

试析博弈困境的两种解决方案^{*}

——以旅行者困境为例

郭美云, 党学哲

(西南大学 逻辑与智能研究中心, 重庆市 400715)

摘要:在博弈困境中, 纳什均衡所预测的博弈解与人们的直观相矛盾, 而且越来越多的博弈实验表明, 这与现实博弈的结果也存在巨大反差。博弈困境的出现引起了人们对经典博弈论理性预设的质疑, 从而提出不同的逻辑模型和解决方案。本文以旅行者困境为例, 对 Halpern & Pass 提出的重复后悔度最小化模型和 Capraro 提出的基于联盟与合作的概率推理模型两种方案进行分析比较, 考察它们是如何成功地解释和预测旅行者困境中选手实际博弈行为, 并对这两种方案的普适性问题和应用前景进行了展望, 同时对博弈论的基础进行了哲学反思。

关键词: 博弈困境; 旅行者困境; 解决方案; 重复后悔最小化模型; 概率推理模型

中图分类号: B815 **文献标识码:** A **文章编号:** 1673-9841(2015)02-0018-08

纳什均衡(Nash Equilibrium)概念的提出和存在性证明^[1]奠定了博弈论这门学科的基础, 为理解和预测人们在策略互动中的行为提供了强而有力的工具。但是, 随着博弈论的发展, 人们普遍意识到, 甚至通过实验研究也发现, 在有些博弈中, 纳什均衡所预测的博弈结果并不符合人们的直观和各种实验研究的结果。人们把这些纳什均衡与直观或现实严重冲突的博弈称为博弈困境, 著名的例子有囚徒困境(Prisoner's Dilemma)、旅行者困境(Traveler's Dilemma)、蜈蚣博弈(Centipede Game)、纳什讨价还价问题(Nash bargaining problem)、伯川德悖论(Bertrand competition)、公共物品供给博弈(Public Good Game)、最后通牒博弈(Ultimatum Game)和独裁者博弈(Dictator Game)等。

旅行者困境是由著名经济学家 Kaushik Basu 于 1994 年^[2]提出来的博弈中的一个新的困境。正如他本人所说:“旅行者困境是一个特殊的并且令人信服的悖论, 在这里, 无情的博弈论理性和直觉观念无法保持一致。”^{[2]391}该困境融合了以往困境中具有代表性的一些主要特征, 从而使博弈论中的根本问题更为集中地得到展现。

旅行者困境的发现和提出, 立刻引起了学术界的广泛关注, 国际上不少博弈论学家和逻辑学家从理论和实验两个方面分别展开研究^[3-7]。与此相反, 国内学者虽然对一般意义上的博弈困境及其产生原因已有所关注^[8-9], 但是对针对博弈困境的各种解决方案缺乏细致而深入的学理分析和研究。对解决方案的深入研究可以加深我们对人类社会中各种博弈困境的理解的同时, 有助于寻找

^{*} 收稿日期: 2014-09-02

作者简介: 郭美云, 哲学博士, 西南大学逻辑与智能研究中心, 副教授。

基金项目: 重庆市人文社会科学重点研究基地重点项目“博弈的逻辑与认知基础研究”(14SKB047), 项目负责人: 郭美云; 国家社会科学基金重点项目“现代逻辑视野的认知研究”(11AZD57), 项目负责人: 何向东; 中央高校基本科研业务费专项资金资助项目“博弈中互动认知的逻辑研究”(SWU1309380), 项目负责人: 郭美云。

新的理论和现实解决方案,还可以避免对博弈论泛泛而谈的批评和指责。本文以旅行者困境为例,对 Halpern & Pass 提出的重复后悔度极小化模型^[3]和 Capraro 提出的基于联盟与合作的概率推理模型^[4]两种方案进行分析比较,以窥它们是如何成功地解释和预测旅行者困境中选手实际博弈行为的,并分析这两种方案各自存在的问题。

一、旅行者困境的产生

关于旅行者困境的描述有很多不同的变种,其中最主要的变化是提供赔偿的范围和罚金的比重不同。以下给出赔偿范围为 2~100、罚金为 2 的版本:

