

关于囚徒困境的几个问题

李伯聪 李军

摘要：囚徒困境是博弈论中的一个重要范例，对这个问题的研究涉及经济学、政治学、社会学、哲学、伦理学等广泛的领域。本文简要地介绍和评述了中外学者在这个问题研究中取得的成果及进行解释时提出的多种见解。囚徒困境提示的个人理性和集体理性、个人主义与道德主义的关系深化了对这些问题的认识。本文最后介绍了用计算机比赛方法研究重复的囚徒困境和合作问题的成果，分析了“一报还一报”程序取得成功和合作得以出现的原因和条件。

关键词：囚徒困境 理性 道德 合作

* * *

一、什么是囚徒困境

1994年10月12日，瑞典皇家科学院宣布把该年度的诺贝尔经济学奖授予约翰·纳什（J.Nash）、约翰·豪尔绍尼（J.Harsanyi）和莱因哈德·泽尔腾（Reinhard Selten）以表彰他们把博弈论（game theory，又译为对策论或游戏论）应用于现代经济分析所作的卓越贡献。以此为契机，我国在1995年出现了一个小小的介绍博弈论的“热潮”。许多读者大概已经注意到：在介绍博弈论的文章中几乎篇篇都会提到所谓的囚徒的困境。

囚徒困境的原文为the Prisoner's Dilemma，又译为囚犯的两难困难，囚犯难题等。这个问题是大约在1950年首先由社会心理学家梅里尔·M·弗勒德（Merril M.Flood）和经济学家梅尔文·德雷希尔（Melvin Dresher）提出来的，后来由艾伯特·W·塔克（Albert W.Tucker）明确地叙述了这种“困境”[1, p.3]。纳什有两篇关于非合作博弈的重要文章分别发表于1950年和1951年。有人认为：塔克的这项工作同纳什的著作一起“基本上奠定了现代非合作博弈论的基石。（2）囚徒困境的重要地位和重要意义由此就不难想象了。

所谓囚徒困境可简要地叙述如下：

两个涉嫌共谋犯罪的嫌疑犯被捕后被警察关在相互隔离的牢房中。他们面临的选择是：或者坦白或者保持沉默（即不坦白）。他们被告知：①如果某个嫌疑犯坦白而其同伙不坦白，则坦白者可获自由而拒不坦白者要被判10年监禁；②如果二人都坦白，则二人都被判5年监禁；③如果二人都都不坦白，则二人皆被判1年监禁。上述情况我们亦可用一支付矩阵表示如下：

		嫌疑犯乙	
		坦 白	沉 默
嫌疑犯甲	坦 白	-5, -5	0, -10
	沉 默	-10, 0	-1, -1

在这种情况下，两个嫌疑犯将如何决策和选择呢？

在囚徒困境的分析和推论中有以下的前提性设定：①每个局中人（在囚徒困境中就是两个囚徒）都知道“博弈的规则”和“博弈结果”的支付矩阵；②每个局中人都是理性的，而这里的所谓“理性”是指只要给出两

种备择策略 (alternatives)，他将总是选择其中对他更有利的那种策略；③局中人是相互隔离不能“串通”，不能“结盟”的。

在这种条件下，两个囚徒怎样分析问题和运用“理性”并得到什么结果呢？

从甲的“立场”来看，共有两种可能情况。第一种可能情况是乙采取坦白的策略，这时，如果他也坦白则要入狱5年，如果他不坦白，则更糟：要入狱10年；两相比较，结论是他应以坦白“对付”乙的坦白策略。第二种可能情况是乙采取沉默的策略，这时他若也沉默，要入狱1年，如果他坦白，则可获得自由；两相比较，结论是他应以坦白对付乙的“沉默”策略。把以上两种可能情况的结论归纳在一起，“总结论”是：无论乙采用坦白策略还是沉默策略，甲都以采取坦白策略为更“佳”的对策；所以，甲要采取坦白策略。

很显然，以上推理对于乙也是适用的。

这样，两个囚徒在“精心推理”之后都采取了坦白策略，结局是两人都被判5年徒刑。

不难看出，这个推理中似乎带有某种悖论性的因素；因为，如果甲、乙两个囚徒都采取沉默策略，则二人都只被判1年徒刑，显然它是一个比二人都入狱5年更好的结果。然而，甲、乙两个富于“理性计算”精神的囚徒在“逻辑上正确”的计算一番之后都采取了坦白策略，获得了都被判5年徒刑的结果。

