

信念修正视域下的悖论研究初探

袁永锋

张建军

摘 要: 悖论研究是一个古老的研究领域, 信念修正理论则是最近三十年发展而成的新型逻辑理论。悖论的形成、发现与解决是一种典型而复杂的信念修正过程。虽然这两个领域的亲缘关系非同一般, 但迄今为止学界对这种亲缘关系缺乏系统的研究。本文从信念修正理论的角度提出一种重新审视悖论相关论题的新思路, 进而提出一种基于信念修正理论的新的解悖方法论, 最后依据这种解悖方法论对罗素悖论给出了一种新的解决方案。这种交叉研究的新思路不仅能推动悖论研究的发展, 也能为信念修正理论指出新的研究方向。

关键词: 信念修正; 悖论研究; 悖论度; 成悖标准; 解悖方法论

中图分类号: B81

文献标识码: A

悖论研究是一个探讨悖论的形成、发现与解决的古老的研究领域。而信念修正 (belief revision) 理论则是近三十年才发展而成的新型逻辑理论, 它对“认知主体是如何理性地改变其信念状态”这一认识论问题做出系统的逻辑刻画。从研究主题来看, 这两者都研究理论或信念状态的改变, 它们的亲缘性非同一般。但从实际发展脉络来看, 它们发展迄今是相互独立的, 学界对它们的亲缘关系缺乏系统的研究。本文将就此作出尝试, 从信念修正理论的角度提出一种重新审视悖论相关论题的新思路, 在此基础上提出一种基于信念修正理论的新的解悖方法论。我们先在第一节简述悖论研究的研究背景和发展趋势, 接着在第二节简介信念修正理论的相关研究现状, 以便于后续深入探讨, 然后在第三节从信念修正理论的角度对悖论研究的相关论题进行细致探析, 最后在第四节提出信念修正视域下的解悖方法论, 并依据这种解悖方法论对罗素悖论给出一种新的解决方案。

收稿日期: 2018-11-13

作者信息: 袁永锋 中山大学哲学系 (珠海)
ganlanshux@126.com

张建军 南京大学哲学系
zhangnju@nju.edu.cn

基金项目: 本文是国家社会科学基金青年项目“信念修正视域下的悖论研究”(16CZX051) 和国家社会科学基金重大项目“广义逻辑悖论的历史发展、理论前沿与跨学科应用研究”(18ZDA031) 的阶段性的成果。

1 悖论研究与解悖方法论

悖论研究是古今中外逻辑学和哲学界的重要阵地,不管是悖论的形成、发现还是解决,都会在逻辑学和哲学界产生深远的影响。我们可以在文献中找到大量相关研究。(参见[9, 14])但正如雷歇尔(N. Rescher)指出,从总体上说,以往的研究基本上是围绕某些具体悖论的解决方案而展开,孤立地、分立地研究某个或某类悖论的形成与解决,对悖论及其解决方法缺乏统一的、全面的处理。([4], 第5页)关于这种统一的处理方法的尝试,汤姆逊(J. F. Thomson)曾富有启发性地揭示了格里灵悖论、理查德悖论、罗素悖论以及理发师悖论的共同结构,并表明这些悖论的根源都在于违背了一个逻辑真理,即对角线引理。根据该引理,解悖者可以揭示出其他结构相同的悖论的形成根源,并可构建一套较为一般化的解悖方法论。([5])不过,国内学者田茂先生则认为二值原则是这些悖论产生的根源。([11])这些研究至少表明:特定种类的逻辑悖论之间具有共同的实质和特征,从而对解悖的一般方法的探求不仅是可能的,而且是必要的。然而,迄今为止对于一般解悖方法论的研究仍处于相对薄弱的状态。([14], 第290页)

由于解悖方法论的研究是对通用解悖方法的普适性探求,故我们不禁要问:怎样的解悖方法才是好的或者说是合理的?罗素认为,解决集合论悖论必须满足下面三个条件:矛盾消失、尽可能使数学原样不动、符合“逻辑的常识”。([10], 第70页)策墨罗(E. Zermelo)也认为,解悖时应寻求必需的足够狭窄的原则,这些原则能够排除所有矛盾,同时这些原则又要充分宽广,保留理论中一切有价值的东西。([8], 第200页)哈克(S. Haack)继而提出,既不应过于宽泛以致于损伤必须保留的推论(勿株连无辜标准),又应充分阻止所有相关的悖论性论证(勿顾此失彼标准);而且她还认为:解悖方案应当提供某些说明,以解释为什么那个前提或原则是可反对的,而不管其表面上如何(哲学上的解决);解悖时应当揭示出被放弃之前提的独立的缺陷,而不应仅仅因其导出悖论而将其放弃以示解决。([3], 第139页)哈克把这种“哲学标准”称为**非特设性标准**。张建军把这三条标准统称为解悖的“罗素-策墨罗-哈克标准”,简称**RZH 解悖标准**。雷歇尔也认为,解悖时不仅要恢复一致性(consistency),还要以最小代价或牺牲来达到这个目的。([4], 第26页)

