



三支决策——基于可解释性研究视角

刘 盾^{1,2}, 叶晓庆³, 李天瑞³

(1. 西南交通大学 经济管理学院, 四川 成都 610031; 2. 服务科学与创新四川省重点实验室, 四川 成都 610031;
3. 西南交通大学 计算机与人工智能学院, 四川 成都 611756)

摘要 该文从可解释性研究视角探讨了三支决策的相关理论与方法。首先, 阐明了三代粗糙集模型的演化过程, 厘清三支决策的发展轨迹; 其次, 聚焦三支决策与粒计算的深度融合, 介绍了序贯三支决策和层次三支决策的相关理论与方法; 再次, 分析三支分类的有效性, 揭示了三支决策能够增强分类效果的内在原因; 最后, 从认知视角诠释了三支决策与人工智能融合的机理, 并给出了三支决策可解释性未来的研究方向。

关键词 三支决策; 粒计算; 粗糙集理论; 决策粗糙集

中图分类号: TP18 **DOI**: 10.16152/j.cnki.xdxbzr.2023-06-009

Three-way decision based on the interpretability research perspective

LIU Dun^{1,2}, YE Xiaoqing³, LI Tianrui³

(1. School of Economics and Management, Southwest Jiaotong University, Chengdu 610031, China;

2. Service Science and Innovation Key Laboratory of Sichuan Province, Chengdu 610031, China;

3. School of Computer and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China)

Abstract This paper mainly discusses the theories and methods with the interpretability research perspective. Firstly, we illuminate the evolutionary process on three generations of rough set models, and clarify the development track of three-way decision. Secondly, we focus on the deep integration between three-way decision and granular computing, and introduce the corresponding theories and methods of sequential three-way decision and hierarchical three-way decision. Thirdly, we analyze the effectiveness of three-way classification, and reveal the inner reason why the three-way decision can enhance the classification effect. Finally, we interpret the fusion mechanism between three-way decision and interpretive AI with a cognitive perspective, and give the research directions of interpretability three-way decision in future.

Keywords three-way decisions; granular computing; roughset theory; decision-theoretic rough sets

作为新一轮科技革命和产业变革的重要驱动力量, 人工智能正在对经济发展、社会进步、国际政治经济格局等诸多方面产生重大而深远的影响。

近年来, 人工智能技术在各行各业得到了快速的发展和广泛的应用。一方面, 它已潜移默化地融入到人们的衣食住行和生活的方方面面中,

收稿日期: 2023-09-30

基金项目: 国家自然科学基金(62276217, 61876157); 四川省科技计划项目(2022JDJQ0034); 中国博士后科学基金(2022M722629); 成都市哲学社会科学规划项目(2023CS124)。

第一作者: 刘盾, 男, 教授, 博士生导师, 从事数据挖掘、知识发现、粗糙集、粒计算等研究, newton83@163.com。

极大地推动了人们智慧化生活方式;另一方面,它也为人们在各领域的生产实践提供了强大的推力,进而催生了众多具有巨大商业价值的 AI 产业和行业。

可解释性问题一直以来是人工智能领域面临的挑战性问题之一。对于复杂决策问题,受数据源的多模态性、数据结构的多样性、数据内涵的复杂性等因素影响,传统的机器学习模型和算法已不能满足大数据决策环境的新需求。随着云计算、云存储、云服务等分布式和并行计算理论的快速发展与广泛普及,以及众多新兴机器学习方法的提出,都为人工智能研究提供了有力的技术支撑和有益的方法借鉴。然而,以深度学习为代表的机器学习模型与算法在执行和计算过程中犹如一个不透明的黑箱,人们只能观测数学模型的数据输入和结果输出,却不了解系统内部的运行机制和工作原理,更难理解最终决策的判断依据和背后机理。这种“说不清,道不明”“只观其果难觅其因”的现象极大地束缚了人工智能相关理论的落地,制约了人工智能相关技术的实施,也阻碍了人工智能与传统行业的进一步深度融合。

作为近二十年不确定性人工智能领域兴起的一种新兴数学工具和方法,粒计算(granular computing)理论通过在现实问题中对粒的抽象与描述、在模型构建中对粒的分析与转换、在算法设计中对粒的合成与分解来处理复杂决策问题,一经提出便引起了国内外研究人员的浓厚兴趣和广泛关注。一方面,粒计算基于“多层次”和“多视角”两个维度,借鉴“粒度粗化”和“粒度细化”两种思想,采用满意近似解替代精确解的策略来求解复杂决策问题,这能够降低问题求解过程中的计算复杂度,大大提升计算能力;另一方面,粒计算通过模拟人类思维方式和行为模式来对原始信息加以理解、概括、描述和操作,并利用信息粒化手段对基本粒子、粒层和粒结构进行抽象、描述和表示,使得决策者能够更加深刻地理解问题的本质特征,从而增强对决策过程的把控和决策结果的可解释性。可以看到,粒计算理论完美地契合了复杂问题求解的一般模式,是人类认知能力和智能思维的重要体现,这也为不确定性人工智能可解释性问题的解决提供了一种有效方法和可行途径。

三支决策(three-way decision)是加拿大学者 Yao Yiyu 教授在近十年来提出并发展起来的一种

处理不确定性决策的粒计算方法^[1]。该方法主要采用“化一为三”“化整为零”和“化繁为简”的策略来简化复杂问题,是一种符合人类认知的“三分而治”理论。三支决策的核心思想是通过 Trisecting-Acting-Outcome 模型(简称 TAO 模型,也称“分治效”模型)将整体(论域)分为 3 个子集或 3 个部分,并对不同子集或部分采取不同的决策行为或分治策略,最后,对相应的行为或策略得到的效果进行评价和反馈^[2]。进一步地,由于三支决策借鉴了粒计算的结构化哲学思维、结构化问题求解方法和结构化信息计算模式去解决复杂决策问题,能够深度诠释粒计算的内涵和外延,且在处理不确定性问题时表现出很好的有效性和可靠性,该方法已逐渐成为不确定人工智能研究领域的一种重要的研究方法。

基于上述分析,在对已有文献进行整理和分析的基础上,本文尝试从可解释性研究视角来系统性研究三支决策理论和方法。首先,从粗糙集理论出发,阐述三代粗糙集模型的演化过程和三支决策发展轨迹;其次,从粒计算理论出发,通过引入时间和空间维度,介绍序贯三支决策(sequential three-way decision)和层次三支决策(hierarchical three-way decision)相关理论与方法;再次,从分类视角出发,探讨三支分类问题,阐述三支决策在分类问题中的有效性;最后,从认知视角出发,深入讨论三支决策的哲学思想,并分析三支决策与可解释性人工智能融合的思路与途径。

1 三支决策与粗糙集理论

三支决策的最早思想来源于粗糙集理论。众所周知,粗糙集理论是波兰数学家 Pawlak 于 20 世纪 80 年代提出的,它是狭义三支决策最经典的模型。其核心思想是通过上近似集和下近似集将论域分为 3 个两两互不相交的子集,并分别命名为正域、边界域和负域,其基本定义如下。

定义 1 假设 U 是一个有限非空子集, $R \subseteq U \times U$ 为论域 U 上的等价关系。记 $\text{apr} = (U, R)$ 为一粗糙近似空间, U 可以通过该等价关系 R 划分成互不相交的子集,形成论域 U 上的一个划分 $U/R = \{[x]_R \mid x \in U\}$ ($[x]_R$ 简记为 $[x]$)。 $\forall X \subseteq U$, 其正域 $\text{POS}(X)$, 边界域 $\text{BND}(X)$ 和负域 $\text{NEG}(X)$ 分别定义为

$$\text{POS}(X) = \{x \in U \mid [x] \subseteq X\};$$

$$\text{BND}(X) = \{x \in U \mid [x] \not\subseteq X \text{ 且 } [x] \cap X \neq \emptyset\};$$

$$\text{NEG}(X) = \{x \in U \mid [x] \cap X = \emptyset\}。(1)$$

可以看到,Pawlak 粗糙集是粗糙集理论最经典的形式,而从 3 个区域推导的决策规则自然形成了三支决策的核心思想,即从正域里获取的正规则用来接受某事物;从负域里获取的负规则用来表示拒绝某事物;落在边界域上的规则表示延迟决策。显而易见,Pawlak 粗糙集存在一定的局限性,因为它没有考虑决策规则的容错性。基于此,概率粗糙集随之被提出来以改进 Pawlak 粗糙集的缺陷,其主要通过两个阈值 α 和 $\beta(0 \leq \beta < \alpha \leq 1)$ 重新定义 3 个决策区域。

