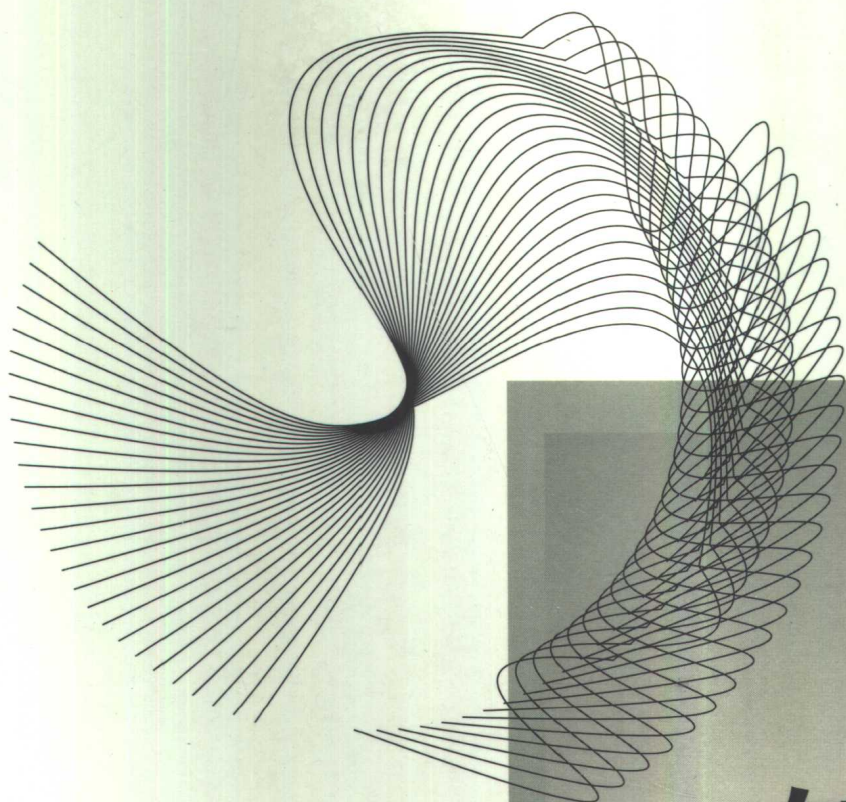


国家九五重点图书

计算机应用技术前沿丛书



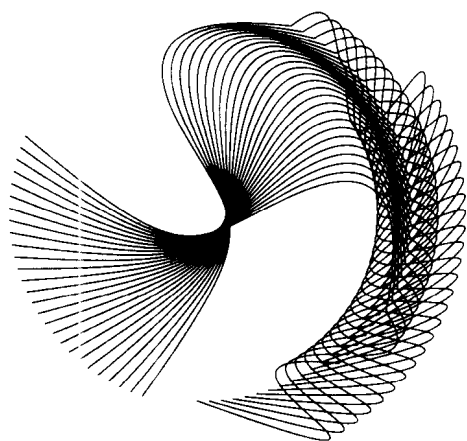
知识库系统原理

孔繁胜 编著

浙江大学出版社



国家九五重点图书



计算机应用技术前沿丛书

知识库系统原理

孔繁胜 编著



图书在版编目(CIP)数据

知识库系统原理 / 孔繁胜编著. —杭州:浙江大学出版社, 2000.9

(计算机应用技术前沿丛书/潘云鹤主编)

ISBN 7-308-02163-7

I. 知… II. 孔… III. 知识库系统-理论
IV. TP311.13

中国版本图书馆 CIP 数据核字 (1999) 第 46564 号

总 责 编 陈晓嘉 梁 兵

封面设计 俞亚彤

版式设计 孙海荣

责任绘图 张作梅

责任出版 李慧华

出版发行: 浙江大学出版社

(杭州浙大路 38 号 邮政编码 310027)

(网址: <http://www.zjupress.com>)

(E-mail: zupress@mail.hz.zj.cn)

责任编辑: 李玲如

排 版 者: 浙江大学出版社电脑排版中心

印 刷: 浙江印刷集团公司

经 销: 浙江省新华书店

开 本: 889mm × 1194mm 1/16

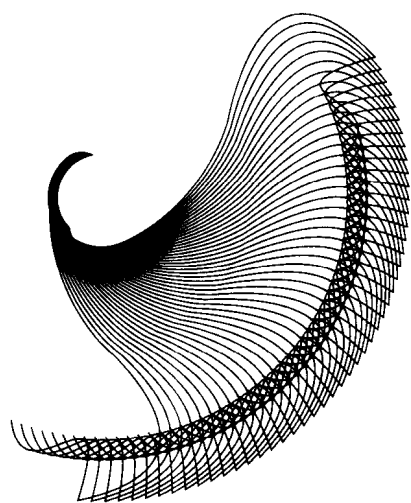
印 张: 15

字 数: 355 千字

版 印 次: 2000 年 9 月第 1 版第 1 次印刷

书 号: ISBN 7-308-02163-7/TP·183

定 价: 30.00 元



主 编 潘云鹤

副主编 陈 纯

顾 问(以姓氏笔划为序)

