



证据网络推理学习 理论及应用

Theory and Applications
of Evidential Network
Reasoning and Learning

姜 江 陈英武 常雷雷 著



科学出版社

证据网络推理学习理论及应用

姜 江 陈英武 常雷雷 著



科学出版社

北 京

内 容 简 介

本书提出并建立了一套完整的证据网络理论和方法体系,对证据网络的定义、结构建模、参数表示、不同参数模型下的推理及证据网络参数和结构学习的相关理论和方法展开了深入论述。全书共分为7章,内容包括:不确定性建模理论,不确定性推理方法,证据网络提出的价值与意义,证据网络模型的基本概念、特点、关键要素和建模流程,证据网络的结构与参数,证据网络的推理问题,不同参数模型下的推理策略与算法,证据网络参数学习模型与计算方法,证据网络信度规则模型库结构学习,以及相关应用研究等。

本书主要面向管理科学与工程、控制科学与工程、信息技术等领域的学者及研究生,也可供相关领域的研究人员阅读参考。

图书在版编目(CIP)数据

证据网络推理学习理论及应用/姜江,陈英武,常雷雷著. —北京:科学出版社,2013

ISBN 978-7-03-038417-1

I. ①证… II. ①姜… ②陈… ③常… III. ①决策论—研究 IV. ①O225

中国版本图书馆CIP数据核字(2013)第194903号

责任编辑:李欣 李静科/责任校对:胡小洁

责任印制:赵德静/封面设计:陈敬

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

北京通州皇家印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2013年8月第 一 版 开本:B5(720×1000)

2013年8月第一次印刷 印张:9 1/2

字数:180 000

定价:48.00元

(如有印装质量问题,我社负责调换)

前 言

不确定性决策是目前管理科学研究和应用中的一个热点问题. 客观世界的实际问题往往涉及众多相互联系又相互影响的因素, 这些因素本身及其相互之间的关系都存在大量的不确定性, 而不确定性可分为两类, 一类是反映客观事物内在本质的随机不确定性, 一类是反映由于人们对客观世界的认识不足、信息缺失或知识缺乏而导致的认知不确定性. 如何描述各种不确定性, 如何在复杂关系分析中对问题有效地建模, 如何综合定量数据和定性知识而作出科学的决策, 都对不确定性管理决策问题的研究提出了新的挑战.

为应对上述挑战, 本书在定性定量综合集成方法论的指导下, 通过对 D-S 证据理论和图模型基础理论的研究, 借鉴贝叶斯网络模型的研究思路, 提出证据网络模型. 证据网络模型是 D-S 证据理论和图模型的结合, 可以充分发挥 D-S 证据理论在不确定性信息处理, 尤其是认知不确定性的建模和分析上的理论优势, 发挥图模型在问题描述、关系分析上的语义优势, 在理论上扩展不确定性建模与分析的研究思路和方法, 为定性经验知识与定量数据的统一建模和综合处理提供技术手段, 在实践上为不确定性管理问题的分析、建模、推理, 以及评估、决策提供技术方法与工具支撑.

为建立一套完整的证据网络理论和方法体系, 本书对证据网络的定义、结构建模、参数表示、不同参数模型下的推理, 以及证据网络参数学习的相关理论和方法开展了深入论述. 全书共分四大部分, 第一部分为证据网络的基本概念 (第 1 章、第 2 章), 包括证据网络模型的定义、关键要素、基本特点和建模流程. 证据网络模型通过定性层面的有向无环图描述变量之间的相互关系, 通过定量层面的信度函数刻画变量之间影响模式和程度, 并综合证据理论和图模型的特点, 为系统分析和建模提供了一种不确定性描述、相互影响关系建模以及综合处理信息的技术手段; 为了构建证据网络的结构和参数模型, 本书提出了基于树模型和基于因果图的证据网络拓扑结构建模方法, 定义了证据理论框架下知识描述的两种规范化证据网络参数模型——条件信度参数模型和信度规则参数模型. 第二部分为证据网络的推理 (第 3 章、第 4 章), 研究并建立了以条件信度和信度规则为参数模型的证据网络推理策略与推理方法. 其中, 为解决条件信度参数模型下的证据网络推理问题, 在条件信度函数计算理论基础上, 提出了证据网络模型的正向因果推理和反向诊断推理方法; 并通过对证据冲突悖论的分析, 提出了一种基于冲突度量的证据网络信度合成算法, 解决了证据网络结点信息融合问题. 在以信度规则为参数模型的证据网络