当然,RZH 解悖标准只是一个宏观的标准,对解悖活动只起到宏观的规范性作用,并没给出指导解悖活动的通用方法论。雷歇尔认为:悖论具有相同的原因,即接受或认可过多的信息,这些信息单独看起来都是似然(plausible)的,但放在一起却导致矛盾;从而也可以通过相同的过程来解悖,即切断导致不一致性的最薄弱的环节。([4], 第5、27页)他还认为,解悖是一个采用多样化战术实行统一策略的过程。([4], 第74页)因此,他提出了一种解悖方法论,其直观思想是:先从疑难簇(aporetic cluster)中找出导致悖论的似然前提,然后排除与不一致性无

关的命题, 获得导致悖论的极小不一致子集, 最后根据外逻辑 (extra-logical) 的认知优先序 (priority ordering) 确定最优的取舍方案, 以最小的牺牲将这个极小不一致集中最弱的似然前提放弃, 从而恢复一致性。([4], 第 29–30 页)

2 信念修正理论

阿尔科荣 (C. E. Alchourrón)、伽登弗斯 (P. Gärdenfors) 和麦金森 (D. Makinson) 于 1985 年共同提出了一种理论改变 (theory change) 的逻辑。([1]) 这标志着信念修正理论的创立。它利用符号逻辑的研究方法和理论工具来研究信念状态的理性改变过程。在这三十多年里, 众学者从不同的认识论角度提出了各具特点的信念修正理论, 使这一领域得到了蓬勃发展, 并对哲学、逻辑学、形式认识论和人工智能等学科的相关领域产生了广泛的影响。(参见 [12]) 它的研究对象主要有三个: 信念状态、接收到的新信息 (或需被放弃的旧信念) 和信念修正过程。其中, 信念修正过程主要分为膨胀 (expansion)、收缩 (contraction) 和修正 (revision) 这三种类型, 它们实质上是从已有信息中挑选出可信信息的理性认知机制。那么怎样的信念修正机制才是理性的呢? 理性的认知主体在发现信念矛盾时会力图发现并放弃不可信的旧信念以恢复一致性。这就是一致性原则。但是必须将这种信念改变尽量控制在最小的范围内, 尽量保留更多有用的信息。这就是最大效用原则 (或最小改变原则)。这二者是学界所公认的两大理性原则, 研究者往往依据它们, 以公理性假定 (postulate) 和函数式构造这两种方式对信念修正过程进行形式刻画。前者对信念修正过程所应满足的性质进行刻画, 后者通过构造可计算函数的方式来刻画这种理性认知机制。而信念修正理论的主要工作即是证明这两种刻画方式之间的等价关系, 即表达定理。

近四年来, 袁永锋和鞠实儿等人系统地研究了理性评价 (rational evaluation) 和膨胀–巩固 (expansion–consolidation) 模式的信念修正理论, 主要探讨了核心信念 (core beliefs) 在理性评价和膨胀–巩固模式的信念修正中的影响机制。(参考 [6, 7]) 他们认为, 当接收到未知可信与否的新信息时, 认知主体会先以核心信念作为可接受性的判据, 对新信息作出理性评价, 将可信与不可信信息区分开来。基于这种直观思想, 他们对理性评价进行了公理性刻画, 并构造了以核心信念集 A 为评价依据, 从新信息 D 中找到不可信信息, 进而获得可信信息的核心评价 (kernel evaluation)。([7]) 核心评价首先找到 D 中的所有与 A 不一致的极小子集, 即 A -不一致核心 (inconsistent kernel)。它们必定包藏有不可信的新信息。然后利用剪切函数从每个 A -不一致核心中删除至少一个不可信的新信息, 剩余的新信息就是该主体所相信的。

袁永锋还认为, 由于有限理性的缘故, 认知主体在评价新信息时可能不能立

即从新信息中找出不可信的信息,此时主体往往会先相信这些新信息,然后再在后续认知实践中发现信念冲突,进而以核心信念作为可接受性的判据,发现并放弃那些不可信的其他信念,以化解信念冲突,使信念状态得以巩固。简言之,这是一种先信念膨胀然后依据核心信念对信念状态进行巩固的过程。袁永锋称之为代谢修正 (metabolic revision), 并对其进行了公理性刻画,还构造了一种核心代谢修正 (kernel metabolic revision)。([6]) 这里信念状态表征为一个二元组 (B, A) , 新信息记为 D , 其中核心信念集 A 是信念库 B 的子集。核心代谢修正首先找到 $B \cup D$ 中的所有 A -不一致核心, 然后利用剪切函数从每个 A -不一致核心中删除至少一个不可信的旧信念, $B \cup D$ 中剩余的信念构成了一致的新信念库, 从而信念状态得到了巩固。

3 悖论研究与信念修正理论

从上面评述中可以看出:矛盾的消除是信念修正理论研究的重点,也是悖论研究的焦点;而且这两个领域都对一般性、通用性的追求报以极大的热情,信念修正理论专注于刻画一般的、形式的信念修正过程,而悖论研究也开始关注一般解悖方法论的研究。由此易知,这两个领域必定有一些共同的论题可供比较和借鉴。