定义 2 假设 $S = (U, R)$ 为一信息系统, $\forall X \subseteq U$, 令 $0 \leq \beta < \alpha \leq 1$, 则概率粗糙集的正域 $\text{POS}_{(\alpha, \beta)}(X)$, 边界域 $\text{BND}_{(\alpha, \beta)}(X)$ 和负域 $\text{NEG}_{(\alpha, \beta)}(X)$ 可分别定义为

$$\text{POS}_{(\alpha, \beta)}(X) = \{x \in U \mid \Pr(X \mid [x]) \geq \alpha\};$$

$$\text{BND}_{(\alpha, \beta)}(X) = \{x \in U \mid \beta < \Pr(X \mid [x]) < \alpha\};$$

$$\text{NEG}_{(\alpha, \beta)}(X) = \{x \in U \mid \Pr(X \mid [x]) \leq \beta\}。(1)$$

式中: $\Pr(X \mid [x]) = |[x] \cap X| / |[x]|$ 表示条件概率; $|\cdot|$ 表示集合的基数。显然地,当 $\alpha = 1, \beta = 0$ 时,概率粗糙集退化成 Pawlak 粗糙集。进一步,如果把 Pawlak 粗糙集视为第 I 代三支决策模型的话,概率粗糙集则为第 II 代三支决策模型。

然而,概率粗糙集仅对不确定性模型的语法进行了描述,而没有对模型的语义进行解释,这表现为两个阈值 α 和 β 的取值都是人为事先给定的。为了更清晰地阐明粗糙集理论的可解释性,Yao 在 20 世纪 90 年代提出了决策粗糙集理论,该理论可以看成是第 III 代三支决策模型^[3]。决策粗糙集主要通过引入贝叶斯决策过程,来实现阈值的自动获取,其基本思想如下。

考虑贝叶斯决策的一个特殊模型,它由 2 个状态集和 3 个行动集构成。其中,状态集 $\Omega = \{X, \neg X\}$ 分别表示某事件属于 X 和不属于 X ,行动集 $A = \{a_P, a_B, a_N\}$ 分别表示接受某事件、延迟决策和拒绝某事件 3 种行动。对于 $x \in U$,考虑到采取不同行动会产生不同的损失,记 $\lambda_{PP}, \lambda_{BP}, \lambda_{NP}$ 分别表示当 x 属于 X 时,采取行动 a_P, a_B 和 a_N 下的损失;类似地,记 $\lambda_{PN}, \lambda_{BN}, \lambda_{NN}$ 分别表示当 x 不属于 X 时,采取行动 a_P, a_B 和 a_N 下的损失。根据贝叶斯

决策理论,采取 a_P, a_B 和 a_N 3 种行动下的期望损失可分别表示为

$$R(a_P \mid [x]) = \lambda_{PP}\Pr(X \mid [x]) + \lambda_{PN}\Pr(\neg X \mid [x]);$$

$$R(a_B \mid [x]) = \lambda_{BP}\Pr(X \mid [x]) + \lambda_{BN}\Pr(\neg X \mid [x]);$$

$$R(a_N \mid [x]) = \lambda_{NP}\Pr(X \mid [x]) + \lambda_{NN}\Pr(\neg X \mid [x])。(2)$$

根据贝叶斯决策准则,需要选择期望损失最小的行动集作为最佳行动方案,可以得到以下 3 条决策规则。

(P):若 $R(a_P \mid [x]) \leq R(a_B \mid [x])$ 和 $R(a_P \mid [x]) \leq R(a_N \mid [x])$ 同时成立,则 $x \in \text{POS}(X)$;

(B):若 $R(a_B \mid [x]) \leq R(a_P \mid [x])$ 和 $R(a_B \mid [x]) \leq R(a_N \mid [x])$ 同时成立,则 $x \in \text{BND}(X)$;

(N):若 $R(a_N \mid [x]) \leq R(a_P \mid [x])$ 和 $R(a_N \mid [x]) \leq R(a_B \mid [x])$ 同时成立,则 $x \in \text{NEG}(X)$ 。

由于 $\Pr(X \mid [x]) + \Pr(\neg X \mid [x]) = 1$,则上述 3 条规则只与条件概率和损失函数这两个因素有关。此外,考虑到接受正确事物的损失不大于延迟接受正确事物的损失,且这两者都小于拒绝正确事物的损失;拒绝错误事物的损失不大于延迟拒绝错误事物的损失,且这两者都小于接受错误事物的损失。有 $\lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}, \lambda_{NN} \leq \lambda_{BN} < \lambda_{PN}$ 。据此,通过联立求解 (P) - (N),可以得到如下结果。

$$\alpha = \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})},$$
$$\beta = \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})},$$
$$\gamma = \frac{(\lambda_{PN} - \lambda_{NN})}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}。(3)$$

显然地,由规则 (B) 可知,不等式 $\alpha > \beta$ 成立,即

$$\frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})} > \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}。$$

进一步地,有

$$\frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})} >$$

$$\frac{(\lambda_{PN} - \lambda_{NN})}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})} > \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}$$

成立,即 $0 \leq \beta < \gamma < \alpha \leq 1$ 成立。基于上述分析,可以重写(P) - (N)的决策规则形式如下。

(P1): 如果 $\Pr(X \mid [x]) \geq \alpha$, 则 $x \in \text{POS}(X)$;

(B1): 如果 $\beta < \Pr(X \mid [x]) < \alpha$, 则 $x \in \text{BND}(X)$;

(N1): 如果 $\Pr(X \mid [x]) \leq \beta$, 则 $x \in \text{NEG}(X)$ 。

显然地,规则(P1) - (N1)给出了论域中每个样本 x 的数学判定条件,满足规则(P1)的样本划分到 X 的正域中,同时做出立即执行决策;满足规则(N1)的样本划分到 X 的负域中,同时做出不执行决策;满足规则(B1)的样本由于决策依据不足,暂缓做出决策。可以看到,决策粗糙集为三支决策提供了一种直观的三划分准则。此外,决策粗糙集为三支决策的两个阈值计算提供了一种具有可解释性的计算方法,即 α 和 β 的取值不是人为主观设定的,其阈值取值与决策损失函数密切相关。正如文献[4]所提到的观点:决策粗糙集是狭义三支决策的典型代表,它既是三支决策研究的原点,也是从语义视角去诠释三支决策可解释性的一次成功尝试。总而言之,从Pawlak粗糙集到概率粗糙集,再到决策粗糙集,粗糙集理论经过三代演化,人们越来越意识到语义解释对于理论模型的重要性,三支决策的可解释性问题也逐渐成为粒计算研究一个新的关注点。

2 三支决策与粒计算方法

为了进一步阐明三支决策的过程和机理,Yao在2018年提出了一个三支决策的一般化模型——“分治效”模型(trisecting-acting-outcome模型,简称TAO模型)^[2]。该模型首次跳出了传统三支决策,只关注粗糙集的狭义研究模式,是广义三支决策最具代表性的理论模型。TAO模型的基本思想如图1所示。

在图1中,TAO模型中将决策过程分为3个阶段:“三分”阶段将论域 U 分为3个决策区域 U_1 、 U_2 和 U_3 ;“治略”阶段针对不同的决策区域采取不同的策略 S_1 、 S_2 和 S_3 ;“效果”阶段主要是

