



中国计算机学会
学术著作丛书

模糊数据库系统

何新贵 著

4.1

1



清华大学出版社
广西科学技术出版社

模糊数据库系统

何新贵 著

清华大学出版社
广西科学技术出版社

(京)新登字 158 号

内 容 简 介

本书全面系统地论述了模糊数据库系统的理论和技术,内容分为三大部分:第一部分首先介绍与数据库有关的模糊性的一般理论,包括模糊集合论、模糊关系理论、模糊图论、模糊逻辑理论以及各种模糊数据的表示和运算等等;第二部分是模糊数据库概论,分别对模糊数据库系统的特性、各种模糊数据模型、模糊数据库语言、模糊数据库管理系统、模糊数据库的物理组织方法、模糊数据的完整性和安全性检查以及数据记录的存储方法等项具体实现技术作了较详细的论述;最后一部分介绍一些各具特色的模糊数据库系统,包括模糊专家数据库系统、主动模糊数据库系统以及模糊知识库系统等。特别在最后一章中,模糊知识库作为模糊数据库的推广和扩充,对模糊知识的表示和运用等重要问题进行了较详尽的讨论,对模糊数据库和模糊知识的结合等问题给出了一些解决方法。本书可作数据库研究人员的研究参考书,也可作为研究生和本专业高年级大学生的教材。

版权所有,翻印必究。

本书封面贴有清华大学出版社激光防伪标志,无标志者不得销售。

图书在版编目(CIP)数据

模糊数据库系统/何新贵著. —北京:清华大学出版社,1994
ISBN 7-302-01654-2

I. 模… II. 何… III. 数据库系统-模糊数据库 IV. TP311.13

中国版本图书馆 CIP 数据核字(94)第 12499 号

出版者:清华大学出版社(北京清华大学校内,邮编 100084)

广西科学技术出版社(南宁市河堤路 14 号,邮编 530021)

印刷者:国防工业出版社印刷厂

发行者:新华书店总店北京科技发行所

开 本:787×1092 1/16 印张:21.25 字数:504 千字

版 次:1994 年 12 月第 1 版 1994 年 12 月第 1 次印刷

书 号:ISBN 7-302-01654-2/TP·704

印 数:0001—4000

定 价:18.00 元

Fuzzy Database System

He Xingui

Abstract

The theory and techniques related to fuzzy database systems are introduced in the book systematically and completely. The contents of the book can be divided into three parts. In the first part, the general theory on fuzziness, including fuzzy sets, fuzzy relations, fuzzy graphs, fuzzy logic, and various fuzzy data, are introduced. The second part of the book presents the theory and techniques on fuzzy databases, which include the characteristics of fuzzy databases, various fuzzy data models, fuzzy database languages, fuzzy database management systems, organization and storage of fuzzy data in fuzzy databases, consistence, integrity and security, as well as some other techniques related to concrete implementations. In the last part, some fuzzy database systems with certain special characteristics are described. The active fuzzy database system is a combination of fuzzy database and active database. The fuzzy expert database and the fuzzy knowledge base are the generalization of the fuzzy databases in the field of Artificial Intelligence. Especially in the last chapter, representation and manipulation of fuzzy knowledge are discussed thoroughly, in the meantime some problems about how to combine fuzzy database with fuzzy knowledge base are discussed. The book can be used as a reference for theoretical researchers and engineers in the database area, as well as a textbook for graduates in the computer science department.

前 言

十年前,我开始对模糊数学在计算机中的实际应用发生兴趣,在模糊逻辑和模糊数据的处理等方面作了一点工作。不久前出版了我的专著《模糊知识处理的理论和技术》,这是我1990年出版的《知识处理和专家系统》一书的续篇,它将模糊数学和知识处理学结合起来,较全面地讨论了知识处理领域的模糊理论和技术。这本《模糊数据库系统》是我的第二本关于模糊理论的书,它把模糊数据库的理论和技术进行了系统化,内容体现了模糊数据库技术的最新进展,其中也包含了我自己近年来的一些工作。