3.1 悖论的本性与信念矛盾

可能有学者认为,悖论仅仅是命题间的逻辑矛盾,而非信念矛盾,因此悖论研究跟信念修正没有直接关系。事实并非如此。如果认知主体既不相信外星人存在,也不相信外星人不存在,那么命题集 {外星人存在, 外星人不存在} 就仅仅是逻辑矛盾,该主体不会认为它包含一个悖论。如果认知主体相信苏格拉底是柏拉图的老师,但不相信它的否定,那么命题集 {苏格拉底是柏拉图的老师, 苏格拉底不是柏拉图的老师} 对于该主体也不包含一个悖论。记 A 为 $\{X \mid X \text{ 是一个集合, 并且 } X \notin X\}$ 。虽然对很多主体而言 A 的定义和 A 是集合构成一个悖论,但是对于不相信 A 是集合的认知主体而言,它们就不构成一个悖论。依此不难看出,对于任何逻辑矛盾:

(1) 如果认知主体不相信导致该逻辑矛盾的每个必要前提,那么该逻辑矛盾对于该主体仅仅是逻辑矛盾而已;

(2) 如果认知主体相信其中某些必要前提,但不相信所有必要前提,那么该逻辑矛盾对于该主体就不仅仅是逻辑矛盾,而且包含一个依据可信前提对不可信前提的归谬 (归谬过程的前提集就属于此情形);

(3) 如果认知主体相信所有必要前提,那么该逻辑矛盾对于该主体就不仅仅是逻辑矛盾而且是信念矛盾,并可能包含一个悖论。

在这三种情形中,为什么只有信念矛盾才可能包含悖论呢?这依赖于悖论之悖的本性。由于逻辑有效的推理是保信念态度的,故似然前提应当蕴涵似然结论。但在悖论的论证中,似然前提的推论却是逻辑矛盾,这就导致了悖谬。悖论之悖也正体现在接受似然前提和反对矛盾结论之间这种命题态度的差异上。如果存在某些必要前提不是似然的,那么这些前提和矛盾结论都不是认知共同体所相信的,也就没有任何命题态度上的反差,该共同体也就不会把该逻辑矛盾称为悖论,而仅仅将其看成是似然前提对非似然前提的归谬。此外,即使悖论最终得到了彻底的解决,解悖者不再相信出错的似然前提,理论危机得到了化解,解悖者依然会因这些似然前提的顽固直观性而继续称它们为悖论,而不会将它们仅仅视为不足道的逻辑矛盾。可见,悖论本质上是一种信念矛盾,而解悖是一种信念修正过程。正如张建军所论证,悖论是相对于认知共同体的背景知识或信念的,而且悖论的发现对这个共同体的共同根本信念提出了挑战;虽然悖论的解决困难重重,但是它们都是可以得到解决的,并且这种解决意味着一个共同体的根本信念得到修正;因此,悖论的发现和解决是人类信念系统或知识体系演化发展的杠杆。([13]; [15], 第 205 页)

3.2 成悖-解悖、膨胀-巩固

在认知过程中,主体有时会由于有限理性的缘故而不能立即从新信息中发现不可信信息,即直观上认为每个命题都是可信的,而这可能会在信念系统中“埋下”不一致性。悖论的形成与这种信念膨胀类似,研究者在构建理论时可能会依据直观而认可过多的信息,并没有意识到理论中所隐含的相互矛盾的内容。正如雷歇尔所言,悖论具有相同的原因,即接受或认可过多的信息,这些信息单独看起来都是似然的,但放在一起却导致矛盾。([4], 第 6、27 页)不过,信念膨胀后往往只是隐性的信念矛盾。在后续认知实践中,认知主体可能会发现它们,并尝试以核心信念为依据寻找似是而非的信息,对信念状态进行巩固,以恢复其一致性。悖论的解决也与此信念巩固类似。研究者在对所构建理论的演绎中可能会偶然发现那些隐性的悖论,即通过逻辑有效的论证将悖论显性化和极小化(即去除无关前提),使认知共同体认识到该理论中隐藏的矛盾,并受理论危机的鞭策而努力寻找该理论中似是而非的命题,进而消除理论中的矛盾。总之,悖论形成的根源在于在信念膨胀过程中认可了隐含矛盾的信息,悖论的发现则是理论内部隐性矛盾被显性化和极小化的过程,而解悖则是一种理论巩固的过程。也就是说,悖论的形成、发现与解决其实是一种膨胀-巩固模式的信念改变过程。不过值得注意的是,悖论的真正解决并不是那种带有主观性的、因人而异的信念修正过程,它需要认知共同体达成共识。虽然对于个人来说,共识的达成并不是悖论得到真正解决的必要条件,但只有认知共同体达成了共识,他们才会认可这个悖论得到了