假设有两名乘客从太平洋的某个岛上度假后乘同一个航班返回自己的国家。不幸的是,他们的行李箱被航空公司弄丢了。两个行李箱装的都是他们在岛上购买的同一种古董。航空公司答应给予他们赔偿,但因为很难评估古董的实际具体价值,两位乘客可以在 2 美元和 100 美元这个区间中任意选取一个数值要求赔偿。但为了保证乘客提出合理的赔偿价格,航空公司特地制定了赔偿规则如下:如果他们要求赔偿的数值一样,那么航空公司就直接按这个数值提供赔偿。如果他们要求赔偿的数值不一样,则按照要求少的那位乘客提出的价格标准进行赔偿,但要求少的那位乘客将获得额外 2 美元的奖励,而要求多的那位乘客则在要求少的那位乘客提出的价格基础上减少 2 美元作为罚金。例如,两位乘客提出的赔偿分别为 98 美元和 100 美元,则他们获得的赔偿各自为 100 美元和 96 美元。

直观地看,两位乘客应该都提出最高的赔偿要求即 100 美元,因为这样的话,他们都可以获得 100 美元的赔偿,或者至少双方都应该提出较高水平的索赔,因为这样的话,即使获得的赔偿没达到 100 美元,仍可以达到较高水平。

但是如果按照经典博弈论的假设,两位乘客都是追求个人的收益最大化的理性人并且是能进行任意有限步推理的高手,那么在给定其中一方提出 100 美元的赔偿的情况下,另一方会选择放弃索赔 100 美元而只索赔 99 美元,因为那样的话,他就可以直接获得 101 美元。但是另一位乘客也知道这一点,则他会放弃索赔 99 美元而提出 98 美元的索赔,因为这样的话,他将获得 100 美元的赔偿,而不是 97 美元。以此类推,按照这个逻辑,他们最终都只会选择 2 美元的赔偿要求,从而陷入每人只能得到 2 美元的困境。实际上这正是博弈论中运用重复剔除被弱占优策略算法(Iterated Removal of Weakly Dominated Strategy,简称为 IRWDS)和纳什均衡所预测的结果。虽然整个推理过程无懈可击,但结果却与人们直观严重冲突,并且与现实中人们的行为选择也不相符合。

国际上有不少研究者针对不同版本的旅行者困境展开实验研究^[5-7]。例如,在赔偿范围为 2~100 美元之间,罚金为 2 美元的情况下,Becker^[6]等在博弈论学会的会员中间所做的调查表明,45 名选手中只有 3 名选择了纳什均衡所预测的博弈结果(2,2),有 38 名选手选择了 90 美元以上的索赔要求,而这 38 名中有 31 名甚至提交了 96 美元以上的索赔。实验还表明,45 名选手中提交 97 美元索赔的选手最后获得的收益最高,平均每人有 85.09 美元。而提交 2 美元索赔的两名所获得的收益是最低的,平均每人只有 3.92 美元。

有意思的是,实验研究还表明,人们实际的选择范围会随罚金比重的变化而不同。Capra^[5]等所做的重复博弈(Repeated Game)实验表明,在罚金很低的情况下,人们倾向于提出较高的索赔要求,并在重复博弈实验中保持这一较高的索赔要求。而在罚金较高的情况下,人们开始也提出较高的索赔,但经过几轮之后,最后会提出纳什均衡所预测的 2 美元。这些实验数据均从不同程度上验证了人们的直观预测,确证了旅行者困境中人的行为选择与纳什均衡的严重偏离,更重要的是它们为新理论提供了模型检验的重要标准。

二、旅行者困境的解决方案

经典博弈论对旅行者困境中人们的行为预测是令人失望的,它所预测的结果不仅不符合直观,

更与实验数据严重偏离。这一理论与实践之间的反差表明,旅行者困境不仅是局中人的困境,也是博弈理论的困境。为此,不少博弈论学家开始从不同的角度提出不同的逻辑模型,试图对实验中看似非理性的行为寻找理性的解释方案,从而对现实中人们的博弈行为作出更好的预测。

