合作演化与网络博弈实验研究进展

荣智海¹, 许雄锐¹, 吴枝喜^{2*}

1. 电子科技大学计算机科学与工程学院, 成都 611731;

2. 兰州大学物理科学与技术学院, 兰州 730000

*联系人, E-mail: eric0724@gmail.com

收稿日期: 2019-04-15; 接受日期: 2019-05-17; 网络出版日期: 2019-09-02

国家自然科学基金(编号: 11575072, 61473060)资助项目

摘要 围绕合作的演化与涌现, 本文基于两人囚徒困境和多人公共品博弈模型, 重点介绍了近年来关于直接互惠和网络互惠的理论与实验研究进展. 重复博弈为直接互惠的产生提供了必要条件. 零行列式策略作为直接互惠中的一个新理论, 近期被从理论与实验方面深入研究. 零行列式策略是一种单步记忆策略, 它可以单方面决定博弈双方的收益关系, 这为探索人类决策行为规律提供了新工具. 网络博弈实验研究发现, 动态网络可以提升并维持人类的合作行为, 而静态网络是否促进合作涌现尚存在争议. 网络博弈实验有助于理解社会网络重要结构的形成机理与复杂系统中的行为涌现.

关键词 网络互惠, 直接互惠, 博弈实验, 复杂网络

PACS: 02.50.Le, 89.75.-k, 89.65.-s

1 引言

从吸血蝙蝠的反哺^[1]到棘鱼之间的互助^[2], 从基因的合作^[3]到人类社会网络的形成^[4], 都与个体之间的互助密切相关. 这些复杂的、大规模的合作现象是人类社会和自然界中组织形成的基石^[5], 也是人类得以繁衍昌盛的动因之一^[6,7].

博弈论(Game Theory)为研究多个体之间的合作行为提供了统一的理论框架^[8]. Von Neumann与Morgenstern^[9]合著《博弈论与经济行为》, 标志着博弈论的诞生. Nash^[10]提出并证明了纳什均衡(Nash Equilibrium)状态的存在性. 在纳什均衡中, 博弈各方不能通

过单方面主动改变自己的行为而增加收益. 群体中均衡状态不是一蹴而就的, Smith和Price^[11]将自然选择和变异理论引入博弈论, 提出了演化博弈理论(Evolutionary Game Theory), 认为个体在重复博弈中, 需要根据局部有限信息通过不断模仿学习和试错, 探索稳定均衡状态. 该状态称为演化稳定策略(Evolutionary Stable Strategy, ESS), 它是纳什均衡的子集, 在该状态下任意少量个体的策略改变都无法侵入整个种群. 进一步, Nowak等人^[12]基于Moran随机过程提出了有限种群中的演化稳定策略. 演化博弈作为一种重要的集群动力学, 近年来引起了统计物理领域学者的广泛关注^[13-20]. 大量理论研究揭示了以囚徒困境为代表的博

引用格式: 荣智海, 许雄锐, 吴枝喜. 合作演化与网络博弈实验研究进展. 中国科学: 物理学 力学 天文学, 2020, 50: 010508

Rong Z H, Xu X R, Wu Z X. Experiment research on the evolution of cooperation and network game theory (in Chinese). Sci Sin-Phys Mech Astron, 2020, 50: 010508, doi: [10.1360/SSPMA-2019-0129](https://doi.org/10.1360/SSPMA-2019-0129)

弈模型中, 种群博弈的均衡状态为背叛占主导, 但这与复杂生物与社会系统中普遍存在的合作现象相悖. 以直接互惠和网络互惠为代表的机制, 为理解远离背叛均衡状态的合作涌现提供了有效的理论工具.

作为一门研究行为与决策的科学, 博弈论除了理论研究外, 还有一个重要的研究范式——博弈实验. 一方面, 通过邀请被试参与博弈实验, 可以验证已有博弈理论的正确性. 另一方面, 通过设计合理的博弈实验, 包括实验室实验(Lab Experiment)和田野实验(Field Experiment), 我们还可以在真实的场景中探索那些在理论研究或日常生活中难以观测到的、驱动人类行为的认知机制, 从而帮助我们更深刻地认识自我, 探索合作行为的涌现机制.

本文主要围绕经典的两人囚徒困境和多人公共品博弈模型, 回顾总结了近年来有关直接互惠和网络互惠这两类重要合作机制的博弈实验研究成果, 希望可以为相关领域的学者带来启发和帮助.

2 重复博弈与直接互惠

2.1 囚徒困境与公共品博弈

囚徒困境(Prisoner's Dilemma, PD)是一个经典的两人两策略博弈, 模型中包含两个参与者, 每个参与者有两个选择: 合作(Cooperation, C)或背叛(Defection, D). 合作者会付出一定代价帮助对手, 背叛者则会不劳而获. 如果他们互相合作, 各自会得到一个奖励(Reward, R); 如果互相背叛, 则分别得到一个惩罚(Punishment, P); 如果一方合作, 另一方背叛, 背叛者会得到较高的诱惑收益(Temptation, T), 合作者会得到较低的受骗的收益(Sucker's Payoff, S). 因此, 我们可以得到如表1所示的收益矩阵.

若四个参数满足 $T > R > P > S$, 期望利益最大化的理性个体会思考: 假如对手是合作者, 由于 $T > R$, 出于贪婪他会选择背叛; 假如对手是背叛者, 由于 $P > S$, 出于恐惧他也会选择背叛. 所以, 无论对手的选择如何, 理性的个体都会选择背叛. 于是双方会陷入互相背叛的“困境”, 仅获得 $2P$ 的总体收益——尽管互相合作可以得到 $2R$ 的总体收益. 理性个体为了最大化自身利益, 陷入了总体收益降低的困境. 在博弈实验中, 常常考虑如下情况: 合作者会花费代价 c 让对手获得收益 b , 背叛者不会付出, 因此可知收益为: $R=b-c$, $S=-c$, $T=b$,

表 1 囚徒困境博弈的收益矩阵

Table 1 The payoff matrix of Prisoner's Dilemma game

收益矩阵	参与者2选C	参与者2选D
参与者1选C	R, R	S, T
参与者1选D	T, S	P, P

$P=0$. 这个特殊的囚徒困境博弈称为捐助博弈(Donation Game).

公共品博弈(Public Goods Game, PGG)是多人版本的囚徒困境博弈模型. 考虑有 n 个参与者, 开始时每个参与者会得到 y 元虚拟货币. 博弈中参与者 i 可以选择将 g_i ($0 \leq g_i \leq y$)元虚拟货币贡献到一个公共池里. 所有参与者做完选择后, 池子里的虚拟货币会被乘以一个强化因子 r ($1 < r < n$)并平均分给所有参与者. 因此, 每个参与者的收益为

$$E_i = y - g_i + \frac{r}{n} \sum_{j=1}^n g_j. \quad (1)$$

由式(1)可以发现, 随着个体贡献的虚拟货币增加, 自己的收益反而会越低. 因此, 对理性且利己的个体来说, 不投入的“搭便车”行为是最优的策略. 然而, 从总体而言, 参与人的贡献越多, 总体的收益就越高. 因此, 公共品博弈构成了一个多人博弈的“社会困境”, 每个个体都期望其他人做贡献, 自己不劳而获. 若每个人都不投入, 会导致“公地悲剧(Tragedy Of The Commons)”的出现^[21].

2.2 重复囚徒困境博弈中的策略

在重复的囚徒困境博弈中合作行为是否会涌现, 一直是科学家感兴趣的课题. 当重复次数已知时, 可以预测理性个体在最后一轮会选择背叛, 根据逆向归纳可知, 理论上有限次重复囚徒困境博弈中合作行为不会出现. 因此, 通常关注博弈轮次未知的无限次重复博弈: 每一轮博弈结束之后以概率 w 继续下一轮, 以概率 $(1-w)$ 结束博弈^[22,23].

