



中国计算机学会学术著作丛书
——知识科学系列 8

非规范知识处理的 基本理论和核心技术

王 珊 刘椿年 主编

清华大学出版社

中国计算机学会学术著作丛书——知识科学系列

非规范知识处理的基本 理论和核心技术

王 珊 刘椿年 主编

清华大学出版社
北 京

内 容 简 介

所谓非规范知识是对不确定的、不完整的、不精确的、非恒常的、不一致的等内涵难处理的知识的总称。本书是国家自然科学基金重大项目“非规范知识处理的基本理论和核心技术”的研究成果,由项目研究概述(总结篇)、非规范知识的基本理论(理论篇)和非规范知识处理的核心技术(系统篇)三大部分组成。其中,总结篇全面总结了项目的研究成果,回顾和分析了研究工作的体会以及得失;理论篇涵盖了非规范知识的数学理论、非规范知识的认知理论以及非规范知识处理的逻辑理论;系统篇主要讨论特定领域的非规范知识处理的核心技术、常识知识及领域本体的构建与应用、语义网和 Web 搜索的研究和应用等内容。

本书适合计算机及自动化专业知识工程和人工智能领域的研究生、教师、工程技术人员和科研人员参考。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

非规范知识处理的基本理论和核心技术/王珊,刘椿年主编.--北京:清华大学出版社,2011.10

(中国计算机学会学术著作丛书.知识科学系列)

ISBN 978-7-302-25334-1

I. ①非… II. ①王… ②刘… III. ①知识工程—研究 IV. ①TP182

中国版本图书馆 CIP 数据核字(2011)第 067997 号

责任编辑:薛 慧

责任校对:赵丽敏

责任印制:

出版发行:清华大学出版社

地 址:北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者:

装 订 者:

经 销:全国新华书店

开 本:185×230 印 张:27.5

字 数:596 千字

版 次:2011 年 10 月第 1 版

印 次:2011 年 10 月第 1 次印刷

印 数:1~ 000

定 价: .00 元

产品编号:

评审委员会

名誉主任委员：张效祥

主任委员：唐泽圣

副主任委员：陆汝铃

委员：（以姓氏笔画为序）

王 珊	吕 建	李晓明
林惠民	罗军舟	郑纬民
施伯乐	焦金生	谭铁牛

“知识科学系列”丛书序

2002 年秋天,由王珏教授策划和组织,复旦大学智能信息处理开放实验室(即现在的上海市智能信息处理重点实验室)举办了一次“机器学习及其应用”研讨会。该研讨会属于实验室的“智能信息处理系列研讨会”之一。十余位学者在综述机器学习各个分支的发展的同时报告了他们自己的成果。鉴于研讨会取得了非常好的效果,而机器学习领域又是如此之广阔,有那么多重要的问题还没有涉及或还没有深入,2004 年秋天王珏教授又和周志华教授联合发起并组织第二届“机器学习及其应用”研讨会,仍由复旦大学实验室举办。这次研讨会又取得了非常好的效果,并且参加的学者比上次更多,报告的内容也更丰富。根据与会者的意见,决定把报告及相关内容编成一本书出版,以便与广大的国内学者共享研讨会的成果。

机器学习是人工智能研究的核心课题之一,不但有深刻的理论内蕴,也是现代社会中人们获取和处理知识的重要技术来源。它的活力久盛不衰,并且日呈燎原之势。对此,国内已经有多种定期和不定期的学术活动。本书的出版反映了机器学习界一种新型的华山论剑:小范围、全视角、更专业、更深入,可与大、中型机器学习会议互相补充。值得赞扬的是,它没有任何学派和门户之见,无论是强调基础的“气宗”,还是注重技术的“剑宗”,都能在这里畅所欲言,自由交流。我很高兴地获悉:第三届“机器学习及其应用”研讨会已经在 2005 年 11 月由周志华教授和王珏教授主持在南京大学成功举行。并且以后还将有第四届、第五届……作为一直跟踪这项活动并从中获得许多教益的一个学习者,我真希望它发展成这个领域的一个品牌,希望机器学习的优秀成果不断地由这里飞出,飞向全世界。