推理研究中,为分析结点之间的相互重要度,提出了不完全信息情况下的证据网络结点权重获取方法;并在证据推理算法的基础上,结合信度结构数据处理方法和信度规则激活算法,实现了数据与证据网络模型的对接,建立了基于信度规则的证据网络推理与结果分析方法.第三部分为证据网络的学习(第5章、第6章),构建了证据网络参数学习的数学模型并设计了基于投影梯度法的证据网络参数学习算法.针对以信度规则为参数的证据网络模型,分析建立了其参数学习问题的非线性目标优化模型,提出以信度结构模型差距度量准则作为优化模型的目标函数;通过推导模型解析表达式函数的梯度,设计基于投影梯度法的证据网络参数学习求解算法;通过提出基于多种维度约减技术的信度规则模型库结构学习方法,解决了证据网络信度规则组合爆炸、关键前件属性选择等问题,为规则模型库结构学习提供了初步研究思路;通过证据网络的学习,建立起了从历史数据和经验知识信息到证据网络模型构建的技术和方法.第四部分(第7章)为总结和展望.

本书得到了国防科学技术大学学术著作出版专项经费、校庆60周年系列学术专著以及国家自然科学基金“证据网络推理、学习方法及应用研究(71201168)”的资助。

在本书的撰写过程中,作者得到了许多学者及专家的无私帮助,其中,特别感谢国防科技大学谭跃进教授、武小悦教授、贺仁杰教授、杨克巍副教授,英国曼彻斯特大学的杨剑波教授、徐冬玲教授、陈玉旺博士、汤大为博士,清华大学的周志杰博士,福州大学的王应明教授等对作者的帮助和支持.此外,本书参考了大量国内外相关文献,书中所附参考文献仅为其中一部分,在此向所有参考文献的作者表示衷心感谢.

限于作者的水平,书中难免有不妥和疏漏之处,敬请读者不吝赐教.

作 者

2013年3月于长沙

目 录

前言

第 1 章 绪论	1
1.1 不确定性建模理论	3
1.1.1 不确定性分类	3
1.1.2 不确定性处理	5
1.2 不确定性推理方法	8
1.2.1 主要的不确定性推理方法	8
1.2.2 贝叶斯网络与影响图	10
1.3 证据理论相关研究现状	12
1.3.1 D-S 证据理论的提出	12
1.3.2 D-S 证据理论综述	15
1.3.3 证据推理	16
1.3.4 证据网络	17
1.3.5 证据理论与风险决策	18
1.4 存在的问题	19
1.5 本书主要内容及安排	20
1.5.1 研究思路	20
1.5.2 研究内容	21
第 2 章 证据网络模型	23
2.1 证据理论与图模型基础	23
2.1.1 D-S 证据理论	23
2.1.2 图模型	29
2.2 证据网络模型的定义	31
2.2.1 证据网络的概念	31
2.2.2 证据网络的特点	32
2.2.3 证据网络建模过程	33
2.3 证据网络模型的结构	34
2.3.1 结点之间的关系	34
2.3.2 基于树模型的证据网络结构建模	35
2.3.3 基于因果关系图的证据网络结构建模	37

2.4 证据网络模型的参数	39
2.4.1 知识表示模型	39
2.4.2 证据网络参数的条件信度表示	40
2.4.3 证据网络参数的信度规则表示	41
2.5 小结	44
第 3 章 条件信度参数模型下的证据网络推理	45
3.1 条件信度参数模型下的证据网络推理问题	45
3.1.1 推理问题	45
3.1.2 研究思路	46
3.2 证据网络条件信度推理方法	47
3.2.1 条件信度函数计算基础理论	47
3.2.2 证据网络条件信度的正向推理	53
3.2.3 证据网络条件信度的反向推理	54
3.2.4 证据网络条件信度的乘积规则	55
3.2.5 证据网络条件信度推理算例	56
3.3 证据网络信度合成方法	58
3.3.1 信度合成悖论分析	59
3.3.2 一种新的证据冲突度量方法	60
3.3.3 基于冲突度量的信度合成方法	64
3.4 证据网络在航天系统安全性分析中的应用	66
3.5 小结	71
第 4 章 信度规则参数模型下的证据网络推理	72
4.1 信度规则参数模型下的证据网络推理问题	72
4.1.1 推理问题	72
4.1.2 研究思路	73
4.2 不完全信息情况下结点权重获取方法	73
4.2.1 偏好关系的定义与表示	74
4.2.2 基于目标规划的权重获取方法	74
4.2.3 结点权重获取的数值算例	77
4.3 基于 ER 的证据网络推理方法	79
4.3.1 信度结构数据转化	79
4.3.2 信度规则的激活	81
4.3.3 证据网络信度规则推理与合成算法	83
4.3.4 证据网络信度规则推理结果分析	86
4.4 证据网络在军事威胁评估与预测中的应用	87