真正的解决。

3.3 成悖标准与置信度

从集合论悖论和理发师悖论可知,悖论间存在悖谬程度的差异。张建军曾从认知主体对似然前提的“公认度”出发,把这种悖谬程度称为悖论度。([15],第210页)我们将利用置信度概念重新定义悖论度概念。上面已表明,悖论之悖体现在必要似然前提和矛盾结论之间命题态度的差异,故悖论度是它们的置信度差异。这种置信度上的落差越大,悖论度也就越高。正因为悖论间有不同的悖论度,人们对“怎样的似然前提才算形成了一个真正的悖论”这一问题持有不同的意见。虽然悖论度可以被形式化为一种清晰的数值标准,但在这问题上任何清晰的数值界线都是难以辩护的。因此,这里提出三条对真正悖论之形成的非数值标准。

逻辑矛盾标准 必要前提必须(逻辑有效地)蕴涵逻辑矛盾,否则不具备形成悖论的基本资格。

高置信度标准 由于简单矛盾命题的置信度最低,故决定悖论度高低的关键在于必要前提的置信度。因此,每个必要前提都必须具有高置信度,使认知共同体难以拒斥,甚至认为它们是无懈可击的;否则,难以形成认知主体对必要似然前提与矛盾结论的命题态度上的巨大落差,从而不被认可为真正的悖论,甚至仅仅是不足道的逻辑矛盾。

等置信度标准 虽然悖论度高低的关键在于必要似然前提的置信度,但必要似然前提间也可能存在置信度差异。如果存在一必要前提,其置信度明显低于其他必要前提,甚至完全不被认知共同体所相信,那么即使其他必要前提的置信度极高,由这些前提构成的逻辑矛盾的悖论度也非常低,认知共同体甚至不会把该逻辑矛盾称为悖论,而更多将其看成是高置信度前提对低置信度前提的归谬。因此,真正悖论的所有必要前提的置信度应当势均力敌。不过在解悖过程中,这些前提的置信度会发生变化而形成较大差异,否则悖论是不可能得到解决的。

我们把这三条标准称为**成悖标准**。符合这三条成悖标准的理论情境,我们才称之为形成了一个真正的悖论。

3.4 RZH 解悖标准与信念修正原则

信念修正理论的一致性原则和最大效用原则,分别与RZH解悖标准的勿顾此失彼标准和勿株连无辜标准不谋而合。但一致性原则与勿顾此失彼标准之间有一些细微差别之处。前者要求恢复信念状态的全局一致性;而后者只要求恢复悖论所涉及论题的局部一致性即可,不考虑其他悖论所造成的不一致性。最大效用

原则和勿株连无辜标准则要求把信念改变控制在尽量小的范围内, 尽可能地保持理论原貌, 尽量多地保留有用的信息。此外, 与信念修正不同, RZH 标准还有一个特殊的标准, 即非特设性标准。它要求解悖方案不能仅仅为消除矛盾这一目的而放弃似然前提, 更需要在根本上揭示出似是而非之前提的独立的内在缺陷, 并因此而放弃这些前提。需要强调的是, “非特设性标准”不能仅从似然前提的置信度作“外在”理解。解悖时似然前提之取舍的根本标准只能是它们是否包含内在缺陷, 尽管对悖论的“解决”终究需诉诸于认知共同体对这种内在缺陷达成共识。因此, 使用“缺陷性标准”一词能够更好地表达出这种哲学标准的内涵。

虽然置信度与内在缺陷相比只是次要的外在性、主观性标准, 但它常常能有效地帮助认知主体定位那些包含内在缺陷的似然前提, 而且不管是彻底还是不彻底的解悖方案, 最后放弃的都是低置信度的似然前提。因此, 我们认为有必要将置信度作为一种新的外在性解悖标准, 它包含三个规范性的解悖功能: 第一, 解悖前对似然前提进行重新评估和比较, 重新赋予置信度, 力求获得置信度差异; 第二, 根据新的置信度差异, 引导认知主体定位那些包含内在缺陷的似然前提; 第三, 确保所放弃似然前提的置信度是最低的, 不能放弃高置信度的似然前提, 反而保留低置信度的似然前提。与这种**置信度标准**相比, 缺陷性标准则是一种更彻底但难以达到的哲学标准。从集合论悖论的解悖史可以看出, 似然前提的内在缺陷往往是非常难发现的, 否则悖论也就不会如此吸引古今智者。

3.5 解悖与信息评价

从上一节可知, 理性评价刻画的是认知主体依据核心信念对新信息可信与否进行评价的过程。解悖活动也会遇到类似情况。在解悖过程中, 往往会遇到这种情形, 即所有似然前提在直观上都难以拒斥甚至无懈可击, 但由于导致矛盾, 它们之中必然包含似是而非的命题, 因此解悖者必须在这些前提当中作出取舍。然而, 这些似然前提在置信度方面往往都势均力敌, 以致于难以发现那些似是而非的命题并作出取舍。因此, 要想解决悖论, 解悖者必须对这些看似都无懈可击的前提进行重新评价, 重新赋予它们新的置信度, 以发现潜藏于其中的似是而非的前提。正如雷歇尔所言, 解悖前对这些似然前提进行认知评价或优先性确定是不可避免的。([4], 第 27 页)