对于不同的策略采取相应的行动 a_1 、 a_2 和 a_3 ,最终得到3种不同的评价结果 O_1 、 O_2 和 O_3 。在TAO模型中,“三分”是“治略”的前提,“治略”是“三分”的目的,而“效果”是“三分”和“治略”的监督保障。对于不同的决策背景,可以根据实际情况来确立对应的划分机理、执行策略和评价效果^[5]。

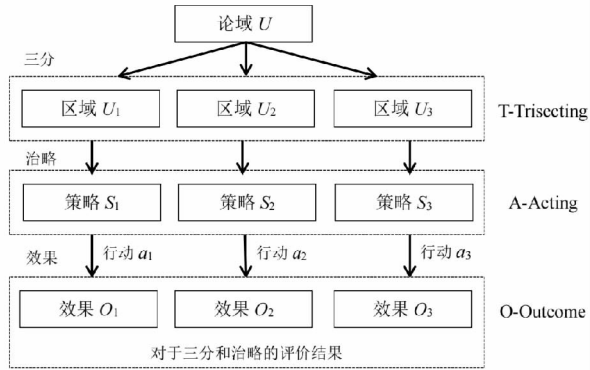


图 1 TAO 模型的一般研究框架

Fig. 1 The research framework of TAO model

进一步地,三支决策对于TAO模型决策机理的探索可分为两个层面。一是关注决策过程的动态性特征。例如:在论文审稿过程中,作者从投稿到接收可能会经历多次修改,而每次修订提交后又可能面临接收、继续修改和拒绝三种结果,如此反复直至完全接收或拒绝,这体现了一种序贯三支决策思想。二是关注决策过程的粒结构特征。例如:在疫情防控过程中,人们可以从宏观搜集全国每日新增确诊数据,接着细化了解各省每日新增数据,然后再细化获知各县市实时数据,直至掌握所住街道/小区的详细数据来判断身边所处疫情情况。如果仅以最粗粒度全国数据来决定某街道/小区的防疫措施,则难以作出微观准确有效的决策,反之亦然。上述两个层面反映了三支决策的时间属性和空间属性。时间属性注重动态决策环境下的“时序性”分析;空间属性则侧重粒度变化情形下的“层次性”建模。进一步地,在很多实际决策问题中,时间和空间两种属性往往是相生相伴、相互依存的。比如论文从投稿到最终被接收或拒绝自然形成一种粒空间结构;而将每天不同粒层的疫情数据汇总起来自然形成一个时序序列。因此,研究三支决策的时空属性自然成为了探索三支决策可解释性的重要手段。

对于三支决策时空属性的研究,粒计算是一种重要的理论和方法^[6]。众所周知,粒计算是当

前智能信息处理领域的一种新的计算范式,它模拟了人类对于世界的多粒度表示、多粒度处理和多粒度决策机制,可看作为一种新的广义机器学习模型和方法。信息粒化是粒计算在处理复杂问题的一种有效手段。通过粒化,人类可以从多个层次和多个视角去认知、描述和理解客观世界和主观世界。进一步地,从知识学习视角出发,对某一事物而言,如果将其知识粒度不断粗化,人们对该事物的认识就越模糊;反之,如果将其知识粒度不断细化,则人们对该事物的认识也就越清晰。可以看到,事物自身知识粒度的动态变化过程反映了人们对事物的认知过程,这与现实世界的结构、人们的渐进式思维模式及行为方式是高度契合的。基于上述分析,可以从“面向粒计算的序贯三支决策”和“面向粒计算的层次三支决策”两个方面来探讨三支决策的时空性问题。

2.1 面向粒计算的序贯三支决策理论与方法

序贯三支决策主要考虑决策过程中的动态特征,它体现了三支决策的时间属性。具体而言,序贯三支决策需要考虑不同时刻2个状态集和3个行动集的决策情形。假设 $\Omega^{(t)} = \{X, \neg X\}$ 表示第 t 时刻属于事件 X 和不属于事件 $\neg X$ 两个状态, $A^{(t)} = \{a_P, a_B, a_N\}$ 分别表示第 t 时刻接受某事件、延迟决策和拒绝某事件3种行动。考虑到采取不同行动会产生不同的损失,记 $\lambda_{PP}^{(t)}, \lambda_{BP}^{(t)}, \lambda_{NP}^{(t)}$ 分别表示当 $x^{(t)}$ 属于 X 时,采取行动 a_P, a_B 和 a_N 时的损失;记 $\lambda_{PN}^{(t)}, \lambda_{BN}^{(t)}, \lambda_{NN}^{(t)}$ 分别表示当 $x^{(t)}$ 不属于 X 时,采取行动 a_P, a_B 和 a_N 时的损失。假设第 t 时刻的损失函数满足数学条件 $\lambda_{PP}^{(t)} \leq \lambda_{BP}^{(t)} < \lambda_{NP}^{(t)}$ 和 $\lambda_{NN}^{(t)} \leq \lambda_{BN}^{(t)} < \lambda_{PN}^{(t)}$ 。令

$$\alpha^{(t)} = \frac{(\lambda_{PN}^{(t)} - \lambda_{BN}^{(t)})}{(\lambda_{PN}^{(t)} - \lambda_{BN}^{(t)}) + (\lambda_{BP}^{(t)} - \lambda_{PP}^{(t)})},$$

$$\beta^{(t)} = \frac{(\lambda_{BN}^{(t)} - \lambda_{NN}^{(t)})}{(\lambda_{BN}^{(t)} - \lambda_{NN}^{(t)}) + (\lambda_{NP}^{(t)} - \lambda_{BP}^{(t)})},$$

可以得到序贯三支决策的规则判定条件。

(P2): 如果 $\Pr(X | x^{(t)}) \geq \alpha^{(t)}$, 则 $x^{(t)} \in \text{POS}^{(t)}(X)$;

(B2): 如果 $\beta^{(t)} < \Pr(X | x^{(t)}) < \alpha^{(t)}$, 则 $x^{(t)} \in \text{BND}^{(t)}(X)$;

(N2): 如果 $\Pr(X | x^{(t)}) \leq \beta^{(t)}$, 则 $x^{(t)} \in \text{NEG}^{(t)}(X)$ 。

基于上述基本模型,国内外众多学者对序贯三支决策理论和方法作了大量研究。Yao 和 Deng 最早讨论了序贯三支决策问题,他们针对延

迟决策问题,将静态三支决策推广到了一个动态序列,并采用粒计算思想提出一种序贯三支决策模型^[7]。Liu 等考虑了不同损失函数变化情形下的动态决策粗糙集模型^[8]。Gao 和 Yao 构建了一种考虑行动策略的序贯三支决策模型^[9]。Jiang 等讨论了一种基于移动的三支决策模型,通过使用比例效用函数来度量三支决策的有效性^[10]。Yang 等从增量学习的视角探讨了多粒度增量学习下的序贯三支决策统一模型^[11]。Xu 等提出了一个基于概率粗糙集的流计算三支决策模型,并将其应用到上下近似集的动态增量学习过程中^[12]。Qian 等提出了一个基于多粒度的序贯三支决策研究框架,详细分析了5种多粒度序贯三支决策模型的联系与区别^[13]。Fang 等建立了一个粒度驱动的序贯三支决策模型来处理动态决策过程中的决策代价和决策成本问题^[14]。Zhang 等提出了一种使用序贯三支决策的成本敏感组合技术 S3WC 来解决情绪分类问题^[15]。Wu 等设计了一种无重叠的三支序列模式挖掘方法,该方法不仅能够挖掘出冗余信息,而且能够高效提取对用户有价值的信息^[16]。进一步地, Li 等提出一种基于序贯三支决策的代价敏感人脸识别模型,来解决错分类代价不平衡和图像信息不充分等问题^[17],并基于深度神经网络探讨了图像特征粒度提取与序贯识别问题,取得了较好的实验效果^[18]。Zhang 等构建了一种序贯三支增强卷积神经网络模型,用以提高分类精度和分类性能^[19]。Savchenko 将序贯三支决策的思想引入到卷积神经网络中,通过在神经网络的不同粒层上提取图像的序贯特征,来实现图像的表征和加速推理^[20]。