4.5 小结	90
第 5 章 证据网络参数学习	92
5.1 证据网络参数学习问题	92
5.1.1 参数学习的研究思路	92
5.1.2 参数学习问题的数学模型	93
5.1.3 多级证据网络的参数学习	94
5.2 参数学习目标函数的计算	95
5.2.1 信度结构模型的距离定义	96
5.2.2 参数学习的目标函数	99
5.3 基于投影梯度法的证据网络参数学习	99
5.3.1 投影梯度法	99
5.3.2 参数学习目标函数的梯度	102
5.3.3 基于投影梯度的证据网络参数学习方法	104
5.4 证据网络参数学习应用	105
5.4.1 石油管线风险预警证据网络模型的参数学习	105
5.4.2 交通事故风险预测证据网络模型的参数学习	107
5.5 小结	110
第 6 章 证据网络信度规则模型库结构学习	111
6.1 证据网络信度规则模型库结构学习问题	111
6.2 基于约减技术的信度规则模型库结构学习	112
6.2.1 维度约减技术	112
6.2.2 基于约减的信度规则模型库结构学习	116
6.3 证据网络信度规则模型库结构学习应用	117
6.3.1 示例背景分析与建模	117
6.3.2 不同约减技术的关键前件属性选择	119
6.3.3 灰靶与主成分分析结果对比	120
6.3.4 多尺度分析与主成分分析结果对比	122
6.3.5 主成分分析结果鲁棒性分析	124
6.4 小结	126
第 7 章 总结与展望	127
7.1 本书总结	127
7.2 展望	128
参考文献	130
索引	139

第1章 绪 论

不确定性决策是目前管理科学研究和应用中的一个热点问题. 管理问题的不确定性主要来源于研究对象所处客观环境及研究问题本身的复杂性, 既包括反映客观事物内在本质的随机不确定性, 也包括反映由于人们对客观世界的认识不足、信息缺失或知识缺乏而导致的认知不确定性. 此外, 解决实际管理问题或系统问题往往需要分析众多相互联系又相互影响的因素及其相互关系, 这些因素之间的关联性成为制约系统功能发挥和管理决策成功与否的关键. 因此, 客观世界的这些不确定性与关联性对不确定性决策问题的研究提出了新的挑战.

在管理决策实践中, 不确定性与关联性无处不在, 如在航天工程、武器装备发展论证、重大基础设施建设等大型系统管理中, 为实现系统管理目标, 需要考虑众多相互联系而又相互影响的子系统或部件之间的关系, 以及在系统内部和外部存在的如费用、风险、人为参与等各种因素联合作用导致的不确定性, 如果不能正确地描述这些不确定性, 不能合理地描述系统内部之间的相互关系, 不能准确地掌握系统运行模式和方向, 就不能在适当的时机作出正确的决策, 从而导致整个工程建设的费用增加、进度拖延, 甚至项目失败和重大损失; 又如在军事作战态势分析与评估、系统故障预测与诊断等控制与决策中, 如果不能对获取的军事情报信息或系统检测数据等定性定量信息及时进行处理, 不能把军事指挥员的经验或系统工程师的判断等经验知识信息综合考虑, 就很有可能错失战机或酿成重大事故, 造成整个战争的失败或重大安全事故等.

因此, 应当从理论和方法上探索在考虑系统关联性和信息不确定性环境下, 如何建立反映系统真实情况的模型, 如何对定性知识与定量数据信息进行统一建模和处理, 如何通过严格的推理计算不断增强人们对系统的认知等. 也正如钱学森曾提出的^[1], “要建立定性和定量综合集成的复杂系统研究方法论, 即要从整体上研究和解决问题, 采用人机结合以人为主的思维方法和研究方式, 对不同层次、不同领域的信息和知识进行综合集成, 达到对整体的定量认识.”

对不确定性的分析、建模和处理建立在四大理论基础之上, 概率论以其坚实的数学理论基础为随机不确定性的研究提供了强大的工具支持; 模糊集理论为处理客观世界中“亦此亦彼”的模糊现象提供了技术手段; 粗糙集可以有效地在不确定、不完备信息中发现并揭示隐含的知识和规律; 证据理论将概率区间化、一般化, 以信度函数的形式处理不完全信息并提供统一的不确定性建模框架. 众多学者指出,

D-S 证据理论较其他不确定性理论能更加有效地处理不完全、不精确信息,克服概率论对不可知性、信念表达、主观认知等不确定知性建模上的不足,提供对各种不确定性的统一建模和有效的信息综合方法。所以, D-S 证据理论为不确定性管理决策问题的建模与求解提供了一种可能的基础理论框架。