对信息进行理性评价, 必定要依据某种判据来进行。悖论的发现表明信念膨胀时用于信息评价的那种判据并没有胜任它的责任, 因此需要一套更严格、更精细、更可靠的判据来对似然前提进行重新评价。但这种判据往往不在所涉及的似然前提及其推论乃至信念系统之中, 因而难以发现。不过这些判据必然离不开且潜藏于该悖论所涉及概念的深层结构之中。它们一旦被发现则可能会成为解悖者的核心信念, 并且继而作为似然前提的缺陷性的判据。如果不对导致悖论的直观

成立的必要前提进行重新评价，是无法找出似是而非的前提的，从而也是无法成功地解决悖论的。因此，将理性评价理论的直观思想和理论工具应用于解悖的想法，不仅是可能的，而且也是必要的。

3.6 雷歇尔解悖方法论与信念修正的基本概念

信念修正理论与以往悖论研究的关注点的一个明显差异在于：前者是对信念修正现象的一般过程的形式刻画，而不研究具体的信念修正活动；但是后者则主要关注具体悖论的解决方案上，在通用解悖方法论方面着墨甚少。雷歇尔的解悖方法论可以说是这一方面的初级尝试，为这方面的研究提供了新起点。他提出了下面这些与信念修正理论非常相似的基本概念。([4])

解悖方法论	疑难簇	极大一致集	极小不一致集	认知优先序	R/A 选择
信念修正理论	信念状态	一致保留	不一致核心	认知牢固序	剪切函数 选择函数

其中，疑难簇是指一种其元素皆为似然命题的不一致命题集。它与信念状态相对应，但后者表征的却不仅仅是不一致命题集，它包括了认知主体所相信的所有命题。极大一致集是指一种增加任何疑难簇内其他元素皆导致不一致性的疑难簇的一致子集，极小不一致集是指一种删除任何元素皆导致一致性的疑难簇的不一致子集。它们分别与一致保留（remainder）和不一致核心（kernel）相对应；只不过一致保留（不一致核心）是指整个信念状态而非局部疑难簇的极大（极小不一致子集。在解悖过程中，要想恢复一致性，必须破坏所有不一致核心，即阻止所有悖论性论证。但是逻辑是价值中立的，它只命令我们必须在悖论的似然前提之间作出取舍，但并没有告诉我们如何作出选择。([4]，第 27 页) 解悖需要一个参考标准为似然前提之取舍提供辩护。对此，雷歇尔认为，解悖的关键在于找到并切断不一致性的证明链中最薄弱的环节。([4]，第 15 页) 他认为这种强弱关系是比较的似然性（或者说置信度）方面的关系，并采用认知优先序关系 $<$ 来表征这种关系。认知牢固序 $\leq$ 与此类似，但它表征的是认知主体对信念的坚持程度； $P \leq Q$ 的意思是说：Q 在认知上至少像 P 一样得到认知主体的坚持。([2]) 虽然将这种序关系应用于信念修正是合适的，但它只是一种主观偏好关系，并不适合于解悖，因为解悖是对真的追求。不过，雷歇尔以相对客观的认知优先序作为取舍标准的做法具有一定的可取之处。最后，R/A 选择方案是指通过保留和删除元素致使不一致疑难簇变得极大一致的选择方案。它的功能与剪切函数和选择函数类似。但 R/A 选择方案是比较简单甚至有瑕疵的，它们只是简单地把最低置信度的似然前提删除。但其实对似然前提的放弃并不是简单地删除该前提即可，因为删除该前提后的疑难簇可能还逻辑地蕴涵该前提（即作为隐性信念），故该悖论可

能仍旧作为隐性信念矛盾潜藏于疑难簇中。所以, 解悖时应像信念修正理论那样, 找到所有蕴涵该悖论性结论的不一致核心, 然后从每个这种不一致核心中放弃掉至少一个似是而非的前提, 从而保证每个不一致核心都遭到了破坏, 使得该疑难簇或信念状态不再蕴涵该悖论。

4 解悖方法论及其案例分析

基于上述基本概念, 雷歇尔提出了一种解悖方法论 ([4]; 第 29–30 页), 这里将他的解悖步骤稍作整理如下。

- (R1) 解悖前详细列出导致悖论的似然前提, 并展示它们如何导致矛盾。
- (R2) 将此不一致集还原为推理上不冗余的极小不一致子集, 去除那些与不一致性无关的“无辜”命题。
- (R3) 详细列出这个疑难簇的所有极大一致子集。
- (R4) 枚举出所有可以避免不一致性的 R/A 选择方案。
- (R5) 评定相关命题的优先序关系, 以作为解悖时信念取舍的依据。
- (R6) 依据优先序关系确定最优的 R/A 选择方案。

