

大语言模型何以推进知识证成?

——基于阿格里帕困境的语用解答策略

How Can Large Language Models Advance Knowledge Justification:
A Pragmatic Solution Strategy of the Agrippa's Trilemma

王彩虹 /WANG Caihong¹ 卢醒寒 /LU Xinghan²

(1. 中共山西省委党校, 山西太原, 030006; 2. 南京师范大学马克思主义学院, 江苏南京, 210023)
(1. Shanxi Provincial Party School, Taiyuan, Shanxi, 030006;
2. School of Marxism, Nanjing Normal University, Nanjing, Jiangsu, 210023)

摘要: 大语言模型对知识证成的推进作用, 可借助其对怀疑论的有效反驳来证明。作为皮浪式怀疑论的代表, 阿格里帕困境否定了命题被证成的可能, 以基础主义、无限主义和融贯主义为代表的传统语义解法无法成为其通解, 且违背了人类的思维模式与经验常识。通过参考威廉姆斯的解法和匹兹堡学派的实用主义思想, 本文建立了以“辩护-质疑”交互系数和时间为自变量的对话分数模型, 它不仅能反映序数最佳意义上的知识证成状态, 而且兼容了三种语义解法, 构成阿格里帕困境的语用解答策略。基于对话分数模型与拟人化因子的适配程度、形式论辩的研究成果与自然语言处理技术的语用特征, 大语言模型可以成为语用策略的良好载体, 继而最终证明了它对知识证成的推进。

关键词: 大语言模型 阿格里帕困境 语义 语用 匹兹堡学派

Abstract: The role of large language models (LLMs) in advancing knowledge justification can be demonstrated by their effective rebuttal of scepticism. As a representative of Pyrrhonian Skepticism, the Agrippan trilemma denies the possibility of propositional justification. Traditional semantic solutions represented by foundationalism, infinitism, and coherence theory cannot provide a general solution and contradict human thinking patterns and common sense. By drawing on Williams's solution and the Pittsburgh School's pragmatism, a dialogue scoring model with the "defence-question" interaction coefficient and time as independent variables was established. This model not only reflects the optimal state of knowledge justification in an ordinal sense but also accommodates the three semantic solutions, thereby constituting a pragmatic resolution strategy for the Agrippan trilemma. Based on the compatibility between the dialogue scoring model and anthropomorphic factors, the findings of formal debate research, and the pragmatic features of natural language processing technology, LLMs can serve as an excellent vehicle for pragmatic strategies, ultimately demonstrating their ability to advance knowledge justification.

Key Words: LLMs; Agrippa's trilemma; Semantics; Pragmatics; Pittsburgh School

中图分类号: N031; G302 DOI: 10.15994/j.1000-0763.2026.05.003 CSTR: 32281.14.jdn.2026.05.003

面对人工智能时代的知识论问题, 学界展开了不同角度的探讨, 如历史唯物主义知识论

产观对生成式人工智能知识的支持、^[1] AI for Science 理念的实践价值、^[2] 笛卡尔认识论语

收稿日期: 2025年7月17日

作者简介: 王彩虹(1989-)女, 山西太原人, 中共山西省委党校副教授, 研究方向为科学哲学、社会哲学。Email: iris_ch@qq.com

卢醒寒(1995-)女, 江苏淮安人, 南京师范大学马克思主义学院博士研究生, 研究方向为知识论、语言哲学。Email: lxh719070855@163.com

境中人工智能之于人类智能的证言效力与交互模式差异^[3]等。然而,上述研究虽不同程度地阐释了技术进步对知识生产的影响,却多采用宏大叙事的手法,微观层面的尝试尚付阙如。事实上,知识证成的另一面是对怀疑主义的解构,即通过推翻怀疑主义的立论基础或内容,使怀疑的结论不成立,继而为知识的正当性提供辩护。因此,本文试从著名的怀疑论命题——阿格里帕困境(Agrippa's Trilemma)入手,借助其解答策略中语用策略对语义解法的超越,印证同样强调前者的大语言模型(Large Language Models, LLMs)在其中发挥的或可能发挥的重要作用,为基于智能技术的知识证成提供新思路。