目前比较成功的解决方案有两种:一种是后悔理性的角度,即一个理性选手总是让自己的后悔最小化,如 Halpern & Pass(2011)提出的重复后悔最小化(Iterated Regret Minimization)解决方案。另一种则从理性选手会试图寻求结成联盟以达成合作的角度,在期望效用理论(Expected Utility Theory)的基础上,Capraro 提出的基于联盟与合作的概率推理模型的解决方案。

(一)重复后悔最小化方案

后悔最小化是决策论中用来解释选手如何进行理性决策的一个重要概念。要把后悔这一概念引入到博弈论,必须考虑如何给出两个不同的后悔概念,一个是在给定其他选手选择的情况下,如何给出一个选手当前选择的后悔程度。另一个是不管其他选手作何种选择,如何比较两个不同策略自身的后悔程度。Halpern & Pass(2011)成功地解决了以上两个问题,从后悔的角度定义出一种与经典理论不同的理性概念,从而给出了一个理性主义的解决方案。

后悔度最小化方案是对经典博弈论中重复剔除被弱占优策略算法的改进。为具体说明这一点,表 1 给出旅行者困境的矩阵式表达:

表 1 旅行者困境的矩阵式表达

		乘客 B						
		100	99	98	97	...	3	2
乘客 A	100	100,100	97, <u>101</u>	96,100	95,99	...	1,5	0, <u>4</u>
	99	<u>101</u> ,97	99,99	96, <u>100</u>	95,99	...	1,5	0, <u>4</u>
	98	100,96	<u>100</u> ,96	98,98	95, <u>99</u>	...	1,5	0, <u>4</u>
	97	99,95	99,95	<u>99</u> ,95	97,97	...	1,5	0, <u>4</u>

	3	5,1	5,1	5,1	5,1	...	3,3	0, <u>4</u>
	2	4,0	4,0	4,0	4,0	...	<u>4</u> ,0	<u>2</u> , <u>2</u>

一个选手如果在给定其他选手选择的情况下,自己单方面偏离这个选择并不能获得更好的收益,那么当前选择就是最佳反应(Best Response)。一个博弈结果是纳什均衡,当且仅当各个选手的当前选择都是最佳反应。

在表 1 中,画线的向量表示其中一个乘客在给定另一个乘客选择的前提下其当前选择是最佳反应。不难看出,只有在(2,2)中,两位乘客的当前选择都是最佳反应,因此是纳什均衡。在纳什均衡中各个选手都安于现状,没有动力偏离当前的选择,因此是个均衡状态。

纳什均衡只是从静态的角度对一个稳定的博弈结果所应该具有的性质的描述。它并没有对各个选手如何通过推理理性地到达这个均衡状态的过程进行说明。因此后来博弈论又发展出求解纳什均衡的重复剔除被弱(强)占优策略算法。

一个选择或者行动又可被称作一个策略。一个当前策略是被弱占优的(Weakly Dominated),当且仅当该选手存在另一个策略,该策略一定和当前策略至少一样好,并且在一些情况下该策略还比当前策略严格好。因此,一个理性的选手是不会选择一个被弱占优的策略的。

我们看到,在旅行者困境中,乘客 A 的 100 美元这个策略就是被弱占优的,因为从表 1 可以看出,在乘客 B 做任意选择情况下,乘客 A 选择 99 美元的收益总是不少于选择 100 美元的收益,并且在乘客 B 选择 99 美元和 100 美元这两种情况下,乘客 A 选择 99 美元的收益严格高于选择 100 美元的收益。因此,作为理性选手的乘客 A 不会选择 100 美元的索赔要求。同理,乘客 B 也不会提出 100 美元的索赔要求。因此,如果两位乘客都足够理性并且知道对方也是理性的话,那么他们都会确信对方不会提出 100 美元的这个索赔要求。

因此,双方都会把 100 美元这个被弱占优策略剔除掉,从而进入表 2 的子博弈中。

表 2 旅行者困境第一个子博弈的矩阵式表达

		乘客 B					
		99	98	97	...	3	2
乘客 A	99	99,99	96, <u>100</u>	95,99	...	1,5	0, <u>4</u>
	98	<u>100</u> ,96	98,98	95, <u>99</u>	...	1,5	0, <u>4</u>
	97	99,95	<u>99</u> ,95	97,97	...	1,5	0, <u>4</u>