个体之间的重复交互不清楚何时结束, 这为跳出困境提供了契机. 人们常常乐于帮助那些曾经向自己施以援手的人, 这类“投李报桃”行为称为直接互惠(Direct Reciprocity). Nowak^[8]基于捐助博弈指出, 在直接互惠中, 若重复博弈的下一轮博弈发生的概率 $w > c/b$ 时, 合作行为就会涌现. 在重复囚徒困境博弈中, 收益

参数还应满足 $2R > T + S$, 以防止博弈双方通过轮流合作和背叛获取比 R 更高的个体平均收益. 当博弈重复进行时, 个体在每一轮博弈中需要思考, 如何根据之前的历史信息在当前做出 C 或 D 的选择, 这种选择的计划称为策略 (Strategy). 通过博弈论可以探索人们决策的过程. 那么, 在重复的囚徒困境博弈中, 什么样的策略会胜出?

20 世纪 70 年代末, Axelrod 等人^[24-27]发起了重复囚徒困境博弈策略锦标赛. 他根据如表 2 的囚徒困境博弈收益矩阵, 邀请来自心理学、经济学、政治学、数学、社会学等不同领域的科学家提交他们认为最好的策略, 每个策略会在计算机上跟其他策略 (包括该策略自己和一个随机策略) 配对后进行重复囚徒困境博弈.

在第一次锦标赛中, Axelrod 收集了 14 个策略, 每个策略与其他策略重复 200 轮的囚徒困境博弈. 第二次锦标赛中, 他收集了 62 个策略, 这次重复博弈的轮数可变, 每一轮博弈后会以固定的小概率结束. 策略可以基于双方之前的历史信息进行设计, 因此可能会变得非常复杂. 然而令人惊讶的是, 两次锦标赛胜出者竟是所有策略中最简单的一个策略——由 Rapoport 提交的“以牙还牙”(Tit For Tat, TFT, 也称为“针锋相对”或者“一报还一报”). TFT 策略在第一轮选择合作, 然后在接下来的每一轮中模仿对手前一轮的选择. 通过这两次策略锦标赛, Axelrod^[27]总结认为, TFT 策略的成功得益于它拥有 4 个重要的特性: (1) 善良性, 不主动背叛; (2) 报复性, 当对方背叛之后会报以背叛; (3) 宽容性, 在对方放弃背叛之后就会立刻跟对方合作; (4) 清晰性, 其简单的策略特性很容易被识别出来. TFT 希望与对手公平分享收益, 而不是试图获得比对手更高的收益.

然而, TFT 策略也有弱点: 在噪声环境中, TFT 因为不能纠正对手的失误而会丧失建立合作的机会. 针对 TFT 策略在噪声环境中的脆弱性, Nowak 等人进一步提出了“慷慨的 TFT”(Generous-Tit-For-Tat, GTFT)^[28]和“赢存输去”(Win-stay, Lost-shift, WSLS)^[29]策略. GTFT 是一种改进的 TFT 策略, 它会以一个与收益参数相关的小概率宽容对手前一轮的背叛行为; 采用 WSLS 策略的个体具有一个期望的收益阈值, 收入高于该阈值则在下一轮坚持当前行为, 反之采取其他策略. 这两个策略都可以在噪音环境中, 有效维持重复博弈中参与者之间的合作行为.

表 2 Axelrod 锦标赛中使用的收益矩阵

Table 2 The payoff matrix used in Axelrod's tournaments

收益矩阵	参与人 2 选 C	参与人 2 选 D
参与人 1 选 C	3, 3	0, 5
参与人 1 选 D	5, 0	1, 1

由于上述策略都是根据前一轮双方的行为决定本轮行为, 因此这类策略称为单步记忆策略 (Memory-one Strategy). 考虑博弈双方 i 和 j 针对前一轮的不同状态 $ij \in \{CC, CD, DC, DD\}$, 参与者 i 的策略向量为 $p^i = (p_{CC}^i, p_{CD}^i, p_{DC}^i, p_{DD}^i)$, 参与者 j 策略向量为 $p^j = (p_{CC}^j, p_{DC}^j, p_{CD}^j, p_{DD}^j)$. 其中 p_{CD}^i 代表参与者 i 和 j 在前一轮分别选择 C 和 D 时, 本轮 i 选择合作的概率. 许多经典的策略都可以用单步记忆策略来表述, 如坚定合作策略 $(p_{CC}, p_{CD}, p_{DC}, p_{DD}) = (1, 1, 1, 1)$ 、坚定背叛策略 $(p_{CC}, p_{CD}, p_{DC}, p_{DD}) = (0, 0, 0, 0)$ 、“以牙还牙”策略 $(p_{CC}, p_{CD}, p_{DC}, p_{DD}) = (1, 0, 1, 0)$ 、“慷慨的 TFT”策略 $(p_{CC}, p_{CD}, p_{DC}, p_{DD}) = (1, q, 1, q)$ (其中遇到背叛者时合作的概率 $q = \min \left\{ 1 - \frac{T-R}{R-S}, \frac{R-P}{T-P} \right\}$)、“赢存输去”策略 $(p_{CC}, p_{CD}, p_{DC}, p_{DD}) = (1, 0, 0, 1)$ ^[8].

2.3 零行列式策略

过去的博弈理论通常认为不存在能够控制对手收益的简单策略. 2012 年, Press 和 Dyson^[30]基于重复囚徒困境博弈提出了零行列式策略 (Zero-determinant Strategy) 理论, 他们根据马尔科夫随机过程理论证明, 不论对手采取何种策略, 存在一种称为零行列式策略的单步记忆策略, 它的实施者总能够单方面保证其期望收益与对手的期望收益满足线性关系, 进而可以通过该策略控制不知情对手的收益. 零行列式策略是直接互惠的一个新理论方向, 它巧妙地由马尔科夫链角度揭示了博弈双方策略与期望收益之间的关系, 这为研究具有单步记忆特征的重要策略提供了统一的理论框架. 通过零行列式策略理论可以刻画更丰富的个体行为, 也为定量研究用户博弈行为提供了有用的工具^[31].

Press 和 Dyson 研究了两类重要零行列式策略子集——设定策略 (Equalizer Strategy) 和敲诈策略 (Extortion Strategy) 的性质. 设定策略指可以单方面让对手长期收益限定在某个预先设定的固定值, 该值可以

在 $[P, R]$ 之间任意设定. 该策略早期曾被Boerlijst等人^[32]独立提出并研究. 使用敲诈策略的个体 i 可以单方面保证他与对手 j 的期望收益满足: $E_i - P = \chi_E(E_j - P)$, 即敲诈者的期望收益与双方背叛收益之差($E_i - P$)是对手的 χ_E 倍. χ_E 称为敲诈因子, $\chi_E = 1$ 意味着二者期望收益相同, 施加该策略的个体期望双方公平, TFT策略是其中一个特例. 而 $\chi_E > 1$ 意味着敲诈行为. 如图1(a)所示, 敲诈者的期望收益不会低于对手的期望收益, 而且敲诈因子 χ_E 的大小由敲诈者单方面控制.

之后的学者重点关注敲诈策略的性质: Adami和

图1 (网络版彩图)蓝线代表不同零行列式策略与对手的期望收益关系: (a) 敲诈策略; (b) 慷慨策略. 横轴为零行列式策略的期望收益 E_i , 纵轴为对手的期望收益 E_j , 菱形虚线框为博弈双方在参数 (R, S, T, P) 下的收益区间

Figure 1 (Color online) Blue lines represent the expected payoff relationship between zero-determinant strategies and its opponent. (a) Extortion strategy; (b) generosity strategy. The x-axis is the expected payoff of zero-determinant strategy (E_i), the y-axis is the expected payoff of its opponent (E_j), and the diamond-shaped dotted box represents the payoff region of two players in terms of the elements of the payoff matrix (R, S, T, P) .