此外,为描述复杂管理问题组成要素或系统部件之间的相互关系及影响,从整体上识别和描述问题,以图模型和网络模型为代表的因果关系分析技术以其结构清晰、语义明确等优势,为结构建模,尤其是系统内部组成要素之间的关系描述提供了一套行之有效的技术手段。

综合上述研究背景分析,为解决不确定性与关联性并存的管理决策问题,本书提出将 D-S 证据理论与图模型相结合,建立证据网络模型。证据网络将综合图模型在问题描述、关系分析上的语义优势,及 D-S 证据理论在不确定性信息处理,尤其是认知不确定性的建模和分析上的理论优势,为不确定性管理决策问题的分析、建模、推理及评估、决策提供技术方法与工具支撑。

证据网络的研究目前还处于萌芽阶段,其建模、推理、学习及应用方面的研究还很少,所以,要形成一套完善的证据网络理论和方法体系,还需要从各个角度去研究证据网络模型:①在证据网络的建模上,需要定义证据网络的关键要素,完善网络拓扑结构的形式及建立方法,提出合理的参数表示模型及描述方法;②在证据网络的推理上,需要探索各种情况下的推理模式与方法,在理论上推导和证明推理方法的合理性和有效性,解决正向因果推理与反向诊断推理等问题;③在证据网络的学习上,需要研究证据网络结构学习和参数学习的技术方法,使得证据网络模型的构建建立在历史数据和经验信息的基础上,更好地反映真实系统特征;④在证据网络的应用上,需要探索其针对实际问题的建模与求解过程,利用证据网络的特点解决管理决策分析领域的实际问题。

综上所述,证据网络模型的提出有其明确的实际背景需求,证据网络建模、推理及学习方法的研究需要进一步完善,其研究在理论和实践上都将具有重要意义。在理论上,把证据理论与图模型结合,利用二者的优势,克服概率论在处理认知不确定性建模和推理上的不足,发挥 D-S 证据理论对不确定性,尤其是不完全性、不精确性建模的优势,将扩展不确定性建模与分析的研究思路和方法,为定性经验知识与定量数据的统一建模和综合处理提供技术手段。对证据网络建模、推理及学习方法的研究,将进一步完善证据网络模型框架,提出合理可操作的证据网络求解策略、思路和算法。在实践上,将探索采用证据网络模型进行定量观察信息与定性经验知识的建模、融合与推理的方法、步骤,解决由不确定性和关联性共同作用下的不确定性管理决策问题、不确定性环境下的系统分析、建模和求解问题,为管理决策分析、系统工程管理等领域的问题分析和求解提供一种有效的技术支持。

1.1 不确定性建模理论

人类在认识客观世界时, 需要面对大量的不确定性. 可以说, 客观世界中的绝大部分现象都是不确定的, 所谓确定的、规则的现象, 只是在一定的前提和特定的假设条件下发生, 只会局部或者较短的时间内存在, 或者说, 不确定性是绝对的, 确定性是相对的^[2]. 随着对不确定性研究的深入, 客观世界的不确定性特征越来越得到学术界的普遍认可, 不确定性建模与分析的研究也逐步得到重视. 从 20 世纪 70 年代开始, 人们就在人工智能领域对不确定性展开研究, 1985 年召开的第一届“人工智能中的不确定性”国际学术会议指出: “不确定性妨碍人们做出最好的决定, 甚至导致做出一个错误的决定.” 所以, 无论是在物理学、数学、生物学等自然科学领域, 还是在哲学、经济学、社会学、心理学、认知学等社会科学领域, 不确定性问题的研究都具有重要意义. 以下从不确定性的分类入手, 归纳总结当前国际学术界处理不确定性的主要基础理论和建模方法, 以及各自的优缺点.

1.1.1 不确定性分类

不确定性一词, 最早出现于 1936 年《政治经济学是否有用》一文中, 诺贝尔经济学奖、图灵奖获得者西蒙从认知科学和行为科学出发, 认为“不可避免的是, 如果经济学家要与不确定性打交道, 就必须理解人类行为面临的不确定性”^[3]. 当前, 不确定性问题研究受到越来越多的关注, 要研究不确定性, 首先要对不确定性有明确的界定, 给出其科学的分类.