然而, 这种解悖方法论有一个致命的不足, 雷歇尔把认知优先序关系作为似然前提的取舍标准, 这是舍本逐末的。虽然认知主体可以根据认知优先序关系 (或置信度差异) 来推测哪个前提是虚假前提, 但主观置信度只是辅助工具, 它无法为放弃似然前提提供有效辩护。即便所有人在所有似然前提上都持共同置信度, 依据此置信度设计的解悖方案也不一定就彻底解决了悖论。解悖所要求的是一种更客观的标准, 即缺陷性标准。只有找到了被放弃前提的内在缺陷, 在相当客观的意义上确定了被放弃前提的虚假之后, 才能真正放心地认为: 悖论确实得到了“解决”。而这个内在缺陷的根本在于违反了某些评价依据, 而可作评价依据的唯一候选就是真理: 逻辑-数学先验真理或经验真理; 否则, 不能称与那些“评价依据”的冲突是一种“违反”, 也就不能称被放弃的似然前提包含某种“内在缺陷”。在纯粹的逻辑悖论的解决过程中, 我们是可以将逻辑数学等先验真理作为缺陷与否的判据的。但是在经验科学或日常生活中的悖论的解决过程中, 我们是无法做到这一点的。由于经验真理只能趋近而无法达到的本性, 在实际解悖过程中, 我们无法把握到经验真理作为评判缺陷与否的判据, 而只能以极高置信度的核心信念作为缺陷与否的判据。如果恰好是核心信念导致悖论, 那么需要提高核心信念的准入门槛, 将更加可靠无疑的、置信度更高的信念作为核心信念, 甚至在解决纯粹逻辑悖论时可以仅仅将逻辑-数学先验真理作为核心信念。这种作为缺陷之判据的核心信念一般有两种来源, 其一是来自于概念体系所蕴涵的先验真理, 其二是来自于实证科学所追求的经验真理。但受直观性干扰, 似然前提的内在缺陷是

难以寻找的,从而其所违背的先验真理或经验真理都是难以发现的。当然有些判据可能是隐性信念,这种情况下解悖是相对容易的。困难的情况是,与似然前提相关的缺陷性的判据往往不是(隐性或显性)信念,而是有待发现的复杂的先验真理和经验真理,此时则需要解悖者在复杂而曲折的解悖过程中,通过艰深的思维实践和科学实践对相关概念和命题进行深刻认识,以形成新的核心信念作为似然前提之缺陷性的判据。

似然前提作为直观成立的前提,它们往往都有一定程度的真实性。而根据成悖的逻辑矛盾标准和命题的二值原则,必然有一似然前提包含“内在缺陷”(虚假性)。这里把该前提称为**似是而非**之前提,直观成立的部分即所是之处,而内在缺陷则是所非之处。依赖于这种所是之处的直观性是相当强大或顽固的,这会掩盖其所非之处,以致于认知主体难以拒斥,悖论才得以形成。因此,解悖不是将置信度最低的似然前提放弃即可,而必须要找到极高置信度的核心信念,依据它们对似然前提进行评价,以发现那些似是而非之前提,并对它们的所非之处进行严厉的反驳,进而表明这些前提在违反那些核心信念的意义上包含一种“内在缺陷”,最后发现并阻断该前提直观成立之判断的形成机制,悖论才能最终得到彻底解决。基于此,我们根据理性评价理论和代谢修正理论提出下面这种信念修正视域下的解悖方法论。

- (Y1) 解悖前从理论中找出导致悖论的不一致核心,排除那些与不一致性无关的命题,并展示这些不一致核心如何导致悖论。
- (Y2) 对不一致核心所涉及概念进行深刻反思和演绎推理,并对经验命题反复进行可靠的科学实践,按严苛的标准从中获得一致的核心信念集 A,作为似然前提缺陷与否的判据。
- (Y3) 依据核心信念集 A 对不一致核心进行评价,枚举出所有与 A 相矛盾的 A-不一致核心,似是而非之前提必然藏于它们之中,从而将怀疑范围缩小至 A-不一致核心。
- (Y4) 如果这个怀疑范围内仍有多个似然前提,以致于不知如何取舍,那么回到 (Y2) 对核心信念集 A 进行扩充,直至怀疑范围内只有一个前提为止,该前提即是似是而非之前提。
- (Y5) 在整个理论范围内对这些似是而非之前提进行信念收缩,把保留下来的似然命题作为解悖后的新理论。
- (Y6) 对每个似是而非之前提,逐一揭示其直观成立之判断的形成机制。
- (Y7) 对每个似是而非之前提,逐一揭示其内在缺陷,并从核心信念集中寻找严厉反驳。
- (Y8) 对每个似是而非之前提,利用对其内在缺陷的反驳来阻断其直观成立之判断的形成机制。

其中, 步骤 (Y1) 实现的是雷歇尔解悖方法论中步骤 (R1) 和 (R2) 的功能, 目的在于剔除与悖论形成无关的命题, 展示不一致核心里的必要前提是如何导致悖论的。其中对无关命题的剔除需要将所有似然前提等价地转换为合取范式, 进而转换为合取支命题的集合, 然后再从中寻找极小的不一致子集, 即不一致核心。步骤 (Y2) 力图找出内在缺陷的评判依据, 这需要在解悖历程中反复进行。步骤 (Y3) 利用这些判据对不一致核心内势均力敌的似然前提进行重新评价, 以此定位似是而非之前提的怀疑范围, 把无辜的似然前提排除在外。步骤 (Y4) 的目的在于进一步缩小似是而非之前提的怀疑范围, 直至找到似是而非之前提为止。步骤 (Y5) 对这些似是而非之前提进行信念收缩, 破坏所有蕴涵这些前提的信念子集, 保证解悖后不再蕴涵这些似是而非的前提, 从而恢复理论的一致性。步骤 (Y6) 和 (Y7) 则揭示这些似是而非之前提直观成立的根源, 并依据核心信念对它们的所非之处进行严厉反驳, 即论证似是而非之前提的否定。此时解悖尚未结束, 解悖者见到此前提依然会认为其显然成立。比如, 依据直观的集合概念认为: 所有非自属集当然构成一个集合。解悖不仅仅是论证似是而非之前提的否定, 还要求阻断该前提直观成立之判断的形成机制, 以持续抵抗那些似是而非之前提的顽固直观性的误导。这就是步骤 (Y8) 的职责。