为什么我们要如此热衷于“模糊”?我认为这完全是由于客观的需要。我们整天生活在一个充满模糊的世界中,对模糊世界的最贴切的认识最好就是模糊的,没有可能也没必要把它弄清楚。自盘古从一片混沌之中开劈了新天地以来,世界似乎是在不断地变得越来越“清楚”,人们正在越来越多地掌握各种客观规律,使越来越多的事物得到了定量的精确数学描述。但是,现实世界的大部分对我们来说还是那样模糊不清,仍有很多很多的事情实在没法精确描述。在我看来,宏观世界中的很多事物也许与海森伯格的关于微观世界的测不准原理类似永远是不可能用完全准确或精确的方法来描述清楚的。而且有些事情也许没有必要费大劲把它描述清楚。总之,如果不用模糊的方法许多客观事物似乎就不大可能十分贴切地反映其本来面貌。60年代中期,美籍伊朗数学家L. Zadeh提出的模糊集合论方法为用数学方法来描述现实世界的模糊性提供了一个途径,但较长一段时间内没有引起大家的注意,甚至还不时地受到一些清高的数学家们的奚落。这不能埋怨那些数学家们的无知,主要还应该怪模糊数学家自己,因为长期以来模糊数学家们花了其主要精力跟在纯数学家的屁股后面做这样和那样的推广工作,企图把纯数学家们做过的事在模糊数学领域再重复地做一遍,多年来都在苦苦地发展一套与纯数学平行的理论,证明一些与纯数学对应的定理,而把现实生活中许多活生生的实际应用问题撇在一边。近年来情况有了很大变化,模糊方法先是在一些不太起眼的应用中获得成功,后来发展到涉及模糊推理、自动控制、模式识别、模糊诊断、综合评判和决策支持等很多应用领域,特别在控制和决策等方面取得了较大的成功,而且越来越多的事例说明各种模糊技术的应用前景是十分广阔而美好的。所以,过去在模糊数学理论方面素质很高的中国人应该在发展和开拓各种应用领域的模糊理论和模糊技术中做出自己的贡献。十年前,我就是在这种责任感的驱使之下开始这方面工作的。惭愧的是做出的成绩太少,目前这些仅仅可算是开始迈出的前几步,今后的历程还很长。不过,在此成书之际,我要感谢我的历届学生们为实践我的许多思想而做的不懈努力和智慧奉献,感谢清华大学出版社对出版拙著的大力支持。

由于本书涉及的领域很新,一些理论或者实践很少或者还缺乏实践的验证,所以书中错误或不妥之处一定不少,诚恳希望同仁和读者们批评指正,以便今后改正和不断完善。

作 者

1993年冬于北京

清华大学出版社 广西科学技术出版社
计算机学术著作出版基金

评 审 委 员 会

主 任 委 员 张效祥

副主任委员 周远清 汪成为

委 员 王鼎兴 杨芙清 李三立 施伯乐 徐家福
夏培肃 董韫美 张兴强 徐培忠

出版说明

近年来,随着微电子和计算机技术渗透到各个技术领域,人类正在步入一个技术迅猛发展的新时期。这个新时期的主要标志是计算机和信息处理的广泛应用。计算机在改造传统产业,实现管理自动化,促进新兴产业的发展等方面都起着重要作用,它在现代化建设中的战略地位愈来愈明显。计算机科学与其它学科的交叉又产生了许多新学科,推动着科学技术向更广阔的领域发展,正在对人类社会产生深远的影响。

科学技术是第一生产力。计算机科学技术是我国高科技领域的一个重要方面。为了推动我国计算机科学及产业的发展,促进学术交流,使科研成果尽快转化为生产力,清华大学出版社与广西科学技术出版社联合设立了“计算机学术著作基金”,旨在支持和鼓励科技人员,撰写高水平的学术著作,以反映和推广我国在这一领域的最新成果。

计算机学术著作出版基金资助出版的著作范围包括:有重要理论价值或重要应用价值的学术专著;计算机学科前沿探索的论著;推动计算机技术及产业发展的专著;与计算机有关的交叉学科的论著;有较大应用价值的工具书;世界名著的优秀翻译作品。凡经作者本人申请,计算机学术著作出版基金评审委员会评审通过的著作,将由该基金资助出版,出版社将努力做好出版工作。

