

观察机制对囚徒困境博弈合作水平的影响

蒋贤惠¹, 蒋玉莲^{2,†}, 黄勤珍²

(1. 西南民族大学 电气工程学院国家民委电子信息重点实验室, 四川 成都 610041;

2. 西南民族大学 电气工程学院, 四川 成都 610041)

摘要: 在演化博弈论中合作有利于增加群体收益. 目前, 虽然大量的研究工作都在关注合作水平的提高, 但忽视了理论和实际的相关性, 同时其策略更新时间仅限于一个点上的收益. 为此, 引入包括完全同步、相对同步、高斯异步四种观察机制模型, 来考虑智能体观察期内所得的平均收益. 随后, 采用蒙特卡洛方法进行实验仿真并分析其对囚徒困境博弈合作水平的影响. 结果表明, 在四种模型中都存在观察时间阈值, 此时合作水平将达到峰值; 最高合作水平在指数异步模型中实现; 高斯异步模型中合作水平与方差成正相关. 同时, 智能体间的异质性也对合作产生了积极影响. 这项研究有助于打破传统研究收益的局限, 促进合作水平的提高. 仿真实验结果为支持社会群体合作提供了一定的理论依据.

关键词: 观察机制; 囚徒困境博弈; 异质性; 合作水平; 平均收益; 蒙特卡洛方法

中图分类号: F224.32

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.0666

引用格式: 作者一, 作者二, 作者三, 等. 控制与决策论文L^AT_EX 模板说明[J]. 控制与决策.

The Impact of the Observation Mechanism on Cooperation Level in the Prisoner's Dilemma Game

JIANG Xianhui¹, JIANG Yulian^{2,†}, HUANG Qinzhen²

(1. Key Laboratory of Electronic Information of State Ethnic Affairs Commission, College of Electrical Engineering, Southwest Minzu University, Chengdou 610041, China; 2. College of Electrical Engineering, Southwest Minzu University, Chengdou 610041, China)

Abstract: In evolutionary game theory, cooperation is beneficial to increasing collective revenues. Presently, although the cooperation level has been raised in abundant studies, the correlation between theory and practice is ignored, while the update time of the strategy is restricted to one point of earnings. Here, we introduce an observation mechanism to investigate the average return of the agent during the observation period, including four observation models: Complete synchronization, Relative synchronization, Gaussian asynchronization and Exponential asynchronization. Subsequently, we employ the Monte Carlo method for the experimental simulation and analyze its effect on the cooperation level of prisoner's dilemma game. The results suggest that there is a time-observing threshold in every model and the cooperation level is optimum at the moment; the highest level of cooperation is realized in the Exponential asynchronous model; the cooperation level in the Gaussian asynchronous model is positively correlated with the square deviation. Meanwhile, the heterogeneity between different agents also produces a positive influence on the cooperation. This work is conducive to eradicating the limitations of benefits in traditional research and promoting the improvement of the cooperation level. Some theoretical basis are provided for supporting social group cooperation.

Keywords: observation mechanism; prisoner's dilemma game; heterogeneity; cooperation level; average income; Monte Carlo method

0 引言

博弈论^[1]是一种利用数学模型研究具有利益冲突的最优决策理论, 为研究个体间合作的出现提供

了理论支持. 合作行为在自然界中无处不在^[2], 这种行为可以增加群体的总回报. 人类的合作行为如何产生和演化以及人类应该如何进行更好地合作

收稿日期: 2022-04-21; 录用日期: 2022-09-03.

基金项目: 国家自然科学基金项目(62073270), 国家民委创新团队、2021中央高校省部级共建平台专项基金(2021NYYXS115), 四川省教育厅创新团队(15TD0050), 西南民族大学2021研究生创新型科研项目(项目编号CX2021SZ36).

†通讯作者. E-mail: jyl-ee@swun.edu.cn.

是合作进化的核心。合作进化的概念^[3]和博弈的动态机制^[4]以及基于策略更新机制的费米过程^[5]表明, 玩家将将自己的收益与邻居的收益进行比较, 并有一定的概率研究邻居的策略。由于现实生活中, 人们的收益并不是简单对时间进行累加, 而是需要参考以往的经验。关于这些经验的记忆也不会迅速遗忘, 因此记忆效应^[6-7]是研究复杂系统博弈模型最优决策时一种很常见的条件。基于记忆的策略更新机制^[8]在进化博弈模型^[9]中普遍存在, 玩家在决策时会对他们先前的记忆和收益的累积效应进行比较。囚徒困境是博弈论中一种常见并具有代表性的例子, 按照纳什非合作博弈理论^[10], 博弈各方陷入囚徒困境很大程度是由于其从个体理性出发, 用最大最小化准则^[11]采取战略。合作博弈^[12-13]和非合作博弈的区别就是集体理性和个体理性的区别。在合作博弈中允许参与者互相协调, 结盟以提高自身利益。本论文与合作博弈相关之处是博弈的参与者会通过相互协调以提高群体的合作水平。关于博弈论的研究, 方法有很多, 比如, 2020年全吉在自愿参与机制^[14]下, 采用多维马尔科夫过程对随机演化公共物品博弈中系统的状态进行建模, 利用马氏过程的极限分布研究了随机系统的均衡和群体合作行为; 2021年饶卫振等研究了合作博弈中的协作配送的问题, 以核仁解为基础, 提出奇偶循环线性规划模型, 得到子联盟满意度均衡程度更高的联盟均衡解^[15]。Leng等利用合作博弈论研究了边际贡献^[16]几乎递减条件下的多人分配问题。2022年Nip^[17]等研究了马尔可夫链选择模型下的两种分类博弈——竞争博弈和合作博弈。而关于研究如何用博弈论去促进合作这个方向, 大多使用改善网络或引入奖惩机制, 以及记忆机制等动力学机制的方法。本文采用蒙特卡洛方法进行实验仿真并分析不同观察机制对囚徒困境博弈合作水平的影响。