	3	5,1	5,1	5,1	...	3,3	0, <u>4</u>
	2	4,0	4,0	4,0	...	<u>4</u> ,0	<u>2</u> , <u>2</u>

同样,不难发现,在这个子博弈中 99 美元对双方来说也是一个被弱占优策略。因此,乘客 A、B 都会剔除索赔 99 美元这一被弱占优策略,从而进入下一个子博弈当中。以此类推,如果两位乘客足够理性并且知道对方也是理性的话,根据重复剔除被弱占优策略算法,最后一个子博弈为表 3:

表 3 旅行者困境最后一个子博弈的矩阵式表达

		乘客 B	
		3	2
乘客 A	3	3,3	0, <u>4</u>
	2	<u>4</u> ,0	<u>2</u> , <u>2</u>

不难看出,最后一个子博弈具有和囚徒困境相同的博弈结构。在这个子博弈中,3 美元是一个被强占优策略(Strongly Dominated Strategy),因为无论一方作何选择,另一方选择 2 美元得到的收益都严格高于选择 3 美元的收益。因此,理性选手不会选择被强占优策略,最后双方博弈的结果为(2,2)。事实上,这也正是纳什均衡理论所预测的结果。

与囚徒困境不同的是,在旅行者困境的纳什均衡求解过程中,融合了博弈论中重复剔除被弱占优策略和重复剔除被强占优策略两种经典算法。因此,旅行者困境是比囚徒困境更为复杂和严重的困境。

在给定其他选手选择的情况下,Halpern & Pass(2011)将当前策略的后悔程度定义为选择最佳反应策略所获得的收益与采取当前策略下所获得收益的差值。据此,我们可以将旅行者困境的博弈模型转换为表 4 所示的后悔度模型:

表 4 旅行者困境的后悔度模型

		乘客 B						
		100(3)	99(3)	98(3)	97(3)	...	3(96)	2(97)
乘客 A	100(3)	1,1	3,0	3,1	3,2	...	3,96	2,97
	99(3)	<u>0</u> ,3	1,1	3, <u>0</u>	3,1	...	3,95	2,96
	98(3)	1,3	<u>0</u> ,3	1,1	3, <u>0</u>	...	3,94	2,95
	97(3)	2,3	1,3	<u>0</u> ,3	1,1	...	3,93	2,94

	3(96)	96,3	95,3	94,3	93,3	...	1,1	2, <u>0</u>
	2(97)	97,2	96,2	95,2	94,2	...	<u>0</u> ,2	<u>0</u> , <u>0</u>

不难看出,一个选手策略的后悔度是随着对方选择的不同而发生变化的,那么如何评估一个策略本身的后悔指数呢?直观来说,一个理性的选手应该尽可能地避免最大后悔事件的发生。因此,可以将一个策略的后悔度定义为它在所有情况下的最大值。例如,选择策略 2 的最大后悔度对于双方来说都是 97(见表 4)。因此,在考虑后悔理性的背景下,选手所面临的形势发生了根本的变化。选择策略 2 成为后悔指数最高的一个策略,任何一个理性的选手都会首先将其剔除掉。因此,博弈双方可以进入一个没有策略 2 的子博弈之中,从而不会陷入纳什均衡所预测的博弈困境当中。

在确定选手不同策略的后悔程度之后,根据 Halpern & Pass(2011)确立的一个理性原则:一个

理性的选手总是选择使得他后悔程度最小的一个策略。因此,在这一理性原则是各个选手公共知识(Common Knowledge)的前提下,两位乘客可以只将后悔程度最小的策略保留,这在表 4 中表现为只将后悔度为 3 的策略保留,从而进入下一个子博弈当中。

值得注意的是,经过剔除后,在子博弈中各个选手的策略的后悔度会产生变化,博弈双方需要重新计算各个策略的后悔度,再次保留后悔度最小的策略,一直重复这一过程直到不能剔除为止,最后达到的状态就是后悔度最小化方案所预测的结果。