基金还支持两社列选的国家高科技重点图书和国家教委重点图书规划中计算机学科领域的学术著作的出版。为了做好选题工作,出版社特邀请中国计算机学会、中国中文信息学会帮助做好组织有关学术著作丛书的列选工作。

热诚希望得到广大计算机界同仁的支持和帮助。

清华大学出版社
广西科学技术出版社 计算机学术著作出版基金办公室

1992年4月

序 言

计算机是当代发展最为迅猛的科学技术,其应用几乎已深入到人类社会活动和生活的一切领域,大大提高了社会生产力,引起了经济结构、社会结构和生活方式的深刻变化和变革,是最为活跃的生产力之一。计算机本身在国际范围内已成为年产值达 2500 亿美元的巨大产业,国际竞争异常剧烈,预计到本世纪末将发展为世界第一大产业。计算机科技具有极大的综合性质,与众多科学技术相交叉而反过来又渗入更多的科学技术,促进它们的发展。计算机科技内容十分丰富,学科分支生长尤为迅速,日新月异,层出不穷。因此在我国计算机科技尚比较落后的情况下,加强计算机科技的传播实为当务之急。

中国计算机学会一直把出版图书刊物作为学术活动的重要内容之一。我国计算机专家学者通过科学实践,做出了大量成果,积累了丰富经验与学识。他们有撰写著作的很大积极性,但相当时期以来计算机学术著作由于印数不多,出版往往遇到不少困难,专业性越强越有深度的著作,出版难度越大。最近清华大学出版社与广西科学技术出版社为促进我国计算机科学技术及产业的发展,推动计算机科技著作的出版工作,特设立“计算机学术著作出版基金”,以支持我国计算机科技工作者撰写高水平的学术著作,并将资助出版的著作列为中国计算机学会的学术著作丛书。我们十分重视这件事,并已把它列为学会本届理事会的工作要点之一。我们希望这一系列丛书能对传播学术成果、交流学术思想、促进科技转化为生产力起到良好作用,能对我国计算机科技发展具有有益的导向意义,也希望我国广大学会会员和计算机科技工作者,包括海外工作和学习的神州学人们能积极投稿,出好这一系列丛书。

中国计算机学会

1992 年 4 月 20 日

目 录

第一部分 关于模糊性的一般理论

第 1 章 模糊集合和模糊关系	3
1.1 引言	3
1.2 模糊集合	3
1.2.1 特征函数与隶属函数	4
1.2.2 格与模糊集	5
1.3 模糊集抽象数据类型	6
1.3.1 一般模糊集合运算	7
1.3.2 λ -截集、支集和核	8
1.3.3 λ 限幅集、 λ 滤波集和 λ -集	9
1.3.4 交型运算和并型运算	10
1.3.5 模运算	11
1.3.6 代数运算	13
1.3.7 投影运算	14
1.4 凸模糊集和模糊数	14
1.5 模糊关系	15
1.5.1 模糊关系的合成运算	16
1.5.2 自反性、对称性和传递性	17
1.5.3 模糊相似关系和模糊等价关系	19
1.5.4 模糊变换	20
1.6 模糊向量	21
1.6.1 数乘和线性组合	22
1.6.2 内积、外积与叉积	22
1.7 贴近度、距离和择近原理	24
1.8 相似度	25
1.9 语义关联度	26
1.9.1 定义和用途	26
1.9.2 近似计算方法	28
1.10 模糊集的数字特征	30
1.10.1 模糊度	30
1.10.2 模糊熵	30
1.10.3 确定度	31
1.10.4 平均隶属度	31
1.10.5 重心或均值	32