在现实生活中, 单个个体之间的互动中存在着一定的协调性^[18], 协调性能够把“都具有的知识”转变为“共同的知识”^[19], 这在很大程度上降低了实际中陷入囚徒困境的概率。因此, 在博弈过程中增加双方的协调性可以让理论研究更贴近于现实^[20-21], 在这一点上, 引入观察机制被认为是一种可行的方法。观察机制为增加个体之间的协调性提供了可能。个体之间协调需要时间条件, 观察机制可以提供这个时间条件。观察机制给博弈各方观察时间, 便于博弈各方多次互动观察来实现协调性的增长。

目前, 已有的研究^[22-25]表明, 引入记忆机制能够在一定程度上有效促进合作水平的提高。同样地, 在现实社会中, 除了记忆力的差异, 每个人的观察力也是各不一样的, 观察力同样对决策有着不可或缺的作用, 最终导致选择策略上的差异。记忆机制假设了每个个体的记忆长度有限, 智能体能记住从上一时刻到这一时刻的历史最佳策略, 然后根据自身的历史记忆进行决策。观察机制让每个个体在决策之前有一段观察期去通过自身的观察力进行观察交互审议获得平均收益, 然后根据观察期间获得的平均收益进行决策。记忆机制虽然考虑到了人的有限记忆的特性, 并且历史记忆的效应经常会在决策中扮演一种重要的角色, 但是其没有考虑到现实中还有除了记忆力还有行为人的其他能力会对网络中合作者的密度产生影响。其次, 当历史记忆长度过长时, 会造成冗长的记忆存在维度爆炸并且计算复杂, 从而导致决策困难。有时即时性的决策还存在考虑不够全面并且有误的问题, 这些都需要从另一个角度考虑影响决策的因素。因此, 我们引入观察机制, 能够从审议到执行, 大大减少错误决策。引入观察机制之后与原始模型相比, 群体中合作水平有很大的提高。之前研究的智能体虽然能记得在之前的博弈中的累积收益和参赛者的策略, 但关于策略更新时间研究^[26-28]只集中在博弈中每一轮的收益。引入观察机制的目的是模拟合作行为更接近现实世界。观察机制由观察时间和观察力两大要素组成。观察机制, 即假定具有不同观察力的玩家在博弈结束后不会立刻更新其策略, 而是会先经过一段观察时间进行观察审议, 后根据在此观察期间获得的平均累积收益情况审议后再做出决策, 与此同时, 由于时间是稀缺资源, 故在观察期间产生的时间成本也在考虑之内。这种通过考虑具有不同观察力和观察时间的玩家所存在的异质性来进行异步更新的方法有助于演化博弈合作机制理论的进一步完善, 也为提高囚徒困境博弈合作水平提供了一种新的视角。

1 模型设计

1.1 基于观察机制的博弈模型

在八邻居制的 $L \times L$ 方格网络中, 每个个体都有一个智能体^[29], 每个个体都将与一个随机选择的邻居包括自己进行囚徒困境博弈。假设每个玩家在自愿完成囚徒困境博弈后算一个时间步骤。每个玩家只能遵循两种策略^[30]: 背叛和合作, 在统一合作中合作者收益回报是1, 在统一背叛中背叛者的收益

回报是0。对于一个玩家选择合作而另一个玩家选择背叛的情况,合作者和背叛者的收益分别是0和 b 。在本文中,我们要求背叛者获利 b 的取值为 $1 < b < 2$ 。在之前,相同的参数及取值^[31]已经被多次采用,并且人们普遍认为这能够体现囚徒困境博弈的所有相关方面^[32]。智能体的收益矩阵为:

$$A = \begin{bmatrix} 1 & 0 \\ b & 0 \end{bmatrix}. \quad (1)$$

在引入观察机制情景下,初始化时对于每个智能体赋予不同的观察力参数 u ,设定参数 u 大于零且采用不规则随机分布。在现实中,个体的决策随着观察力的变化而不断变化,而观察力则随着个体经过观察行为后所获收益的变化而变化。在本研究中,假设所有个体都具有不同的初始观察力值。智能体 i 在每个策略更新前一轮的收益记为 $Y_i(t)$,在完成每个策略更新后的下一轮博弈中智能体 i 获得收益记为 $Y_i(t+1)$ 。智能体观察力的改变规则依据每个智能体在观察时间结束后即经过策略更新后再经过一轮博弈所获的收益 $Y_i(t+1)$ 和策略更新前的智能体收益 $Y_i(t)$ 。若个体观察后更新所获得的收益 $Y_i(t+1)$ 大于个体策略更新之前的收益 $Y_i(t)$,我们判定该观察行为有效,则该个体的观察力改变;反之,智能体的观察行为并未达到预期效果,则智能体的观察力保持不变。可以描述如下:

$$\begin{cases} u_i = u_i + \Delta, Y_i(t+1) > Y_i(t), \\ u_i = u_i, Y_i(t+1) \leq Y_i(t). \end{cases} \quad (2)$$