2.2 面向粒计算的层次三支决策理论与方法

层次三支决策主要考虑决策过程中的多层级结构特征,它体现了三支决策的空间属性。与序贯三支决策模型类似,层次三支决策同样需要考虑第 s 粒层的2个状态 $\Omega^{(s)} = \{X, \neg X\}$ 和3种行动 $A^{(s)} = \{a_P, a_B, a_N\}$ 。记 $\lambda_{PP}^{(s)}, \lambda_{BP}^{(s)}, \lambda_{NP}^{(s)}$ 分别表示当 $x^{(s)}$ 属于 X 时,采取行动 a_P, a_B 和 a_N 时的损失;记 $\lambda_{PN}^{(s)}, \lambda_{BN}^{(s)}, \lambda_{NN}^{(s)}$ 分别表示当 $x^{(s)}$ 不属于 X 时,采取行动 a_P, a_B 和 a_N 时的损失。假设第 s 粒层的损失函数满足数学条件 $\lambda_{PP}^{(s)} \leq \lambda_{BP}^{(s)} < \lambda_{NP}^{(s)}$ 和 $\lambda_{NN}^{(s)} \leq \lambda_{BN}^{(s)} < \lambda_{PN}^{(s)}$ 。令

$$\alpha^{(s)} = \frac{(\lambda_{PN}^{(s)} - \lambda_{BN}^{(s)})}{(\lambda_{PN}^{(s)} - \lambda_{BN}^{(s)}) + (\lambda_{BP}^{(s)} - \lambda_{PP}^{(s)})},$$

$$\beta^{(s)} = \frac{(\lambda_{BN}^{(s)} - \lambda_{NN}^{(s)})}{(\lambda_{BN}^{(s)} - \lambda_{NN}^{(s)}) + (\lambda_{NP}^{(s)} - \lambda_{BP}^{(s)})},$$

可以得到相应层次三支决策的规则判定条件。

(P3): 如果 $\Pr(X \mid x^{(s)}) \geq \alpha^{(s)}$, 则 $x^{(s)} \in \text{POS}^{(s)}(X)$;

(B3): 如果 $\beta^{(s)} < \Pr(X \mid x^{(s)}) < \alpha^{(s)}$, 则 $x^{(s)} \in \text{BND}^{(s)}(X)$;

(N3): 如果 $\Pr(X \mid x^{(s)}) \leq \beta^{(s)}$, 则 $x^{(s)} \in \text{NEG}^{(s)}(X)$ 。

进一步地,可认为层次三支决策主要采用了多粒度计算和多尺度计算两种方法。

对于多粒度三支决策,Qian 等从信息融合视角出发,提出了多粒度粗糙集模型,通过粒化获得的信息粒集和多个粒结构进行数据分析,并从挖掘的知识中融合集成有效的决策知识^[21]。苗夺谦等提出了粒度空间的三层模型概念,探讨了覆盖、基和粒结构在不同粒度空间下的关系^[22]。王国胤等从粒计算模型的角度详细探讨了模糊集、粗糙集、商空间理论模型及其他扩展粒计算模型中知识的不确定性问题^[23]。Zhang 和 Miao 从粗糙集正域、边界域和负域三个层次出发对三支决策约简作了系统性研究^[24]。Yue 等讨论了基于模糊粗糙变换的阴影邻域三支分类问题^[25]。Chen 等构建了一种基于三支决策的自适应分层特征表示模型来处理边界域问题,并通过自适应学习方法来搜寻边界域处理的最优粒度空间^[26]。Yang 等人采用自顶向下的造粒方法构建了多层次复合粒结构,并建立了一种基于不同粒度融合的三支决策研究框架,以解决粒计算中的数据融合问题^[27]。

对于多尺度三支决策,Wu 和 Leung 提出多尺度决策表等概念,并从粒计算角度研究了多尺度决策表中最优尺度的选择问题^[28]。Li 和 Hu 讨论了多尺度问题中最优尺度的选择问题^[29]。Hao 等人提出了基于多尺度信息表的序贯三支决策模型,以解决动态多尺度决策中的最优尺度选择问题^[30]。Huang 等人构建了从多尺度直觉模糊信息表中获取信息的多粒度决策粗糙集方法,并进一步探讨了最优尺度选择和约简问题^[31]。Luo 等通过对属性值分类进行细化和粗化,提出了在不完全多尺度信息系统中的决策规则增量更新方法,为增量三支决策提供了理论依据和语义解释^[32]。Zhang 等人建立了基于尺度空间的序贯三支决策模型,设计了结合三支决策和 Hasse 图的

最优尺度组合选择算法,用以提高多尺度决策系统的搜索效率^[33]。Yu 等探讨了三支决策的聚类分析模型和聚类数的自动学习算法^[34]。Shen 等研究了多层次分类视角下三支决策的分块约简问题^[35]。She 等讨论了多尺度决策表的泛化约简方法^[36]。Huang 和 Zhan 从多尺度视角出发,提出了一种基于后悔理论的三支决策模型,旨在解决肝病诊断中患者的分类和排序问题^[37]。

可以看到,国内外众多学者分别从不同的研究视角对三支决策理论和粒计算方法作了大量系统性研究,并在多个方面取得了一系列重要成果,这极大地推动了可解释性三支决策与粒计算研究领域的发展。

3 三支决策与三支分类

本小节在前面分析的基础上,进一步从分类视角来阐明三支决策的有效性。

首先,以决策粗糙集为研究对象,将二支决策模型作为参照基准,来验证三支决策的有效性。值得一提的是,二支决策是式(3)中当 $\beta \geq \alpha$ 成立时的情形。在二支决策中,参数 γ 成为划分决策区域的关键因素,它将论域划分为两个区域。此时,决策规则(P) - (N) 的退化为

(P4): 如果 $\Pr(X \mid [x]) \geq \gamma$, 则 $x \in \text{POS}(X)$;

(N4): 如果 $\Pr(X \mid [x]) < \gamma$, 则 $x \in \text{NEG}(X)$ 。

为了更清晰地说明问题,下面引入混淆矩阵的概念。最经典的混淆矩阵用于解决二分类问题,表 1 是一个 2×2 的混淆矩阵,其行代表预测类别,其列代表真实类别。

表 1 混淆矩阵
Tab. 1 The confusion matrix

混淆矩阵	真实类别	
	P (正)	N (负)
	T (真)	TN
预测类别	F (假)	FN

在表 1 中,TP(true positive, 真阳性)表示样本的真实类别是正例,并且模型预测的结果也是正例的情形;TN (true negative, 真阴性)表示样本的真实类别是负例,并且模型将其预测为负例

的情形;FP (false positive, 假阳性)表示样本的真实类别是负例,而模型将其预测为正例的情形;FN (false negative, 假阴性)表示样本的真实类别是正例,而模型将其预测为负例的情形。根据表 1,可以定义评价分类模型优劣的 4 个测度:准确率 (precision)、召回率 (recall)、精度 (accuracy)和 F_1 值。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}};$$
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}};$$
$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}};$$
$$F_1 = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}。 \tag{4}$$

将表 1 中的混淆矩阵推广到二支决策和三支决策中,可以分别定义它们在不同行动下的混淆矩阵,如表 2 所示。

表 2 二支决策与三支决策的混淆矩阵

Tab.2 The confusion matrix of two-way decision and three-way decision

决策类型	行动	状态 $P(X)$	状态 $N(\neg X)$
二支决策	a_P	$n_{PP}^{(2)}$	$n_{PN}^{(2)}$
	a_N	$n_{NP}^{(2)}$	$n_{NN}^{(2)}$
三支决策	a_P	$n_{PP}^{(3)}$	$n_{PN}^{(3)}$
	a_B	$n_{BP}^{(3)}$	$n_{BN}^{(3)}$
	a_N	$n_{NP}^{(3)}$	$n_{NN}^{(3)}$