Kiureghian 等^[4] 指出不确定性有 7 个来源: 基本随机变量内在的不确定性, 选择错误的概率模型导致的不确定性, 建模错误导致的不确定性, 在概率模型参数估计中的统计不确定性, 在物理模型参数估计中的统计不确定性, 数据观察中的错误导致的不确定性, 变量描述、近似表示中的建模不确定性. 李德毅院士^[3] 把不确定性分为随机性、模糊性、不完全性、不稳定性和不一致性五个方面, 其中, 随机性和模糊性是最基本的. 刘宝碇在《不确定理论教程》^[5] 中也指出: 不确定性的表现形式是多种多样的, 如随机性、模糊性、粗糙性及其他的多重不确定性, 其中随机性指事件发生的条件不充分, 使得条件与结果之间没有明确的因果关系, 在事件的结果上表现出的不确定性质; 而模糊性的出现是由于概念本身模糊, 一个对象是否符合这个概念难以确定, 在质上没有明确定义, 在量上没有明确界限, 这种性质不是由人的主观认识造成的, 而是事物的一种客观属性.

Oberkampf 和 Paté Cornell 等在研究风险及复杂系统不确定性分析与建模当中, 通过多年的总结, 在概念和计算结构上把不确定性划分为两类, 随机不确定性 (aleatory uncertainty) 和认知不确定性 (epistemic uncertainty), 以下简称 A 类和 E

类不确定性^[6-9]. 美国 Sandia 国家实验室 (Sandia National Laboratories, SNL) 于 2002 年 8 月在新墨西哥州 Albuquerque 城召开了一届不确定性研讨会, 围绕随机不确定性和认知不确定性展开研讨, 提出了五个挑战性的问题: ① E 类不确定性应当如何表示? ② E 类和 A 类不确定性如何组合及推理计算? ③ 不确定性的多源信息在计算之前如何集成? ④ 在实际运算时如何处理不确定性参数? ⑤ 如何调整各种处理方法用于实际计算^[10]? 研讨会上的优秀论文在国际期刊《可靠性工程与系统安全性》(*Reliability Engineering & System Safety*, RESS) 2004 年的一期专刊上发表^[11], 从而掀起了一股从分析认知不确定性建模、多种认知不确定性集成及认知与随机不确定性融合等问题的不确定性研究热潮.

单词 Aleatory 来自拉丁文 Alea, 意为“掷骰子”, Epistemic 来自希腊语 Episteme, 意思为知识^[4]. 近几年, 众多学者对两类不确定性展开研究, 并分别给出相关的定义与内涵, 表 1.1 选取几个代表性的研究进行比较分析.

表 1.1 Aleatory 与 Epistemic 不确定性的比较

	Aleatory	Epistemic
Paté Cornell 1996 ^[6]	已知或可观察事物的可变性, 样本的随机性等	来源于缺少对基本现象的知识
Oberkampff 2004 ^[7]	可变性, 不可约性, 固有性, 随机性; 物理系统或环境本质的变化; 系统可以在确定的范围内随机变化, 概率分布适合描述这种不确定性	可约的, 主观的, 知识状态; 来自对系统一定程度的无知; 建模分析过程中知识的缺乏; 不完全信息, 不完全知识; 数学表示更加的困难; 随着知识的增长和信息的获取, 这种不确定性得以降低
Soundappan 2004 ^[12]	不可约性; 物理现象内在的可变性	可约性; 如果收集到信息, 不确定性可以被降低或去除
Kiureghian 2009 ^[4]	现象本质的随机性; 建模人员不可能预知减低它的可能性	知识或数据的缺乏所导致; 通过收集更多的数据和改进模型来降低这种不确定性
Samson et al. 2009 ^[13]	随机性, 可变性, 不可约性; 系统内在的可变性, 来自于系统有很多潜在的不同行为可能; 数学上可用概率论建模	主观的, 知识的状态, 可约的; 知识的缺乏; 不完全信息; 没有足够的实验数据; 难以量化建模
Limbouurg et al. 2010 ^[14]	不可约的, 内在的, 可变的, 随机的; 真实特征的变化; 不能通过获取数据和知识而消除	可约的, 知识缺乏, 含糊的; 随着更多数据的获取, 不确定性降低

总而言之, A 类随机不确定性是客观事物内在的本质, 只能通过科学的方法认识它、描述它, 但不可以消除它. 而 E 类认知不确定性是主观的不确定性, 它是由于人们对客观世界的认识不足、信息缺失或知识缺乏而产生的不确定性, 随着获取信息的不断增多, 人们认识水平的不断增长, 可以逐渐减少这种不确定性. 比如

投硬币实验, 假如有一个特殊的硬币, 开始时我们不知道其出现正反的可能性是多少, 这是因为对这一枚硬币缺乏知识, 但做足够多次试验后, 我们获取了足够的信息以确定出现正反的可能性, 这就消除了所谓认知不确定性。但是, 即使知道了硬币的足够信息, 我们仍无法确定下一次到底是出现正或反, 这是不能消除的随机不确定性。