其中,参数 $\Delta=1$ 表示智能体观察力的单位增量。

每个玩家自愿完成博弈后都会获得收益,并被其他玩家选择继续进行博弈。每个玩家在每个时间步都可以进行多次博弈,玩家的虚拟累计收益是所有交互的总和。假设智能体 i 在观察时间 M_i 后更新其策略。在 M_i 个蒙特卡洛^[33]时间步长期间,智能体 i 的虚拟累积收益为 $P_i(M_i)$,可以描述如下:

$$P_i(M_i) = \sum_{t=0}^{M_i} Y_i[M_i(t)]. \quad (3)$$

智能体在更新下一个策略之前都有一个观察时间间隔,会根据其观察力 u 和观察时间损失成本 R 得到从上次更新时间到现在期间内观察到的总收益。假设观察期间内的时间损失成本为 $R(M)$,本文中所考虑的观察时间损失成本函数满足:

$$R(x) = \frac{1}{1 + e^{-x}}. \quad (4)$$

为了更好的满足网络中真实的博弈行为,本文提出了改进的费米更新函数。首先定义了智能体最

终观察后所得的平均收益 $K(M)$ 如下:

$$K_i(M_i) = \frac{u_i P_i(M_i) - R_i(M_i)}{Q_i(M_i)}; \quad (5)$$

$$K_j(\tau) = \frac{u_j P_j(\tau) - R_j(\tau)}{Q_j(\tau)}, \quad (6)$$

其中, u_i 表示智能体 i 的观察力, $P_i(M_i)$ 表示智能体 i 在 M_i 时间步长期间的虚拟累积收益, $R_i(M_i)$ 表示观察期间内的时间损失成本, $Q_i(M_i)$ 表示该智能体在 M_i 时间步长期间玩的博弈次数; u_j 表示智能体 j 的观察力, $P_j(\tau)$ 表示智能体 j 在上次更新其策略后在 τ 时间步中的虚拟累积收益, $R_j(\tau)$ 是智能体 j 上次更新其策略后在 τ 时间步长内的观察时间损失成本, τ 的值范围为 $1 \leq \tau \leq M_i$, $Q_j(\tau)$ 表示智能体 j 上次更新其策略后在 τ 时间步长内玩的博弈次数。式(5)代表智能体 i 在经过观察时间 M_i 后所得的平均收益,式(6)代表智能体 j 在经过观察时间 τ 后所得的平均收益。

其次将(5)、(6)代入得,在时间步长 t 处,该智能体 i 以相同的概率选择其邻居 j 之一,并且采用具有以下概率的邻居策略:

$$W_{i \rightarrow j} = \frac{1}{1 + e^{\frac{K_i(M_i) - K_j(\tau)}{k}}}, \quad (7)$$

其中, k 代表环境噪声,当 k 趋近于0表示群体中任意一个个体都是绝对理性的,也就是说,只要 $K_j(\tau)$ 大于 $K_i(M_i)$,个体 i 的策略将替换为它某个邻居 j 的策略。当 k 趋近于 ∞ 表示群体中所有个体将不受邻居收益和策略的影响,完全随机采取决策。因此尽管邻居 j 的收益可能低于个体 i 的收益,但是 i 的策略仍可能替换为 j 的策略,而这会导致更低的收益。当 k 发生变化时,结果的变化趋势不受影响^[34-35],因此在本文中将 k 设置为0.1。

考虑了智能体每次的观察时间 M_i ,以下是我们提出的四种不同的观察方式模型。

第一种“完全同步”模型,每个智能体同时开始进行观察并具有相同的观察时间 M_i 。即 $M_i \equiv M, i = 1, 2, \dots, L^2, M_i(0) = M$ 。

第二种“相对同步”模型,是每个智能体都有相同的观察时间, $M_i \equiv M_i, i = 1, 2, \dots, L^2$,但不同的智能体在不同时刻开始观察。 $M_i(0)$ 其值是任意的, $M_i(0) \in [1, M], i = 1, 2, \dots, L^2$ 。

第三种“高斯异步”模型让每个智能体的观察时间长度不同。由于自然界大多数变量的概率分布都大致遵循高斯分布^[36]。该模型使 M_i 的值遵循高斯分布且 $M_i \in [1, 2M - 1]$ 。

第四种“指数异步”模型,使 M_i 的值遵循指数

分布, 即 $M_i \sim E(1/M)$ 。指数分布是一种连续概率分布并具有无记忆性^[37], 用来表示事件发生的时间间隔具有一定的可靠性。

1.2 演化博弈过程

每轮博弈由以下四部分组成:

1) 每个智能体随机选择邻居包括自己;

2) 计算在观察期内被选择的智能体和自己的虚拟总收益, 将观察力与虚拟总收益结合后减去观察期内的时间损失成本, 最终得到智能体所观察到的总收益;

3) 根据改进的费米准则更新智能体的策略;