事实证明,这是一个非常高效率的算法,并且最后保留的策略是和罚金的相关性与经验研究相符合。若设罚金为 p ,则在第一轮当中,两位乘客只会保留 $[100 - 2p, 100]$ 这一范围的索赔要求。本例中罚金 $P=2$,因此经过第一轮剔除后两位乘客只会保留 $[96, 100]$ 范围的索赔要求,重复这一过程到不能剔除为止,最后双方博弈的结果为 97 美元,这与 Becker 和 Capra 等人的实验数据是一致的。

(二)基于联盟与合作的概率推理模型

基于联盟与合作的概率推理方案从大量实验数据偏离纳什均衡的现象中总结出人不是绝对自私的,而是具有联盟和合作倾向这一重要特征。

为采用联盟与合作的思想研究像旅行者困境那样的非合作博弈中的合作式均衡,Capraro 提出一个合作原则作为预设^{[4]6}:

如果选手在博弈中结成一个联盟的话,那么他们会据此来预测博弈的走向并且按照那个最好的预测进行博弈。

在旅行者困境中,作为一个理性的选手,他们可以清醒地意识到,如果双方都绝对自私的话,那么他们会意识到自己会逐步陷入到纳什均衡所预测的困境当中。因此,考虑结成各种联盟以寻求潜在的合作是他们摆脱困境的出路之一。

在博弈论中,任何一部分选手都可以组成一个联盟,但一个选手只能加入一个联盟,并且所有联盟合起来应该刚好是全体选手。特殊地,每个选手本身也可以看作一个联盟。当然,所有选手也是一个联盟。值得注意的是,这里所说的联盟都是一种虚拟的联盟,联盟内部不允许交换任何信息或者达成任何转移支付的协议。在旅行者困境中,因为只有两个选手,恰好无外乎这两种联盟,我们不妨把它们分别称作自私联盟和合作联盟。如果形成合作联盟的期望效用高于形成自私联盟的期望效用,则显然选手会倾向于形成合作联盟。

根据合作原则,在旅行者困境中,选手们如果按照自私联盟,则显然他们可以预测并且按照这个预测进入纳什均衡(2,2)的博弈结果当中。(2,2)是纳什均衡,意味着没有选手有动力改变当前选择,因此自私联盟是一个稳定的联盟。按照自私联盟博弈的结果,最后双方的收益都是 2 美元。换句话说,2 美元是他们结成自私联盟的期望效用。因此,我们可以把 2 看做是自私联盟对于这两位乘客的价值。

现在假设两位乘客结成合作联盟,即他们共同提出 100 美元的索赔要求,则他们在结成合作联盟时的最大收益是 100 美元。但 100 美元并不能成为合作联盟对于他们的价值。因为这是一个不稳定的联盟,其中任何一个选手如果提出 99 美元的赔偿要求,则他可以得到 101 美元的收益。因此,我们可以说他脱离联盟的动力是 1 个单位。但是作为一个理性的选手,他除了计算自己脱离联盟的动力外,还须评估脱离联盟所产生的风险。其中最大的风险是如果对方依然坚持自私原则或者他预测到自己会脱离联盟的话,则对方可以提出 98 美元的索赔要求,在这种情况下自己只能得到 96 美元的索赔,从而比原来结成联盟的情况下损失 4 美元。因此,我们可以说他脱离联盟的风险是 4 个单位。

从乘客 A 的角度看,如果他选择坚守合作联盟,在他对乘客 B 脱离联盟的动力和风险进行上述评估之后,他就可以据此对乘客 B 是否也会坚守合作同盟这两个事件发生的概率进行评估。因

此,根据概率理论,乘客 A 在自己坚守联盟的条件下,他可以推断出乘客 B 脱离合作联盟这一事件的概率是 $1/(1+4)=1/5$,而乘客 B 坚守合作联盟这一事件的概率则是 $4/5$ 。