Hintze^[33]指出, 由于敲诈者与背叛者博弈时收益相同, 它容易受背叛者入侵, 因此在无限种群中敲诈策略的演化是不稳定的. Hilbe等人^[34]在有限种群中详细分析了敲诈策略的演化稳定性, 发现它在小种群中可以稳定存在, 而在大种群中虽与背叛策略之间存在中性漂移, 但敲诈策略却可以扮演催化剂作用, 促使合作行为在均匀混合种群中涌现.

Akin^[35]研究了一类扩展的零行列式策略——善良策略(Good Strategy). 它保证对手为了获得最高收益只能采取合作行为. 在此框架下Stewart和Plotkin针对囚徒困境博弈研究了零行列式策略和善良策略的关系和性质, 指出这两类策略的交集是敲诈策略的对偶策略, 称为慷慨策略(Generosity Strategy). 它可以单方面保证双方期望收益为 $E_i - R = \chi_G(E_j - R)$, 其中 χ_G 称为慷慨因子. 如图1(b)所示, 慷慨者的收益总是不会高于对手的期望收益, 但对手为了最大化自己收益必须与慷慨者合作. 在捐助博弈模型中, GTFT策略和WSLS策略分别是慷慨策略和善良策略的特例^[36]. Hilbe等人^[37]探讨了零行列式策略的记忆性对于动力学过程的影响, 理论分析了WSLS的善良纠错特征.

零行列式策略理论也可以由两人囚徒困境博弈扩展到多人公共品博弈中^[38,39], 在 $n(>2)$ 人博弈中同样存在零行列式策略, 可以单方面保证其收益与其他 $n-1$ 人的收益之和满足线性关系. 多人博弈中也存在设定、敲诈、慷慨等零行列式策略子集. 以零行列式策略为代表的单步记忆策略为研究随机博弈中的合作演化提供了新的途径^[40].

3 基于零行列式策略的博弈实验

Hilbe等人^[41]基于零行列式策略开展了真人博弈实验, 以验证其对被试收益控制的有效性. 他们召集60名被试与预先设计好的零行列式策略(计算机程序)进行重复60轮的囚徒困境博弈(收益参数 $R = 0.3$, $S = 0.0$, $T = 0.5$, $P = 0.1$). 实验共分四组, 采用的策略参数分别为严厉的敲诈(Strong Extortion, 敲诈因子 $\chi_E = 3$) ($p_{CC}, p_{CD}, p_{DC}, p_{DD}$)=(0.692, 0.0, 0.538, 0.0); 温和的敲诈(Mild Extortion, 敲诈因子 $\chi_E = 1.5$) ($p_{CC}, p_{CD}, p_{DC}, p_{DD}$)=(0.857, 0.00, 0.786, 0.0); 严厉的慷慨(Strong Generosity, 慷慨因子 $\chi_G = 3$) ($p_{CC}, p_{CD}, p_{DC}, p_{DD}$)=(1.0, 0.182, 1.0,

0.364); 温和的慷慨(Mild Generosity, 慷慨因子 $\chi_G = 1.5$) ($p_{CC}, p_{CD}, p_{DC}, p_{DD}$)=(1.0, 0.077, 1.0, 0.514). 敲诈策略初始会选择背叛, 而慷慨策略初始会选择合作. 研究人员没有告知被试其对手是计算机程序或人类, 也没有告知博弈重复的轮次.

如图2(a)所示的博弈实验结果表明, 零行列式策略在真人博弈实验中可以实现对被试收益的控制, 实验结果与理论预测一致: 敲诈策略获得了比人类对手

更高的收益, 慷慨策略的收益则比人类对手低, 并且实验数据都落在理论值的周围. 然而有趣的是, 从图2(b)中可以发现, 慷慨策略获得了比敲诈策略更高的收益. 从理论而言, 如果人类被试的目的是想最大化自己的收益, 无论遇到敲诈策略或者慷慨策略, 其最好的选择都是无条件的合作. 但是已有的研究表明, 人类行为除了追求利益最大化之外, 还受其他因素的影响, 包括对公平性的追求. 当遇到敲诈策略时, 选择合作虽然可以最大化自己的收益, 但却造成了不公平性——双方的收益差越来越大. 因此, 这降低了被试的合作意愿, 抑制了合作的涌现, 最终导致了双输的结局. 而在面对慷慨策略时, 最大化自己的收益与追求公平性是一致的, 所以人类被试以更高的频次进行合作, 达到了共赢的局面.

Wang等人^[42]采用与文献^[41]相同的严厉零行列式策略和博弈参数设置, 通过控制博弈实验重复的轮次, 以及是否告知被试对手是计算机程序这两个重要因素, 进一步展开了对比实验. 如图3他们发现, 当且仅当同时告知被试对手是计算机程序并且博弈轮次比较多(500轮)时, 敲诈策略的收益不仅高于人类对手, 也高于慷慨策略. 一方面, 当被试知道自己在跟计算机程序博弈, 而非与人类博弈时, 其对公平性的追求会降低, 因此进行选择的主要动机是最大化自我收益. 另一方面, 由于零行列式策略包含随机性, 被试需要一定的学习时间, 而较多的重复轮次提供了足够的学习机会, 帮助被试找到对自身收益最优的方式——与敲诈策略合作, 这反过来促进了敲诈策略的收益增长, 这一结果与Press和Dyson^[30]提出敲诈策略时的理论预期相符.

4 网络演化博弈

经典的演化博弈理论中通常假设个体以均匀混合(Well-mixed)的方式交互, 即任意两个个体之间接触的可能性都是一样的. 然而, 现实生活中个体之间的接触并非是全耦合或者完全随机的, 一般具有某些特定的网络组织结构, 这种网络结构与合作演化行为存在着密切的联系. 网络互惠(Network Reciprocity, 也称为空间互惠(Spatial Reciprocity))是一类重要的合作演化机制, 它关注博弈模型、网络结构与策略演化规则之间关系. 网络演化博弈可以帮助人们理解社会网络的结

图2 (网络版彩图)在重复囚徒困境博弈中, (a) 四种不同的零行列式策略与被试的实际收益(圆点)与理论预测(黑色实线); (b) 四组实验中人类被试(Hum)和零行列式策略(ZD)的平均收益^[41]

Figure 2 (Color online) (a) The empirical payoff data (circles) and the theoretical prediction (black full lines) of four different zero-determinant strategies and their opponents; (b) the average payoffs of human subjects (Hum) and zero-determinant strategies (ZD) in four experiments in repeated Prisoner's Dilemma games ^[41].

图 3 (网络版彩图)(a) 博弈实验重复 500 轮且被试被告知 (500A); (b) 重复 500 轮且被试不被告知 (500U); (c) 重复 60 轮且被试被告知 (60A); (d) 重复 60 轮且被试不被告知 (60U) 对手为计算机程序的重复囚徒困境博弈中, 严厉的敲诈策略 (E) 与严厉的慷慨策略 (G) 及人类被试 (Hum) 的平均收益^[42]

Figure 3 (Color online) The average payoffs of strong extortion (E), strong generosity (G) strategies, and human subjects (Hum) while (a) game repeats for 500 rounds and subjects are informed (500A), (b) game repeats for 500 rounds and subjects are not informed (500U), (c) game repeats for 60 rounds and subjects are informed (60A), and (d) game repeats for 60 round and subjects are not informed (60U) that they were playing with computer programs^[42].

构与群体合作行为的形成机理, 近年来引起了广泛关注^[13–19,43–46].