值得一提的是王珏教授有一篇颇具特色的综述文章为本书开道。长期以来,许多有识之士为国内学术界缺少热烈的争鸣风气而不安。因为没有争鸣就没有学术繁荣。细心的读者可以看出,这篇综述的观点并非都是传统观点的翻版,并且很可能不是所有的同行都认同的。作者深刻反思了机器学习这门学科诞生以来走过的道路,对一些被行内人士几乎认作定论的观点摆出了自己的不同看法。其目的不是想推出一段惊世骇俗的宏论,而是为了寻求真理、辨明是非。在这个意义上,王珏教授也可算是一位“独孤求败”。如果有人能用充分的论据指出其中可能存在的瑕疵,他也许会听到一片鼓掌之声更感到宽慰。

随着本书的出版,中国计算机学会丛书知识科学系列也正式挂牌了。在衷心庆贺这个系列诞生的同时,我想重复过去说过的一段话:“二十多年来,知识工程主要是一门实验性

科学。知识处理的大量理论性问题尚待解决。我们认为对知识的研究应该是一门具有坚实理论基础的科学,应该把知识工程的概念上升为知识科学。知识科学的进步将从根本上回答在知识工程中遇到过,但是没有很好解决的一系列重大问题”。本系列为有关领域的学者提供了一个宽松的论坛。衷心感谢王珏、周志华、周傲英三位编者把这本精彩的文集贡献给知识科学系列的首发式。我相信今后机器学习著作仍将是这个系列的一个常客。据悉,第四届机器学习研讨会将于今秋在南大举行,届时各种观点又将有进一步的发展和碰撞。欲知争鸣烽火如何再燃,独孤如何锐意求败,且看本系列下回分解。

陆汝铃

2006 年 1 月

序

第一台电子计算机诞生于 20 世纪 40 年代。到目前为止,计算机的发展已远远超出了其创始者的想象。计算机的处理能力越来越强,应用面越来越广,应用领域也从单纯的科学计算渗透到社会生活的方方面面:从工业、国防、医疗、教育、娱乐直至人们的日常生活,计算机的影响可谓无处不在。

计算机之所以能取得上述地位并成为全球最具活力的产业,原因在于其高速的计算能力、庞大的存储能力以及友好灵活的用户界面。而这些新技术及其应用有赖于研究人员多年不懈的努力。学术研究是应用研究的基础,也是技术发展的动力。

自 1992 年起,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与产业的发展,推动计算机科技著作的出版,设立了“计算机学术著作出版基金”,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。时至今日,本套丛书已出版学术专著近 50 种,产生了很好的社会影响,有的专著具有很高的学术水平,有的则奠定了一类学术研究的基础。中国计算机学会一直将学术著作的出版作为学会的一项主要工作。本届理事会将秉承这一传统,继续大力支持本套丛书的出版,鼓励科技工作者写出更多的优秀学术著作,多出好书,多出精品,为提高我国的知识创新和技术创新能力,促进计算机科学技术的发展和进步作出更大的贡献。

中国计算机学会

2002 年 6 月 14 日

知识实力是国家的软实力(代序)

在科学上,基础研究往往被一些决策者认为是“远水解不了近渴”。但事实并不是这样。基础研究对人类社会影响极大,甚至能改变历史的发展进程。1944—1945年,“二战”进入关键时刻,交战双方都在全力准备最后一搏。结果大家是知道的:同盟国取得了最后胜利。世人也熟知原子弹的出现和使用在其中起到了重要作用。据说,根据美国计算,如果没有原子弹,最后登陆日本本岛将要付出百万伤亡的惨重代价。然而,一些历史资料的透露也使我们捏了一把冷汗。法西斯德国和日本均早已开始原子弹的研究。尤其是德国,拥有像海森伯这样的世界一流物理学家。如果不是希特勒的决策失误,如果不是德国科学家走了一些弯路,如果不是同盟国采取了坚决的破坏措施,一旦希特勒德国首先研究出了原子弹,则人类社会的灾难将难以想象。须知向美国总统建议研制原子弹的是从德国逃往美国的爱因斯坦,须知首先发现原子核裂变现象的是德国的科学家,还须知德国法西斯上台以后,在反犹的口号下驱赶了大批的优秀科学家,其中就有许多著名的物理学家。这不仅是“为丛驱雀、为渊驱鱼”,而且是向敌方输送关键的知识财富,我们为法西斯政权做了大蠢事而感到庆幸。