在没有人会脱离合作联盟的情况下,乘客 A 的预期收益显然为 100 美元。而乘客 B 脱离合作联盟这一事件可以有两种情况,即乘客 B 可以提出 99 和 98 美元的索赔要求,因为在这两种情况下,他都可以获得不少于坚守合作联盟中 100 美元的收益。其中最坏的情况是乘客 A 选择 100 美元索赔而乘客 B 选择 98 美元索赔这种情况,这时乘客 A 的收益仅为 96 美元。因此在乘客 B 脱离合作联盟这一情况下,乘客 A 的预期收益最坏且为 96 美元。

使用概率工具计算选手的期望效用并使其最大化是博弈论中用来预测选手理性行为的另一个非常有力的工具。根据乘客 A 对于乘客 B 坚守合作联盟和脱离合作联盟这两个事件发生的概率评估,乘客 A 不难计算出自己在坚守联盟条件下的期望效用,即 $100 \times (4/5) + 96 \times (1/5) \approx 99$ 。同理,乘客 B 计算出自己在坚守联盟条件下的期望效用也是 99。因此,我们可以把 99 看做是合作联盟对于这两位乘客的价值。显然,99 远远大于 3,这就是在旅行者困境中人们倾向于合作的原因所在。

目前为止,在基于联盟的背景下,我们只是得出了在旅行者困境中形成合作联盟的价值远远高于形成自私联盟的价值,从而人们会倾向于合作这一结论。但这并不意味着,两位乘客一定都会提出 100 美元的索赔要求,因为我们前面提到,合作联盟并不是一个稳定的联盟。但是,我们可以把合作联盟中计算出来的期望效用看做是两位乘客的公共信念(Common Belief)。根据形成合作联盟所计算的期望效用,两位乘客会合理地预期博弈只会发生在一个各方收益都不会低于合作联盟价值的子博弈当中,从而聚焦^①于如表 5 所示的一个子博弈之中。

表 5 合作联盟公共信念下诱导出的子博弈

		乘客 B			
		100	99	98	97
乘客 A	100	100,100	97, <u>101</u>	96,100	95,99
	99	<u>101</u> ,97	99,99	96,100	95,99
	98	100,96	<u>100</u> ,96	98,98	<u>95,99</u>
	97	99,95	99,95	<u>99</u> ,95	<u>97,97</u>

在选手根据合作联盟计算出期望效用诱导出来的子博弈中的纳什均衡就被称作是合作式均衡。在本文给出的旅行者困境的例子中,可以看出(97,97)正是合作式均衡。另外,随着罚金的提高,双方坚守合作联盟事件的概率变低,从而使坚守合作联盟条件下的期望效用变低,进而导致合作式均衡越来越接近纳什均衡。可见,无论是对结果的预测还是罚金的相关性而言,合作式均衡方案的预测结果和 Becker 与 Capra 等人的实验数据也是一致的。

三、两种解决方案的评价

与其他方案将选手在博弈困境中偏离纳什均衡归因于犯错^[11]和有限理性^[12]不同,这两种方案依然高举理性主义的旗帜,在坚持经典博弈论理性人假设的同时,还在此基础上增加了诸如后悔与合作等更多的理性概念。因此,这两种方案都是在坚持理性主义路线的基础上,在原来非合作博弈理论框架下,对经典博弈理论的丰富和完善。

在重复后悔最小化方案中,选手通过分别计算给定对方选择前提下当下选择的后悔度和策略本身的后悔度,从而将整个博弈中各个博弈结果收益之间的差异纳入理性决策的考虑之中。而在基于联盟与合作的概率推理方案中,选手则综合考虑对方脱离联盟的额外收益和风险,通过比较合作联盟与自私联盟的价值,对对方关于合作联盟的忠诚度进行全面评估,从而形成自己的合理预期,最后做出理性决策。因此,与传统博弈论对选手收益只侧重于质的研究不同,这两种方案都将选手收益在博弈中量的差别考虑进来,这在模型中是通过求差、取最大值和最小值等方式来实现

① 焦点(Focal Points)理论是 2005 年诺贝尔经济学奖获得者 Thomas C.Schelling 提出的在博弈论中用来解释选手在实际博弈中选择均衡的一个重要方法。

的,从而将各个博弈结果中收益的差距在模型中体现出来,更能全面反映博弈中选手理性决策的实际情况。可以说,这两种方案是在经典博弈论质的方法的基础上,质的方法与量的方法相结合的两个成功案例。