对不确定性进行这种分类, 近几年得到国际学术界的广泛认可, 这一点通过 2009~2010 年在 RESS, *Risk Analysis*, *Artificial Intelligence* 等国际期刊杂志上出现的相关文章可以说明。这种分类方法可以辅助进行更加清晰和透明的建模和分析研究, 比如在工程系统分析中, 由于知识缺乏导致的不确定性可以在模型上通过引入辅助非物理变量表示, 这些非物理变量可以通过收集更多的信息或采用更先进的科学方法加以描述和解决。当然, 在实际问题建模中, 经常不容易区分 A 类和 E 类不确定性, 还是需要建模人员的实际经验和认识加以确认。

1.1.2 不确定性处理

不确定性处理是通过选择科学的数学方法完成对不确定性的描述、分析和处理, 主要包括不确定性建模、不确定性推理和敏感性分析等基本内容, 具体讲, 就是实现对上述 A 类和 E 类不确定性进行统一建模, 实现集成和推理。在历史上, 采用概率论提供的数学方法来描述 A 类和 E 类不确定性可以追溯到 17 世纪^[15], 本书主要关注 E 类不确定性的分析、处理、建模及推理问题。

SNL 的《不确定性专题研讨》总结了除概率论和贝叶斯理论外, 用以处理 E 类不确定性的新兴理论, 包括概率理论 (probability theory), 区间分析 (interval analysis), 概率盒理论 (probability boxes theory), 模糊集理论 (fuzzy set theory), 可能性理论 (possibility theory)^[16], D-S 证据理论 (Dempster-Shafer theory of evidence)^[17], 可转移信度函数模型 (transferable belief model)^[18], 不精确概率论 (imprecise probability theory), 信息距决策论 (information-gap decision theory)^[19] 及一般信息论 (generalized information theory) 等, 并指出这些理论都处于发展的初期阶段, 都与经典概率论和贝叶斯理论有所联系^[7], 但每种理论各有优劣, 对处理 E 类不确定性的使用还存在争议^[12]。虽然这些理论还有一些有待研究的问题, SNL 的专题研究的目标就是评价、发展和应用这些理论, 使它们更加有效地去处理可靠性工程、风险分析、系统安全性评估中存在的 uncertainty 问题。

在 RESS 2004 年的专刊中, 对研讨会中有特别贡献的 22 篇文章进行了总结^[20], 指出每篇文章针对不同角度问题, 采用不同的理论方法, 解决了不同的问题, 回答了挑战性问题中提出的一个或几个问题; 虽然没有提出通用的处理不确定性的建模方法, 也没有达成一致意见, 但是这些争论对推动不确定性研究的发展具有积极的作用。下面介绍使用最多的几种理论。

1) 概率理论

概率方法是最自然也是最早被尝试用于建模不确定性的方法之一, 因为概率论本身是关于随机现象和不确定性的数学理论. 概率论用概率测度来描述事件 (对象) 的不确定性, 以贝叶斯理论作为概率推理的基本方法. 概率论使用概率论方法处理不确定性能保证推理结果的正确性, 被认为是数学基础最强的不确定性处理理论^[21].

关于对概率的认识, 也分为两大学派: 频率主义学派 (客观主义) 和贝叶斯主义学派 (主观主义). 前者认为概率是事件发生的频率, 用试验次数趋于无穷时的极限来刻画, 是完全客观的; 后者认为概率是主观信念, 是人对不确定性事件的主观意愿决定的结果. 理论上, 当有充足的数据来估算一个事件的概率时, 主观概率应当收敛于客观概率. 概率论有坚实的数学基础和完备的理论体系, 几个世纪以来, 在科学研究的各个领域发挥了重要的作用. 概率论能很好地描述 A 类随机不确定性, 但是对一些 E 类认知不确定性, 即随着信息量的增加而消除的不确定性在概率模型里很难表示.

此外, 概率论在应用中也受到了一些质疑, S. Pender^[22] 曾指出概率论在项目管理中的运用有六大局限: “概率理论建立在随机这一假设的基础上, 而项目执行的是有人计划的人为行为; 项目的一次性减少了源自概率分析得到的统计数据的正确性和可靠性; 概率理论假设未来是可知的, 实际上项目的不确定性和不可知性无法避免, 特别是人的行为无法预知; 人类处理所有信息以达到优化决策的能力有限; 因为不确定性和不可知性的存在, 知识的时间特性对项目计划来讲非常重要, 概率理论的过去-将来两时段模型忽略了知识变化的连续性; 项目的有关参数和结果应及时和项目的利益相关者交流, 人的意思有时很难用概率来表达.”