如今,世界大战的硝烟已经远去半个多世纪,国家之间的较量主要转入经济层面。强国一般不会轻易地采用军事入侵手段。经济掠夺更为重要和有效,还不用担负道义的名。而且经济掠夺也不需要像开拓殖民地时期那样采取赤裸裸的方式。各种各样的知识产权就是强国手中的看家武器。创造了“高通神话”的美国高通公司,不需要开动一部机器,每年靠着与CDMA一类技术有关的专利,坐收百亿美元的“渔利”。我国企业生产的DVD机,国外市场销售红火,然而绝大部分利润被拥有专利的国外公司收走,生产厂家每台DVD机获益不到一美元,服服帖帖地为他人作嫁衣。老一辈的人有所谓“鬼打墙”的说法:夜晚,人在野外走路,到处碰壁,始终在原地打圈,怎么也走不出去,就说是鬼在捣乱。这当然只是一种想象。但是现实生活中却真的可以看到“鬼打墙”。垄断企业把某个领域的专利组合起来,形成所谓专利池。创业者一旦进入,任你左冲右突,也难以突破这道专利墙。无论你想从哪里突围,总有一款专利挡着你。

然而,仅仅是专利之类的知识产权还不能完全说明问题。专利是凝固的创新,不是进行中的创新,更不是点燃创新的火种。专利意味着今天的财富,并不意味着明天的辉煌。为了可持续发展,必须要有持续的创新。具有在关键领域持续创新能力的国家是具有知识

实力的国家。一个国家的实力体现在很多方面。经济实力、国防实力和科技实力是其硬的方面,而知识实力和文化实力则是其软的方面。知识实力不等于拥有很多知识。爱因斯坦说过:想象力比知识更重要。知识实力也不等于科技实力,但是两者之间有着紧密的联系。知识实力要经过艰苦转化才能成为科技实力,而科技实力没有知识实力的推动就会停滞不前。许多有识之士重视教育实际上也是重视知识实力的体现。现在大家都承认,一个重视经济实力的国策必定也是重视科技实力的。但是我们还要指出:一个重视科技实力的国策应该也是重视知识实力的。

知识实力对一个国家的强盛极其重要。在知识实力上远远超过其他国家的强国拥有的是制知识权。你有的知识,我都有;你没有的,我也有。你今天有的,我早就有了。我现在有的,你何时能有还遥遥无期。而且我会采取一切手段不让你有,最好是你永远没有,至少是等我遥遥领先了你才有。你需要的,我可以给你,也可以不给你。你拥有之后会不利于我的,我绝对不给。我愿意给你的,你必须付出让我满意的高价。这就是目前世界上国与国之间的行事规则。不服也得服。说什么制空权、制海权、制信息权,什么也离不开制知识权。

要掌控制知识权(掌控-制知识权)必须从积累知识实力开始。各行各业都有自己的知识领域,而知识实力研究关心的是超越于各行各业之上的元知识。研究也有不同的角度。从政府的角度来研究是培养国家知识实力的政策问题,而从科学工作者来研究则是处理各类知识的一般性原理、理论和技术。表面上看这是两条平行的直线。实际上并不完全平行。它们会相交的。一系列的“新潮”概念吸引着世界各国政府和学者双方的注意。知识社会、知识经济、知识产业、知识企业、知识管理、知识工人、知识服务等等的理论和实践,过去总觉得离我们那么遥远,而如今看来又和我们这样地贴近。但是近了不等于就抓住了。种种问题至今还令我们困惑。我们什么时候才能为知识产业正名,并且列之为第四产业?我们什么时候才能把知识服务提高到国策的水平,并且列于各类服务之首?我们什么时候才能使遍布全国的网吧不再是培养“问题少年”的温床,而是成为向青少年普及知识的校外课堂?最新的统计表明我国的网民已超过4.5亿,其中又有多少是在网上学习知识的呢?