重复后悔最小化方案继承了经典博弈论中重复剔除被弱占优策略算法的一些缺点。在重复后悔最小化方案的剔除过程中,我们是一次性选择后悔度最小的策略。另外一种选择是只剔除后悔度最大的策略,因为在不同的子博弈中后悔度会发生改变,因此这会导致不同的博弈结果。也就是说,剔除顺序的不同,该方案会预测出不同的博弈结果。这表明这种方案所预测的结果并不具有唯一性,从而使其精确性受到影响。

更为严重的是,这揭示出重复后悔最小化方案并不具有“无关选项独立性”(Independence of Irrelevant Alternatives)这一重要性质,即一个博弈从后悔理性的角度看,如果加上一个完全应该被剔除的选项,反而会对博弈的结果产生影响,其根本原因在于在这种后悔度计算方式下,无关选项的加入会让原本的策略的后悔度发生变化。Halpern 等人虽然也意识到这一内在缺陷,但他们将其归因于人们在实际博弈中同样会受到无关选项的心理干扰^{[3]18}。这种求助于心理学的解释当然是不能让人满意的,因为这与该方案所秉承的理性主义风格相冲突。

重复后悔最小化方案和基于联盟与合作的概率推理方案都想将其解决方案应用到更多的博弈困境当中。但是,这两种方案对于囚徒困境都是失效的。因为在囚徒困境中,背叛对于双方都是强占优的策略,而强占优策略的后悔度为零。因此,双方选择背叛正是重复后悔最小化方案所预测的博弈结果,从而不能解释人们在囚徒困境中会有合作的现象。而根据基于联盟与合作的概率推理方案,所计算出来的合作联盟的价值并不具有参考意义,因为所得到的子博弈和原博弈是一样的。如果将子博弈中的纳什均衡当作合作式纳什均衡,则同样不能解释囚徒困境中的合作现象。

相比较而言,基于联盟与合作的概率推理方案普适性更强。例如,重复后悔最小化方案不能解释协同博弈(Coordination Game)中人们为何偏向于选择对双方结果都最好的纳什均衡,而基于联盟与合作的概率推理方案却能很轻松做到这一点。因为在协同博弈中,合作联盟的价值远高于自私联盟的价值,而博弈双方脱离合作联盟的动力为零。因此,基于联盟与合作的概率推理方案更能解释协同博弈中的合作现象。

基于联盟与合作的概率推理方案也存在不足之处。从前面的分析中可以看到,该方案没有将概率的思想贯彻到底。在求解均衡的前期,我们通过概率手段计算出合作联盟对双方的价值,从而使博弈进入一个双方收益都不低于合作联盟价值的子博弈中,但在子博弈中又使用纳什均衡来求解合作式均衡。更大的问题是,在基于联盟与合作的概率推理方案中,Capraro 对所计算出来的合作联盟的价值作多种解读,他有时甚至直接将合作联盟的价值看作该方法所预测的博弈结果^{[13]8}。同时也为了处理囚徒困境一类的困境,Capraro 后来提出在子博弈中计算混合策略纳什均衡^①(Mixed Strategy Nash Equilibrium)^{[3]42}。求解子博弈混合策略纳什均衡的方法虽然在囚徒困境中有一定作用,因为在囚徒困境中,每位选手只有合作与背叛两个选项,但在旅行者困境中的子博弈中,往往每位选手都有多种索赔策略,所以在旅行者困境中是难以计算的。这种不一致的做法表明基于联盟与合作的概率推理方案不具有统一性,从而对其普适性产生影响。这也是过分依赖概率这种量的方法的一个内在缺陷。

总的来看,概率方法是更具普适性的一种方法。因为纳什均衡的存在性定理表明,任何一个有穷策略式博弈总是存在混合策略的纳什均衡^[1]。Halpern 等人后来也不得不结合概率的方法,将混合策略考虑进来^{[3]20},从而得以将其方法应用到协同博弈当中。