4) 比较更新策略后智能体的收益有无增加。

2 实验仿真与讨论

本文利用Pycharm软件在规则网络上进行蒙特卡洛仿真模拟演化博弈过程。实验中取 L 的值为100, 合作者和叛逃者都随机分布在正方形格子上, 则初始合作率为 $P_c(0) = 0.5$ 。为了获得稳定的结果, 每次实验均由 1×10^5 个蒙特卡洛时间步组成, P'_c 是由接下来1000个蒙特卡洛时间步中 $P_c(t)$ 的平均值取的。最后, 重复十次, 得到 P'_c 的平均值作为最终稳定值 P_c 。

2.1 “完全同步”模型

在“完全同步”模型下, M_i 是一个常数, 每个玩家开始同时进行观察后更新其策略。对于 $b=1.3$ 、 $b=1.5$ 、 $b=1.7$, 图1详细显示了 P_c 如何随 M 的变化而变化。当 $b=1.3$ 和 $b=1.5$ 时, 随着 M 值的增大, 合作方的比例增大。在 M 的值达到某个阈值后, 合作率几乎不再增加并保持不变或轻微下降。也就是说, 当玩家的思考观察时间增加时, 会对合作产生积极的影响。但是, 过长的观察时间会产生很大的观察期时间损失成本并且有观察力的限制就不会产生更积极的影响。联系实际生活中, 若在认真观察后如果还是一味地犹豫不决, 会导致结果并不好。无论 M 的取值如何, 当 b 的取值达到1.7时, 网格中几乎每个智能体都选择了背叛, 这是因为当 b 的取值过大时, 背叛是每个智能体的最佳选择。

从图1(a)(b)可以看出 b 值对 P_c 的影响。当 $M=1$ 时, 一方面大多数人都想在没有观察机制的情况下研究背叛者的策略。另一方面, 随着背叛诱惑的存在, 背叛者的平均报酬普遍高于博弈开始时的合作者, 那时背叛是每个智能体的最佳选择。当 M 的值从1增加到2时, 由于引入观察机制, P_c 突然跳变。当 $M>1$ 且 b 为固定值时, 合作者的比例随 M 的增大

而增大, 但增大的幅度会越来越小。另一个结果是, 当 $M>1$ 且 $b<1.3$ 或者 $b>1.7$ 时, M 的取值对合作者比例影响不大。当 $1.3<b<1.7$ 时, 随着 M 值的增大, 合作者的比例增大。当 $M>1$, 观察机制被引入模型时, 一些个体会在高峰期保持原有的策略来研究背叛者, 这为合作率的提升奠定了基础。同时智能体在观察期间内观察到的累积平均收益的影响更有效地阻碍了背叛的传播, 从而在很大程度上提高了最终合作率。

2.2 “相对同步”模型

此时, 每个智能体都具有相同的观察时间间隔长度 M_i , 任意值 $M_i(0)$ 表示每个智能体开始观察的时间, $M_i(0) \in [1, M]$ 。智能体以 M_i 为周期周期性观察并进行更新策略, 即初始观察时间上引入异质性。在这种情况下, 随着时间的变化, 合作者的比例是稳定的。图1(c)表示当 $b=1.3$, $b=1.5$, $b=1.7$, 且当 $M \in [1, 100]$ 时, P_c 随 M 的变化情况。图1(d)展示了 $M=1, 2, 3, \dots, 10, 100$ 时, P_c 随 b 的变化情况且 $b \in [1, 2]$ 。

当 $M=1$ 时模型二策略更新机制与模型一中策略更新机制相同。图1(b)和图1(d)证实了这一点, 如果每个智能体开始观察时刻不同, 则 $M=1$ 的结果与模型一中出现的结果是一致的。需要注意的是当 M 从1增加到2时, P_c 的值变化很大, 当 M 从2增加到100时, 的值开始稳定变化。图1(d)也表明在 $M=1$ 的情况下, 合作者的比例仍然处于非常低的水平, 即使 b 的值非常小, 大多数人还是宁愿背叛; 当 $b>1.3$ 时, 每个智能体无一例外都选择背叛。对于 $b=1.7$, 图1(b)与图1(d)相比, 图1(b)显示无论 M 的值如何变化, 选择合作的智能体寥寥无几, 图1(d)表明当 $M>3$ 时, 总是会有一些合作者。当 $1.3<b<1.7$ 时, 观察时间间隔 M 值对合作的促进作用在模型一中相对较大。而在模型二中, $1.3<b<2.0$ 时, 随着 M 值的增加, 这种效应一直存在, 且比模型一更为明显。这也说明了智能体之间靠不同时刻开始观察产生的这种异质性可以促进合作。

2.3 “高斯异步”模型

每个智能体更新策略的时间步长为 M_i , M_i 值呈高斯分布。假设 $M_i \sim N(M, \sigma^2)$ 且 $M_i \in [1, 2M - 1]$ 。每个智能体在策略更新的观察时间间隔长度上都是异构的。取 $\sigma=1$, 当时间变化时, 合作者的比例保持稳定。