当然,强调软实力不是为了贬低硬实力。经济实力、国防实力和科技实力还是非常重要的。没有硬实力,软实力也难以发挥作用,难以迅速转化为更强的硬实力。我们只是想说:知识实力不是万能的,但是没有知识实力是万万不能的。

我们是从科学和技术的角度研究知识和知识处理的。我们有义务关心和推动国家知识实力的增长。位卑未敢忘忧国。值此国家自然科学基金重大项目“非规范知识处理研究”顺利结题并出版文集之际,写下以上这些话以体现“国家兴亡,匹夫有责”。由于我们工作在这个领域,如果我们不说,就是失职了。

陆汝铃

庚寅年岁末

编者的话

《非规范知识处理的基本理论与核心技术》作为同名国家自然科学基金重大研究课题(课题编号 60496320-60496327)的成果汇编,终于由清华大学出版社出版了。

所谓非规范知识是对不确定的、模糊的、不完整的、不精确的、非恒常的、不一致的等等内涵难处理的知识的总称。因此,本课题的研究对象是非常宽泛的,这也给项目的组织和实施,以及研究力量的聚焦都带来很大的挑战。但是,从另一个角度讲,在还不清楚这些不规范知识之间存在何种关联的情况下,就人为地割裂它们,对于探索和建立非规范知识处理的基本理论并没有好处。

正是由于这样的原因,本论文集并没有在论文的组织上花太多的精力,除第一篇文章外,只是将其余论文粗略地划分为非规范知识的基本理论和非规范知识处理的核心技术两大部分。这也客观地反映了本课题的组织思路。

在“非规范知识处理的基本理论和核心技术——重大项目研究概述”一文中,陆汝钤作为项目负责人总结了国家自然科学基金重大项目“非规范知识的基本理论和核心技术”的研究成果,并回顾和分析了研究工作的体会,以及得失。这是本论文集中比较全面反映本项目研究进展的部分。将其独立出来作为本论文集的第一部分是合适的。作为一个整体本项目获得了评审专家的一致肯定,评为“特优”。陆汝钤的文章可以帮助读者了解本项目的主要成果。

理论篇包含 14 篇论文。可分为三组。

第一组是非规范知识的数学理论。包括 3 篇论文。陆汝钤在“知识、范畴论、结构主义”中回顾了历史上一些专家对知识本质的分析,提出知识的本质是结构化的信息,最适于表述知识结构性的数学工具是范畴论。提出带类型的范畴论作为知识的结构性理论的数学基础,并给出了一些初步的定义、结论和例子。最后回顾了结构主义作为一种思潮在历史上的影响,并指出了结构主义和功能主义的互补性。柳欣欣在“Combining λ and π Through Continuation”中将 CCS 语法和 λ 算子组合,得到了一种无类型的一般传值过程算子,该算子集中了 λ 算子、CCS 和 π 算子的有用属性。应明生在“通信并发系统中的拓扑结构、随机性与噪声”中较为系统地研究了进程演算中的拓扑、概率进程演算及带噪声的 π 演算,以刻画通信并发系统的近似行为、随机性、噪声等复杂特征,对这方面的工作进行了总结。

第二组是非规范知识的认知理论,包括 6 篇论文,主要讨论非规范知识表达的“云模

型”、“Agent 模型”以及“信息转换模型”等。李德毅等人在“云模型研究进展”中认为人类认知中的概念也客观存在随机性,云模型以概率理论为基础,利用认知物理学方法,刻画人类认知中定性概念与定量实现之间的转换。在“网络化数据挖掘与群体智能”中同一作者群将云模型应用于复杂网络数据挖掘中。基于云模型和数据场等学术思想,提出了“拓扑势”概念,将网络节点质量和边属性(如流量、带宽)等物理世界中的重要特征映射到网络模型中,用来刻画节点主体性及相互间影响的局域性。李三江在“Multi-agent Coordination Using Nearest-neighbor Rules: Revisiting the Vicsek Model”中研究了利用最近邻规则的多 Agent 协调问题,证明了在比前人结果相对更弱的条件下,所有移动自治 Agent 将同向移动。张伟等人在“多 Agent 系统合作求解机制”中综述了 Agent 合作求解机制并讨论了联盟、协商、拍卖等多种合作求解形式,介绍了课题组主要研究工作和贡献。苏开乐等在“模型检测多智能体系统”中提出了一种实用有效的完全自动的模型检测算法 CKLn,给定一个多智能体的逻辑模型 M 的语义结构和用户所关心的逻辑规范 α ,该算法能测试 $M \models \alpha$ 是否成立。钟义信在“‘信息-知识-智能’生态链意义下的知识问题研究”中从信息、知识和智能三者之间的内在联系及其所形成的“生态链”视角,阐述了知识的内涵和度量,借此揭示信息、知识和智能转换(简称为“信息转换”)的机制,在此基础上阐明“信息转换”的重要学术意义。