王 选 李三立 李国杰 李衍达

沈昌祥 张 钺 张效祥 汪成为

何志钧 杨芙清 金怡濂 陈火旺

胡启恒 唐泽圣 戴汝为



序

20 世纪是一个科技、经济空前发展的时代,从世纪初相对论、量子理论的创立到今天以信息产业为龙头的高科技产业成为经济发展的第一支柱,人类社会发生了根本性的变革。而在这场以科学技术为社会发展直接动因的变革中,意义最深远、影响最广泛的的就是计算机及其相关技术的发展和应用。

在过去的 50 年里,计算机已从最初的协助人类进行精密和复杂运算的单一功能的运算器发展成为能够模拟人类智慧、感觉,可以适应环境、具有多种功能的设备;计算机及其技术的应用已从尖端科学、国防工业延伸到了人类活动的各个领域。它不仅充实和革新了传统的科学与技术研究手段,带来了新的方法和理念,成为科学技术谋求新发展的沃壤,而且作为信息处理的核心工具和技术支撑起庞大的信息产业,进而改变了全球的产业格局。

目前,一个由遍布全世界的计算机连接起来的跨越国界,巨大而高速的多媒体信息网络正在形成,因此,不仅信息产业本身,而且各行各业都将面临计算机应用技术研究和发展的重大课题。我们针对国内外有关计算机应用技术的最新进展,根据浙江大学及国内其他学者在计算机辅助设计系统和方法、计算机动画、设计思维模拟、自主式智能系统、智能决策系统、软件智能化技术、工程数据库管理系统、图象和声音

压缩技术等方面取得的成果,组织编撰了《计算机应用技术前沿》丛书。我们希望,通过这套书,为从事相关领域研究工作的专家和学者提供一些有参考价值的方法和技术,为普及、推广进而开拓新的应用技术作出一点贡献。

潘云鹤

一九九九年十月



前言

“知识库系统”是人工智能的一个重要分支,是近年来发展起来的一门综合性的新兴学科。随着计算机科学技术和人工智能技术的迅速发展,基于知识的智能系统已在工业、农业、军事和科教等国民经济领域得到广泛应用;与此同时,人们对智能系统的要求和期望也越来越高,迫切希望它们能有效地解决更加复杂的问题。实践证明,任何一个智能系统的行为都离不开知识,如何合理地存贮、组织、管理、应用和共享知识是所有智能系统深入发展所面临的共同问题。“知识库系统”正是在这种背景下发展起来的。一方面,知识库系统的技术集成了知识表示、知识获取和问题求解等人工智能重要领域的最新成果;另一方面,知识库系统又与数据库的最新技术息息相关。可以说,知识库系统是人工智能和数据库技术发展的结晶。因此,知识库系统的研究始终受到人们的高度重视,我国 863 高科技计划始终把知识库系统的研究作为智能计算机主题的重要部分;美国国防部 DARPA(Defence Advanced Research Project Agency)自 1997 年 6 月开始组织实施 HPKB(High Performance Knowledge Base)重大研究项目,参加该项目研究的有 Stanford, MIT, CMU 全美十余所著名的大学和研究机构。

本书以知识库系统的发展为主线,结合知识处理和数据库技术的最新成果,全面地介绍了知识库系统的基础理论、技术进展、现有成就

和今后的发展趋势。本书第1章是知识库系统的概述,简要介绍了知识和知识库的定义、知识库系统的结构和相关的技术,使读者对本书有一个全面的了解。一阶谓词逻辑语言是最基本的知识表示工具,也是逻辑程序设计的基础,第2章在扼要阐述一阶谓词逻辑原理的基础上介绍了近十余年发展起来的非经典逻辑及其在知识库系统中的应用。关系和面向对象技术是目前数据库技术的主流,也是知识库系统和数据库技术的结合点之一,在研究大容量知识库系统中尤为重要。本书第3,4章分别介绍了知识的关系表示和面向对象的知识表示方法。第3章中重点讨论了二元关系知识库,因为这是进行知识库一致性检查的成功范例。在第4章中介绍了基于框架和基于案例的知识库系统,这是因为我们认为面向对象技术和框架的概念尽管来自不同的领域,强调不同的特征,但它们的本质是相同的。第5章讨论了目前应用最广泛的基于规则的知识表示方法,其中在规则中使用变量和 Rete 算法等内容都是近年来的研究成果。第6章主要介绍了近年来研究并得到采用的基于模型的知识表示方法,基于模型的知识表示方法在表示结构化知识和解决复杂的问题中十分重要。不确定知识的表示与推理是知识库系统的关键技术之一,也是近年来研究的热点,第7章叙述了6种重要的处理不确定性知识的理论和方法,并比较了它们各自的得失。