图2展示了 M 、 b 、 σ 对合作率 P_c 稳定值的影响。

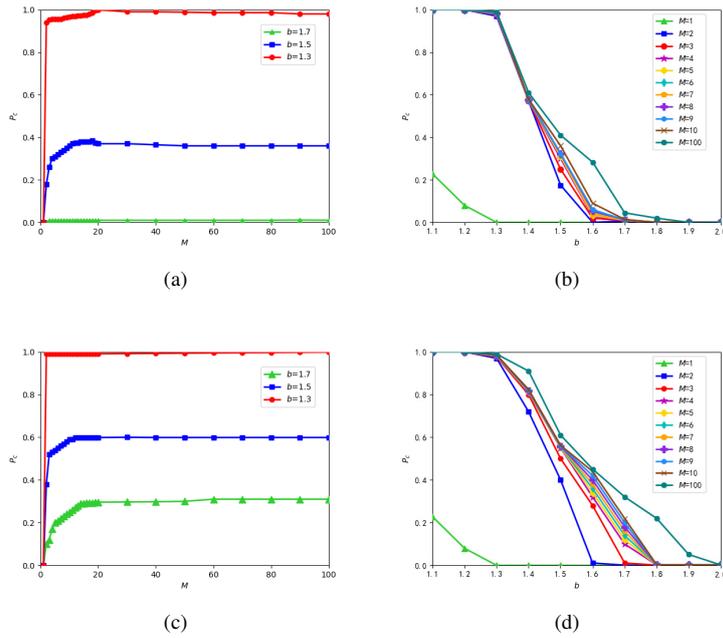


图 1 完全同步模型和相对同步模型中观察时间对合作率的影响

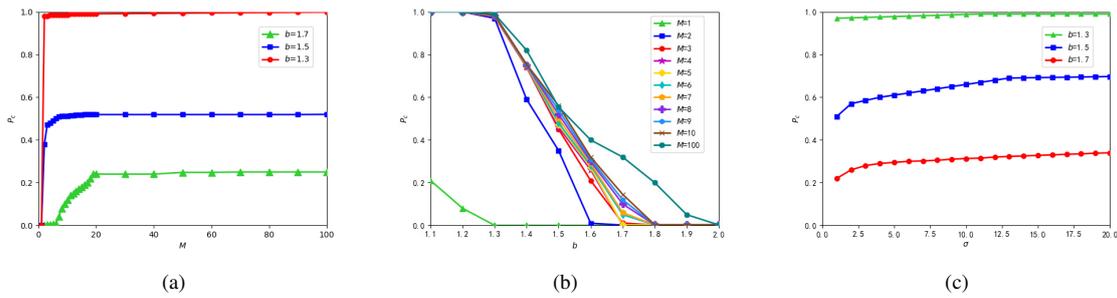


图 2 高斯异步模型中观察时间和参数σ对合作率的影响

从图2(a)(b)可以看出，在模型三中，随着 M 的增加，合作值显著增加。在 $\sigma=1$ 下，合作者的比例与模型二几乎相同。当 $M=1$ 时，模型三没有引入观察机制，所以此时实验结果与之前模型并无差异。当 $1.3 < b < 2.0$ 时，随着观察时间间隔 M 值的增加，合作者的比例也会有增加。对于策略更新前有观察时间间隔($M \neq 1$)的群体，当 $\sigma=1$ 时，观察时间间隔 M 的增加会对 P_c 的值产生的影响稍逊于模型二的观察方式下所产生的影响。

进一步研究了标准差 σ 对 P_c 值的影响。当 $M=100$ 时， $\sigma \in [1, 20]$ ， P_c 作为 σ 函数的稳定值如图2(c)所示。图2(c)表明，当 b 和 M 都在固定值时，合作者的比例随 σ 的增加而增加。简而言之，每个智能体在更新策略的观察时间间隔差越大，群体中合作者的比例越高。从这个角度出发，我们证明了智能体间策略更新时间长度的异质性对合作有正向的影响。

2.4 “指数异步”模型

指数异步模型的观察方式与上一种模型不同。每个智能体更新策略的时间步长为 M_i ， M_i 值呈指数分布。假设 $M_i \sim E(1/M)$ ，每个智能体在策略更新的观察时间间隔上是无记忆性的。图3展示了 M 和 b 对合作率 P_c 稳定值的影响。 $M=1$ 的结果与其他模型中出现的结果是一致的。从图3可以看出，在模型四中，随着观察时间间隔 M 值的增加，合作者比例显著增加，直到 $M=18$ 时，不同诱惑值 b 下的合作者比例达到峰值。当 $M > 18$ 时，群体中的合作者比例保持稳定不变或轻微下降。

当 $b=2$ 时，图3(b)显示无论 M 的值如何变化，没有智能体选择合作，图3(b)表明当 $M > 1$ 时，在 $1.1 < b < 2.0$ 条件下群体中的合作者总是存在的。当 $1.3 < b < 2.0$ 时，随着观察时间间隔 M 值增加会促进合作涌现，这种效应一直存在且比其他模型更为明显。这一点说明了智能体之间用指数异步观察方式比相对同步观察方式对合作有更好的促进作用。