在表 2 中, $n_{PP}^{(2)}$ 、 $n_{PN}^{(2)}$ 、 $n_{NP}^{(2)}$ 和 $n_{NN}^{(2)}$ 分别表示在两种状态下,采取两种不同行动所对应的的样本数;类似地, $n_{PP}^{(3)}$ 、 $n_{PN}^{(3)}$ 、 $n_{BP}^{(3)}$ 、 $n_{BN}^{(3)}$ 、 $n_{NP}^{(3)}$ 和 $n_{NN}^{(3)}$ 分别表示在两种状态下,采取 3 种不同行动所对应的的样本数,令 U 表示论域中的所有样本。据此,可以定义二支决策和三支决策在式 (4) 中不同测度的数学表示形式。

$$\text{Precision}^{(2)} = \frac{n_{PP}^{(2)}}{n_{PP}^{(2)} + n_{PN}^{(2)}};$$
$$\text{Recall}^{(2)} = \frac{n_{PP}^{(2)}}{n_{PP}^{(2)} + n_{NP}^{(2)}};$$
$$\text{Accuracy}^{(2)} = \frac{n_{PP}^{(2)} + n_{NN}^{(2)}}{|U|};$$
$$F_1^{(2)} = \frac{2n_{PP}^{(2)}}{2n_{PP}^{(2)} + n_{PN}^{(2)} + n_{NP}^{(2)}};$$

$$\text{Precision}^{(3)} = \frac{n_{PP}^{(3)}}{n_{PP}^{(3)} + n_{PN}^{(3)}};$$
$$\text{Recall}^{(3)} = \frac{n_{PP}^{(3)}}{n_{PP}^{(3)} + n_{BP}^{(3)} + n_{NP}^{(3)}};$$
$$\text{Accuracy}^{(3)} = \frac{n_{PP}^{(3)} + n_{NN}^{(3)}}{|U|};$$
$$F_1^{(3)} = \frac{2n_{PP}^{(3)}}{2n_{PP}^{(3)} + n_{PN}^{(3)} + n_{BP}^{(3)} + n_{NP}^{(3)}}。 \tag{5}$$

接下来,分别定义二支决策和三支决策的损失函数矩阵,如表 3 所示。

表 3 二支决策与三支决策的损失函数矩阵

Tab.3 The loss function matrix of two-way decision and three-way decision

决策类型	行动	状态 $P(X)$	状态 $N(\neg X)$
二支决策	a_P	$\lambda_{PP}^{(2)}$	$\lambda_{PN}^{(2)}$
	a_N	$\lambda_{NP}^{(2)}$	$\lambda_{NN}^{(2)}$
三支决策	a_P	$\lambda_{PP}^{(3)}$	$\lambda_{PN}^{(3)}$
	a_B	$\lambda_{BP}^{(3)}$	$\lambda_{BN}^{(3)}$
	a_N	$\lambda_{NP}^{(3)}$	$\lambda_{NN}^{(3)}$

根据表 2 和表 3,可以分别计算二支决策和三支决策的总体损失为

$$\text{cost}_{\text{DEC}}^{(2)} = \begin{pmatrix} \lambda_{PP}^{(2)} & \lambda_{PN}^{(2)} \\ \lambda_{NP}^{(2)} & \lambda_{NN}^{(2)} \end{pmatrix} \otimes \begin{pmatrix} n_{PP}^{(2)} & n_{NP}^{(2)} \\ n_{PN}^{(2)} & n_{NN}^{(2)} \end{pmatrix} =$$
$$\lambda_{PP}^{(2)} \times n_{PP}^{(2)} + \lambda_{PN}^{(2)} \times n_{PN}^{(2)} + \lambda_{NP}^{(2)} \times n_{NP}^{(2)} +$$
$$\lambda_{NN}^{(2)} \times n_{NN}^{(2)} \tag{6}$$

$$\text{cost}_{\text{DEC}}^{(3)} = \begin{pmatrix} \lambda_{NP}^{(3)} & \lambda_{PN}^{(3)} \\ \lambda_{BP}^{(3)} & \lambda_{BN}^{(3)} \\ \lambda_{PP}^{(3)} & \lambda_{NN}^{(3)} \end{pmatrix} \otimes \begin{pmatrix} n_{NP}^{(3)} & n_{BP}^{(3)} & n_{PP}^{(3)} \\ n_{PN}^{(3)} & n_{BN}^{(3)} & n_{NN}^{(3)} \end{pmatrix} =$$
$$\lambda_{NP}^{(3)} \times n_{NP}^{(3)} + \lambda_{PN}^{(3)} \times n_{PN}^{(3)} + \lambda_{BP}^{(3)} \times n_{BP}^{(3)} +$$
$$\lambda_{BN}^{(3)} \times n_{BN}^{(3)} + \lambda_{PP}^{(3)} \times n_{PP}^{(3)} + \lambda_{NN}^{(3)} \times n_{NN}^{(3)} \tag{7}$$

进一步地,在上述分类问题中,假设在把正例预测为真,以及把负例预测为假(即预测类别和真实类别一致)这两种情形下,都不会产生决策损失。有 $\lambda_{PP}^{(2)} = \lambda_{NN}^{(2)} = 0$, $\lambda_{PP}^{(3)} = \lambda_{NN}^{(3)} = 0$,则式(6)和式(7)可以简化为

$$\text{cost}_{\text{DEC}}^{(2)} = \lambda_{PN}^{(2)} \times n_{PN}^{(2)} + \lambda_{NP}^{(2)} \times n_{NP}^{(2)} \tag{8}$$

误分类损失

$$\text{cost}_{\text{DEC}}^{(3)} = \lambda_{PN}^{(3)} \times n_{PN}^{(3)} + \lambda_{NP}^{(3)} \times n_{NP}^{(3)} +$$

误分类损失

$$\lambda_{BP}^{(3)} \times n_{BP}^{(3)} + \lambda_{BN}^{(3)} \times n_{BN}^{(3)} \tag{9}$$

延迟损失

可以看到,相对于二支决策,三支决策引入了延迟决策策略,其损失是由误分类损失和延迟损失两部分构成。进一步地,要使得式(8)的总体损失最小,即是要分别去求解阈值 γ^* ,使得 $\text{cost}_{\text{DEC}}^{(2)}$ 最小;同理,要使得式(9)的总体损失最小,即是要去求解阈值 α^* 和 β^* ,使得 $\text{cost}_{\text{DEC}}^{(3)}$ 最小。

$$\begin{aligned} \min \text{cost}_{\text{DEC}}^{(2)} &= \underset{\gamma^*}{\operatorname{argmin}} (\lambda_{PN}^{(2)} \times n_{PN}^{(2)} + \\ &\quad \lambda_{NP}^{(2)} \times n_{NP}^{(2)}); \\ \min \text{cost}_{\text{DEC}}^{(3)} &= \underset{(\alpha^*, \beta^*)}{\operatorname{argmin}} (\lambda_{PN}^{(3)} \times n_{PN}^{(3)} + \\ &\quad \lambda_{NP}^{(3)} \times n_{NP}^{(3)} + \lambda_{BP}^{(3)} \times n_{BP}^{(3)} + \\ &\quad \lambda_{BN}^{(3)} \times n_{BN}^{(3)}). \end{aligned}$$

为了解决上述问题,Liu 讨论了在两种样本分布情形下,三支决策在不同测度和总体损失优于二支决策的判定条件,并通过大量实验验证了在不同数据集和不同损失函数设定情形下,三支决策在分类问题中的有效性^[38]。此外,Yao^[39]和刘盾等^[40]分别深入探讨了在 $[0, \beta]$ 、 (β, γ) 、 $[\gamma, \alpha]$ 和 $(\alpha, 1]$ 4个决策区间产生误分类损失(包括拒真错误造成的损失和采伪错误造成的损失)和延迟损失产生的原因,详细分析了概率粗糙集(第Ⅱ代三支决策模型)优于Pawlak粗糙集(第Ⅰ代三支决策模型)和 γ -二支决策的数学机理,并给出了严格的数学证明。此外,Yin等基于概率论的视角,深刻阐述了三支决策优于二支决策的内在机理^[41]。上述研究都尝试从可解释性的视角去探讨三支决策在分类问题中的优势,这本身也推动了三支决策在机器学习领域中的发展。