我们认为, 罗素的类型论就仅仅给出了“ $\{X \mid X \text{ 是集合且 } X \notin X\}$ 不是一个集合”的特设性论证, 但是没有阻断“ $\{X \mid X \text{ 是集合且 } X \notin X\}$ 是一个集合”这个直观判断的形成机制。这里按照上述解悖方法论 (Y1)–(Y8) 对罗素悖论给出一个新的解悖方案。

(Y1') 首先, 我们找到导致罗素悖论的不一致核心 $O = \{\alpha, \beta\}$, 其中

α : $\{X \mid X \text{ 是集合且 } X \notin X\}$ 是一个集合 (记此集合为 H),

β : 如果 X 是一个集合, 那么要么 $X \in X$ 要么 $X \notin X$ 。

(Y2') 然后, 我们通过下面的反思来获得一致的核心信念集 A 。

令 $P(X)$ 为任一关于集合 X 的谓词。显然, 按 $B = \{X \mid P(X) \text{ 且 } X \notin X\}$ 方式定义的对象 B 都不满足 $P(X)$, 而是该谓词外延之外的对象。比如, 令 $P(X)$ 的外延为 $\{\{1\}, \{2\}, \{3, 4\}\}$, 按 B 的定义可得 $B = \{\{1\}, \{2\}, \{3, 4\}\}$, 显然 $B \notin \{\{1\}, \{2\}, \{3, 4\}\}$, 即 B 超出了 $P(X)$ 的外延。故我们可以把这种定义方式称为超越性定义方式。由于这种超越性定义方式只会产生 $P(X)$ 外延之外的对象, 当 $P(X)$ 表示“ X 是一个集合”时, 按此方式定义的对象 H 也在该谓词外延之外, 即 H 不是一个集合。此时, 不能用符号“ $\{$ ”和“ $\}$ ”来暗示它是集合。即使故意把 H 看作一个集合并将其添加进谓词“ X 是一个集合”的外延之中, 然后按照上述超越性定义方式重新定义 H' , 这个 H' 也仍然不是一个集合。不过, 这个 H 可能是其他种类的数学对象, 可用其他符号来表示这种定义, 比如 $[X \mid X \text{ 是集合且 } X \notin X]$ 。

此外,这种非集合对象是无穷无尽的。证明如下:令 $C = \{n\}$ (n 为某整数), $P(X)$ 表示“ X 是一个集合但不是 C ”,并且 $S = \{X \mid P(X) \text{ 且 } X \notin X\}$ 。假设 S 是一个集合。由于 $\{n+1\} \in S$, 故 S 不是 C 。故 $P(S)$ 成立。此时易得 $S \in S \leftrightarrow S \notin S$, 矛盾。故假设错误, S 不是一个集合。再者,对于任意给定的 $C = \{n\}$, $\neg P(\{n\}) \wedge (\{n\} \notin \{n\})$ 但 $P(\{m\}) \wedge (\{m\} \notin \{m\})$ (其中 $m \neq n$), 即 $\{n\}$ 不满足 S 的定义条件但 $\{m\}$ 满足 S 的定义条件。因此,只要给定的 n 不同,所定义出的 S 就各不相同。因此,存在无穷无尽的非集合对象。

总之,这种超越性定义方式具有适用范围:对于某些谓词 $P(X)$, 它会正常地产生集合;但对于诸如“ X 是一个集合”的另一些谓词 $P(X)$, 它并不构成一个集合,而产生非集合的数学对象。因此,我们得到下列核心信念:
 γ : 定义 $\{X \mid P(X) \text{ 且 } X \notin X\}$ 是超越性定义,所定义的对象总是超出谓词 $P(X)$ 的外延。

另外,根据集合概念,如果 X 是一个集合,那么它的所有元素都是确定的,因此要么 $X \in X$ 要么 $X \notin X$ 。故 β 也是一个核心信念。因此,我们可以获得关于集合的核心信念集 $A = \{\gamma, \beta\}$ 。