1.10.6	方差	32
1.10.7	高阶矩	32
1.11	模糊集间的映射与扩张原理	33
1.12	隶属函数的产生	34
1.12.1	统计方法	34
1.12.2	选择比较法	36
1.13	模糊序列和模糊元组	37
1.14	广义模糊集	38
1.14.1	语言值	39
1.14.2	语言模糊集和语言值代数	39
1.14.3	格模糊集和偏序模糊集	40
1.14.4	高阶模糊集	41
1.14.5	广义模糊集	41
第2章	模糊数据和模糊运算	43
2.1	模糊数的表示	43
2.1.1	正规模糊数	43
2.1.2	模糊区间数	44
2.1.3	模糊中心数	44
2.1.4	集合模糊数	45
2.1.5	模糊分布数和正态模糊数	45
2.1.6	模糊相似数	46
2.2	模糊整数和模糊实数	46
2.3	模糊复数	47
2.4	模糊字符和模糊字符串	47
2.5	模糊布尔量	48
2.5.1	模糊数值布尔量	48
2.5.2	模糊区间布尔量	48
2.5.3	模糊语言值布尔量	49
2.6	模糊结构类型	49
2.6.1	模糊结构量	50
2.6.2	模糊元组	51
2.7	模糊算术运算	51
2.7.1	正规模糊数上的模糊算术运算	51
2.7.2	模糊中心数上的模糊算术运算	53
2.7.3	模糊区间数上的模糊算术运算	53
2.7.4	其它模糊数上的模糊算术运算	54
2.8	模糊关系运算	54
2.8.1	基于数字特征的模糊关系运算	54

2.8.2	基于集合运算的模糊关系运算	55
2.8.3	各种模糊数上的模糊关系运算	58
2.9	模糊字符串运算	58
2.9.1	拼接运算	58
2.9.2	λ 截和 λ 压缩运算	59
2.9.3	取子串运算	59
2.10	模糊逻辑运算	59
2.11	模糊表达式和模糊函数	60
2.11.1	模糊表达式	60
2.11.2	模糊函数	61
2.12	各种模糊数间的语义距离	61
2.12.1	正规模糊数间的语义距离	61
2.12.2	模糊区间数间的语义距离	62
2.12.3	模糊中心数间的语义距离	62
2.12.4	精确数与模糊数间的语义距离	62
第3章	模糊逻辑和统计逻辑	64
3.1	三值逻辑	64
3.1.1	Kleene 的强三值逻辑	65
3.1.2	Lukasiewicz 逻辑	66
3.1.3	Bochvar 逻辑	66
3.1.4	三值代数——计算三值逻辑	67
3.2	狭义模糊逻辑	69
3.3	模糊区间值逻辑	72
3.4	模糊语言值逻辑	73
3.5	广义模糊逻辑	75
3.5.1	I 型推理问题	77
3.5.2	II 型推理问题	78
3.5.3	III 型推理问题	79
3.6	模糊分布值逻辑	79
3.7	加权模糊逻辑	81
3.8	模糊计算逻辑	83
3.9	非单调模糊逻辑	85
3.10	概率逻辑	88
3.11	统计归纳逻辑	91
3.12	归纳与概念的形成	100
第4章	模糊图论	102
4.1	模糊图	102
4.2	有向模糊图	104

4.3	多重模糊图	105
4.4	模糊图上的操作	106
4.4.1	λ -截操作和 λ -截图	106
4.4.2	λ -过滤和 λ -限幅	106
4.4.3	模糊交、并操作	107
4.5	模糊路径和连通性	108
4.6	模糊树	109
4.7	模糊与/或树	109
4.8	模糊 Petri 网	111
4.9	H 网	113
4.10	模糊 H 网	116

第二部分 模糊数据库概论

第 5 章	模糊数据库的特性	121
5.1	数据的模糊性	121
5.2	数据间联系的模糊性	122
5.3	约束条件的模糊性	122
5.4	数据上操作的模糊性	123
5.5	查询语言的模糊性	124
5.6	子模式和用户视图的模糊性	124
5.7	模糊数据的冗余性	125
5.8	数据间依赖关系的模糊性	125
5.9	小结	126
第 6 章	模糊数据模型	127
6.1	引言	127
6.2	模糊关系数据模型	127
6.2.1	元组模糊的关系数据模型	127
6.2.2	模糊值关系数据模型	130
6.2.3	集合值关系数据模型	130
6.2.4	模糊值模糊关系数据模型	131
6.2.5	属性加权的模糊值模糊关系数据模型	132
6.2.6	模糊投影与模糊视图	134
6.2.7	元组间的语义距离	135
6.2.8	结语	136
6.3	模糊网络数据模型	137
6.3.1	数据结构与限制	137
6.3.2	操作与运算	138
6.3.3	模式与子模式	139