知识库的作用基于它具有丰富的、不断更新的知识。如何采用各种技术,从不同途径获取知识已成为近年来人工智能和知识库系统的热门研究领域。第8章详细阐述了这一领域的最新研究成果,如 KDD (基于数据库的知识发现)等。本书第9章——高性能知识库系统——HPKB,以美国军方正在研制的大型知识库为背景,全面分析了当前知识库系统的三大主要研究目标:

1. 大容量、可重用、分布式知识库的建立;
2. 领域知识的获取;
3. 有效的问题求解能力。

为实现上述目标,HPKB 采用并集成了多项最新技术:

1. 基于本体的知识表示；
2. 大容量本体知识库；
3. 分布式本体服务器；
4. 大容量知识库的推理技术；
5. 知识多视角建模；
6. 大容量知识库的一致性检查；
7. 基于自然语言的知识获取技术；
8. 从文本、WWW 和知识库中抽提和集成知识等。

通过上述内容的介绍使读者对知识库系统的发展和趋势有进一步的了解。

目 录

1 知识库系统概述	1
1.1 引言	1
1.2 知识及其特性	1
1.2.1 数据和信息	2
1.2.2 知识的定义	2
1.2.3 知识的层次	3
1.2.4 知识的类型	3
1.3 知识表示	4
1.3.1 自然语言	4
1.3.2 知识的形式化表示	4
1.3.3 不确定性知识的表示	5
1.4 知识的运用	6
1.4.1 知识的搜索技术	6
1.4.2 知识的推理方法	7
1.5 知识库和知识库系统	10
1.5.1 定义	10
1.5.2 知识库的推理和维护	11
1.6 知识库系统构造技术	12

1.6.1	知识表示模式	12
1.6.2	问题求解方法	13
1.6.3	知识获取	14
1.7	小结	16
2	知识的逻辑表达	17
2.1	关于推理的简要讨论	17
2.1.1	命题和断言	17
2.1.2	一致性检查	18
2.1.3	演绎	19
2.1.4	形式化方法	19
2.2	一阶谓词逻辑	20
2.2.1	引言	20
2.2.2	一阶谓词逻辑语言的句法	20
2.2.3	一阶谓词逻辑语言的语义	22
2.2.4	命题逻辑中的归结推理方法	28
2.3	非经典逻辑在知识处理中的应用	31
2.3.1	引言	31
2.3.2	多分类逻辑	32
2.3.3	模态逻辑	37
2.3.4	非单调逻辑	50
2.3.5	时序逻辑	53
2.3.6	模糊逻辑	57
2.4	小结	60
3	基于关系的知识表示	62
3.1	知识的关系表示	62
3.1.1	关系结构	62
3.1.2	关系	63
3.1.3	关系模型	64

3.1.4	关系代数	66
3.1.5	关系计算	70
3.2	二元关系知识库	71
3.2.1	系统的描述	71
3.2.2	语义一致性检查	72
3.2.3	存在的问题	75
4	面向对象的知识表示	76
4.1	对象类及其特性	76
4.1.1	对象类的定义	76
4.1.2	继承性	78
4.1.3	封装性	81
4.1.4	动态联编	81
4.1.5	消息发送	83
4.1.6	对象类和例示的关系	83
4.2	基于框架的知识表示	84
4.2.1	基本概念	84
4.2.2	基于框架的推理	85
4.3	基于案例的推理——CBR	86
4.3.1	CBR 循环	86
4.3.2	案例的检索	87
4.3.3	案例的修改	89
4.3.4	错误结论的处理	90
4.3.5	基于案例的推理和学习系统——CREEK	90
4.4	小结	96
5	基于规则的知识表示	98
5.1	产生式和产生式系统	98
5.2	基于规则的知识表示	99
5.2.1	基于规则的系统	99