在网络演化博弈中, 每个个体占据网络中一个节点. 每一轮中, 个体会根据特定博弈模型与周围邻居博弈并获取收益, 然后根据策略演化规则更新自己的策略. 理论上被大家广泛研究的策略演化规则是模仿动力学 (Imitation Dynamics), 即个体 i 会随机选择一个邻居 j , 模仿邻居 j 当前策略的概率与二者收益差 $(E_j - E_i)$ 相关: $p = f(E_j - E_i)$, 其中 $E_i(E_j)$ 是本轮个体 $i(j)$ 与邻居博弈后的收益^[47,48]. 此外, 其他常用的策略演化规则还有个体期望规则 (Aspiration-driven Rule)^[49] 和近视最优反应 (Myopic Best Response)^[50]. 个体期望规则借鉴“赢存输去”思想, 个体通过比较期望收益与实际收益, 决定是否更新策略: 如果实际收益比期望收益高则坚持当前策略, 反之则会随机选择其他策略. 近视最优反应是指在策略演化时, 个体倾向于选择可以让它获得最高收益的策略. 模仿动力学与个体期望/近视最优反应规则区别在于, 前者策略更新与邻居密切

相关, 而后者策略更新主要取决于自身的推断.

网络演化博弈理论关注静态网络上不同结构对于合作行为的作用机理. 基于模仿动力学, Nowak 和 May^[51] 在规则网络上发现, 参与囚徒困境博弈的个体很容易结成合作簇, 这促进了合作在网络中的维持. 而 Hauert 和 Doebeli^[47] 基于雪堆博弈发现, 与均匀混合情况相比, 规则网络抑制了合作的涌现. Santos 等人^[48,52] 的研究指出, 无论对于两人囚徒困境/雪堆博弈模型还是多人公共品博弈模型, 在基于累计收益的策略演化情况下, 无标度网络比规则网络更能促进合作行为的涌现. Rong 等人^[53] 研究了度相关性对于两人囚徒困境和多人公共品博弈^[54] 合作演化的影响, 指出同配网络的中心节点由于倾向相互连接, 减弱了坚持合作的能力; 而异配网络上中心节点间沟通的隔绝使合作者不易湮灭. 而在基于平均收益的策略演化情况下, Wu 等人^[55] 的研究表明, 无标度网络上合作水平比规则网络要低. Chen 和 Wang^[56] 发现, 个体期望规则可以促进均质网络中合作行为的涌现; Wu 和 Rong^[57] 指出, 同步和异步更新机制下网络合作涌现情况可能是不同的. 基于死生过程和捐助博弈, Ohtsuki 等人^[58] 发现合作涌现的理论阈值与网络平均度 k 密切相关: $b/c > k$, 即益损比超过邻居平均度时, 合作容易涌现; 进一步, Allen 等人^[59] 利用随机游走和溯祖理论, 给出了任意结构网络上合作涌现的理论阈值, 这为分析网络群体上合作演化提供了新的理论依据.

近年来, 直接互惠和网络互惠之间的关联被学者逐渐关注, 基于零行列式策略的网络演化博弈动力学被深入研究. Szolnoki 和 Perc^[60] 基于囚徒困境博弈发现, 在近视最优反应规则下, 敲诈策略可以在规则网络和无标度网络上稳定存在. Rong 课题组^[61,62] 根据不同演化规则, 系统研究了各种网络上的零行列式策略演化规则: 基于模仿动力学指出, 规则网络和无标度网络中敲诈者的引入提供了促进合作的不同演化路径; 基于近视最优反应演化规则, 研究了策略演化时间尺度的多样性对复杂网络系统中敲诈策略的演化机理和促合作机制^[63]; 基于个体期望演化规则发现, 在规则格子上适当的个体期望可以促进敲诈/慷慨策略的涌现^[57,64].

博弈行为一方面受网络结构影响, 另一方面也会反作用于底层网络. 研究博弈行为与底层网络之间协同演化的动态网络博弈近年来被学者广泛关注. Zim-

mermann和Eguíluz^[65]基于囚徒困境博弈模型研究了动态网络演化博弈, 发现若允许不满意的个体拥有重新选择邻居的能力, 则随着网络度分布变得异质, 合作行为也会提升. Fu等人^[66]将声誉引入网络动态演化博弈, 发现声誉有助于促进合作水平. Li等人^[67]提出了一个可以生成无标度特性的动态网络演化博弈模型. Lee等人^[68]研究了一个多重适应的动态网络演化博弈模型, 发现博弈动力学的多尺度演化导致了层次结构的涌现.

5 网络博弈实验

5.1 静态网络博弈实验

过去20多年的理论研究显示, 参与博弈的个体通过有效结簇, 可以促进网络中群体合作行为的涌现. 然而网络博弈实验方面的研究近年来才逐渐开展^[69].

Traulsen等人^[70]通过在方格网络上进行重复囚徒困境博弈实验, 研究了个体的更新规则以及网络结构对合作行为的影响. 该实验共召集了400名被试, 分成25组, 每组16名被试. 博弈实验在4×4带周期边界的方格网络上开展. 在网络实验组(15组)中, 16名被试被分别放置在方格网络的每个节点上, 并在整个实验过程中保持位置不变, 因此其邻居关系固定. 而在均匀混合的实验组(10组)里, 被试每一轮都会被随机安排到方格网络上的其他位置上, 从而消除网络结构的影响. 博弈模型采用重复囚徒困境博弈(博弈参数 $R=0.3$, $S=0.0$, $T=0.4$, $P=0.1$). 图4实验结果表明, 在博弈开始时, 有62%的被试遵循了“模仿收益最高的邻居”这一更新策略, 剩下38%的被试行为则无法用“模仿规则”来解释. 随着时间推移, 无法被模仿规则解释的策略比例以每轮约4%的速度下降, 这说明人类被试的策略演化规则是动态变化的, 而非理论所预设的稳定不变. 研究还发现, 网络实验组和均匀混合实验组的合作水平相似. 其他几个相关的研究中也得到了相似结果^[71,72].

西班牙学者后续开展了包含1229名被试的大规模网络囚徒困境博弈实验^[73]. 该研究同时在平均度为4的带周期边界的方格网络(625个节点)和异质的无标度网络(604个节点)上进行. 被试首先在固定网络上进行51轮的网络实验组博弈, 之后再进行59轮均匀混合情况的控制组博弈实验, 博弈参数为 $R=7$, $P=S=0$,

图4 (网络版彩图)在固定邻居关系的方格网络博弈中, 大部分被试的策略改变可以用“模仿收益最高的邻居(Imitate The Best)”规则来解释, 包括当邻居的策略不同(Heterogeneous)时转换为收益最高的策略以及当邻居的策略都相同(Homogeneous)时坚持当前的策略. 但仍有一部分策略的改变(Others)无法被这个规则解释, 这些无法被解释的比例以大约每轮4%的速度降低^[70]

Figure 4 (Color online) While playing games on fixed lattice, most strategy changes can be explained by “imitation of the best neighbors (imitate the best)”, including changing to the best available strategy in a heterogeneous environment or sticking to the strategy when everyone does the same in a homogeneous environment. However, there is a portion of strategy changes which cannot be explained by imitation rule, and decays approximately by 4% per round [70]

$T=10$. 如图5所示, 实验发现, 无论是在方格网络或无标度网络上, 开始有大约60%的被试选择合作, 而该比例在博弈过程中不断下降, 最终结束时只有约30%的合作被试, 这个比例与均匀混合的控制组结果相似. 进一步, 通过对实验过程中被试的合作频率分布分析发现, 在两种网络上被试的合作行为非常相似, 均呈现正态分布的特征: 大部分个体的合作频率处于中等水平, 少部分个体有较高或较低的合作水平.

上述研究得出对网络互惠较为悲观的结论. 然而, 在对文献^[70]的评论中, Helbing与Yu^[74]指出理论与实验不符的一个重要原因是理论结果大多基于模仿动力学规则获得, 而实验研究并没有对策略演化规则作进一步限制. Ohtsuki等人^[58]的理论研究指出, 在网络捐助博弈中当益损比 $b/c > k$ (网络平均度)时, 合作可以在网络中涌现. Rand等人^[75]基于此开展了网络博弈实验验证工作. 他们首先在线下招募了109名大学生, 通过对比随机交互与满足 $b/c > k$ 的最近邻网络博弈实验, 发现益损比与网络平均度的关系会影响了网络上合作行为. 进一步, 他们在亚马逊土耳其机器人(Amazon Mechanical Turk)平台上招募了1163名线上

图 5 (网络版彩图) (a) 实验组(Experiment)与(b) 控制组(Control)博弈实验中, 在方格(Lattice)网络和异质(Heterogeneous)网络上每一轮的合作比例^[73]
Figure 5 (Color online) The fraction of cooperation per round on the lattice and the heterogeneous network in (a) the experiment and (b) the control game experiments ^[73].