一、阿格里帕困境的FIC语义解答及其局限性

根据皮浪式怀疑论(Pyrrhonian Skepticism),我们在面对难以解决的分歧时,应当采取一种悬置(suspend)判断的态度,即坚称“我无法说出眼前的观点中哪一个是应当相信或不相信的”,^[4]以此抵达心灵的宁静(ataraxia)。由五个不同模式组成的阿格里帕困境则是悬置判断的常见手段,其中又以独断假设(辩护停止于不再继续提供理由的 R_n)、无穷倒退(理由R被无限回溯)、循环论证(P既是前提又是结论)三种错误最具代表性,构成“阿格里帕三难”问题。通过宣称任何知识证成都必将归属于以上三种谬误,皮浪式怀疑论排除了“命题被证明为知识”的可能。

对此,主流解法多针对独断假设、无穷倒退、循环论证中的一个问题展开辩护,形成了三种解答策略,即基础主义(Foundationalism)、无限主义(Infinitism)和融贯主义(Coherentism),可合称为FIC方案。它们的定义如下:(1)基础主义对应独断假设,它认为推理可以被自明的基础命题支撑起来,即 R_n 不能继续提供理由,

是因为它背后的基础命题使它不用继续提供理由。比如在数学题的求解中,我们往往将某一个不可证的数学定理作为运算或证明的基石。^[5](2)无限主义对应无穷倒退,它肯定了“无穷”的本体论价值,并认为它是人类理性精神和认知模式的体现。譬如黑格尔意义上的“真正的无限”,即精神借助反思(如内在矛盾的发展、自我批判)将有限性纳入无限的自我超越,与之对应的是旧形而上学中机械重复的“恶无穷”,即无意义的无穷倒退。^[6](3)融贯主义对应循环论证,它将知识视为多个命题相互支持的整体,以整体的辩护成功支撑单个命题的证成。^[7]新本格推理流派是比较生动的例子,作者通过构建一个各事件环环相扣、相互印证,却不一定存在于现实世界的推理体系,实现逻辑自治,如《钟表馆事件》中连续的时间诡计^①。^[8]

阿格里帕困境的提出和FIC解法的回应,不仅是怀疑论与反怀疑论的简单论争,更将知识证成的路径限制在封闭的语言系统内,侧重于讨论语言的内部逻辑、形式规则、既定意义即语义(semantics)是否成立。通过应景特设(ad hoc),一个具备自明性的前提、一个自我发展的无穷链条或一个相互融贯的体系被规定为知识证成的必要条件。该做法是否成立暂且不论,至少FIC解法在结果上的局限性表明了它无法成为知识证成的有力支持。纵观三种解法,不难发现它们都采取了“并非最差”的解答策略,即表明某一条路径优于另外两条路径,却无法得出同时处理三重困境的通解。

究其原因,首先在于语义本身。考虑到同等条件下三种困境的理论变量差异较大甚至相互矛盾,譬如独断假设情境中的 R_n 与无穷倒退情境中的无限辩护完全相反,导致它们本就难以共存。相对应地,反驳与解答只能分开进行,这直接决定了FIC无法成为三个问题的通解。此外,即使从整体角度审视FIC解法,我们会发现三条路径,特别是基础主义和无限主义,都是针对论证前提与结论之间的先后顺序、

^①在绫辻行人(Ayatsuji Yukito)的新本格推理小说《钟表馆事件》中,作者通过一系列证据链,构建了不同于外部世界的独立时间体系,是借助命题间的整体协调建立闭合的信念系统、达到内部融贯的生动体现。

单独或整体对应、连续关系的质疑与反驳,并未彻底走出休谟怀疑论所批判的先因后果、单因独果、因果连续的线性逻辑。但符合人类认知模式的推理并非总是简单的、首尾相接的线性关系,还可能出现更加复杂的逻辑框架,比如融贯主义语境中各命题形成的集合或网络结构。因此我们可以认为,停留于语义层面的FIC解法被它们赖以生存的语义反对,局限性由此产生。