5.2.2	基于规则系统的结构	99
5.2.3	基于规则系统的特点	100
5.2.4	基于规则系统的实例	100
5.3	规则和知识库	102
5.3.1	一致性维护	102
5.3.2	封闭性世界假定	103
5.3.3	在规则中使用变量	103
5.4	数据驱动和目标驱动求解方法	104
5.4.1	数据驱动方法	104
5.4.2	规则中变量的例示	106
5.4.3	Rete 算法	107
5.4.4	目标驱动求解方法	108
5.4.5	混合方法	109
5.5	搜索策略	110
5.5.1	冲突解决方案	110
5.5.2	深度优先和宽度优先搜索	111
5.5.3	启发式搜索	112
5.5.4	规则关系网络	112
5.6	解释机制	114
5.7	小结	115
6	基于模型的知识表示	116
6.1	深层模型和浅层模型	116
6.1.1	浅层知识模型	116
6.1.2	深层知识模型	117
6.1.3	混合知识模型	118
6.2	基于模型的知识表示	119
6.2.1	规则的局限性	119
6.2.2	功能、结构和状态建模	119
6.2.3	模型的使用	121

6.2.4	监控	123
6.2.5	试探性诊断	123
6.2.6	故障模拟	124
6.2.7	使用问题树	125
6.3	小结	125
7	不确定知识的表示与推理	127
7.1	引言	127
7.2	概率论	128
7.2.1	基本概念	128
7.2.2	贝叶斯(Bayes)规则	128
7.2.3	贝叶斯规则的另一种表示方法	129
7.2.4	一个实例	129
7.2.5	概率论处理不确定性的缺陷	130
7.3	置信度理论	131
7.3.1	引言	131
7.3.2	条件部分置信度值小于 1.0 的规则	132
7.3.3	复合条件	132
7.3.4	实例	132
7.3.5	置信度理论和概率论的相互关系	134
7.3.6	置信度理论的优点	134
7.4	证据理论	134
7.4.1	命题和集合	135
7.4.2	基本概率赋值函数	135
7.4.3	信任函数	136
7.4.4	似然函数和信任区间	136
7.4.5	组合概率赋值函数和信任函数	137
7.4.6	应用实例	138
7.5	可能性理论	140
7.6	发生率计算	143

7.6.1	纯数值机制的限制	144
7.6.2	事件	144
7.6.3	发生率计算	145
7.6.4	加权概率	145
7.6.5	事件的表示	146
7.6.6	事件的使用	147
7.7	INFERNO	147
7.7.1	引言	148
7.7.2	INFERNO 中的不确定性的传播	149
8	知识获取	152
8.1	知识获取及其过程	152
8.2	知识分析	153
8.3	知识获取中的模型	154
8.4	自动知识获取	155
8.4.1	归纳法符号学习	155
8.4.2	优化算法	160
8.4.3	人工神经网络	162
8.5	基于数据库的知识发现-KDD	173
8.5.1	引言	173
8.5.2	用数据库作为知识源	174
8.5.3	知识发现的过程	175
8.5.4	知识发现的方法	179
8.5.5	AKS 原理及总体结构	181
8.6	小结	187
9	高性能知识库系统——HPKB	189
9.1	HPKB 概述	189
9.1.1	开发 HPKB 的目的	189
9.1.2	HPKB 的关键技术	190

9.1.3	HPKB 的任务和功能模块	190
9.2	本体知识库	192
9.2.1	什么是本体	192
9.2.2	本体知识表示语言 MELD	193
9.2.3	HPKB 本体知识库的层次结构	203
9.2.4	本体知识库推理机制	209
9.3	HPKB 的知识获取	211
9.3.1	知识获取工具 HAWK	211
9.3.2	从领域本体中生成获取知识的工具	216
9.4	小结	219



1 知识库系统概述

“知识库系统”是人工智能的一个重要分支,是近年来发展起来的一门综合性的新兴学科。本章简要介绍知识和知识库的定义、知识库系统的结构和相关的技术,使读者对本书有一个全面的了解。

1.1 引言

进入 90 年代以来,随着计算机和人工智能技术的迅速发展,大型智能系统或基于知识系统(knowledge-based system)的研制开发已成为计算机应用的研究热点之一。目前,智能系统已开始应用于工业、农业、军事等国民经济的每一部门,其涉及的领域包括了设计、规划、决策、诊断和控制等各个方面。与此同时,人们对智能系统提出了更高的要求,迫切要求它们能更有效地解决更复杂的问题,代替人完成各种困难的任务。一个公认的事实是:智能系统解决问题的能力在很大程度上取决于它们拥有知识的质量和数量。因此,如何有效地利用存贮、管理、组织、维护和更新大规模的知识,如何有效地利用存贮的知识进行推理和问题求解,即知识库系统技术便应运而生,并成为人工智能技术的重要分支。应当看到,知识库系统技术是一门综合性的学科。它将计算机技术(大容量的存贮器)、数据库技术、网络技术和人工智能技术的最新成果集于一体,而知识库技术的发展也将进一步推动这些相关技术的深入发展。

1.2 知识及其特性

本节从数据、信息和知识的基本概念着手,介绍了知识的定义、知识的层次和知识的类型。