6.4	模糊层次数据模型	140
6.5	模糊实体-联系数据模型	140
6.5.1	数据结构与限制	141
6.5.2	操作与运算	142
6.5.3	模糊 E-R 图	142
6.6	面向对象的模糊数据模型	143
6.6.1	面向对象的方法学	143
6.6.2	模糊对象类的定义方式	144
6.6.3	模糊对象的构成方式	144
6.6.4	系统预定义的原子对象类	146
6.6.5	模糊结构对象类	147
6.6.6	一个具体的面向对象的模糊数据模型	148
6.7	模糊逻辑数据模型	153
6.7.1	数据库的外延形式与内涵形式	153
6.7.2	模糊谓词与模糊关系	153
6.7.3	模糊规则与模糊关系	154
6.7.4	模糊逻辑数据模型	156
6.8	模糊演绎数据模型	157
6.9	模糊知识库模型	158
6.10	结语	159
第 7 章	模糊数据库语言	160
7.1	引言	160
7.2	语言的模糊模型	160
7.3	模糊语法	161
7.3.1	0 型模糊语法	164
7.3.2	1 型模糊语法	164
7.3.3	2 型模糊语法	165
7.3.4	3 型模糊语法	165
7.4	模糊语义	166
7.4.1	单词	167
7.4.2	语言值	167
7.4.3	词组	167
7.4.4	语气算子	168
7.4.5	模糊化算子	168
7.4.6	模糊陈述句	169
7.4.7	模糊蕴含句	169
7.5	模糊程序设计语言	170
7.5.1	处理对象的模糊化	170

7.5.2	类型的模糊化	170
7.5.3	语句的模糊化	170
7.6	模糊关系型数据库语言	172
7.6.1	模糊数据定义语言 FDDL	172
7.6.2	模糊数据操作语言 FDML	175
7.7	面向对象的模糊数据库语言	177
7.7.1	面向对象的模糊关系型数据库语言	177
7.7.2	结语	179
第 8 章	模糊数据库管理系统	180
8.1	系统的功能	180
8.1.1	模糊数据的物理组织、存储和管理	180
8.1.2	模糊数据库语言的实现	180
8.1.3	缓冲区的调度和管理	181
8.1.4	模糊存取控制	181
8.1.5	模糊完整性检查	181
8.1.6	模糊一致性控制(并发控制)	181
8.1.7	安全性控制	182
8.1.8	模糊数据字典的维护	182
8.1.9	垃圾收集和数据库重组	183
8.1.10	事务处理	183
8.1.11	出错和例外处理	183
8.1.12	与其它软件和运行环境间的接口	183
8.1.13	应用支撑和服务功能	183
8.2	系统的结构	183
8.2.1	层次结构	184
8.2.2	模块结构	185
8.3	系统的接口	189
8.3.1	与操作系统的接口	189
8.3.2	与主语言的接口	189
8.3.3	与运行环境的接口	189
8.3.4	与应用程序和最终用户的接口	190
8.4	系统的实现问题	190
8.5	模糊演绎数据库和模糊演绎数据库管理系统	191
8.5.1	模糊演绎数据库管理系统	191
8.5.2	模糊 Prolog 和模糊数据库管理系统的结合	191
8.6	模糊数据库应用系统开发环境	192
8.6.1	模糊数据库设计工具	192
8.6.2	应用系统测试工具	193