进言之,对线性逻辑的过分执着又会导致对现实空间及其规定的忽略。即使假设FIC的任一解法为真,但倘若它无法在日常生活中被复现,甚至挑战了人类自身的思维习惯与常识,那便是无意义的,也会遭遇不同程度的限制。从经验角度考察,FIC解法至少存在以下两点问题:一方面,就经验常识而言,无限的回溯语境或成为伪命题。随着辩护与质疑的反复开展,命题的确定性会不断提高,并最终接近后期维特根斯坦所定义的确定性相对最高的“枢轴”状态,追问和回溯随之终止;另一方面,现实的人的交流无法达到纯粹的问答模式,而会在合作原则的作用下,被会话参与人有意无意地引向最易达成共识的方向。正如日常语言学派代表人物奥斯丁(John Langshaw Austin)所言,任何语言活动都至少有一部分属于惯例问题并受到惯例约束。^[9]通过会话实现共识则是惯例之一。因而在该目标的驱使下,当知识论辩的次数逼近人类行为模式的上限时,我们追求的往往不是绝对的“实数正义”(唯一正确的),而是优先级最高的“序数正义”(相对正确的),形成一种知识论让位于实用论的局面。显然,上述观点中的大多数,都无法通过来自怀疑论的最严格的检验,但过度的理性主义和理想化形式,往往导致错误与幻相。逻辑分析并非某种神秘的“先验分析”的结果,相反,它始终在实践行为及其空间中展开。因而不尊重日常经验的回答即使在逻辑上达到完美,却难免是冗余的,甚至是无效的。这构成阿格里帕困境传统语义解法局限性的另一成因。

总之,FIC路径并未实质性地推翻阿格里帕困境,而更多是停留于内部视角的、为了应

对不利条件而特设的补全,甚至仅仅是将原有问题换了一种表述方式,并不能算作真正的解答。同时,语义解法所导致的对线性逻辑的过分坚持,不仅违背了人类思维模式的基本形式,也轻视了日常生活经验对知识证成的规范与约束。这意味着它未能调和理性主义与实用主义的分歧,从根本上引发了失败。为了找到更加通用而有效的解法,应当尽可能规避以上误区,将视角由纯粹的逻辑分析转向实用诊断,实现从语义到语用的调整。

二、基于匹兹堡学派实用主义的语用解答策略

1. 语用转向与威廉姆斯的解答尝试

不可否认的是,语义分析是逻辑分析的重要模式,但理论并不能独立于实践而存在,抽象的推演会被置于实践应用,进而获得具体的个体性特征。诚如匹兹堡学派代表人物布兰顿(Robert Brandom)所认为的那样,在判断的过程中,独立于语用的纯粹语义分析是不可能实现的,^[10]应认识到特定场景下语用之于语义的优势,并将前者作为调和理性主义与实用主义的助力。

有别于语义,语用(pragmatics)关注的是符号与使用者(包括人的意图、行为和社会文化背景)之间的关系,强调符号在具体情境中的实际应用和效果。换言之,语用代表的是语言动物在长期的意向性活动中形成的用语言阐释、评价事物的共识,常表现为某一历史或文化背景下,人们应该在“何时何地怎么说何话”,而非说“正确却不合时宜的话”的规则。就阿格里帕困境而言,其辩护与质疑的载体是语言,其参与者通常是人类,因此不可避免地受到语用要素及其规则的影响。相应地,针对阿格里帕困境的反驳,需要充分考虑既有规则和论辩实践。威廉姆斯(Michael Williams)的解法便体现了这种语用转向。他认为在知识证成的场景中,辩护者的表现受到其教育水平和辩护能力的限制,且面对的质疑应是合情合理的而非任意的,这一点未被阿格里帕困境的提出者明

确规定。更重要的是,在阿格里帕困境的交互场景中,辩护者与质疑者的论辩责任是不对等的,前者需要持续提供有力的理由,而后者只需要不断提问即可达到质疑的目的。

威廉姆斯认为,这种由阿格里帕困境预设的责任不对等使知识证成沦为不合理的“理由索取与给予活动”,呈现出规则适用和语言使用上的错误,进而导致支撑困境发生的整体场景的失效,因此皮浪式怀疑论借三重困境达成的判断悬置不具备正当性。^[11]这种实践意义上的诊断得到了学界的认可,比如布兰顿就曾在论文中公开支持威廉姆斯对知识证成发生环境的重视及其对推翻怀疑论的价值,并将其代入到自己的推论主义体系之中。^[12]国内学者王聚结合了后期维特根斯坦的观点,指出威廉姆斯的解法本质上是一种建基于基础命题的“流变的基础主义”,即认识到基础命题在不同语境中适用性各异的事实,鼓励辩护者与质疑者基于双方对基础命题的分歧展开认知审查和语境转换,以此推进共识达成。^[13]基于上述结论,我们认为威廉姆斯的解法之所以有可能从源头上消解阿格里帕困境,是因为它既承认基础命题及其自明性理由 R_n 的重要性,又认识到它并非绝对的不可错,需要根据应用场景、参与人的交互反应等综合判断。这昭示了实用论语境中去绝对化的知识证成理念,更符合现实空间中的真实论辩。受此启发,我们将尝试给出阿格里帕困境的语用解答策略。