这里说“逻辑上正确”，是因为甲和乙都是依据以下的推理形式进行推理的：

(1) P或Q

(2) 如果P，则R

(3) 如果Q，则R

所以：R

显然，这是一个正确的逻辑推理形式。

需要指出，上述案例中的那些具体数字究竟是多少，并不具有重要意义，甚至四种可能的结果组合能否“数字化”也不具有关键性意义；问题的关键是四种可能的结果组合对于当事人来说，其“好”“坏”程度构成一个有序的“序列”。即当事人（上述案中的“囚徒”）可在互相比较后将四个结果排出他所认为的“最好”、“次好”、“第三好”和“最糟”的次序来。

为了把囚徒困境中的“两难 (dilemma)”性甚至“悖论 (paradox)”性因素更突出、更尖锐地显示出来，有人虚构了一个更“典型”的情景 [1, p.6]（按：以下的中文转述，为了便于理解，略有改动）：假定在前述“案例”中，法庭向两个囚犯“宣布”了新的“判决规则”：如果两人都坦白认罪，将被判为快刀砍头的死刑；如果两人都沉默，则两人将被宣布无罪释放并获赔偿1000元；如果一人坦白，一人沉默，则坦白者可获自由并加10000元奖赏，而沉默者将被凌迟处死。在这个“案例”中，“无情”而“严格”的逻辑将使两个“理性”的囚徒共同选择坦白策略，从而走上高悬一把快刀的断头台，而不选择共同沉默从而获得自由并且还能在皮包中装上1000元货币。难道这竟然会是“理性人”的“理性”选择吗？

二、囚徒困境的若干分析和实例

在庫恩的“范式”理论中，范式的一个重要含义就是范例。可是，许多人都只注意了作为“理论框架”和“科学方法”的范式，而忽略了作为范例的范式。其实，在科学理论体系和科学活动中，范例所起的重要作用也是不可低估的。不难看出，囚徒困境正是博奕论中的一个重要范例。

有人可能会认为：囚徒困境只是“象牙之塔”中的理论家虚构的一种“思想游戏”，或是一种难得一见的特例。这实在是一种错觉。

应该承认：囚徒困境最初的确是“象牙之塔”中的“抽象理论思维”的“产物”。如同其他的许多“象牙之塔”中的产物一样，人们愈来愈深刻地感觉到它同“现实世界”实际上是存在着密切联系的。

已经发现，在经济学、政治学、社会学及现实生活中类似囚徒困境的事例是大量存在的。

对于囚徒困境的原因、条件、性质、意义等问题，博奕论专家、经济学家、政治学家、社会学家、哲学家、伦理学家、社会心理学家已从许多方面进行了大量的研究。

有人注意到在囚徒困境的“案例”中两个局中人不能互相通信是一个限制性条件，因此他们设想这就是造成

囚徒困境的“原因”。例如，有人说：“囚徒的苦恼在于他们不能商量沟通。”（3）于是，有人便提出了如下的解决问题的途径：两个局中人可以“进行通信，然后合作”。

反对者认为这是于事无补、不解决问题的。“显然，一个囚徒要保持沉默的意图不可能引起另一个囚徒有同样的意图——如果另一方忽视这个意图的话。不准通信的假定对于囚徒困境来说不是关键性的。可以增加一个通信条件，但这仍然不能造成什么不同，只要假定每个人对可能结果的偏爱顺序与以前相同并且每个人可以自由地选择与另一个人的意图相反的策略。每个囚徒都可以对另一个囚徒说：‘如果你将保持沉默，那么我也将保持沉默’；可是，只要每个囚徒可自由地说一套而做另一套，那么困境就没有什么变化”。“有人也许会假定可能存在着强迫性协议条款，比方说对违反协议进行特定的处罚，这就会改变囚徒的偏好顺序。这样的假定会改变原有的决策问题的结构。而我们一直假定每一方只对造成他有可能获得的最有利的结果有兴趣，并且他的偏好顺序是严格给定的。在这个假定之下似乎没有任何合作‘协议’能造成什么不同。”