- (Y3') 不一致核心 O 的子集中与 A 相矛盾的 A -不一致核心只有 $K = \{\alpha\}$, 故似是而非之前提必然藏于 K 中。
- (Y4') 由于 K 中只有一个元素,故 α 即为似是而非之前提。
- (Y5') 对集合论进行信念收缩,使集合论不再蕴涵 α , 而且限制超越性定义方式在集合论中的应用,形成新的集合论。
- (Y6') 揭示出似是而非之前提 α 这个直观判断的形成机制,即:认为满足定义条件的对象聚合 $\{X \mid X \text{ 是集合且 } X \notin X\}$ 就是集合。
- (Y7') 似是而非之前提 α 的内在缺陷在于:由于任何采用超越性定义方式 $\{X \mid P(X) \text{ 且 } X \notin X\}$ 得到的对象都超出了谓词 $P(X)$ 的外延,故按此方式定义的 H 也超出谓词“ X 是一个集合”的外延,即它不是一个集合,但该定义却采用了集合的表示方式伪装成一个集合。
- (Y8') 根据 (Y6') 和 (Y7') 可知,满足定义条件“ $P(X)$ 且 $X \notin X$ ”的对象聚合不一定是集合,因为由超越性定义所构造之对象超出了谓词“ $P(X)$ ”的外延。这就阻断了“ $\{X \mid X \text{ 是集合且 } X \notin X\}$ 是一个集合”这个直观判断的形成机制。

从上面论述不难发现,与罗素的类型论相比,这个解决方案不需要因解悖的缘故而对集合给出特设性的类型区分,故满足解悖的非特设性标准或缺陷性标准。这个案例也表明:上述解悖方法论具有确实可行性,对悖论研究具有一定的指导意义。

5 结语

综上所述, 悖论研究和信念修正理论这两个亲缘领域, 包含大量的相关主题可进行深入的比较研究。但由于文章目的和篇幅的关系, 这里并不对更多相关论题作更深入细致的探讨, 这是未来的工作方向。我们相信, 这种跨领域交叉研究的新思路, 不仅能给悖论研究领域带来新的活力, 推动悖论研究尤其是解悖方法论的进一步发展, 而且也能信念修正理论领域指出新的研究方向, 从而进一步推动逻辑学和哲学的发展。

参考文献

- [1] C. E. Alchourrón, P. Gärdenfors and D. Makinson, 1985, “On the logic of theory change: Partial meet contraction and revision functions”, *Journal of Symbolic Logic*, **50**(2): 510–530.
- [2] P. Gärdenfors and D. Makinson, 1988, “Revisions of knowledge systems using epistemic entrenchment”, in M. Y. Vardi (ed.), *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, pp. 83–95, Burlington, Massachusetts: Morgan Kaufmann Publishers.
- [3] S. Haack, 1978, *Philosophy of Logics*, Cambridge: Cambridge University Press.
- [4] N. Rescher, 2001, *Paradoxes: Their Roots, Range, and Resolution*, Chicago: Carus Publishing.
- [5] J. F. Thomson, 1962, “On some paradoxes”, in R. J. Butler (ed.), *Analytical Philosophy*, pp. 104–119, New York: Barnes & Noble.
- [6] Y. Yuan, 2017, “Rational metabolic revision based on core beliefs”, *Synthese*, **194**(6): 2121–2146.
- [7] Y. Yuan and S. Ju, 2015, “Rational evaluation in belief revision”, *Synthese*, **192**(7): 2311–2336.
- [8] E. Zermelo, 1967, “Investigation in the foundation of set theory I”, in J. van Heijenoort (ed.), S. Bauer Mengelberg (tran.), *From Frege to Gödel*, pp. 199–215, Cambridge, Massachusetts: Harvard University Press.
- [9] 陈波, 悖论研究, 2014 年, 北京: 北京大学出版社。
- [10] 罗素 (著); 温锡增 (译), 我的哲学的发展, 1982 年, 北京: 商务印书馆。
- [11] 田茂, “论二值原则——部分逻辑悖论产生的根源”, *逻辑学研究*, 2018 年第 1 期, 第 92–112 页。
- [12] 袁永锋, “基于核心信念的非优先复合修正理论”, 2014 年, 博士学位论文, 中山大学。
- [13] 张建军, “A study of the definition of ‘logical paradox’”, *逻辑与哲学 (第三届两岸逻辑教学与学术会议文集)*, 2009 年, 文富文化出版社。
- [14] 张建军, *逻辑悖论研究引论*, 2014 年, 北京: 人民出版社。
- [15] 张建军等, *当代逻辑哲学前沿问题探讨*, 2014 年, 北京: 人民出版社。

(责任编辑: 罗心澄)

A Preliminary Study on Paradox Research based on Belief Revision

Yongfeng Yuan Jianjun Zhang

Abstract

Paradox research is an ancient research field, and belief revision is a new logic theory developed in the latest three decades. Paradox formation, discovery, and resolution are some kinds of canonical and complex belief revision processes. Although there are many extraordinary interrelations between the two fields, there are not any systematical researches about the interrelations. This article proposes one new research sight for paradox-related problems from the points of view in belief revision, and then proposes one new methodology for paradox resolution based on belief revision theory, and at last we propose a new resolution for Russell's Paradox based on this new methodology. This new interdisciplinary research view will not only promote the development of paradox research, but also point out a new research direction in belief revision.

Yongfeng Yuan Department of Philosophy (Zhuhai), Sun Yat-sen University
ganlanshux@126.com

Jianjun Zhang Department of Philosophy, Nanjing University
zhangnju@nju.edu.cn