2. 基于布兰顿计分模式的语用解答提出

有别于FIC对线性逻辑的强调,威廉姆斯的解法更关注辩护者与质疑者的互动行为及其结果。他的尝试事实上将知识证成从FIC的限制中解放出来,将它化约为以说服谈话者为目的的 R_n 搜寻^①。这与匹兹堡学派特别是布兰顿所追求的去绝对化和实践意义上的理性不谋而合。通过对比他和威廉姆斯的理论,会发现二者分享了以下关键词:(1)布兰顿计分(scorekeeping)模式中的对话进程与威廉姆斯强调的阿格里帕困境的交互场景类似。(2)布

兰顿模式对对话分数(最大值)的形式化描述,和“流变的基础主义”对自明性理由 R_n 及其良好交互表征的追求相契合。基于此,我们可以尝试将二者结合并建构新的语用解答模型,以求形式化地表示辩护者和质疑者的互动行为和知识证成的状态:

设定对话分数(Epistemic Confidence Index)为 s ,用于表征在知识论辩过程中质疑者对对话(辩护)进展程度的评分,以此表示知识证成的进度。具体而言,该指数依赖于时间变量 t (每个理由 R 被提出的时间)与“辩护-质疑”交互过程系数(Argumentative Dialectic Process)为 c ,从时间 t_1 到 t_2 之间,通过交互过程或者说对话进程 c_1 获得新的对话分数 s_2 。其中,“辩护-质疑”交互过程是一个多因素作用的交互复杂性系数,受到辩护者的辩护能力、质疑者的批判理性程度,以及对话双方的共识需求强度等要素共同影响,因此形成了函数关系 $s = f(c, t)$,且 $c = g(a_1(t), a_2(t), a_3(t), \dots)$,这里 $a_1(t), a_2(t), a_3(t) \dots$ 就是指辩护者的辩护能力、质疑者的批判理性等要素随时间的变化,进而表达式展开后是 $s = f(g(a_1(t), a_2(t), a_3(t), \dots), t)$ 。同时,模型采用一种去绝对化的知识证成模式,虽然存在达到某一阈值后视为绝对知识证成的情境,但在实际认知交互场景中,更加追求序数最佳(ordinal optimization)意义上的成功,即在特定域内或交互过程中,追求对话分数的最大化,而非必须跨越绝对阈值。此外,每次单独的“辩护-质疑”交互过程都会产生一个单独的对话分数 s ,所有这些单次评估值共同构成对话分数集合 S ,它反映了整个辩护进程中质疑者认知确信程度的变化趋势。

该策略可以兼容FIC解法。在基础主义的场景下,“辩护-质疑”交互过程 c 的各项作用因素均处于正常状态,那么当具有自明性的 R_n 出现时,特定域域的 s_{max} 随之出现。因为最适应双方交互语境且确定性最高的理由往往具备最强的说服效力;而无限主义则要求对话分数 s 呈现整体单调递增趋势。依旧以黑格尔意义上

^①出于建模需要,本研究合并了基础命题和与之相关的自明性理由,将二者统称为 R_n 。

的“真正的无限”为例,随着交互次数和时间变量 t 的增长,理由 R 在不断的自我超越中实现了说服力的提升,每一次“辩护-质疑”进程 c 都会产生正反馈,从而使对话分数 s 处于持续增加状态;就融贯主义而言,则需要假定参与辩护的命题集合及相关参与主体保持稳定。在此前提下,如果相邻对话分数集合(例如集合 S_{m-1} 与 S_m)之间的末位与首位元素呈现高度相似性,则可合理推定这两个集合所对应的命题间存在一致或融贯关系。综上所述,原本只能独立展开的基础主义、无限主义、融贯主义三条FIC路径,在语用解答策略的框架下有可能实现共存。

需要说明的是,一方面,实践场景中理由 R 的确定性及其说服力与对话分数 s 之间的实际关系较为复杂,甚至可能构成另一个独立的加权函数。本文仅基于经验假设了较为普遍的正相关情况;另一方面,鉴于缺乏充足的实证数据支持,本模型并不具备严格的绝对量化意义,仅供阿格里帕困境的分析与反驳参考。