2) 模糊集理论

1965 年, 美国著名的控制论专家 L.A.Zadeh 发表了著名的开创性论文 *Fuzzy sets*, 标志着模糊数学的诞生^[23]. 模糊数学不是让数学变得模糊, 而是用数学来描述客观世界存在的模糊现象. 模糊现象与随机现象不同, 随机现象研究随机事件, 而随机事件的结果是“非此即彼”; 模糊现象研究模糊事件, 而模糊事件的结果是“亦此亦彼”, 这是随机事件与模糊事件的根本区别. 模糊集合论的贡献在于引入了集合中元素对该集合有一个“隶属度”, 从而将经典集合论里的特征函数取值范围由集合 $\{0,1\}$ 推广到 $[0,1]$ 上的一个子区间, 将经典二值逻辑推广至多值逻辑, 使得模糊性可以用 $[0,1]$ 上的区间来度量^[24].

3) 粗糙集理论

粗糙集理论由波兰数学家 Pawlak 于 1982 年提出^[25]. Pawlak 把那些无法确认的个体都归属于边界区域, 而这种边界区域被定义为上近似集和下近似集之差集. 该理论是一种刻画不完整性和不确定性的数学工具, 能有效地分析和处理不精确、

不一致、不完整等各种不完备信息,并从中发现隐含的知识,揭示潜在的规律^[26].

4) 证据理论

D-S 证据理论,是在 Arthur Dempster 于 20 世纪 60 年代后期提出的上下概率及其合成规则的基础上^[27,28],由 Glenn Shafer 在 1976 年发表的专著《证据的数学理论》(*A Mathematical Theory of Evidence*) 中正式建立并逐步发展起来的^[17]. 证据理论用取值在区间 $[0,1]$ 内的信度函数 (belief function) 与似然函数 (plausibility function) 两个数值组成的区间表示决策者在给定证据下对假设或命题的信念,并用 Dempster 合成规则对不同证据产生的信念进行综合. 证据理论可以视为对概率方法的一种改进,这种改进主要体现在四个方面: ① 证据理论对无知 (ignorance) 的表示更加明确合理; ② 证据理论具有综合不同信度函数的 Dempster 合成规则,而概率论框架不便处理不同概率函数的综合; ③ 证据理论与概率论关于信念 (belief) 的基本观点是不同的,证据理论把信念视为主体基于证据产生的认识,而概率论认为主体的信念是先验的,证据再作用仅是修改信念; ④ 它适合于处理不完全信息,即缺乏明确性而带来的不确定性;

5) 几大理论比较分析

各种不确定理论框架解决问题的重点不同,模糊集理论侧重于表示和管理模糊信息;贝叶斯概率理论具有严格的数学理论基础,并且已经发展为较为完善的表示管理随机性知识的方法,但需要精确的概率判断. D-S 证据理论采用不精确的信度判断来抓住不精确的证据,即结果的可能性通过一个信度区间来判断,而非点概率.

虽然模糊集和粗糙集都可以用来处理模糊和不确定问题,但二者的着眼点不同. 模糊集在处理不确定性问题时,主要着眼于知识的模糊性,强调的是集合边界的不分明性;而粗糙集理论在处理模糊和不确定性问题方面着眼于知识的粗糙性,强调的是集合对象间的不可分辨性. 粗糙集理论是为开发自动规则生成系统而提出的,而 D-S 理论是用于证据中,这两种理论的动机是不同的,在相容性观点下,他们具有相似性:由粗糙集理论中的上下近似可以导出内外概率函数是 D-S 理论中信任和似然函数的特例;但在 D-S 理论的概率分配观点下,他们是不同的,在 D-S 理论中需要对基本概率分布进行假设,而在粗糙集中不需要数据之外的任何附加信息. 粗糙集理论与证据理论虽有一些相互交叠的地方,但本质不同,粗糙集使用集合的上下逼近,而证据理论使用信度函数作为主要工具. 表 1.2 给出四大不确定性处理理论的比较.

表 1.2 概率理论、模糊集理论、粗糙集理论、证据理论的比较

比较	概率理论	模糊集理论	粗糙集理论	证据理论
对象间关系的基础	数据的随机性	概念边界的 不分明性	对象间的不可分辨 关系	部分信息已知,部分 信息未知
不确定的刻画方法	概率	隶属程度	粗糙度	信度函数

续表

比较	概率理论	模糊集理论	粗糙集理论	证据理论
研究方法	概率分布函数	隶属函数	对象的分类	mass 函数
对知识的近似描述	概率	隶属度	上下近似集	信度函数
先验知识	需要	需要	不需要	需要
与普通集合的联系	概率	截集	上下近似	信度函数
计算方法	数学期望和方差	连续特征函数产生	粗糙度函数与上下近似	信度和似然函数

1.2 不确定性推理方法

1.2.1 主要的不确定性推理方法

不确定性推理的研究一直是人工智能研究领域的一个热点, 不确定性推理方法的选择很大程度上取决于不确定性建模的理论基础, 当前, 比较常用的不确定性推理方法主要有贝叶斯推理^[29]、模糊推理^[23]、证据推理^[17,27]等, 以下介绍几种主要的推理方法.