被试, 在最近邻网络上通过控制不同的益损比 b/c 和网络平均度 k , 研究了两者关系对被试合作行为的影响. 如图6(a)所示, 当 $b/c > k$ 时, 经过一段时间的调整之后, 合作行为在网络上可以稳定存在; 而在 $b/c \leq k$ 的情况下, 合作比例会持续下降. 对最后一轮的被试行为分析(图6(b))可以看出, 在 $b/c > k$ 的组里有更高的合作比例.

网络演化博弈理论研究表明, 在固定交互网络上合作者可以通过结簇抵御背叛者的入侵. 因此, Rand 等人利用同配因子研究了网络博弈实验中合作簇演化. 同配因子是网络中合作者的合作邻居平均数与背叛者的合作邻居的平均数之差. 在图7中可以看到, 当 $b/c > k$ 时, 同配因子的平均值在博弈开始后快速升高, 说明合作簇可以在网络上涌现, 这与理论预测相符; 而该现象在 $b/c \leq k$ 时却没有被观察到. 这揭示了静态网络中合作维持和涌现与合作簇密切相关.

5.2 动态网络博弈实验

在实际中人们的社交关系经常是动态变化的. 随着时间的推移, 一些已有的关系会消失, 另一些新的关系会建立^[76].

Rand等人^[77]召集了785名被试进行20人为一组的

图 6 (网络版彩图)(a) 在 $b/c > k$ 和 $b/c \leq k$ 的情况下, 每一轮合作被试的比例; (b) 不同的 b/c 与 k 值组合下, 最后一轮选择合作的被试比例^[75]

Figure 6 (Color online) (a) The fraction of cooperation per round for $b/c > k$ and $b/c \leq k$; (b) the fraction of subjects choosing cooperation in the last round with different values of b/c and k ^[75].

图 7 (网络版彩图) b/c 与 k 关系不同时, 每一轮博弈中同配因子(Assortment)的平均值^[75]

Figure 7 (Color online) The average value of assortment per round for different relationships between b/c and k ^[75].

动态网络博弈实验. 实验考虑如下4种情况: 随机交互、固定网络、缓慢演化动态网络和快速演化动态网络. 实验初始时, 被试之间有20%比例的边被随机连接. 之后被试在网络上根据连边进行重复捐助博弈($b=100$, $c=50$). 对于随机交互实验, 在每轮博弈之后, 被试间会被重新随机连接. 在固定网络实验中, 初始邻居确定后就保持不变. 对于动态网络实验, 每一轮博弈

之后, 有 $f\%$ 的“个体对”会被选中进行更新邻居: 如果两个个体之间已经存在连边, 则随机选中一端被试, 他有权决定是否要断开这条边; 如果他们之前并没有相连, 则随机选中的被试可以选择是否要建立连边. 无论哪种情况, 被选中的一方在做决定之前, 可以看到另一方在前一轮博弈中的选择. 对于缓慢演化动态网络, “个体对”比例 $f\%=10\%$, 快速演化动态网络中 $f\%=30\%$.

如图8所示, 虽然初始不同网络上的合作比例非常接近, 但在随机、固定和缓慢演化的网络上, 合作比例都会在博弈过程中不断下降. 而在快速演化动态网络中, 合作可以一直保持在较高的比例. 进一步的分析发现, 快速演化网络比其他网络有更异质的度分布, 并且C-C连边关系最为稳定, C-D/D-C次之, D-D最脆弱. 被试们倾向于与合作者建立连边, 而与背叛者断开连边. 因此某种程度上, 连断边可以被视为一种奖惩机制.

同期, Wang等人^[78]通过召集被试进行了重复12轮的动态网络囚徒困境博弈实验研究, 其中博弈参数 $R=4$, $S=-1$, $T=7$, $P=1$. 初始时, 他们使用了含5个邻居的六派系网络和随机正则网络. 每 r 轮博弈之后, 被试可以选择更新 k 个邻居关系. 与文献^[77]不同的是, 在这个实验中, 断边是单方面的, 无需对方的同意, 但连边必须经过对方的同意. 如图9所示, 结果表明在动态网络上, 初始的合作比例显著高于静态网络; 当静态网络上的合作水平在下降的时候, 动态网络上的合作比例可以维持在80%–100%的水平; 由于被试一开始就被告知了博弈重复的轮数, 因此随着博弈接近

图8 (网络版彩图) 在随机交互(Random)、固定网络(Fixed)、缓慢演化动态网络(Viscous Dynamic)和快速演化动态网络(Fluid Dynamic)上每一轮的合作被试比例^[77]

Figure 8 (Color online) The fraction of cooperation subjects per round under four conditions: random interaction, fixed network, viscous dynamic network, and fluid dynamic network ^[77].

结束, 动态网络上的合作比例会快速下降, 最后与静态网络的合作比例一致. 作者在动态网络中发现了合作簇的涌现, 这是由于合作者避开了主动与背叛者连边, 而背叛者倾向于与背叛者连边, 没有积极寻求与合作者进行连边导致的.

5.3 通过网络博弈实验探索复杂社会系统的形成

除了关注网络结构对博弈行为的影响以外, 网络博弈实验也被用于探索复杂社会系统结构的演化机理. 2012年, 哈佛大学的研究人员深入坦桑尼亚北部, 通过公共品博弈实验对哈扎人(Hadza)之间的社会合作网络展开了研究^[4]. 他们对17个营地的205名成年哈扎人(男性102名, 女性103名)开展一轮公共品博弈实验, 构建了馈赠社会网络. 他们发现社会网络的度与捐献程度有高度的正相关性, 形成的社会合作关系具有现代社会网络的重要特征, 如度异质、高聚类、同质性和互惠性等. 这表明社会网络的重要结构特征在合作机制驱动下可能在人类早期就已经形成.

社团结构在我们的日常生活中扮演着重要的角色, 长期以来吸引了许多学者的关注. 然而, 在社会网络中, 形形色色的个体之间如何涌现出社团结构, 一直是一个悬而未解的问题. Han等人^[79]通过网络上的最后通牒博弈实验, 研究了社团结构涌现的问题. 他们召集了4组被试进行重复60–70轮的重复博弈实验, 每组50人, 分为T1, T2和C1, C2组. 其中T1和T2是静态网络组, 这两个组的被试初始会被分配到一个规则二分网络(T1, 每个节点有4个邻居)或随机二分网络(T2, 平均每个节点有4个邻居)上. C1和C2作为控制组, 参与的被试每轮会在保持度不变情况下被重新随机配对.

被试在两两配对之后会进行最后通牒博弈(Ultimatum Game, UG), 其中一方是提议者(Proposer), 另一方是响应者(Responder). 每一轮博弈中, 两者共同决定如何分配一笔100元的虚拟货币. 提议者会提出将总额中 p 元 ($0 \leq p \leq 100$) 分给对方, 响应者有一个最低接受金额 q ($0 \leq q \leq 100$), 提议者事先不知道响应者本轮的最低接受值 q . 如果提议者的提议金额 $p < q$, 响应者会拒绝该提议, 双方将一无所得. 反之若 $p \geq q$, 响应者会接受该提议, 双方按照提议分配这100元虚拟货币. 如果双方都是理性的, 那么在单轮最后通牒博弈中, 理论上提议者会提出最小分配单位的钱分给对手, 而响应者

图 9 (网络版彩图)初始为六派系网络(a, c和e)或随机正则网络(b, d和f)情况下, 网络博弈实验中每一轮的合作比例. 被试每隔 r 轮可以更新 k 个邻居关系. (a, b) $r=1$; (c, d) $r=3$; (e, f) $r=6$. 不同线代表不同的 k 值^[78]

Figure 9 (Color online) The fraction of cooperation per round in networked game experiments for cliques (a, c, and e) and random (b, d, and f) initial conditions. The subjects can make up to k partner updates after every r rounds. (a, b) $r=1$; (c, d) $r=3$; (e, f) $r=6$. Different lines represent different values of k [78].