3. 语用解答策略对FIC语义解法的超越

如前所述,无论是追求绝对的真理阈值,还是有限时域内序数最佳意义上的证成,对话分数 s 均被“辩护-质疑”交互过程系数 c 决定。有鉴于 c 是受到辩护者的辩护能力、质疑者的批判理性程度、共识需求强度等要素共同影响的复杂系数,被 c 决定的 s 与变量 t 之间必然无法形成如 $f(x)=ax+b$ 一般的线性关系,而是出现由 c 的加权函数所涉及的多个要素构成的非简单线性的路径。这意味着即使从简化现实并高度抽象的模型来看,对话分数尚且无法与“辩护-质疑”的交互进程之间形成广义的线性关联,遑论发生在现实空间中的论辩实践。至此,我们可以得出以下结论:以语义无法与语用相分离为前提,在现实场景中,FIC所规定的先因后果、单因独果或因果连续的线性证成逻辑是难以实现的,取而代之的是涉及相互关联的多重外部因素的复杂体系。相应地,正因为语

用策略对现实的尊重,它在阿格里帕问题的局部语境中构成了对语义解法的超越,有效克服了后者的局限性。

这种超越可以被思想史证实。移借布兰顿本人对交互活动的评价,“交流不是一起拥有某个共同的东西,我们要放弃这一范式,转向另一种范式:交流是一种实践中的合作。”^[14]通过将奥斯丁意义上的会话合作原则化约为塞拉斯(Wilfrid Sellars)意义上的规范性(normativity),他坚定了实用主义对于语言意义的外指取向,强调语言不但是重视内部逻辑的“自我对话”,更是连接人与他者的“问性媒介”。我们认为,这既是对匹兹堡学派规范性传统的再强调,也体现了与人之活动相伴的社会性与实践性。这事实上反映了旧形而上学已失能的实在论讨论被现实或实用取向超越。就本文的研究主题而言,前者所代表的是皮浪式怀疑论对知识证成的全盘否定,以及FIC解答所追求的反驳之共相,后者则是接受了命题、理由、规则在不同环境具有不同适应水平的事实,如其所是地探讨解法之殊相。阿格里帕困境解法的语用转向,实质上反映了强调普遍与统一的第一哲学到强调不同学科与具体知识优先性的第二哲学的转向。

当然,上述结论多默认了布兰顿计分模式及其理论倾向的正确性,并将其作为重要的立论依据。然而即使在匹兹堡学派内部,也不乏布兰顿理论的批评者。麦克道威尔(John McDowell)就曾以诠释学为基础,公开指责计分模式可应用视野的狭小,认为它虽然成功地将逻辑分析从主体之中转向主体之间,但仍缺乏针对更为基础和更为宽广语境的适用性,难以从个体走向世界。^[15]事实上,麦氏基于诠释学的批判视角并不完美,存在滥用实用主义的问题,即为了追求更普遍的实用而将计分模式的应用场景过度延伸,甚至抵达“世界”这一泛指人类生活方式发生空间的高度^①。这显然偏离了语用主义的建立初衷:作为语义无法摆

①伽达默尔区分环境(Umwelt)和世界(Welt)的目的是界定人的理性与动物性的本能,类似于阿奎那意义上的理性灵魂和感性/植物灵魂。我们认为,这是从人类整体的抽象角度思考其生活方式,比本文所探讨的以及布兰顿计分模式所应用的个别场景宽泛得多,会导致研究重心过度转向外部世界而忽略语词的后果。

脱的底色，语用的出现是为了更好地探讨语义而不是偏废之。或者说，语用策略将知识证成引向命题外部并不意味着对内部视角的颠覆，而是为了二者的融合与一致。即使世界这一要件经常在语言哲学、心灵哲学等领域被讨论，但这并不意味着它适用于所有议题，特别是本文所关涉的相对微观的问题。如果一味沿用麦氏的思维，那么解答阿格里帕困境的最终目的将脱离知识证成（知识信息的本体更偏向语义）的初衷，落入有关某一解法外延性高低的争执，这是无意义的。因此确定应用场景、以一种取长补短的工具论视角借鉴布兰顿的理论即可。