第三组是非规范知识处理的逻辑理论,包括 5 篇论文。陈仪香在“近似推理与证据推理”中阐述了关于模糊推理的王国俊的 3I 推理机制,并介绍了非独立证据推理、突变证据推理以及模糊可能推理。林作铨等人在“语义 Web 矛盾知识的处理”中指出标准语义 Web 语言基于二值描述逻辑,不能处理矛盾。论文基于二值描述逻辑 $\mathcal{ALC}$ 提出具有含矛盾推理能力的四值描述逻辑 $\mathcal{ALC}_4$ 。 $\mathcal{ALC}_4$ 提供了一种能够使用经典推理系统来处理语义 Web 中矛盾知识的方法。刘椿年等人在“约束(归纳)逻辑程序设计中不确定知识的处理”中就约束逻辑程序设计 CLP 和约束归纳逻辑程序设计 CILP 中如何处理不确定知识的问题展开讨论。刘大有等人在“时空知识表示与推理”中指出“时空知识”是“知识”的重要组成部分,总结了时空推理的主要热点研究内容。朱朝晖在“择优型非单调逻辑简介”中就非单调后承、语义模型及表示定理等方面对择优型非单调逻辑做了介绍。

论文集的第三部分是系统篇。共有 11 篇论文。大致也可以分为三组。

第一组是特定领域的非规范知识处理的核心技术,试图结合特定领域的应用背景深化非规范知识处理的研究,包括生物信息学领域、软件需求管理领域、金融市场投资、可信任网络及数据流挖掘等。共有 6 篇文章。卜东波等在“非编码 RNA 及其参与的复杂生物网络”中介绍了非编码 RNA 功能研究、复杂生物网络分析以及非编码 RNA 参与的生物网络的构建等工作。王飞等在“调控网络和代谢网络集成分析综述”中综述了调控网络和代谢网络集成分析方面的研究工作。周志华在“二次学习”中介绍了寻求既具有强泛化能力又具有良好可理解性的“二次学习”(twice-learning)方法,并讨论了二次学习在生物信息学方面的应用。金芝等在“非规范软件需求管理的逻辑方法和管理框架”中系统介绍了所提出

的非规范软件需求处理的一种解决方案,包括非规范软件需求处理过程、软件需求非规范性的统一处理框架、软件需求矛盾性的度量方法以及软件需求不一致归结策略。邓小铁等在“算法交易及金融市场的套利分析”中针对算法交易这一实际应用,从市场均衡有效的角度出发,对市场的套利模型以及套利分析的相关进展给予了介绍,并就若干新的套利方法以及在算法交易策略研究中的应用进行了探讨,可为后金融危机时期的金融市场的投资决策给出现实的指导建议,对维护金融市场健康稳定发展具有一定的积极意义。毛国君等的论文“数据流挖掘研究现状与关键问题”讨论了数据流挖掘的研究现状和发展趋势。

第二组论文讨论常识知识及领域本体的构建与应用,包括2篇文章。曹存根等在“社会角色心理常识的获取方法”中对社会角色及其相关概念进行了仔细的分析和定义,并抽取出角色的本体特征,提出了一种在特定社会领域建立角色本体并以此为基础进行该领域人类心理常识获取和组织的有效方法,论文通过家庭领域的实例分析验证了该方法的可行性和正确性。陆汝占在“汉语概念内涵模型及其应用研究”中基于数理逻辑背景,提出用于分析语义的汉语概念内涵逻辑模型原型理论,在当前处于强势的结构主义语法及概率统计方法的情势下,冷静地逆势回归到重视汉语本体研究。