也将接受这个提议. 然而实验研究表明, 响应者会拒绝不公平的提议^[5]. 最后通牒博弈可以看作含惩罚的公共品博弈^[80], 它被广泛地应用于研究公平感与利他性.

在网络最后通牒博弈实验中, 被试初始会被随机分配到一个角色(提议者或响应者), 这个角色在整个博弈过程中会保持不变^[79]. 在二分网络上, 提议者的邻居都是响应者, 反之亦然. 在每轮博弈中, 每位提议者会向响应者邻居提出一个提议金额 p , 而每位响应者需要向提议者邻居提出一个最低接受金额 q . 被试

根据最后通牒模型依次与邻居博弈得到累计收益, 将之除以邻居度得到的平均收益作为其本轮收益. 被试博弈前知道前一轮自己和邻居的提议值或者最低接受值, 以及自己的平均收益.

研究发现, 对比控制组(C1和C2), 在静态网络组(T1和T2)中提议金额有非常大的方差, 这表明在网络上出现了更大的提议多样性. 通过如图10所示的空间分布可以看出, 在静态网络上的提议者之间涌现了局部的社团结构, 社团内的提议者都具有相似的提议

图 10 (网络版彩图)提议者的时空模式. 通过将具有共同邻居的提议者放在一起, 可以看到在静态网络上涌现了提议者的社团结构(a, b), 但这在随机重连的控制组上却没有被观察到(c, d). 颜色的深浅代表提议金额 p 值大小^[79]
Figure 10 (Color online) Spatiotemporal patterns of proposers. By making the proposers with common neighbors adjacent to each other, it can be observed that local communities of proposers emerge on the static network (a, b), which are not found in the control groups with random interactions (c, d) ^[79].

金额. 而该现象在控制组中并没有发现, 这表明底层网络组织对社团的涌现起了重要作用.

通过对响应者的行为分析, 作者发现响应者会做出非理性的决定, 但提议者们会根据响应者以往的行为, 采用近视最优反应规则来学习和优化自己的提议, 并最终根据响应者的特性形成局部的社团结构. 这也揭示了在网络上提议的多样性是由响应者接受程度的多样性引起的. 这个机制需要被试间存在网络关系才能发挥作用. 反之在随机交互的控制组中, 虽然响应者的行为和静态网络组非常相似, 但提议者无法形成局部社团. 最后作者通过计算机进行了更大规模的仿真研究, 验证了实验结果的鲁棒性.

通过网络博弈实验也可以探究社会系统中财富分配的问题. 通过亚马逊土耳其机器人在线平台召集 1462 名被试, 进行重复 10 轮的捐助博弈实验($b=100$, $c=50$), Nishi 等人^[81]在动态网络系统中研究了“初始财富不平等性(Inequality, 用基尼系数刻画)”和“邻居财富透明性(Visibility)”的演化. 被试被分为 80 组, 初始时每组的参与者之间通过随机网络连接. 在每一轮博弈之后, 被试可以选择连边或断边. 研究结果发现, 如果财富不透明(被试仅知道自己的收益), 基尼系数会在博弈演化中逐渐降低, 个体财富趋于平等化. 但当财富透明时(被试不但知道自己的收益, 还知道邻居的收益), 初始的财富不平等会一直保持.

如图 11 所示, 当财富透明时, 合作比例和网络平均度都显著低于财富不透明的情况. 这都跟初始财富平等与否无显著相关性. 更进一步研究发现, 当存在初

图 11 (网络版彩图)初始时财富不平等(Initial Inequality)程度以及财富是否透明(Wealth Visibility)对每一轮的平均财富(a), 合作比例(b), 网络平均度(c)和网络传递性(d)的影响^[81]
Figure 11 (Color online) The influence of initial inequality and wealth visibility on the (a) average wealth, (b) proportion of cooperation, (c) network degree, and (d) network transitivity ^[81].

始的财富不平等时, 相对富有的个体更倾向于选择背叛, 而相对贫穷的个体更倾向于选择合作, 这进一步加强了富有者对贫穷者的剥削, 最终导致较高的不平等性. 阶层固化和财富不平等是许多社会遇到的问题, 这个实验研究对如何解决这一问题有重要的参考价值.

6 总结

“合作行为如何演化”被*Science*列为未来需要解决的25个最重要科学问题之一^[82]. 博弈论为探索理性个体之间合作的涌现提供了有力的理论框架. 本文主要基于两人囚徒困境和多人公共品博弈模型, 系统地介绍了近年来围绕直接互惠和网络互惠这两个重要合作涌现机制开展的理论和实验研究成果.

直接互惠关注重复交互的熟人之间的信任建立和长期合作形成的条件. 零行列式策略作为直接互惠的一个新理论, 提供了单方面控制博弈双方收益关系的途径. 近期基于两人囚徒困境的博弈实验也验证了零行列式策略对收益控制的有效性, 并且发现其可以在博弈实验中引起人类参与者微妙而深刻的行为变化. 传统的真人博弈实验主要研究不同人类参与者之间的博弈, 博弈双方的不可控因素会增加实验结果的不确定性. 而基于零行列式策略开展的人-机博弈实验, 为在更加可控的条件下探索个体的决策机理提供了潜在的工具.

网络互惠作为促进合作涌现的重要机制, 探讨了网络结构、策略演化和博弈模型之间的相互关系. 一方面, 网络博弈实验有助于揭示重要的网络结构特征(如异质性、互惠性、社团结构等)在群体中的形成机

制. 另一方面, 网络结构对于合作演化的影响近年来也在博弈实验领域备受关注. 动态网络中的连边/移边机制可以看作特殊的奖励/惩罚手段. 实验研究表明, 动态网络可以促进合作行为的涌现, 这与前期网络博弈理论预测一致. 而静态网络是否促进合作尚存在争议. 网络博弈理论研究通常假设个体基于特定演化规则进行策略的学习与更新, 然而在真实场景中, 人类参与者的策略学习与演化规律是复杂且动态变化的, 甚至包含一定的非理性成分. 因此, 如何基于已有的研究理论, 通过巧妙的博弈实验设计, 探索个体博弈决策的规律与群体行为的涌现, 仍是充满机遇与挑战的问题.

重复博弈中, 奖励和惩罚对于合作演化的作用机理也是近年来直接互惠理论和实验研究的重要课题^[7]. 真人博弈实验发现, 在多人公共品博弈中, 通过惩罚搭便车的背叛者, 可以促进多人博弈中的合作涌现^[83]. 然而惩罚本身也是有代价的, 基于两人囚徒困境的博弈实验表明, 惩罚的引入并不会提高博弈双方总体收益^[84], 甚至可能抑制合作^[85]. 虽然惩罚可以抑制“搭便车”的现象, 但“惩罚”本身又构成了二阶的“搭便车”问题^[86]——“惩罚者”惩罚背叛者, 其他个体会搭惩罚者的便车. 为此, Sigmund等人^[87]通过理论分析, 指出采用集体惩罚(Pool Punishment)替代同伴惩罚(Peer Punishment), 即在公共品博弈中, 并不是由惩罚者直接惩罚背叛者, 而是将惩罚代价投入第二个公共池(惩罚池), 由惩罚池而不是个人去惩罚背叛者, 该机制可以维持合作行为的长期稳定性. Chen等人^[88]提出了一种称为“先萝卜后大棒”的机制, 将奖励与惩罚结合在一起, 以较低的代价对合作起到了极大的促进效果. 这些关于奖惩的成果值得在未来博弈实验中进行验证和进一步研究.