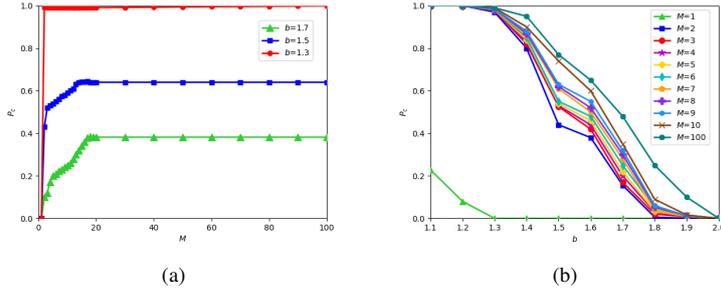


图3 指数异步模型中观察时间对合作率的影响

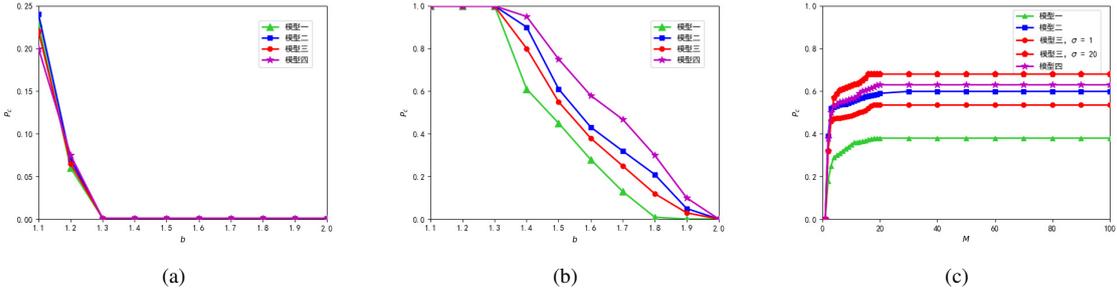


图4 四种观察机制模型对合作水平提高的效果对比

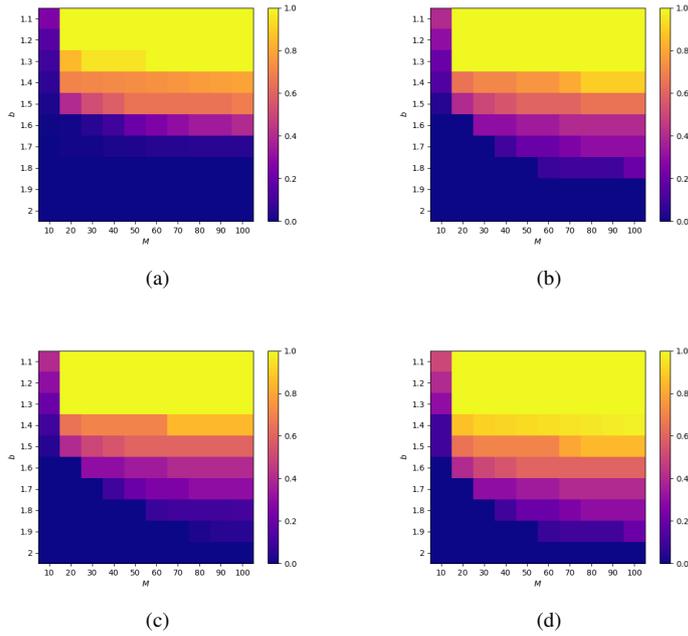


图5 各模型合作率在 (b, M) 坐标分布

2.5 四种模型对比验证

为了更好地反映各个模型的观察方式对群体合作者比例的影响，我们比较了四种模型的结果。在高斯异步模型中分别取 $\sigma=1, M=1$ 和 $\sigma=1, M=100$ ，图4(a)(b)显示了这四种模型下 P_c 随 b 的变化情况。

由图4(a)可知，当 $M=1$ 时，四种模型的情况基本相同，这是因为策略更新没有经过观察思考的时间，即不存在观察机制。在这种情况下，没有时间间隔的观察方式对结果没有影响。在图4(b)中，当 $M=100, b < 1.3$ 时所有智能体都选择合

作。在 $b > 1.3$ 状态下，每种情况下的背叛诱惑对合作有不同程度的影响。只有当我们改变智能体初始观察时刻或观察时间时，我们才能发现与模型一相比，合作者的比例有很大的增加。由图4(c)可知，四种模型存在的观察时间间隔阈值为18。当 $M > 18$ 时，合作水平将不会再有很大提升。

为了清楚地理解 M 和 b 的作用，对于四种模型我们都提出了一个 (b, M) 坐标平面热力图如图5所示。通过热力图可以发现本文讨论的四种模型中， P_c 的值都是随着观察时间间隔 M 的增加而增加，并

随着 b 值的减小而增加。从图4(b)可以看出, 模型一的合作率低于本文讨论的所有其他模型, 并且模型二、四和模型三的合作率提升的效果是与 σ 相关的。当 $\sigma=1$ 时, 模型二、四的合作率高于模型三的合作率。而当 $\sigma=20$ 时, 模型三有较高的合作率。进一步表明, 个体间异质性越大, 其效应越大。相对同步模型和指数异步模型都从初始观察时刻和观察时间间隔的不同方面考虑个体间的异质性。从这些结果来看, 异质性也可以促进合作, 但其影响程度取决于参数的取值。

3 总结与展望

现实中个体先观察后决策是普遍存在的, 本文定义了四种不同观察机制并研究其对囚徒困境博弈合作水平的影响。在本文讨论的四种模型下, 适当增加观察时间对促进合作都有积极的影响, 但这些积极影响在不同的情境中稍有差异。由于本文考虑了观察期间的的时间损失, 一旦观察时间的长度达到一个阈值, 进一步增加将不再有积极影响。影响的程度取决于每个参数的取值, 准确地说, 是本文中服从高斯分布的均值方差和观察时间。引入观察机制能让由异构智能体比同质智能体组成的群体具有更高的合作水平。个体间的异质性在模型二、模型三和模型四中都能得到体现, 且其都促进了合作。实验结果表明异步更新比同步更新更容易促进合作行为的涌现。