第三组论文是关于语义网和 Web 搜索的研究和应用,包括3篇文章。王珊等在“示范语义网及其知识服务平台研究”中从本体库管理、资源抓取系统、标注系统、本体构建、本体学习、个性化服务和示范应用等方面介绍了最新的研究进展,以及他们以经济学科和法学为专业领域,构建可提供网上知识和信息服务的示范语义网的工作。周傲英等在“Dolphin——学术搜索引擎的技术研究和系统实现”中指出,学术搜索引擎如何为用户提供更多有意义信息和更加多样化的查询分析工具,是一个重要的课题。论文介绍了自己带领团队设计开发的学术搜索引擎系统 Dolphin,可以为用户提供符合查询意图的更多知识性结果,并为用户发掘出搜索结果中更多的有意义信息。孙吉贵等人在“Web 下的语义搜索技术与实现”中综述了语义搜索门户网站进展情况,介绍了语义搜索相关技术及实现,并重点介绍了他们课题组开发的语义搜索原型——Crab 智能搜索系统。

林惠民院士、陈润生院士、史忠植研究员、李昂生研究员、刘志勇研究员、朱梧楨教授、姜云飞教授、周立柱教授、张松懋研究员、王国胤教授、吕建教授、石勇研究员、陈红教授、陈恩红教授、刘挺教授、瞿裕忠教授、李娟子教授、于剑教授等在百忙中对稿件进行了认真的评审,提出了许多中肯的意见。根据这些意见,有些稿件进行了合并,有些进行了修改,有些甚至作撤稿处理。他们对本论文集的贡献是不能忽视的。

本论文集的出版经过了很长时间的酝酿。在2009年初项目结题总结会前,就决定要将课题成果集结出版。当时商定的方针为:(1)不与已发表文章重复;(2)是课题组代表性成果;(3)在项目总结的基础上增加技术性内容,并加强相关工作综述和比较的力度,使各自的研究成果放在一个大视野下展示;(4)每个“独立”课题(三级课题或以下)应根据各自的研究成果的具体情况,提交不少于一篇的文章。同时决定将项目执行过程中拟出版而因为种种原因没有出版的《知识工程与知识科学研讨会》稿件合并出版。我们还商定所有的论

文都需要由没有参加重大项目的第三方专家审稿,要求所有作者要根据审稿意见,认真修改稿件。有关论文编排的框架和出版进度也经过了慎重的讨论研究。

中国人民大学的杜小勇、胡鹤,北京工业大学的王洁在收集整理稿件、组织审稿、提醒进度等方面做了大量实际工作,特此表示感谢。

最后感谢全体项目组的成员和本论文集的作者,没有他们的研究成果和一贯的支持与理解,本论文集是不可能出版的。从课题酝酿立项到论文集出版前后历经十年,这是一群执著的研究者在这个重要的领域辛勤耕耘的一段历史记录。相信他们的工作会在中国的科学史上留下重要的记忆。我们以陆汝钤院士在 2009 年元旦前发给课题组全体成员的新年贺词作为结束语。

《沙漠斗士》

在广阔无垠的沙漠里耕耘,
让升腾的汗水托住亘古晴空,
把生命刻印在从未有人走过“路”上,
这条路指向的是永远不能达到的目的地。



王 珊
刘椿年
2011 年春节
于北京

简明目录

第一部分 总 结 篇

1	非规范知识处理的基本理论和核心技术——重大项目研究概述	陆汝铃	1
---	-----------------------------------	-----	---

第二部分 理 论 篇

2	知识、范畴论、结构主义	陆汝铃	22
3	Combining λ and π Through Continuation	Xinxin Liu(柳欣欣)	41
4	通信并发系统中的拓扑结构、随机性与噪声	应明生	52
5	云模型研究进展	刘玉超 陈桂生 李德毅	58
6	网络化数据挖掘与群体智能	陈桂生 张海粟 李德毅	70
7	Multi-agent Coordination Using Nearest-neighbor Rules: Revisiting the Vicek Model	Sanjiang Li(李三江)and Huaiqing Wang(王怀清)	84
8	多 Agent 系统合作求解机制	张 伟 童向荣 刘惊雷 等	95
9	模型检测多智能体系统	苏开乐 陈清亮	109
10	“信息-知识-智能”生态链意义下的知识问题研究	钟义信	116
11	近似推理与证据推理	陈仪香	132
12	语义 Web 矛盾知识的处理	林作铨 马 跃	150
13	约束(归纳)逻辑程序设计中不确定知识的处理	刘椿年 王 洁 郑 磊	166
14	时空知识表示与推理	刘大有 王生生 欧阳继红 等	181
15	择优型非单调逻辑简介	朱朝晖	198