参考文献

- 1 Wilkinson G S. Reciprocal food sharing in the vampire bat. *Nature*, 1984, 308: 181–184
- 2 Milinski M. TIT FOR TAT in sticklebacks and the evolution of cooperation. *Nature*, 1987, 325: 433–435
- 3 Vaidya N, Manapat M L, Chen I A, et al. Spontaneous network formation among cooperative RNA replicators. *Nature*, 2012, 491: 72–77
- 4 Apicella C L, Marlowe F W, Fowler J H, et al. Social networks and cooperation in hunter-gatherers. *Nature*, 2012, 481: 497–501
- 5 Fehr E, Fischbacher U. The nature of human altruism. *Nature*, 2003, 425: 785–791
- 6 Nowak M A, Highfield R. *Supercooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*. New York: Free Press, 2011
- 7 Wang L, Cong R, Li K. Feedback mechanism in cooperation evolving (in Chinese). *Sci Sin Inf*, 2014, 44: 1495–1514 [王龙, 丛睿, 李昆. 合作演化中的反馈机制. *中国科学: 信息科学*, 2014, 44: 1495–1514]
- 8 Nowak M A. Five rules for the evolution of cooperation. *Science*, 2006, 314: 1560–1563

- 9 Von Neumann J, Morgenstern O. Theory of Games and Economic Behavior. Princeton: Princeton University Press, 1944
- 10 Nash J F. Equilibrium points in n -person games. *Proc Natl Acad Sci USA*, 1950, 36: 48–49
- 11 Smith J M, Price G R. The logic of animal conflict. *Nature*, 1973, 246: 15–18
- 12 Nowak M A, Sasaki A, Taylor C, et al. Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 2004, 428: 646–650
- 13 Hauert C, Szabó G. Game theory and physics. *Am J Phys*, 2005, 73: 405–414
- 14 Szabó G, Fáth G. Evolutionary games on graphs. *Phys Rep*, 2007, 446: 97–216
- 15 Gross T, Blasius B. Adaptive coevolutionary networks: A review. *J R Soc Interface*, 2008, 5: 259–271
- 16 Roca C P, Cuesta J A, Sánchez A. Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. *Phys Life Rev*, 2009, 6: 208–249, arXiv: [0911.1720](#)
- 17 Perc M, Gómez-Gardeñes J, Szolnoki A, et al. Evolutionary dynamics of group interactions on structured populations: A review. *J R Soc Interface*, 2013, 10: 20120997
- 18 Adami C, Schossau J, Hintze A. Evolutionary game theory using agent-based methods. *Phys Life Rev*, 2016, 19: 1–26
- 19 Szabó G, Borsos I. Evolutionary potential games on lattices. *Phys Rep*, 2016, 624: 1–60, arXiv: [1508.03147](#)
- 20 Perc M, Jordan J J, Rand D G, et al. Statistical physics of human cooperation. *Phys Rep*, 2017, 687: 1–51, arXiv: [1705.07161](#)
- 21 Hardin G. The tragedy of the commons. *Science*, 1968, 162: 1243–1248
- 22 Roth A E, Murnighan J K. Equilibrium behavior and repeated play of the Prisoner's Dilemma. *J Math Psychol*, 1978, 17: 189–198
- 23 Murnighan J K, Roth A E. Expecting continued play in Prisoner's Dilemma games. *J Conflict Resolut*, 1983, 27: 279–300
- 24 Axelrod R. Effective choice in the Prisoner's Dilemma. *J Conflict Resolut*, 1980, 24: 3–25
- 25 Axelrod R. More effective choice in the Prisoner's Dilemma. *J Conflict Resolut*, 1980, 24: 379–403
- 26 Axelrod R, Hamilton W D. The evolution of cooperation. *Science*, 1981, 211: 1390–1396
- 27 Axelrod R M. The Evolution of Cooperation (Revised Edition). New York: Basic Books, 2006
- 28 Nowak M A, Sigmund K. Tit for tat in heterogeneous populations. *Nature*, 1992, 355: 250–253
- 29 Nowak M, Sigmund K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, 1993, 364: 56–58
- 30 Press W H, Dyson F J. Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. *Proc Natl Acad Sci USA*, 2012, 109: 10409–10413
- 31 Hilbe C, Chatterjee K, Nowak M A. Partners and rivals in direct reciprocity. *Nat Hum Behav*, 2018, 2: 469–477
- 32 Boerlijst M C, Nowak M A, Sigmund K. Equal pay for all prisoners. *Am Math Mon*, 1997, 104: 303–305
- 33 Adami C, Hintze A. Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nat Commun*, 2013, 4: 2193, arXiv: [1208.2666](#)
- 34 Hilbe C, Nowak M A, Sigmund K. Evolution of extortion in Iterated Prisoner's Dilemma games. *Proc Natl Acad Sci USA*, 2013, 110: 6913–6918, arXiv: [1212.1067](#)
- 35 Akin E. What you gotta know to play good in the iterated Prisoner's Dilemma. *Games*, 2015, 6: 175–190
- 36 Stewart A J, Plotkin J B. From extortion to generosity, evolution in the iterated Prisoner's Dilemma. *Proc Natl Acad Sci USA*, 2013, 110: 15348–15353
- 37 Hilbe C, Martínez-Vaquero L A, Chatterjee K, et al. Memory- n strategies of direct reciprocity. *Proc Natl Acad Sci USA*, 2017, 114: 4715–4720
- 38 Pan L, Hao D, Rong Z, et al. Zero-determinant strategies in iterated public goods game. *Sci Rep*, 2015, 5: 13096, arXiv: [1402.3542](#)
- 39 Hilbe C, Wu B, Traulsen A, et al. Cooperation and control in multiplayer social dilemmas. *Proc Natl Acad Sci USA*, 2014, 111: 16425–16430
- 40 Hilbe C, Šimsa Š, Chatterjee K, et al. Evolution of cooperation in stochastic games. *Nature*, 2018, 559: 246–249
- 41 Hilbe C, Röhl T, Milinski M. Extortion subdues human players but is finally punished in the Prisoner's Dilemma. *Nat Commun*, 2014, 5: 3976
- 42 Wang Z, Zhou Y, Lien J W, et al. Extortion can outperform generosity in the iterated Prisoner's dilemma. *Nat Commun*, 2016, 7: 11125
- 43 Wang L, Fu F, Cheng X J, et al. Evolutionary games on complex networks (in Chinese). CAAI Trans Intel Sys, 2007, 2: 1–10 [王龙, 伏锋, 陈小杰, 等. 复杂网络上的演化博弈. 智能系统学报, 2007, 2: 1–10]
- 44 Wu Z-X, Rong Z H, Wang W X. Games on complex networks (in Chinese). Adv Mech, 2008, 38: 794–804 [吴枝喜, 荣智海, 王文旭. 复杂网络上的博弈. 力学进展, 2008, 38: 794–804]
- 45 Perc M, Szolnoki A. Coevolutionary games—A mini review. *Biosystems*, 2010, 99: 109–125
- 46 Rong Z H, Wu Z-X, Wang W X. Research on the networked cooperative dynamics of coevolutionary games (in Chinese). J Univ Elec Sci