研究发现, 引入观察机制对个体的学习起一定的促进作用。随着观察时间变长, 观察期的时间损失成本会越来越大, 并干扰观察信息的有效传播, 因此存在一个最佳观察时间阈值。现实生活中人们遇到事情时, 深思熟虑做决策是正确的。但在深思熟虑之后若还是一味的焦虑不果断并且反复思虑忧愁, 就会引起反应迟钝, 记忆力下降等不好的情况。这与天生的性格, 家庭环境教育环境等因素也是息息相关的。这些分析结论对博弈论应用到当前社会大环境具有一定的理论指导和实践的价值。正如政府机构的审议机制一样, 漫长的观察、评估、审议可以让立法更加权威, 政策更加利民。

本文存在一定的局限性。例如现实中观察期长短是根据每个人的观察力不同发生不规律的变化, 进一步可以研究建立观察时间间隔在进化过程中跟随每个智能体自身的观察力发生变化的进化博弈模型。除此之外, 本文只考虑了观察期的时间损失成本, 实际上还会存在多种其他不可预测的额外损失等因素会影响博弈模型改变。这些不足有待今

后的研究改进。

参考文献(References)

- [1] Smith J M. Evolution and the theory of games[M]. Cambridge University Press, 1982.
- [2] Cheng D, He F, Qi H. Modeling, analysis and control of networked evolutionary games[J]. IEEE Transactions on Automatic Control, 2015, 60(9): 2402-2415.
- [3] Axelrod R, Hamilton W D. The evolution of cooperation[J]. Science, 1981, 211(4489): 1390-1396.
- [4] 章超,徐名海. 基于Agent网络化博弈的建模与分析[J]. 计算机应用与软件, 2022, 39(01): 301-307.
(Zhang C, Xu M H. Modeling and analysis based on agent networked game[J]. Computer Applications and Software, 2022, 39(01): 301-307.)
- [5] 王先甲. 复杂网络上的演化博弈及其学习机制与演化动态综述[J]. 阅江学刊, 2021, 13(03): 70-84.
(Wang X J. Evolutionary game on complex networks, its learning mechanism and evolutionary dynamics review[J]. Yuejiang Journal, 2021, 13(03): 70-84.)
- [6] Wang W X, Ren J, Chen G, et al. Memory-based snowdrift game on networks[J]. Physical Review E, 2006, 74(5): 056- 113.
- [7] Challet D, Zhang Y C. Emergence of cooperation and organization in an evolutionary game[J]. Physica A: Statistical Mechanics and its Applications, 1997, 246(3-4): 407-418.
- [8] 谢逢洁. 复杂网络上博弈行为演化的合作激励[J]. 上海交通大学学报, 2015, 49(8): 1256-1262.
(Xie F J. Cooperative incentives for the evolution of game behavior on complex networks[J]. Journal of Shanghai Jiaotong University, 2015, 49(8): 1256-1262.)
- [9] Wang S Y, Liu Y P, Zang F, et al. Super-rational aspiration induced strategy updating promotes cooperation in the asymmetric prisoner's dilemma game[J]. Applied Mathematics and Computation, 2021, 403(9): 126-135.
- [10] Elsadany A A. Dynamics of a cournot duopoly game with bounded rationality based on relative profit maximization[J]. Applied Mathematics and Computation, 2017, 294(12): 253-263.
- [11] 全吉, 周亚文, 王先甲. 社会困境博弈中群体合作行为演化研究综述[J]. 复杂系统与复杂性科学, 2020, 17(1): 1-14.
(Quan J, Zhou Y W, Wang X J. A review of research on the evolution of group cooperative behavior in social dilemma game[J]. Complex Systems and Complexity Science, 2020, 17(1): 1-14.)
- [12] 张明, 朱建军. 考虑供应商技术截断的“主-供”合作机制演化博弈分析[J]. 控制与决策, 2021, 36(10): 6.
(Zhang M, Zhu J J. Evolutionary game analysis of "main manufacturer-supplier" collaboration mechanism considering supplier's technology truncation[J]. Control and Decision, 2021, 36(10): 6.)
- [13] 王治莹, 赵宏丽, 聂慧芳. 考虑公众预期后悔的突发