（[1]，pp.9—10）

从上述分析中可以看出，囚徒困境的“要害”不是通信问题。有人可能会认为囚徒困境的“要害”在于：每个局中人只关心自己的利益而对他人的利益漠不关心，为了自利的目的他甚至不惜违背“协议”。这就是囚徒困境的“要害”。于是，有人提出了相应的解决问题的答案：不要做一个利己主义者。

“这个回答，至少就其现在的叙述方式而言，由于立足于误解而陷于失败。在这个上下文中‘利己主义者’的最自然的解释是指那些主要——如果不是唯一地——只关心个人利益的人。因此，只关心他自己将在监狱中关多长时间的囚徒是一个利己主义者。但是，很清楚，囚徒困境无须受限于这种意义上的利己主义者。可以构造出一些例子其中的每个人在其所处的选择环境中都与利他主义的。读者可以回想起，对于每个局中人对结果之偏好的全部要求只是存在着一定的顺序。在此顺序背后的动因可以是利己的，或是利他的，或者是兼有二者。”比方说，我们可以很简便地构造出另一个例子：情景和条件与本文开头所举的例子皆相同，只是假定两个囚徒在进行决策选择时，决定偏好顺序的标准不是完全与唯一的利己标准，而是完全是唯一的利他标准，这时的结果会有什么变化吗？没有。所谓的囚徒困境在这个“利他主义者”的“情景”中依然是一个“困境”。所以，囚徒困境的前提条件是偏好的发散（divergence）而不是自私自利或对他人漠不关心

[1, p.10]

值得特别注意的是，我国学者在介绍、解释和评论囚徒困境时许多人都特别强调囚徒困境凸现了个人理性和集体理性的矛盾。例如，张维迎说：“囚徒困境反映了一个很深刻的问题，这就是个人理性与集体理性的矛盾。”（[2]，p.123）郑也夫说：“理性的个人，加在一起成了非理性的社会。”（3）茅于軾说：“囚犯难题具有极深刻的含义，它解释了何以短视地以利益为目标将导致对大家都不利的结局。”（4）李子江使用了专业性的学术用语，指出在囚徒困境中，“双方都会选择对自己收益最大，风险最小的策略，从而导致了对策的稳定均衡点的出现。”这就是所谓纳什均衡点（其定义为：假使其他参加者不变换策略，任何单个参加者不能以单方面变换策略来提高他的支付）。但这一纳什均衡点却不是全局的最优解（帕累托最优）。”由此可见，双方都采取最优策略时，整个对策系统却处于较差的品质。”（5）这些专门术语的含义换成“易懂”的术语也是认为囚徒困境揭示“个人理性与集体理性之间存在着一种内在的矛盾”。

我国学者已举出了许多为我国读者所熟悉的属于囚徒困境的事例。

在人民公社制度下，“如大家都好好干，粮食打得多些，大家不是都可以过得好些吗？问题在于：我好好干，你不好好干，我吃亏，你好好干，我不好好干，我占便宜，于是，纳什均衡就是大家都不好好干，都饿肚子。”（[2]，p.124）

在囚徒困境中，两个囚徒都选择了坦白认罪策略。“这个结果与他们是否真的有罪（按：本段为叙述方便，转引时对原文略有改动，但对原文之原义绝无修改）无关，他们即使无罪，也会做出承认有罪的选择。‘文革’中坦白从宽、抗拒从严的政策之所以能使一大批人承认自己的‘罪行’，其奥妙即在于此。”（4）此外，大家常见的挤公共汽车，大家痛恨的行贿成风，也都是囚徒困境的“表现”。

实际的现实生活是复杂的。人们发现：囚徒困境的“社会效果”也并不总是负面效果，在特定条件下，它也

可能造成某种对社会有益的效果。

在经济学中最常提到的事例是卡特尔欺骗和寡头垄断的定价问题〔6〕〔7〕。设想有两个寡头垄断厂家，若遵守协议的垄断价格，则两个寡头皆可获得高额利润。若一方违反协议降价促销，则可获得更高额的利润，对方因产品滞销，只能获得很少的利润。两个寡头都按这个“思路”考虑定价策略，结果是大家都不遵守协议，都降低销价，结果都不能获得高额利润，而只能得较少的利润。容易看出，在这种情况和条件下的卖方垄断寡头陷入囚徒困境是有利于消费者、有益于社会的。所以许多国家都在法律上不允许卖方垄断寡头订立垄断价格协议。