- Technol China, 2013, 42: 10–22 [荣智海, 吴枝喜, 王文旭. 共演博弈下网络合作动力学研究进展. 电子科技大学学报, 2013, 42: 10–22]
- 47 Hauert C, Doebeli M. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, 2004, 428: 643–646
- 48 Santos F C, Pacheco J M. Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys Rev Lett*, 2005, 95: 098104
- 49 Macy M W, Flache A. Learning dynamics in social dilemmas. *Proc Natl Acad Sci USA*, 2002, 99: 7229–7236
- 50 Matsui A. Best response dynamics and socially stable strategies. *J Economic Theor*, 1992, 57: 343–362
- 51 Nowak M A, May R M. Evolutionary games and spatial chaos. *Nature*, 1992, 359: 826–829
- 52 Santos F C, Santos M D, Pacheco J M. Social diversity promotes the emergence of cooperation in public goods games. *Nature*, 2008, 454: 213–216
- 53 Rong Z, Li X, Wang X. Roles of mixing patterns in cooperation on a scale-free networked game. *Phys Rev E*, 2007, 76: 027101
- 54 Rong Z, Wu Z X. Effect of the degree correlation in public goods game on scale-free networks. *Europhys Lett*, 2009, 87: 30001
- 55 Wu Z X, Guan J Y, Xu X J, et al. Evolutionary Prisoner's Dilemma game on Barabási-Albert scale-free networks. *Phys A-Statist Mech its Appl*, 2007, 379: 672–680
- 56 Chen X, Wang L. Promotion of cooperation induced by appropriate payoff aspirations in a small-world networked game. *Phys Rev E*, 2008, 77: 017103
- 57 Wu Z X, Rong Z. Boosting cooperation by involving extortion in spatial Prisoner's Dilemma games. *Phys Rev E*, 2014, 90: 062102, arXiv: 1408.1793
- 58 Ohtsuki H, Hauert C, Lieberman E, et al. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 2006, 441: 502–505
- 59 Allen B, Lippner G, Chen Y T, et al. Evolutionary dynamics on any population structure. *Nature*, 2017, 544: 227–230, arXiv: 1605.06530
- 60 Szolnoki A, Perc M. Evolution of extortion in structured populations. *Phys Rev E*, 2014, 89: 022804, arXiv: 1401.8294
- 61 Xu X, Rong Z, Wu Z X, et al. Extortion provides alternative routes to the evolution of cooperation in structured populations. *Phys Rev E*, 2017, 95: 052302
- 62 Mao Y, Xu X, Rong Z, et al. The emergence of cooperation-extortion alliance on scale-free networks with normalized payoff. *Europhys Lett*, 2018, 122: 50005
- 63 Rong Z, Wu Z X, Hao D, et al. Diversity of timescale promotes the maintenance of extortioners in a spatial Prisoner's Dilemma game. *New J Phys*, 2015, 17: 033032
- 64 Rong Z H, Zhao Q, Wu Z X, et al. Proper aspiration level promotes generous behavior in the spatial Prisoner's Dilemma game. *Eur Phys J B*, 2016, 89: 166
- 65 Zimmermann M G, Eguíluz V M. Cooperation, social networks, and the emergence of leadership in a Prisoner's Dilemma with adaptive local interactions. *Phys Rev E*, 2005, 72: 056118
- 66 Fu F, Hauert C, Nowak M A, et al. Reputation-based partner choice promotes cooperation in social networks. *Phys Rev E*, 2008, 78: 026117
- 67 Li W, Zhang X, Hu G. How scale-free networks and large-scale collective cooperation emerge in complex homogeneous social systems. *Phys Rev E*, 2007, 76: 045102
- 68 Lee S, Holme P, Wu Z X. Emergent hierarchical structures in multiadaptive games. *Phys Rev Lett*, 2011, 106: 028702, arXiv: 1012.5408
- 69 Rand D G, Nowak M A. Human cooperation. *Trends Cogn Sci*, 2013, 17: 413–425
- 70 Traulsen A, Semmann D, Sommerfeld R D, et al. Human strategy updating in evolutionary games. *Proc Natl Acad Sci USA*, 2010, 107: 2962–2966, arXiv: 1001.3768
- 71 Grujić J, Fosco C, Araujo L, et al. Social experiments in the mesoscale: Humans playing a spatial Prisoner's Dilemma. *PLoS ONE*, 2010, 5: e13749
- 72 Suri S, Watts D J. Cooperation and contagion in web-based, networked public goods experiments. *PLoS ONE*, 2011, 6: e16836, arXiv: 1008.1276
- 73 Gracia-Lázaro C, Ferrer A, Ruiz G, et al. Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma. *Proc Natl Acad Sci USA*, 2012, 109: 12922–12926
- 74 Helbing D, Yu W. The future of social experimenting. *Proc Natl Acad Sci USA*, 2010, 107: 5265–5266, arXiv: 1012.4182
- 75 Rand D G, Nowak M A, Fowler J H, et al. Static network structure can stabilize human cooperation. *Proc Natl Acad Sci USA*, 2014, 111: 17093–17098
- 76 Palla G, Barabási A L, Vicsek T. Quantifying social group evolution. *Nature*, 2007, 446: 664–667, arXiv: 0704.0744

- 77 Rand D G, Arbesman S, Christakis N A. Dynamic social networks promote cooperation in experiments with humans. *Proc Natl Acad Sci USA*, 2011, 108: 19193–19198
- 78 Wang J, Suri S, Watts D J. Cooperation and assortativity with dynamic partner updating. *Proc Natl Acad Sci USA*, 2012, 109: 14363–14368
- 79 Han X, Cao S, Shen Z, et al. Emergence of communities and diversity in social networks. *Proc Natl Acad Sci USA*, 2017, 114: 2887–2891
- 80 Sigmund K, Hauert C, Nowak M A. Reward and punishment. *Proc Natl Acad Sci USA*, 2001, 98: 10757–10762
- 81 Nishi A, Shirado H, Rand D G, et al. Inequality and visibility of wealth in experimental social networks. *Nature*, 2015, 526: 426–429
- 82 Pennisi E. How did cooperative behavior evolve? *Science*, 2005, 309: 93
- 83 Fehr E, Gächter S. Cooperation and punishment in public goods experiments. *Am Eco Rev*, 2000, 90: 980–994
- 84 Dreber A, Rand D G, Fudenberg D, et al. Winners don't punish. *Nature*, 2008, 452: 348–351
- 85 Wu J J, Zhang B Y, Zhou Z X, et al. Costly punishment does not always increase cooperation. *Proc Natl Acad Sci USA*, 2009, 106: 17448–17451
- 86 Fehr E, Rockenbach B. Detrimental effects of sanctions on human altruism. *Nature*, 2003, 422: 137–140
- 87 Sigmund K, De Silva H, Traulsen A, et al. Social learning promotes institutions for governing the commons. *Nature*, 2010, 466: 861–863
- 88 Chen X, Sasaki T, Brännström Å, et al. First carrot, then stick: How the adaptive hybridization of incentives promotes cooperation. *J R Soc Interface*, 2015, 12: 20140935

Experiment research on the evolution of cooperation and network game theory

RONG ZhiHai¹, XU XiongRui¹ & WU Zhi-Xi^{2*}

¹ School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China;

² Institute of Computational Physics and Complex Systems, Lanzhou University, Lanzhou 730000, China

Focusing on the mechanisms of the evolution and emergence of cooperative behavior, this paper reviews the recent development of direct reciprocity and network reciprocity from perspectives of both theoretical and experimental research based on two-player Prisoner's Dilemma and multi-player Public Goods games. Repeated interactions provide the necessary condition for the emergence of direct reciprocity. The zero-determinant strategy, which is a new theory in direct reciprocity, has been deeply investigated from both theory and experiment. The zero-determinant strategy is a kind of memory-one strategies, which can unilaterally determine the payoff relationship in repeated games and may provide a novel tool to explore human decision behavior. Investigations on networked game experiments revealed that dynamic networks can facilitate and maintain human cooperation, while the influence of static networks on cooperation remains controversial. Furthermore, it is summarized that networked game experiments provide an efficient approach to studying the formation mechanism of network structure and the emergence of cooperation in complex systems.

network reciprocity, direct reciprocity, game experiments, complex networks

PACS: 02.50.Le, 89.75.-k, 89.65.-s

doi: 10.1360/SSPMA-2019-0129