三、LLMs是语用解答策略的有效载体

在完成了质疑、反思与正解后，我们需要回归最初的疑问：如何以阿格里帕困境的有效反驳为切入点，证明人工智能技术对知识证成的推动作用？二者存在一种互证关系，即借助语用解答策略与智能模型相契合的事实，证明前者的可及性与后者的有用性。不过，无论是鉴于神经网络算法的可解释性问题，还是考虑到哲学学科的职能，本研究完整地呈现以LLMs为载体的论辩过程分析或定量统计，是相当困难的。但仍能从定性的角度，证明LLMs与语用策略的相关性。根据前文对语用策略的形式化描述，以及人工智能与逻辑学领域的交叉成果，大体可以有以下三条证明路径：（1）LLMs可以直接沿用以对话分数 s 为核心的语用模型（2）LLMs可以抽象出人类的论辩模式并建模（3）LLMs本身就是轻语义却重语用的技术实体。其中（1）可以直接成为LLMs与语用模型相适配的充分必要条件，（2）至少构成必要条件，（3）更强调双方的整体相关性。

首先讨论路径（1），以语用模型 $s=f(c, t)$ 为基础，我们的主要任务是证明智能体具备类似于人的参与“辩护-质疑”交互的能力。换言之，需要证明智能体的加入不会破坏威廉姆斯意义上的“辩护者-质疑者”权力平等的交互格局。仅从LLMs的交互表现和AI技术强大的计算能力出发，它能在不超出对话长度限制

的情况下开展与人的日常交互，是不言自明的，但LLMs能否实现与人的平等论辩，仍待讨论。对此，我们引入一个定义域为 $[0, 1]$ 的拟人化因子（Anthropomorphism Factor, M ），其中， $M=1$ 表示完全符合人类的语用规则， $M<1$ 表示语用策略的某种“去人类化”执行，即目前多数LLMs的水平。基于此，将LLMs参与后的对话分数模型表达为 $s'=M \cdot f(c, t)$ 的加权函数，即 $s'=M \cdot f(g(a_1(t), a_2(t), a_3(t), \dots), t)$ 。

抽象地看，该加权函数的成立意味着LLMs对于 $f(c, t)$ 来说是可介入的，而在 $s=f(c, t)$ 已相对完备的情况下，需要证明的是因子 M 的正当性。考虑到LLMs对“辩护-质疑”过程的参与主要通过模仿人类语言来实现，因此我们选取自然语言处理（Natural Language Processing, NLP）为切入点，讨论机器语言与人类语言的拟合。一般来说，通过将给定语料库的词例（token）概率分布记作 P_{data} 、将智能模型输出的概率分布记为 P_{θ} ，自回归语言模型把联合概率建模成条件概率的乘积，以追求模型的最优收敛，即达到 $P_{data}=P_{\theta}$ 的极大似然状态。^[16]作为NLP领域公认的训练原理， P_{θ} 与 P_{data} 的差值事实上是机器拟人化程度的数理表征，结合拟人化因子 M 的定义域， P_{θ}/P_{data} 可以在数值上被近似为 M ，这也赋予了 M 可及性与正当性。

具象地看，LLMs拥有规避逻辑陷阱的辩护能力。作为新图灵测试的重要内容，威诺格拉德模式挑战（Winograd Schema Challenge, WSC）旨在测试人工智能系统的语义理解和常识推理能力，并在不断的句对扩充中证成了LLMs具有理解回指消解、共指消解等复杂表述的能力。^[17]据此，我们可以认为在现实的论辩过程中，LLMs可以接近一个合格乃至强力参与者的水平，避免与WSC挑战原理相似的红鲱鱼谬误、叙述性诡计等具象的逻辑陷阱。这意味着随着智能技术的不断发展，LLMs参与阿格里帕困境的语用策略模型是可能的，且拟人化因子有望无限接近1甚至超过1，呈现出拟合或适度的过拟合状态，使机器具有更高的论辩水平。

路径（2）则涉及人工智能与逻辑学的交

又领域,主要表现为对形式论辩的研究。相较于为形式逻辑自身发展服务的非形式论辩,形式论辩致力于探索数理逻辑、人工智能与传统语言论辩的结合,“通过以各种形式逻辑为基础,构建形式系统,以实现多种类型的非单调推理”。^[18]譬如杜恩(Phan Ming Dung)提出的抽象论辩框架(Abstract Argumentation Framework, AAF),就是将论辩抽象为原子概念,并依据评价标准,得到集体可接受的论证集合,进而展开可接受论证子集寻找、论辩语义的求解与论辩系统的修正。随着NLP技术的发展,AAF得到了充分讨论,廖备水、^[19]阿尔法诺(Gianvincenzo Alfano)^[20]等人均对此展开深入研究。总之,虽存在由果溯因之嫌,但相关研究的长足发展确证了以AI技术抽象并建模人类论辩模式的可行性。