1) 定性推理

定性推理是通过对 (物理) 系统的结构、行为、功能及它们之间的关系和因果性进行研究, 以探索人类常识 (定性) 推理机制为目的, 从而有效地完成各项求解任务的一种跨领域的推理方法体系^[30]. 其研究源于对物理现象的研究, 逐步发展到机械机构、控制系统、社会经济, 直到现在的生物系统和认知模型; 研究定性推理的起源或诱因是现实世界的常识推理 (定性知识, 基于经验的知识), 特定领域的定性推理, 启发式知识, 动态与时态推理, 因果推理等.

2) 概率推理

人们根据不确定性信息作出推理和决策, 需要对各种结论的概率作出估计, 这类推理成为概率推理. Kahneman 和 Tversky 开辟了概率推理这一重要的研究领域^[31]. 所谓概率推理, 就是通过一些变量的信息来获得其他变量的概率信息的过程. 贝叶斯推理的问题是条件概率推理问题. 目前在概率推理的理论和应用方面有很多广泛深入的研究. 使用概率理论处理知识的不确定性的一种可视化的方法是贝叶斯网络.

3) 模糊推理

模糊逻辑是一种用逻辑表达式描述模糊集合中的隶属关系的方法, 概率论假定一个事件要么是正确要么是, 但是不能确定它的状态, 而模糊逻辑假定一个事件可以在不同程度上是正确的, 本质上, 模糊逻辑是经典二值逻辑的推广^[23]. 模糊集理论用语言变量和隶属度函数来建模语言不确定性. 模糊逻辑推理建立在模糊

集合论基础之上,通过模仿人的推理过程,传递模糊信息.其优点是:①不要求精确的信息;②对一些语言描述的问题很好求解;③模糊逻辑提供了一种从输入到输出空间的很好的推理方法;④可以处理不完全信息.其不足是:①由于隶属度函数是主观的,依赖于背景问题,所以很难构建合理的隶属度函数;②其不精确概念本身就有很多缺点;③缺少正式的对函数化构建规则的定义.

近年来,模糊推理技术的研究取得了不少成果,在方法上出现了合成推理规则、量化命题模糊推理、区间值模糊推理、格值模糊推理等各种改进方法,也出现了 BL 系统、Lukasiewicz 系统等成功的模糊推理系统.当然,模糊逻辑并非无懈可击,也存在很多不足和问题.

4) 基于规则的推理

20 世纪 70 年代中期发展起来的基于规则的 (rule-based) 的方法也被尝试用来处理不确定性.这种方法希望建立在基于规则的逻辑系统的基础上.与基于概率的方法不同的是,基于规则的方法中,知识工程师试图获得一个完整的推理过程.

5) 基于案例的推理

基于案例的推理来源于人类认知心理活动,属于类比推理方法和机器学习方法,是帮助人们获取实例样本性知识的手段.最早由 R. Schank 教授提出,是根据目标案例的提示而得到历史记忆中的源案例,并由源案例来指导目标案例求解的一种策略. CBR 系统通过回忆并修改过去曾解决过的相似问题及其解法来求解当前问题,符合人类求解问题的思维方式,是对传统的基于规则的推理系统的一个重要补充^[32].

6) 可信度推理

可信度是指人们对于一事物或现象的认同程度,这是一个根据经验主观确定的参数.基于可信度的不确定推理是 Shortliffe 等在确定性理论的基础上提出的一种不确定推理方法,并首先在医疗诊断专家系统中得到成功的应用.可信度有时称确认度,也称为可信度因子或确定性因子,反映了领域专家对不确定知识的信任程度,可信度以数量化的度量方式表示知识的可靠性,通常采用产生式规则表示不确定性知识^[32].

7) 粗糙集推理

基于粗糙集的推理首先是知识的获取与预处理,然后形成决策表经过约简后得到规则,再根据规则进行推理.粗糙集理论比较适合处理以数据表形式表示的信息,因此可以把客观世界抽象为一个知识表达系统,该系统中的基本成分就是所要研究的对象.将知识表达系统采用表格来实现,表列是属性,表行是对象,表中的某一行内容表示了系统中某个对象的所有信息.容易看出,一个属性对应一个等价关系,一个表可以看成是定义的族等价关系,即知识库.知识库和知识表达系统通过