有的中国读者也许会认为：西方经济学著作中经常讲述的这个例子同中国的经济现实相距太远。很凑巧，最近的新闻媒介报道了一个中国的真实事例〔8〕。1996年北京市百货大楼、西单商场等八大商场与小天鹅、海棠等九大洗衣机厂签订了一个名为《关于统一首都洗衣机市场零售价的联合协议书》而实为制定垄断价格的协议。不久，有的商场就开始悄悄地违反这个协议了。可以看出：这表明囚徒困境的“现象”已经开始显露端倪了。2月8日，北京市工商局明确表示：八大商场与九大洗衣机厂的行为属不正当竞争和损害消费者权益的违法行为，应立即制止。

在这里需要说明的是，上举各例严格地说起来并不是“两人”的囚徒困境问题，而是属于多人或大规模囚徒困境（an n-person or large scale Prisoner's Dilemma）类型的问题。目前许多中国经济学家常常谈到的“免费搭车”（free riders，又译为“搭便车”或“白搭车”）现象就是在这种“大规模囚徒困境”中产生的。

（关于大规模囚徒困境的简要论述和分析，可见文献〔1〕，pp.13—14）

三、理性、道德和囚徒困境

理性和道德都是哲学的基本范畴，是许多哲学家竞相关注和研究的对象。可是，在西方哲学史上，虽然对理性的研究和对道德的研究也有某些联系，但总的来看，仍需认为它们是“两股道上跑的车”。以康德为例，他是近代欧洲哲学史上罕见的同时在理性研究和道德研究中所获“超一流”成就的哲学大师，他的《纯粹理性批判》（以研究理论理性为主题）和《实践理性批判》（以研究道德为主题）都是“头等”的哲学名著。可是，他的关于理论理性的理论与关于道德哲学的理论仍然是割裂的，二者之间有一条巨大的鸿沟。

从哲学的观点来看，囚徒困境的重要启示之一就是它要求哲学家把对理性的研究同对道德的研究“内在的”、“不可分割”地联系起来。

传统上，哲学家主要是致力于在认识论领域中对理性问题进行研究的；现在，可以看出：为了深化对理性的哲学研究，哲学家已经必须努力借鉴和积极参与经济学和伦理学中对理性问题的相关研究了。

前文言及囚徒困境提出的一个重要问题是个人理性与集体理性的相互关系问题。这显然不但是一个重要的经济学问题而且是一个重要的哲学问题。

随着经济学的发展，经济学家对这个问题的认识是有所变化和不断深化的。张维迎指出，现代“经济学越来越重视人与人之间关系的研究，特别是经济学开始注意到理性人的个人理性行为可能导致的集体非理性。这一点和传统经济学形成明显对照，在传统经济学里，价格可以使个人理性和集体理性达到一致。现代经济学开始注意到个人理性和集体理性的矛盾与冲突，但是解决这个问题的办法并不是象传统经济学主张的那样通过政府干预来避免市场失败所导致的无效状态，而是认为，如果一种制度完排不能满足个人理性的话，就不能贯彻下去。所以解决个人理性与集体理性之间冲突的办法不是否认个人理性，而是设计一种机制，在满足个人理性的前提下达到集体理性。”（〔2〕，p.121）“从囚犯困境中，我们可以引出一个很重要的结论：一种制度（体制）安排，要发生效力，必须是一种纳什均衡。否则，这种制度安排便不能成立。”（〔2〕，p.124）

个人理性和集体理性的关系是一个大问题，要在这方面的研究中取得新的进展，还有待于哲学家、经济学家和伦理学家的协力研究。

本文第二节中曾谈到造成囚徒困境的“要害”不是利己主义问题。很显然，其含义决不是说囚徒困境同伦理学无关，也不是说利己主义同道德主义的分歧对囚徒困境问题没有影响。只是这种影响很不简单罢了。