4 三支决策与认知计算

认识的本质是主体对客体的能动反映,它是动态的、可计算的,能够依据实际任务知识化、智能化、自主化地呈现出“感知—分析—决策—执行”的递进行为。一般而言,人类的认知规律主要体现在两个方面:一是通过认知将复杂事物简单化;二是人类处理信息的能力是有限的。为了更好地刻画人类的认知规律,Yao系统地研究了基于认知计算的三支决策理论,并指出认知计算是研究可解释性人工智能的关键技术和三支决策发展的重要方向^[42]。

最早对三支决策与认知计算融合的思考来源于粒计算的三元论^[43]。该理论认为粒计算的核心由粒子、粒层和粒结构3部分组成。①粒子是

一个最基本的抽象概念,它可以通过分解/聚合得到一个粒化/抽象的理解和描述,粒本身具有内部、外部和环境依赖3个特征。②粒层是由相同性质或相同大小的粒子构成,每一层次由多个相互影响、相互关联的粒子构成,不同粒层给出不同粒度或尺度的描述。③粒结构将粒子和粒层有序地组织起来,其形成的多层次结构可以给定粒度的全局描述,它体现了粒计算所倡导的多粒度、多尺度、多层次、多视角的结构化方法。基于此,粒计算的三元模型以粒结构为基础,包括3个要素:结构化哲学思维、结构化问题求解方法和结构化信息计算模式,它们构成了粒计算三角形。进一步地,针对每一个要素,Yao给出了一些具体的思路和方法。他认为结构化思维可以从“还原论、系统论、层次论”3种理论去诠释粒计算哲学思想;结构化问题求解可以从“自顶向下、自底向上、自中向外”3种策略去设计粒计算处理算法;结构化信息处理可以从“Pandemonium架构、渐进式计算、序贯决策”3种方式去揭示粒计算运行机制^[43]。粒计算三元论的核心思想如图2所示。

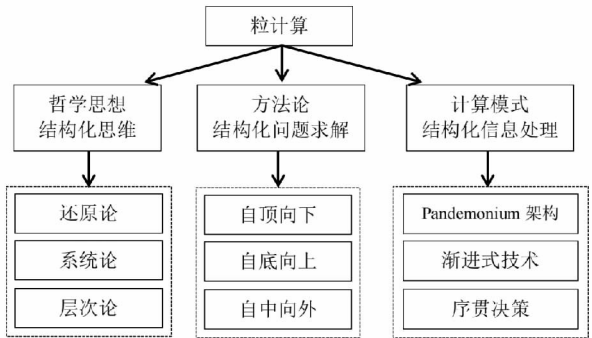


图2 粒计算的三元论

Fig. 2 The triarchic theory of granular computing

随着大数据时代的来临,数据科学和认知科学已成为近年来快速增长的研究热点领域。三支决策与数据科学、认知科学的融合也成为国内外学者重点关注的研究方向。基于此,Yao在符号/形式、意义/内容和价值/效用3个层面上给出了数据科学研究的三级框架,提出了SMV空间(symbols-meaning-value space,“符号-意义-价值”空间)的概念^[44]。该框架借助计算机模拟人类的感知、注意力和学习等思维过程,从信息科学、认知科学和计算机科学不同视角出发,实现对思考、记忆等人脑认知过程的模拟,最终描述、理解和构建了“我们感知什么和如何感知”“我们知道什么和如何知道”,以及“我们做什么和如何做”3个根

本性问题,并明确指出三支决策是一类有效的认知学习启发式方法,解释了三支决策的 3 种处理模式,即“自下而上”的数据驱动,“自上而下”的决策驱动,以及“自中而外”的知识型方法。具体而言,SMV 空间可从以下 3 个视角去理解、诠释和分析。

4.1 基于信息科学视角下的 SMV 空间

从信息科学视角出发,SMV 空间主要体现为 DIKW (data-information-knowledge-wisdom,数据-信息-知识-智慧) 层级模型。该模型构建了一个金字塔形的多层次知识结构。其中,数据层是最基本的层次,它包含了最底层的知识;信息层加入了内容,人们可以分析数据间的关系来获得信息;知识层需要思考“如何去使用”,在实际决策中使用信息产生了知识;智慧层体现了“什么时候才用”,它包含了最高层的知识。Yao 利用三支决策思想,将 DIKW 层级模型中的信息层和知识层合并提出了一种广义的 DKW 层级模型^[45]。在该模型中,数据是输入层(最底层),是未经处理的原材料;信息和知识被组合成一个单一的中间层,它们可以从分析数据中获得;而智慧是输出层(最高层),用于指导人们科学地使用知识。

4.2 基于认知科学视角下的 SMV 空间

从认知科学视角出发,Yao 在 DKW 三层级模型的基础上做了更进一步的思考,认为 SMV 空间体现为观察、知悉和行动(seeing, knowing and doing) 3 个步骤。从本质而言,SMV 空间和 DKW 层级模型可视为一个 IPO (input-process-output,输入-处理-输出) 三支决策框架,它实现了机器/系统与人的联系与交互。一方面,从机器/系统视角,可以构建一个基于 CAD (collection-analysis-decision,采集-分析-决策) 的三层次模型来整合数据的搜集、分析和决策。人们通过数据采集进行输入;借鉴数据分析产生信息和知识;利用决策最终形成智慧。另一方面,从人的视角,DKW 层级模型可以形成一个 PCA (perception-cognition-action,感知-认知-行动) 三层级结构,感知层关注信号(符号),负责数据的搜集和输入;认知层关注信息和知识(意义),负责将数据转化为有用知识;行动层关注价值,负责结果的输出。为了更加清晰地阐述上述思想,Yao 构建一个 PCA 模型的基本框架,如图 3 所示。

4.3 基于计算机科学视角下的 SMV 空间

从计算机科学视角出发,SVM 空间可看为是

一个以数据为中心、以知识为基础、以智慧为导向的计算机系统。它包含了“数据”“数据中蕴涵的知识”,以及“从数据中获得知识所产生的智慧”3 个要素,它们分别对应了 DKW 三层级结构中的底层数据(无知识)层次、中层知识层次和顶层智慧(更高级知识)层次。具体而言,在数据/符号级别,是以数据为中心的 system 对原始数据进行操作,其主要功能是收集、存储和检索数据,而不去考虑数据的语义和含义;在知识/意义层面,是利用基于知识的 system 处理不同类型的知识;在智慧/价值层面,是以智慧为导向的 system 去支持和帮助人们做出明智的决策并采取正确的行动。通过上述 SVM 三要素和 DKW 三层级结构,可以构建一个以底层数据为基础、以中层知识为支撑、以顶层智慧和价值为目标的决策系统,并形成 一个 3×3 的数据科学概念框架,如表 4 所示。

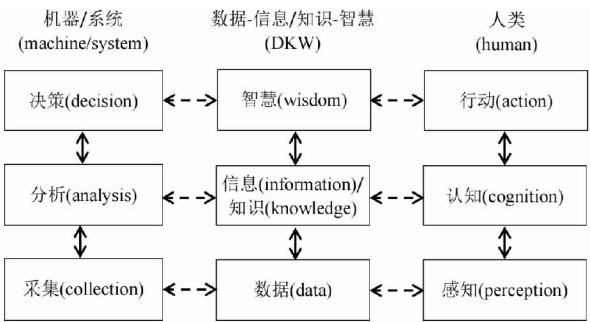


图 3 感知-认知-行动模型

Fig. 3 Perception-cognition-action (PCA) model

表 4 数据科学的 3×3 概念框架

Tab. 4 3×3 conceptual framework of data science

SMV 空间	层级	特征	功能
符号	形式/数据	数据是资源	使数据可获取
意义	内容/知识	知识是力量	使数据有意义
价值	价值/智慧	智慧是价值	使数据有价值

进一步地,Yao 近期思考了三支决策与可解释性人工智能融合的可能性,指出三支决策可以利用 SMV 空间和 WWH (why-what-how,为什么-是什么-怎么做) 法则相结合,去探讨可解释性人工智能相关问题^[46]。表 5 给出了一个 SMV × WWH 的三支决策研究框架。