8.6.3	模糊数据转换工具	193
8.7	结语	193
第9章	模糊数据库的物理组织	195
9.1	模糊数据的物理表示和存储	195
9.1.1	隶属函数的物理表示	195
9.1.2	模糊区间数和模糊中心数的物理表示	196
9.1.3	模糊分布数的处理	196
9.1.4	模糊相似数的物理表示	196
9.1.5	模糊字符串的物理表示	196
9.2	顺序文件	197
9.2.1	文件的组织	197
9.2.2	记录的插入和检索	197
9.3	散列文件	198
9.3.1	线性散列文件	199
9.3.2	随机散列文件	200
9.3.3	溢出散列文件	201
9.3.4	拉链散列文件	202
9.3.5	指标散列文件	205
9.3.6	散列码的产生方法	208
9.4	索引文件	209
9.4.1	索引顺序文件	209
9.4.2	索引无序文件	211
9.4.3	索引的组织	211
9.5	树结构文件	212
9.5.1	一般树结构	212
9.5.2	二叉树结构	213
9.5.3	二叉平衡树结构	213
9.5.4	B 树	214
9.5.5	B ⁺ 树	214
9.5.6	B* 树	215
9.5.7	带稠密索引的 B ⁺ 树	215
9.6	数据记录的存储方法	216
9.6.1	定长字段法	216
9.6.2	标明字段长度法	216
9.6.3	特殊字符分隔法	216
9.6.4	首部索引法	216
9.6.5	链接法	217
9.6.6	指针字段法	217

9.6.7 超文本存储法	218
第 10 章 其它实现技术	219
10.1 引言	219
10.2 安全性检查	219
10.3 并发控制	221
10.3.1 锁的种类	221
10.3.2 “两阶段锁”协议	222
10.3.3 死锁问题	222
10.4 完整性检查	222
10.5 查询优化	223
10.6 恢复	223
10.7 数据库语言的翻译	224
10.7.1 数据库语言的解释执行方式	224
10.7.2 数据库语言的编译执行方式	224
10.7.3 编译程序的组成和结构	224
10.7.4 状态矩阵法	227
10.7.5 递归子程序法	232

第三部分 模糊数据库系统

第 11 章 模糊专家数据库系统	239
11.1 系统的功能	239
11.2 系统的结构	239
11.3 模糊数据库	240
11.3.1 模糊数据的类型	240
11.3.2 模糊数据库语言	241
11.4 模糊规则库	245
11.4.1 模糊规则的形式	245
11.4.2 模糊规则库语言	246
11.5 应用	247
第 12 章 模糊主动数据库系统	248
12.1 引言	248
12.2 主动数据库系统	249
12.3 常用基本事件	250
12.4 事件代数	252
12.5 模糊的主动数据库系统	253
12.5.1 模糊的数据	254
12.5.2 模糊的事件	254
12.5.3 模糊的事件运算	255

12.5.4	模糊的事件驱动规则·····	255
12.6	模糊的主动数据库管理系统·····	256
12.6.1	模糊主动数据库管理系统的主动性功能·····	256
12.6.2	模糊主动数据库管理系统的结构·····	256
12.7	模糊事件监视器的实现·····	257
12.7.1	实现途径·····	257
12.7.2	实现模式·····	257
12.8	面向主动对象的模糊数据库系统·····	258
12.8.1	面向对象数据库的不足·····	258
12.8.2	主动的面向对象数据库模型·····	259
12.8.3	主动性和面向对象方法结合的优点·····	262
第 13 章	模糊知识库系统 ·····	264
13.1	引言·····	264
13.2	模糊知识的表示·····	264
13.2.1	定性概念和定性知识的表示·····	264
13.2.2	模糊关系·····	266
13.2.3	模糊逻辑表示·····	269
13.2.4	模糊产生式·····	269
13.2.5	模糊框架·····	272
13.2.6	模糊语义网络·····	281
13.2.7	模糊与/或语义图·····	284
13.2.8	模糊过程·····	290
13.2.9	混合的模糊知识表示模式·····	293
13.2.10	知识表示模式的选择和设计·····	294
13.3	模糊知识的运用·····	295
13.3.1	模糊关系的处理和运用·····	295
13.3.2	模糊逻辑推理·····	296
13.3.3	模糊产生式的处理和运用·····	301
13.3.4	模糊框架的处理和运用·····	304
13.3.5	模糊语义网络的处理和运用·····	307
13.3.6	模糊过程的处理和运用·····	310
13.4	模糊知识库·····	310
13.5	模糊推理机·····	311
13.6	模糊知识库管理系统·····	311
13.6.1	模糊知识库管理系统的功能·····	311
13.6.2	模糊知识库管理系统的结构·····	312
英中对照索引 ·····		314
参考文献 ·····		322