从伦理学的角度来看，利己主义者（在其所有的行动中都试图使他的偏好所得最优化）同道德主义者（在其所有的行动中都试图使那些受其行动影响的那些人的偏好所得最优化）的区别是根本性的，是不容混淆的。约翰·沃特金斯承认这两种人都是聪明和有理性的。他曾经通过研究得出三个结论：①两个理性的利己主义者在囚徒困境的情景中是处于真正的两难困境（dilemma）中；②两个道德主义者在这样的情景中大概也完全处于两难困境中；③道德主义渗入其中有时使情景变得更好，有时则变得更糟；道德主义在这里对利己主义没有明显的优势。后来，通过更详细而深入的研究，他认为需要放弃论点①，坚持论点②，弱化论点③。值得特别注意的是他的研究还得出了如下的结论：“假如在社会中道德主义者占一个很大的多数，比如说95%或更多，那么，对一个人转向道德主义的激励就会相当强。”“将会出现一种带头羊式的效果（bandwagon effect）：如果道德主义已经在这个社会传播得相当广的话，则可望它的传播会更广。”（[1]，p.74）

由于囚徒困境同伦理学有密切的关系，所以当代美国最著名的伦理学家罗尔斯和诺齐克在他们的伦理学名著（9）（10）中都涉及了囚徒困境，这就不足为奇，而是很自然的了。

四、合作、进化和囚徒困境

在囚徒困境的研究和分析中，合作问题居于某种枢纽性的地位。这个问题同上述的理性问题、道德问题有着十分密切的联系，但我们也可把它当成一个相对独立的问题进行研究。

本文以上谈到的囚徒困境都是指“一次性博弈”条件下的困境。在这种条件下，局中人对合作都没有兴趣。可是，如果博弈不是一次性的，而且可以多次重复的博弈，情况就要发生变化了。例如，在所举两个囚徒的例子中，如果不同的判决结果仅仅是立即释放条件下罚金数目的多少，那么，这两个囚徒就完全有可能在以后的日子里多次重复地遇到同样的困境了。这就产生了重复的囚徒困境问题。

在这方面的研究中试图回答的一个关键问题是：在一个没有核心权威的利己主义者的世界上什么条件能使合作出现？（[1]，p.320）这个问题在政治哲学、国际政治、经济学、进化论等许多领域中都是非常重要的。在彻头彻尾自私自利的生物群落中能发展出一种合作战略吗？如果答案是肯定的，它对进化论的重要意义是不言而喻的。

为了研究这个问题，罗伯特·阿克塞尔洛德（Robert Axelrod）组织了两次有关的计算机比赛，得到令人惊讶的结果。（[11]、[1]，pp.320—338）

计算机比赛的方式是征集一些编成计算机程序的战略，每一特定的战略按一定的规则以自己的合作策略或欺骗策略对付博弈对手的合作策略或欺骗策略。例如，全骗战略（ALLD）是每一回合的博弈都进行欺骗；随机战略是以随机方法决定在一次博弈中是进行合作还是进行欺骗，等等。在比赛时让这些战略进行“随机赛”，然后以比较每个战略得分的多少来衡量其优劣。

第一次比赛时共有15个程序参加比赛。这15个程序中有简单的，也有复杂的。令人惊讶的是，所有程序中最短的一个程序（只有四行）取得了胜利。这个程序是一位心理学家兼哲学家安纳托尔·拉波波特（Anatol Rapoport）设计的，它被称为一报还一报（TIT FOR TAT）。这个程序的战略原则是：第一回合进行合作，以后不管对方怎么走，均采取对方上一个回合中的策略。

在总结了第一次比赛的经验教训后，阿克塞尔洛德组织了更大规模的第二次比赛，他不仅邀请了第一次的所有参加者，而且广为宣传争取有更多的人参加。他还给每位参赛者寄去了一份关于第一次比赛的分析。他特别着重分析了善良（不首先进行欺骗，但允许以欺骗报复欺骗）和宽恕（报复之后不再怀恨在心）这两个战略性概念。“一报还一报”就兼具这两种特性，第二次比赛有6个国家、11个专业的62位参赛者。这次提出的程序总的来看要比第一次复杂得多。最复杂的一个程序是用FORTRAN语言写的，共152行。拉波波特再次提出“一报还一报”参加比赛，这个程序在第二次比赛的所有程序中仍然是最短的。

结果真是让人目瞪口呆——“一报还一报”再次取得了胜利。

第二次比赛的参赛者中，有的人不愿相信善良战略，他们宁愿相信可以用比较巧妙的花招和诡计战胜对手。而第二次比赛的结果却再次证明了善良战略的成功；在前十五名中只有第八名不是善良程序，而最后十五名中只有一名是善良程序。第二次比赛结果提示出的第三个重要战略概念是可激怒性，即应该很快地被欺骗者