此外,三支决策与形式概念分析和概念认知等不确定性理论的深度融合也成为三支决策与认知科学研究的一个重要研究方向,国内外学者近年来也在这一领域取得了重大研究进展。Li 等

从认知计算的角度讨论了基于粒计算的概念学习,并将粒计算与认知概念结构相结合,提高概念学习效率^[47]。进一步地,Li 等人基于概念学习和序贯三支决策的思想研究了多粒度的三支概念学习,从认知的角度揭示了三支概念学习在解决决策问题时的核心思想,并系统总结了概念的渐进式认知理论与方法^[48]。Xu 等给出了一个新型认知系统模型,并设计了信息粒度变换的算法^[49]。Wang 和 Yao 提出一种基于数学形态学的三支聚类方法^[50]。Wu、Leung 和 Mi 研究了在形式背景下的粒度计算和知识约简问题^[51]。Zhi 和 Li 针对不完备形式背景,提出了基于粒描述的决策规则获取方法^[52]。Qi 等讨论了三支决策与概念格的理论模型,这赋予了形式概念分析和概念格新的语义解释^[53]。Ren 和 Wei 考虑了三支概念格的约简理论与方法^[54]。值得一提的是,张�钹院士在《后深度学习时代的人工智能》的报告中提到“要解决当前人工智能所面临的挑战,可以考虑两种解决办法:其一,是把人工智能中‘知识驱动’与‘数据驱动’这两个方法结合起来;其二,是回到神经网络的本源,借助于人脑神经的工作机制研究,进一步推动深度神经网络模型的深入发展。”^[55]。“知识驱动”跟语义挂钩,是可理解的;“数据驱动”是黑箱的方法,是不可理解的。如果能把两种方法整合起来,有可能推动人工智能技术的发展与应用。上述成果和观点为未来三支决策可解释性研究指明了明确的方向。

表 5 SMV × WWH 三支决策研究框架

Tab. 5 SMV × WWH research framework of on three-way decision

SMV 空间	可解释性层级	问题(WWH)
符号	结论是什么	为什么需要一些假设?
		结果如何?
		如何得到这些结论?
意义	结论的意义	为什么结论有意义?
		结论的意义是什么?
		怎么去解释结论?
价值	结论的价值	为什么结论是合理的?
		结论的价值如何?
		怎么去使用结论?

5 结语

人工智能的可解释性问题是当前数据科学与信息技术关注的热点课题和前沿方向。以三支决策与粒计算理论为代表的理论和方法通过模拟人类思维方式和行为模式,通过信息化、分层递阶等技术手段对复杂问题本质特征进行剖析和分解,实现了对决策问题的简化、增强了对决策结果的解释,并已逐渐成为人工智能的一个重要研究方向。本文从粗糙集理论、粒计算方法、分类问题和认知计算 4 个层面出发,来探讨三支决策的可解释性问题,体现了三支决策在处理不确定性人工智能及复杂决策问题的有用性和有效性,具有重要的理论意义和广泛的应用前景。虽然近十年来,三支决策在理论、方法和应用上取得了较大进展,但其一般化理论还未完全构建,很多课题还有待深入探究。2022 年中国人工智能学会组织编纂的《中国人工智能系列白皮书——粒计算与知识发现》分册中已明确把三支决策列为粒计算的主要研究方向^[56],基于分类/聚类/社团发现的三支决策方法、基于推荐系统的三支决策方法、基于自然语言处理和情感分析的三支决策方法、基于图像识别的三支决策方法,以及基于多标签/多示例/多模态的三支决策方法也有望成为探究不确定性人工智能黑箱模型的重要途径。可以预见,三支决策与行为科学、三支决策与计算智能、三支决策与机器学习、三支决策与认知计算等前沿研究领域,将成为未来三支决策可解释性研究的重要问题和热点方向。

参考文献

[1] YAO Y Y. Three-way decisions with probabilistic rough sets[J]. Information Sciences, 2010, 180(3): 341-353.

[2] YAO Y Y. Three-way decision and granular computing [J]. International Journal of Approximate Reasoning, 2018, 103: 107-123.

[3] YAO Y Y, WONG S K M. A decision theoretic framework for approximating concepts [J]. International Journal of Man-Machine Studies, 1992, 37(6): 793-809.

[4] 刘盾, 梁德翠. 广义三支决策与狭义三支决策[J]. 计算机科学与探索, 2017, 11(3): 502-510.

LIU D, LIANG D C. Generalized three-way decisions and special three-way decisions[J]. Journal of Fron-

- tiers of Computer Science and Technology, 2017, 11 (3): 502-510.
- [5] 刘盾, 李天瑞, 杨新, 等. 三支决策: 基于粗糙集与粒计算研究视角[J]. 智能系统学报, 2019, 14(6): 1111-1120.
- LIU D, LI T R, YANG X, et al. The temporality and spatiality of three-way decisions [J]. CAAI Transactions on Intelligent Systems, 2019, 14 (6): 1111-1120.
- [6] 刘盾, 李天瑞, 梁德翠, 等. 三支决策的时空性[J]. 智能系统学报, 2019, 14(1): 141-149.
- LIU D, LI T R, LIANG D C, et al. Three-way decisions: The perspective of rough sets and granular computing[J]. CAAI Transactions on Intelligent Systems, 2019, 14(1): 141-149.
- [7] YAO Y Y, DENG X F. Sequential three-way decisions with probabilistic rough sets [C] // Proceeding of 10th IEEE International Conference on Cognitive Informatics and Cognitive Computing, Banff, AB, Canada: IEEE, 2011: 120-125.
- [8] LIU D, LI T R, LIANG D C. Three-way decisions in dynamic decision-theoretic rough sets [C] // Proceeding of 8th International Conference on Rough Sets and Knowledge Technology. Berlin, Heidelberg: Springer, 2013: 291-301.
- [9] GAO C, YAO Y Y. Actionable strategies in three-way decisions [J]. Knowledge-Based Systems, 2017, 133: 141-155.
- [10] JIANG C M, YAO Y Y. Effectiveness measures in movement-based three-way decisions [J]. Knowledge-Based Systems, 2018, 160: 136-143.
- [11] YANG X, LI T R, FUJITA H, et al. A unified model of sequential three-way decisions and multilevel incremental processing [J]. Knowledge-Based Systems, 2017, 134: 172-188.
- [12] XU J F, MIAO D Q, ZHANG Y J, et al. A three-way decisions model with probabilistic rough sets for stream computing [J]. International Journal of Approximate Reasoning, 2017, 88: 1-22.
- [13] QIAN J, LIU C H, MIAO D Q, et al. Sequential three-way decisions via multi-granularity [J]. Information Sciences, 2020, 507: 606-629.
- [14] FANG Y, GAO C, YAO Y Y. Granularity-driven sequential three-way decisions: A cost-sensitive approach to classification [J]. Information Sciences, 2020, 507: 644-664.
- [15] ZHANG Y B, MIAO D Q, WANG J Q, et al. A cost-sensitive three-way combination technique for ensemble learning in sentiment classification [J]. International Journal of Approximate Reasoning, 2019, 105: 85-97.
- [16] WU Y X, LUO L F, LI Y, et al. NTP-Miner: Non-overlapping three-way sequential pattern mining [J]. ACM Transactions on Knowledge Discovery from Data, 2022, 16(3): 51.
- [17] LI H X, ZHANG L B, HUANG B, et al. Sequential three-way decision and granulation for cost-sensitive face recognition [J]. Knowledge-Based Systems, 2016, 91: 241-251.
- [18] LI H X, ZHANG L B, ZHOU X Z, et al. Cost-sensitive sequential three-way decision modeling using a deep neural network [J]. International Journal of Approximate Reasoning, 2017, 85: 68-78.
- [19] ZHANG L B, LI H X, ZHOU X Z, et al. Sequential three-way decision based on multi-granular autoencoder features [J]. Information Sciences, 2020, 507: 630-643.
- [20] SAVCHENKO A V. Fast inference in convolutional neural networks based on sequential three-way decisions [J]. Information Sciences, 2021, 560: 370-385.
- [21] QIAN Y H, ZHANG H, SANG Y L, et al. Multigranulation decision-theoretic rough sets [J]. International Journal of Approximate Reasoning, 2014, 55 (1): 225-237.
- [22] 苗夺谦, 徐菲菲, 姚一豫, 等. 粒计算的集合论描述 [J]. 计算机学报, 2012, 35(2): 2351-2363.
- MIAO D Q, XU F F, YAO Y Y, et al. Set-theoretic formulation of granular computing [J]. Chinese Journal of Computers, 2012, 35(2): 2351-2363.
- [23] 王国胤, 张清华, 马希骛, 等. 知识不确定性的粒计算模型 [J]. 软件学报, 2011, 22(4): 676-694.
- WANG G Y, ZHANG Q H, MA X A, et al. Granular computing models for knowledge uncertainty [J]. Journal of Software, 2011, 22(4): 676-694.
- [24] ZHANG X Y, MIAO D Q. Three-layer granular structures and three-way informational measures of a decision table [J]. Information Sciences, 2017, 412/413: 67-86.
- [25] YUE X D, ZHOU J, YAO Y Y, et al. Shadowed neighborhoods based on fuzzy rough transformation for three-way classification [J]. IEEE Transactions on Fuzzy Systems, 2020, 28(5): 978-991.
- [26] CHEN J, XU Y, ZHAO S, et al. AH3: An adaptive hierarchical feature representation model for three-way decision boundary processing [J]. International Journal

