务管理。

本书的使用

斯坦福大学实行每学年4个学期的制度，所以采用本书的主要课程CS245的教学时间仅为10周。在1999年冬季学期，Hector Garcia-Molina使用了本书的“试用”版，教学内容包括以下部分：2.1～2.4节，整个第3章和第4章，5.1节和5.2节，6.1～6.7节，7.1～7.4节，整个第8章，第9章去掉9.8节，10.1～10.3节，11.1节，以及11.5节。

第6章和第7章的剩余部分（查询优化）在高级课程CS346中讲授。在该课程中，要求学生实现他们自己的DBMS。本书中未包括在CS245中的其他部分可以在另一门高级课程CS347中讲授，该课程讨论分布式数据库和高级事务处理。

实行学期制的学校可以将本书与前一本教科书《数据库系统入门教程》结合使用。我们建议将《数据库系统入门教程》用于第一个学期，同时进行数据库应用程序设计实习。第二学期可以讲授本书的大部分或全部内容。将数据库的学习分为两门课程的好处是，不打算致力于DBMS研究的学生可以仅选修第一门课程，然后可以将数据库技术应用于他们所进入的计算机科学的任何分支。

选修要求

学生一般不会在大学的第一学年选修使用本书的课程，所以我们期望本书的读者具有计算机科学的传统领域中相当广泛的背景知识。我们假定读者已经学习过数据库程序设计，特别是SQL。读者最好了解关系代数，并且对于基本数据结构有一定程度的熟悉。同样地，关于文件系统和操作系统的知识也是很有帮助的。

习题

本书包括大量习题，几乎每一节都有习题。我们用惊叹号标记出难度较大的习题，或习题中难度较大的部分。对于特别难的习题，我们用两个惊叹号标记。

某些习题或习题中的部分用星号标记。对于这些习题，我们将努力通过本书的Web页面提供解答。这些解答向公众发布，读者可以用来进行自我测试。注意，在有些情况下，习题B要求你对习题A的解答进行修正或改编。如果习题A的某些部分有Web发布的解答，那么在Web页面上

也会有习题B的相应部分的解答。

WWW支持

本书的主页为

http://www-db.stanford.edu/~ullmam/dbsi.html

在主页上你可以找到标注星号的习题的解答、勘误表以及辅助材料。我们打算在每一次讲授CS245和其他数据库课程的相关部分时，将注释也提供到主页中，包括作业、考试和解答等。

H. G.-M
J.D.U.
J. W.
加州，斯坦福大学

目　录

第1章 DBMS实现概述

数据库对于当今的任何部门都是至关重要的。数据库被用来保存内部记录，在WWW中向顾客和客户展示数据，以及支持许多其他的商务处理。同样地，在许多科学研究的核心中也需要数据库。数据库被用来表示天文学家、基因组研究人员、探索蛋白质的医药特性的生物化学家以及许多其他的科学研究人员收集到的数据。

数据库的能力是由于知识和技术的结合，这是数十年研究开发的结果，并且已经嵌入到专门的软件中，这个软件称作*数据库管理系统*或DBMS，或更口语化地称为“数据库系统”。DBMS是一个强有力的工具，用于高效地创建和管理大量的数据，并使得数据能够安全地长期保存。这些系统是当前最复杂的软件类型之一。DBMS为用户提供的能力包括：

1) *持久存储*。与文件系统类似，DBMS支持对非常大量的数据进行存储，这些数据独立于使用数据的任何处理程序而存在。然而，DBMS在提供灵活性方面比文件系统做得更好，例如，它提供支持对非常大量的数据进行高效存取的数据结构。
2) *编程接口*。DBMS使得用户能够通过强有力的查询语言访问和修改数据。DBMS在数据操纵的灵活性方面也比文件系统强。与文件读写相比，DBMS提供的对存储的数据进行操作的方式要复杂得多。
3) *事务管理*。DBMS支持对数据的并发存取，即多个不同的进程（称作“事务”）同时对数据进行存取。为避免同时访问所造成的不良后果，DBMS支持*隔离*，即看起来事务是一次一个地在执行，以及*原子性*，即要求事务或者完全执行；或者完全不执行。DBMS还支持*回复性*(resilience)，即能够从多种类型的故障或错误中恢复的能力。

核心技术回顾

本书的读者应已学习过Ullman和Widom的《数据库系统入门教程》这一层次的数据库系统方面的内容，并了解SQL编程方面的知识。他们应该熟悉下列术语：

- *数据*：值得保留的任何信息，一般来说是电子形式的。
- *数据库*：为了访问和修改而组织的、在长时期内保留的数据的集合。
- *查询*：从数据库中抽取特定数据的操作。
- *关系*：将数据组织到二维表中的组织方式，表中的行（元组）表示基本的实体或某种事实，表中的列（属性）表示实体的特性。
- *模式*：数据库中数据结构的描述，通常称作“元数据”。

1.1 Megatron 2000数据库系统介绍

如果你使用过某个DBMS，或许是一个支持常见的SQL查询语言的DBMS，你可以想象实现这样

的一个系统并不困难。你心里可能会想到Megatron系统公司最近（虚构）提供的一个实现：Megatron 2000数据库管理系统。该系统在UNIX和其他操作系统上运行，采用关系方法，支持SQL语言。

1.1.1 Megatron 2000实现细节

首先，Megatron 2000 采用文件系统来存储它的关系。例如，关系`students(name, id, dept)`会存储在文件`/usr/db/students`中。对于每个元组，在文件`Students`中有一行。一个元组中各个分量的值存储成由特殊的标记字符`#`互相分开的字符串。例如，文件`/usr/db/students`可能像下面这样：

```
Smith#123#CS
Johnson#522#EE
  ...
```

数据库模式存储在特定的文件`/usr/db/schema`中。对于每一个关系，文件`schema`中有一个以该关系的名字起始的行，在该行中，属性名和类型交替出现。字符`#`作为行中元素间的分隔符。例如，文件`schema`中可能包含如下的行

```
Students#name#STR#id#INT#dept#STR
Depts#name#STR#office#STR
  ...
```

在这里对关系`Students(name, id, dept)`进行了描述；属性`name`和`dept`的类型是字符串，属性`id`的类型是整数。同时还有对模式为`Depts(name, office)`的关系的描述。

例1.1 下面是使用Megatron 2000 DBMS的一个对话的例子。我们运行在一台称作`dbhost`的机器上，通过UNIX层次的命令`megatron 2000`启动DBMS。

```
dbhost> megatron2000
```

产生如下响应

```
WELCOME TO MEGATRON 2000!
```

现在我们与Megatron 2000的用户界面对话，对于Megatron 提示符（&），我们可以健入SQL查询[⊖]作为响应。用#结束查询。例如，

```
& SELECT *
  FROM Students #
```

产生下表作为回答

name	*id*	*dept*
Smith	123	CS
Johnson	522	EE

Megatron 2000还允许我们执行一个查询，然后把结果存在一个新的文件中，做法是用一条竖线

⊖ 在1.4.2节中对于SQL有一个简单的回顾。