激怒并进行报复。

第二次比赛后，阿克塞尔洛德又进行了一些重赛，其中的模拟生态适应性的“生态比赛”尤为重要。在这种比赛中，失败的程序最终将遭到生态灭绝的厄运。饶有趣味的是前述的第八名不善良程序在前200代生态比赛中成绩颇佳，随后由于比较弱的程序逐步灭绝，整个生态环境中优良程序比例增大，这个不善良程序在大约1000代比赛时同那些被它欺骗过的程序一样也遭到了灭绝的命运。而“一报还一报”在生态比赛中愈来愈遥遥领先。

阿克塞尔洛德认为关于合作的演化有三个基本问题：①初始存活性问题：在一个普遍进行欺骗的世界上，合作怎样才能开始？答案是需有一小批合作的生物体侵入，即使只有一小撮也足以使合作有一个立足之地；②强壮性问题：具有善良性、可激怒性、宽恕性、可识别性特点的战略才是具有强壮性的战略；③稳定性问题：欺骗的世界可被一伙合作者侵入，而合作者的世界却能不让欺骗者侵入。一旦合作建立起来，它将永远持续下去。

在对囚徒困境的研究中还有一些其它的重要结果，例如：克雷普斯（Kreps）等四位学者在囚徒困境中增加了一些假定，证明了短期合作的可能性。（12）

在本文的最后，我还想为霍布斯作一点辩护。在我国许多学者的笔下，霍布斯成了一个只知武断地认定“人对人象狼一样”的凶神恶煞。这其实是对霍布斯的人云亦云的误解。霍布斯说：“每一个人只要有获得和平的希望，就应该力求和平；在不能得到平时，他就可以寻求并且利用战争的一切帮助和利益。”前者就是“寻求和平，信守和平”，后者就是“利用一切可能的办法来保卫我们自己。”（13）不难看出，霍布斯的这个观点同“一报还一报”在精神实质上是颇为相通的。

总而言之，囚徒困境提出了许多饶有趣味而重要的问题，但愿今后能有更多的中国学者加入到研究这些问题的行列中来。

参考文献

〔1〕 R.Campbel and L.Sowden (ed.),Paradoxes of Rationality and Cooperation,Prisoner's Dilemma and Newcomb's Problem, Vancouver, The University of British Columbia Press,1995.

〔2〕 张维迎：从九四年诺贝尔经济学奖看主流经济学的新发展（1—7），原载《经济学信息报》，转载于复印报刊资料《理论经济学》，1985年第6期，第119—131页。

〔3〕 郑也夫：《走出囚徒困境》，光明日报出版社，1995年，第201页。

〔4〕 茅于軾：《生活中的经济学》，上海人民出版社，智慧出版有限公司，1993年，第254—256页。

〔5〕 李子江：《关于纳什均衡点的研究》，《学术研究》，1995年第2期，第43—49页。

〔6〕 劳埃德·雷诺兹：《微观经济学》，马宾译，商务印书馆，1986年，第218—219页。

〔7〕 范里安：《微观经济学：现代观点》费方域等译，上海三联书店，1992，第599页。

〔8〕 月阴：“失败的垄断”，《中国经济时报》，1996年2月23日，第1版。

〔9〕 约翰·罗尔斯：《正义论》，何怀宏等译，中国社会科学出版社，第260、564页。

〔10〕 罗伯特·诺齐克：《无政府、国家与乌托邦》，何怀宏等译，中国社会科学出版社，1991年，第124页。

〔11〕 D.R.Hofstadter：“对囚徒疑难所做的计算机比赛提示合作是如何演化出来的”，《科学》，1983年第9期，第91—98页。

〔12〕 北京大学外国哲学史教研室编译：《西方哲学原著选读》上卷，商务印书馆，1982年，第397—398页。

〔13〕 张欣：“对策论”，载汤敏、茅于軾主编：《现代经济学前沿专题》第一集，商务印书馆，1989，第235页。

〔作者简介〕：李伯聪，1941年生。中国科学院研究生院教授。

李军，1952年生，河南省图书馆馆员。

(本文责任编辑 张祖贵) *