- 性抢购事件演化博弈分析[J]. 系统工程学报, 2022, 37(1): 11.
(Wang Z Y, Zhao, H L, Nie H F. Evolutionary game analysis of sudden panic buying events considering the anticipated regret of public[J]. Journal of Systems Engineering, 2022, 37(1): 11.)
- [14] 全吉, 储育青, 王先甲. 自愿参与机制下的公共物品博弈与合作演化[J]. 系统工程学报, 2020, 158(02): 188-200.
(Quan J, Chu Y Q, Wang X J. Public goods game under voluntary participation mechanism and the evolution of cooperation[J]. Journal of Systems Engineering, 2020, 158(02): 188-200.)
- [15] 饶卫振, 段忠菲, 朱庆华. 一种均衡协作配送子联盟满意度的成本分摊方法[J]. 系统工程学报, 2021, 166(04): 476-494.
(Rao W Z, Duan Z F, Zhu Q H. A cost allocation method for balancing the satisfaction of cooperative distribution sub-alliances[J]. Journal of Systems Engineering, 2021, 166(04): 476-494.)
- [16] Leng M, Luo C, Liang L. Multiplayer allocations in the presence of diminishing marginal contributions: Cooperative game analysis and applications in management science[J]. Management Science, 2021, 67(5): 2891-2903.
- [17] Nip K, Wang C, Wang Z. Competitive and cooperative assortment games under Markov chain choice model[J]. Production and Operations Management, 2022, 31(3): 1033-1051.
- [18] Jian Q, Li X, Wang J, et al. Impact of reputation assortment on tag-mediated altruistic behaviors in the spatial lattice[J]. Applied Mathematics and Computation, 2021, 396: 125928.
- [19] Li X, Sun S, Xia C. Reputation-based adaptive adjustment of link weight among individuals promotes the cooperation in spatial social dilemmas[J]. Applied Mathematics and Computation, 2019, 361: 810-820.
- [20] Shi L, Shen C, Geng Y. Winner-weaken-loser-strengthen rule leads to optimally cooperative interdependent networks[J]. Nonlinear Dynamics, 2019, 96(1): 49-56.
- [21] Shu F, Liu X, Li M. Impacts of memory on a regular lattice for different population sizes with asynchronous update in spatial snowdrift game[J]. Physics Letters A, 2018, 382(20): 1317-1323.
- [22] 邓云生, 杨洪勇. 基于囚徒困境策略的改进HK网络上的合作博弈[J]. 智能系统学报, 2018, 13(03): 479-485.
(Deng Y S, Yang H Y. Improved cooperative game on HK network based on prisoner's dilemma strategy[J]. Journal of Intelligent Systems, 2018, 13(03): 479-485.)
- [23] Sun P, Zhang J. Group decision under uncertain information[J]. Chaos, Solitons & Fractals, 2020, 131(20): 109-131.
- [24] Hu X, Wu M. The analysis of evolutionary prisoner's dilemma game based on weighting effect[J]. International Journal of Modern Physics C, 2021, 32(05): 215-226.
- [25] Danku Z, Perc M, Szolnoki A. Knowing the past improves cooperation in the future[J]. Scientific Reports, 2019, 9(1): 1-9.
- [26] Shu F, Liu X, Fang K, et al. Memory-based snowdrift game on a square lattice[J]. Physica A: Statistical Mechanics and its Applications, 2018, 496(14): 15-26.
- [27] 张建磊, 李智琦, 曹明. 多智能体系统中基于演化博弈的群体状态控制[J]. 控制理论与应用, 2018, 35(05): 601-609.
(Zhang J L, Li Z Q, C M. Group state control based on evolutionary game in multi-agent systems[J]. Control Theory & Applications, 2018, 35(05): 601-609.)
- [28] 邢海云, 赵建立. 变异机制在网络演化博弈中的应用[J]. 山东大学学报(理学版), 2016, 51(12): 103-107.
(Xing H Y, Zhao J L. Application of variation mechanism in network evolution game[J]. Journal of Shandong University (Science Edition), 2016, 51(12): 103-107.)
- [29] Perc M, Jordan J J, RAND D G, et al. Statistical physics of human cooperation[J]. Physics Reports, 2017, 687(2): 1-51.
- [30] Lozano S, Moreno P, Adenso-Diaz B, et al. Cooperative game theory approach to allocating benefits of horizontal cooperation. European Journal of Operational Research, 2013, 229(2): 444-452.
- [31] Li F, Cao X, Sheng P. Impact of pollution-related punitive measures on the adoption of cleaner production technology: Simulation based on an evolutionary game model[J]. Journal of Cleaner Production, 2022, 33(9): 130-143.
- [32] Nowak M A, May R M. Evolutionary games and spatial chaos[J]. Nature, 1992, 359(6398): 826-829.
- [33] Wang M, Pan Q, He M. Individuals with the firm heart are conducive to cooperation in social dilemma[J]. Chaos, Solitons & Fractals, 2020, 137(11): 109-129.
- [34] Dong Y, Xu H, Fan S. Memory-based stag hunt game on regular lattices[J]. Physica A: Statistical Mechanics and its Applications, 2019, 519: 247-255.
- [35] Szabó G, Vukov J, Szolnoki A. Phase diagrams for an evolutionary prisoner's dilemma game on two-dimensional lattices[J]. Physical Review E, 2005, 72(4): 047107.
- [36] Zou Q, Hu K. Heterogeneous aspiration resolves social dilemma in structured populations[J]. Chaos, Solitons & Fractals, 2020, 134(16): 109-115.
- [37] Akhter Z, M S, Ormoz E. On the order statistics of exponentiated moment exponential distribution and associated inference[J]. Journal of Statistical Computation and Simulation, 2022, 92(6): 1322-1346.

作者简介

蒋贤惠(1997—), 女, 硕士研究生, 从事复杂网络演化博弈论、多智能体系统的研究, E-mail: jxh971028@163.com;
蒋玉莲(1977—), 女, 副教授, 硕士生导师, 从事智能控制、博弈、深度学习的研究, E-mail: jyl-ee@swun.edu.cn;
黄勤珍(1965—), 女, 教授, 硕士生导师, 从事智能控制、智能信息处理的研究, E-mail: 13548002166@139.com.