路径(3)的重点在于,明确LLMs与人类存在认知层面的非同一性,因而前者的语言或思维习惯并不必然等同于人。这意味着部分人类本位观点的失能或争议的消解,比如语言学中先验的唯理主义(人具备有别于其他物种的先天语言能力)^[21]与经验建构主义(人的语言能力由后天经验塑造)^[22]之争。之所以举以上例子,是为了引出以下问题:鉴于LLMs学习人类语言的独特方式,它是否从一开始就与人类不完全相同,不仅不可能具备先天语言能力,而且它自然地重语用习惯而轻语义习惯?基于语言学领域的经典分歧,我们可以认为LLMs的成功证明了人类并非唯一具有语言能力的存在,进而推论出语言的习得可以依赖于经验,因为LLMs明显是通过后天大量的语料训练学会了自然语言,即接近后期维特根斯坦及其拥护者的经验建构主义观点。那么LLMs是凭何经验、以何方式建构了语言能力?如前所述,NLP技术十分重要。一方面它提供了以字段、句段为单位的语料标注,另一方面帮助构建与人类语言习惯拟合的概率分布和抽象模型,进而借助样本尽可能大、代表性尽可能强的语言数据,展开基于深度学习的经验获取,并用一种计算机更易理解的抽象化、形式化手段建构经验,形成“基于模型构建的实验系统”。^[23]

通俗地说,相比于逐字逐句地深入理解语义,LLMs更多是学习不同语言的排列规则与使用场景,将自身塑造为一种“拥有几何空间特征和概率统计特征的数学模型”^[24]而非纯粹的语言模型或逻辑模型,因此在面对诸如古希腊语的“死语言”时,LLMs能表现出极强的统计和理解能力。以上特质意味着LLMs偏向于事件(所做的)而不是内容(所说的),具备较强的语用特征,这构成了LLMs与阿格里帕困境语用策略相适配的底层原理。

至此,我们完成了三条路径的阐释,不同程度地证明了LLMs与阿格里帕困境的语用解答相适应,在该方面存在应用潜能。当然,即使LLMs可以套用本研究的语用策略模型或威廉姆斯的解答策略,抑或符合匹兹堡学派的实用主义模式,我们并不能断言它能彻底打破传统怀疑论的所有诘难。本研究所强调的,仅仅是LLMs可以凭借语用取向与高度拟人的语义理解能力平等地参与命题的“辩护-质疑”过程,进而帮助反驳阿格里帕困境,并从该维度促进知识证成。

结 论

综上所述,鉴于任何辩护与质疑都依赖于现实场景,脱离语用的语义分析是无法完成的。就阿格里帕困境的反驳工作而言,语用策略同样值得重视。基于此,本文的基本结论如下:

第一,FIC解法的表面局限性源于它对日常生活经验的忽略,此时语用策略的职能是弥补传统解法的不足而非完全取代。

第二,威廉姆斯的语用解法与布兰顿的计分模式之间存在可被文献证实的一致性。考虑到前者超越FIC的解答效果,以及后者重视实践活动的正确态度,本研究在计分模式的基础上构建出函数 $s=f(c, t)$,并将 s 的取值作为自明性理由 R_n 及其背后的基础命题是否出现的判断标准。该模型能兼容FIC的全部三种解法,构成阿格里帕困境语用解答策略的形式化描述。

第三,LLMs与语用解答策略之间存在较高适配度,因而有可能实质性地参与阿格里帕

困境的反驳。这证明了AI技术可以将反驳怀疑论作为有效切入点,促进知识证成。

总之,知识论与认识论密不可分,而人的认识模式从不是片面且单一的。无论是面对理性主义与实用主义、内在论与外在论,或是语义与语用的矛盾关系,都不应将双方看作纯粹互斥的对象,而是发掘彼此与不同场景的适配程度与运作方式,以实现对立中的统一。^[25]

[参考文献]