① 在没有纯策略纳什均衡的博弈中,可以将选手的选择看做是纯策略上的概率分布,因而非决定性的。混合策略纳什均衡的提出就是为这些博弈的稳定状态构建模型。

四、结 语

针对人们在博弈困境中并不是如经典纳什均衡所预测的那样,而是存在大量偏离纳什均衡的选择行为的现象。我们有必要对博弈论的哲学基础进行反思,一方面,这涉及博弈论这门学科的性质和定位:博弈论是一门规范性(Normative)学科,还是一门描述性(Descriptive)学科?将这些博弈困境定义为困境的一个预设是博弈论至少应该具有描述性学科的特点,即它应该能够对博弈中的选手做出的选择进行解释,从而对人们应该如何理性选择提供决策参考。另一方面,博弈困境的产生使得我们必须重新反思非合作博弈论中的理性人假设,即每个选手都是只追求自身利益最大化的,并且能够进行任意有限步的推理。但选手不仅是一个理性人还是一个社会人,人们在实际博弈中的行为选择还受选手类型、收益敏感度、期望、相互间的信任和社会规范等理性因素甚至犯错等非理性因素的影响。但是,一个理论往往出于简单或便于研究的需要,又不可能考虑所有这些因素的影响。这也是许多社会科学研究中面临的理论困境。

进入 21 世纪以来,世界形势和利益格局变得日益纷繁复杂,特别是像中国这样的发展中国家又处于社会的剧烈转型之中,各个国家、社会群体和利益阶层的冲突和斗争日益激烈,人们越来越意识到在国家战争、地区冲突和利益纷争中,通过相互合作达到共赢往往是解决问题的最佳选择。对博弈中的合作进行研究一直是博弈论研究中的难题。目前博弈论对合作博弈论的研究还非常有限,围绕博弈困境的解决主要还是在非合作博弈的理论框架下进行的。因此,如何在非合作博弈的理论框架下对选手博弈困境中实际存在的合作行为提供理论解释和更为准确的预测模型是博弈论中的一个巨大挑战。在这个意义上说,重复后悔最小化方案和基于联盟与合作的概率推理方案不仅是博弈论理论研究中取得的重要进展,还为人们在社会行为中普遍存在的合作和共赢现象提供了理论支撑和理论指导。

参考文献:

- [1] Nash J F. Equilibrium points in n-person games[J]. Proceedings of the National Academy of Science of the United States, 1950,36(1): 48-49.
- [2] Basu K. The Traveler's Dilemma: Paradoxes of Rationality in Game Theory[J]. American Economic Review, 1994,84(2): 391-395.
- [3] Halpern J Y, Pass R. Iterated Regret Minimization: a new solution concept[J]. Games and Economic Behavior, 2012,74(1):184-207.
- [4] Capraro V. A solution concept for games with altruism and cooperation[EB/OL]. <http://arxiv.org/pdf/1302.3988v2.pdf>.2013.
- [5] Capra M, Goeree J K, Gomez R, Holt C A. Anomalous Behavior in a Travelers Dilemma? [J].American Economic Review, 1999(89)3: 678-690.
- [6] Becker T, Carter M, Naeve J. Experts Playing the Travelers Dilemma[R].Discussion Paper 252, Institute for Economics, Hohenheim University,2005.
- [7] Basu K, Becchetti L, Stanca L. Experiments with the Travelers Dilemma: welfare, strategic choice and implicit collusion[J]. Social Choice and Welfare ,2011,37(4): 575-595.
- [8] 潘天群. 博弈论中理性人假设的困境[J]. 经济学家,2003(4):99-104.
- [9] 袁艺,茅宁. 从经济理性到有限理性:经济学研究理性假设的演变[J]. 经济学家,2007(2):21-26.
- [11] McKelvey R, Palfrey T. Quantal response equilibria for normal form games[J]. Games and Economic Behavior. 1995(10) 1, 6-38.
- [12] Stahl D, Wilson P. Experimental evidence on player's models of other players[J]. Economic Behavior and Organization, 1994(25) 3: 309-327.
- [13] Capraro V. A Model of Human Cooperation in Social Dilemmas[R]. PLoS ONE 2013,8(8):e72427. doi: 10.1371/journal.pone.0072427.