第三部分 系 统 篇

16	非编码 RNA 及其参与的复杂生物网络	刘长宁 孙世伟 赵 屹 等	211
17	COBRA 在代谢网络中的研究现状综述	奚燕萍 王 飞	243
18	二次学习	周志华	268
19	非规范软件需求管理的逻辑方法和管理框架	金 芝 牟克典 魏 博	280
20	算法交易及金融市场的套利分析	邓小铁 王 峰	302
21	数据流挖掘研究现状与关键问题	毛国君 刘椿年	320

22	社会角色心理常识的获取方法	雍 兮 曹存根 张冬蕾 等	331
23	汉语概念内涵模型及其应用研究	陆汝占 陈玉泉	343
24	示范语义网及其知识服务平台研究	王 珊 杜小勇 何 军 等	357
25	Dolphin——学术搜索引擎的技术研究和系统实现		
		王 晔 王晓玲 周傲英	378
26	Web 下的语义搜索技术与实现	孙吉贵 欧阳丹彤 叶育鑫 等	391

目 录

非规范知识处理的基本理论和核心技术——重大项目研究概述	陆汝钤	1
1 本项目的重大科学问题和国内外研究背景		1
2 重大项目概况及其代表性成果		4
3 回顾之一：我们的研究成果在非规范知识领域占什么位置		13
4 回顾之二：非规范知识如何界定？如何分类？		15
5 回顾之三：本项目涉及的基本理论和核心技术		17
6 回顾之四：跨领域和跨课题的交叉研究		18
参考文献		20
知识、范畴论、结构主义	陆汝钤	22
1 知识是什么		22
2 带类型范畴论		26
2.1 范畴论的基本定义		26
2.2 有向图的可交换性		27
2.3 带类型函子和带类型子范畴		29
2.4 泛适性及其应用		29
2.5 抽象推理能力和领域知识的结合		30
2.6 有缺陷知识和类范畴		32
3 关于结构主义		34
3.1 结构主义是一种思潮		34
3.2 数学结构主义和范畴论		36
3.3 结构主义和功能主义		37
参考文献		38
Combining λ and π Through Continuation	Xinxin Liu(柳欣欣)	41
1 Introduction		41
2 From λ_{cps} to a Value-passing Process Language		42
3 Reduction Semantics and Observations		44
4 Labelled Semantics and Bisimulation		46
5 Conclusion and Future Works		50
References		51

通信并发系统中的拓扑结构、随机性与噪声	应明生	52
1 引言		52
2 进程演算中的拓扑结构		52
3 概率进程演算		55
4 基于概率逻辑的概率顺序程序的语义		56
5 带噪声的 π 演算		56
6 结束语		57
参考文献		57
云模型研究进展	刘玉超 陈桂生 李德毅	58
1 背景		58
2 云模型		59
2.1 正态分布和重尾分布		60
2.2 正向云发生器和高斯云分布		61
3 云模型中概念的操作方法		63
3.1 逆向云发生器		64
3.2 云变换		65
4 云模型在云安全中的运用		67
5 总结		68
参考文献		69
网络化数据挖掘与群体智能	陈桂生 张海粟 李德毅	70
1 网络化数据挖掘研究		71
1.1 复杂网路度量指标		72
1.2 节点的重要性排序(偏序集)		74
1.3 社区结构与骑墙节点		75
1.4 网络骨干成员挖掘算法		76
2 群体智能研究		76
2.1 社区大众交互中常识知识的自然缺省		77
2.2 大众交互中群体智能的涌现		79
2.3 基于复杂网络统计挖掘的群体智能应用		80
3 小结		81
参考文献		82
Multi-agent Coordination Using Nearest-neighbor Rules: Revisiting the Vicsek		
Model	Sanjiang Li(李三江)and Huaiqing Wang(王怀清)	84
1 Introduction		84
2 The Vicsek Model and the Nearest-neighbor Rule		85