- [1] Walton, N., Nayak, B. S. 'Rethinking of Marxist Perspectives on Big Data, Artificial Intelligence (AI) and Capitalist Economic Development'[J]. *Technological Forecasting and Social Change*, 2021, (166): 120576.
- [2] Van Noorden, R., Perkel, J. M. 'AI and Science: What 1,600 Researchers Think'[J]. *Nature*, 2023, 621(7980): 672-675.
- [3] Freiman, O. 'AI-Testimony, Conversational AIs and Our Anthropocentric Theory of Testimony'[J]. *Social Epistemology*, 2024, 38(4): 476-490.
- [4] 塞克斯都·恩披里克. 悬搁判断与心灵宁静: 希腊怀疑论原典[M]. 包利民等译, 北京: 中国社会科学出版社, 2004, 39.
- [5] Hamami, Y., Morris, R. L. 'Understanding in Mathematics: The Case of Mathematical Proofs'[J]. *Noûs*, 2024, 58(4): 1073-1106.
- [6] 黑格尔. 小逻辑[M]. 贺麟译, 北京: 商务印书馆, 2011, 96-99.
- [7] Thorstad, D. 'The Complexity—Coherence Trade-Off in Cognition'[J]. *Mind*, 2025, 134(534): 422-457.
- [8] 绫辻行人. 钟表馆事件[M]. 刘羽阳译, 北京: 新星出版社, 2016.
- [9] Austin, J. L. 'Performative Utterances'[A], Urmson, J. O., Warnock, G. J. (Eds.) *Philosophical Papers*[C], London: Oxford University Press, 1961, 220-240.
- [10] Brandom, R. *Perspectives on Pragmatism: Classical, Recent, and Contemporary*[M]. Cambridge: Harvard University Press, 2011, 52.
- [11] Williams, M. 'Avoiding the Regress'[A], Turri, J., Klein, P. D. (Eds.) *Ad Infinitum: New Essays on Epistemological Infnitism*[C], New York: Oxford University Press, 2014, 227-242.
- [12] Brandom, R. 'Fighting Skepticism with Skepticism: Supervaluational Epistemology, Semantic Autonomy, and Natural Kind Skepticism'[J]. *Facta Philosophica*, 2000, 2(2): 163-178.
- [13] 王聚. 基础流变与责任对称——论推论语境主义对阿格里帕难题的消解[J]. 世界哲学, 2023, (2): 117-125.
- [14] 孙宁. 匹兹堡学派研究——塞拉斯、麦克道威尔、布兰顿[M]. 上海: 复旦大学出版社, 2018, 161.
- [15] McDowell, J. *The Engaged Intellect: Philosophical Essays*[M]. Cambridge: Harvard University Press, 2009, 148-149.
- [16] 袁毓林. ChatGPT等大模型的语言处理机制及其理论蕴涵[J]. 外国语(上海外国语大学学报), 2024, 47(4): 2-14.
- [17] Browning, J., Yann, L. 'Language, Common Sense, and the Winograd Schema Challenge'[J]. *Artificial Intelligence*, 2023, (325): 104031.
- [18] 余喆、廖备水. 基于形式论辩系统的滑坡论证分析——人工智能如何理解自然语言论辩[J]. 科学技术哲学研究, 2019, 36(2): 1-7.
- [19] Liao, B., Jin, L., Koons, R. C. 'Dynamics of Argumentation Systems: A Division-based Method'[J]. *Artificial Intelligence*, 2011, 175(11): 1790-1814.
- [20] Alfano, G., Greco, S., Parisi, F., Trubitsyna, I. 'Complexity of Credulous and Skeptical Acceptance in Epistemic Argumentation Framework'[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, 38(9): 10423-10432.
- [21] Chomsky, N. *Aspects of the Theory of Syntax*[M]. Cambridge: The MIT Press, 1965, 56.
- [22] 维特根斯坦. 哲学研究[M]. 李步楼译, 北京: 商务印书馆, 2000, 20.
- [23] 任思腾. 计算机模拟的模型观与实验观辨析[J]. 自然辩证法通讯, 2025, 47(8): 47-55.
- [24] 郭贵春、梁德柱. 大语言模型应该有“内在的语言生成能力”吗?——论乔姆斯基批评ChatGPT的局限性[J]. 科学技术哲学研究, 2025, 42(1): 1-9.
- [25] 费多益. 知识的确定与心灵的限度[J]. 自然辩证法研究, 2015, 31(11): 19-25.

[责任编辑 王巍 谭笑]