- of Approximate Reasoning, 2021, 130: 259-272.
- [27] YANG X, LI T R, LIU D, et al. A temporal-spatial composite sequential approach of three-way granular computing[J]. Information Sciences, 2019, 486: 171-189.
- [28] WU W Z, LEUNG Y. Optimal scale selection for multi-scale decision tables[J]. International Journal of Approximate Reasoning, 2013, 54(8): 1107-1129.
- [29] LI F, HU B Q. A new approach of optimal scale selection to multi-scale decision tables[J]. Information Sciences, 2017, 381: 193-208.
- [30] HAO C, LI J H, FAN M, et al. Optimal scale selection in dynamic multi-scale decision tables based on sequential three-way decisions[J]. Information Sciences, 2017, 415/416: 213-232.
- [31] HUANG B, WU W Z, YAN J J, et al. Inclusion measure-based multi-granulation decision-theoretic rough sets in multi-scale intuitionistic fuzzy information tables[J]. Information Sciences, 2020, 507: 421-448.
- [32] LUO C, LI T R, HUANG Y Y, et al. Updating three-way decisions in incomplete multi-scale information systems[J]. Information Sciences, 2019, 476: 274-289.
- [33] ZHANG Q H, CHENG Y L, ZHAO F, et al. Optimal scale combination selection integrating three-way decision with Hasse Diagram[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(8): 3675-3689.
- [34] YU H, LIU Z G, WANG G Y. An automatic method to determine the number of clusters using decision theoretic rough set[J]. International Journal of Approximate Reasoning, 2014, 55(1): 101-115.
- [35] SHEN W, WEI Z H, LI Q W, et al. Three-way decisions based blocking reduction models in hierarchical classification[J]. Information Sciences, 2020, 523: 63-76.
- [36] SHE Y H, QIAN Z H, HE X L, et al. On generalization reducts in multi-scale decision tables[J]. Information Sciences, 2021, 555: 104-124.
- [37] HUANG X F, ZHAN J M. A three-way decision approach based on regret theory in multi-scale decision information systems[J]. Information Sciences, 2021, 581: 711-739.
- [38] LIU D. The effectiveness of three-way classification with interpretable perspective[J]. Information Sciences, 2021, 567: 237-255.
- [39] YAO Y Y. The superiority of three-way decisions in probabilistic rough set models[J]. Information Sciences, 2011, 181(6): 1080-1096.
- [40] 刘盾, 姚一豫, 李天瑞. 三枝决策粗糙集[J]. 计算机科学, 2011, 38(1): 246-250.
- LIU D, YAO Y Y, LI T R. Three-way decision-theoretic rough sets[J]. Computer Science, 2011, 38(1): 246-250.
- [41] YIN L J, ZHANG Q H, ZHAO F, et al. Superiority of three-way decisions from the perspective of probability[J]. Artificial Intelligence Review, 2023, 56(2): 1263-1295.
- [42] YAO Y Y. Three-way decisions and cognitive computing[J]. Cognitive Computations, 2016, 8(4): 543-554.
- [43] YAO Y Y. A triarchic theory of granular computing[J]. Granular Computing, 2016, 1(2): 145-157.
- [44] YAO Y Y. Symbols-Meaning-Value (SMV) space as a basis for a conceptual model of data science[J]. International Journal of Approximate Reasoning, 2022, 144: 113-128.
- [45] YAO Y Y. Tri-level thinking: Models of three-way decision[J]. International Journal of Machine Learning and Cybernetics, 2020, 11(5): 947-959.
- [46] YAO Y Y. The Dao of three-way decision and three-world thinking[J]. International Journal of Approximate Reasoning, 2023, 162: 109032.
- [47] LI J H, MEI C L, XU W H, et al. Concept learning via granular computing: A cognitive viewpoint[J]. Information Sciences, 2015, 298: 447-467.
- [48] LI J H, HUANG C C, QI J J, et al. Three-way cognitive concept learning via multi-granularity[J]. Information Sciences, 2017, 378: 244-263.
- [49] XU W H, PANG J Z, LUO S Q. A novel cognitive system model and approach to transformation of information granules[J]. International Journal of Approximate Reasoning, 2014, 55(3): 853-866.
- [50] WANG P X, YAO Y Y. A three-way clustering method based on mathematical morphology[J]. Knowledge-Based Systems, 2018, 155: 54-65.
- [51] WU W Z, LEUNG Y, MI J S. Granular computing and knowledge reduction in formal contexts[J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(10): 1461-1474.
- [52] ZHI H L, LI J H. Granule description of incomplete data: A cognitive viewpoint[J]. Cognitive Computation, 2022, 14(6): 2108-2119.
- [53] QI J J, QIAN T, WEI L. The connections between three-way and classical concept lattices[J]. Knowl-

edge-Based Systems, 2016, 91: 143-151.

[54] REN R S, WEI L. The attribute reductions of three-way concept lattices [J]. Knowledge-Based Systems, 2016, 99: 92-102.

[55] 张钊. 后深度学习时代的人工智能[EB/OL]. [2016-10-21]. [https://baike.baidu.com/tashuo/](https://baike.baidu.com/tashuo/browse/content?id=7aefb4f4090e5877f2dc759f)

browse/content?id=7aefb4f4090e5877f2dc759f.

[56] 中国人工智能系列白皮书——粒计算与知识发现[EB/OL]. [2022-07-01]. <https://www.caii.cn/index.php?s=/home/article/detail/id/2290.html>.

(编辑 张 欢)

作者简介:



刘盾,博士,教授,博士生导师,清华大学博士后,美国 Carnegie Mellon 大学和加拿大 Regina 大学访问学者,四川省杰出青年科技人才,四川省学术和技术带头人,CCF 杰出会员。2022 年和 2023 年入选全球前 2% 顶尖科学家终身成就榜,主要研究方向为大数据决策分析、粒计算与知识发现。任中国人工智能学会粒计算与知识发现专委会常务委员、中国人工智能学会知识工程与分布智能专委会常务委员、四川省人工智能学会理事、四川省电子商务教指委委员。在 *IEEE TFS*、*IEEE SMCA*、*INS*、*I & M*、*IJPR* 等国内外信息科学和管理科学期刊发表学术论文 160 余篇,其中 SCI 检索 80 余篇,Google Scholar 引用次数 6 000 余次,H-index 指数 43。主持国家自然科学基金 4 项,主编专著 2 部。曾获得四川省社会科学优秀成果二等奖 2 项、三等奖 2 项。