

## 囚徒困境中贴标签控制的促合作设计

董 瑞<sup>1,2†</sup>, 陈 琳<sup>2</sup>, 王先甲<sup>2</sup>

(1. 河南科技学院 数学科学学院, 河南 新乡 453003; 2. 武汉大学 系统工程研究所, 湖北 武汉 430072)

**摘要:** 研究囚徒困境中促进合作演化的控制设计及控制下的合作演化路径问题. 基于标签捐赠机制, 设计对个体贴标签的规则, 增加个体的认知能力, 协助个体决策. 建立反馈控制下的合作演化动态模型, 分析贴标签控制促进合作演化的实现途径, 论证容许控制上界与控制效果之间的一致性. 通过基于合作识别和基于背叛识别两种控制的比较, 得到两个结论. 一是合作率较低时, 基于背叛识别的控制更容易促进合作率提升, 反之, 基于合作识别的控制则更容易实现. 二是这两种识别方式需要相同的博弈重复次数. 仿真分析各参数和控制上界对控制律设计和合作演化轨迹的影响.

**关键词:** 囚徒困境; 演化稳定策略; 机制设计; 促合作控制; 基于标签捐赠

**引用格式:** 董瑞, 陈琳, 王先甲. 囚徒困境中贴标签控制的促合作设计. 控制理论与应用, 2019, 36(7): 1104 – 1112

DOI: 10.7641/CTA.2018.80166

## Promoting cooperation design of the tag-based donation control in prisoner's dilemma

DONG Rui<sup>1,2†</sup>, CHEN Lin<sup>2</sup>, WANG Xian-jia<sup>2</sup>

(1. School of Mathematics Sciences, Henan Institute of Science and Technology, Xinxiang Henan 453003, China;

2. Institute of Systems Engineering, Wuhan University, Wuhan Hubei 430072, China)

**Abstract:** The problem of designing a control to promote the evolution of cooperation in the “prisoner's dilemma” has been investigated and the evolutionary path has been discussed. Design the rules for labeling players as cooperator or defector based on the “tag-based donation” mechanism to improve players' cognitive ability and help them decision. The replicator dynamics restrained by the feedback control to promoting cooperation has been modeled and the approach of tag-based control to promote the evolution of cooperation has been analyzed. The consistency between the upper bound of the admissible control and the control effect is demonstrated. By comparing the difference between the controls based on recognizing cooperation and based on recognizing defection, the following two results have been obtained. One result is that the control based on recognizing defection is more implementable if the degree of cooperation is low and vice versa. The other result is that the two controls demand same repeated encounter times. The simulations are employed to analyze the effects on the controller and trajectories of the replicator dynamics caused by the parameters of the games and the upper bound of the admissible control.

**Key words:** prisoner's dilemma; evolutionarily stable strategy; mechanisms-design; promoting cooperation control; tag-based donation

**Citation:** DONG Rui, CHEN Lin, WANG Xianjia. Promoting cooperation design of the tag-based donation control in prisoner's dilemma. *Control Theory & Applications*, 2019, 36(7): 1104 – 1112

### 1 引言

合作困境反映个人理性与集体理性的矛盾, 在演化博弈框架下, 背叛将战胜合作, 造成的结果是群体平均适应度持续下降, 直至合作消亡. 但是, 生物学中的合作是伴随着生命出现而出现的, 合作现象广泛存在. 这种矛盾引发了对支持合作的特定机制的研究,

在《科学》杂志上被列为全世界科学家提出的125个科学挑战中的25个核心问题之一<sup>[1]</sup>.

Nowak<sup>[2]</sup>指出, 直接互惠<sup>[3-4]</sup>、间接互惠<sup>[5-6]</sup>、亲缘选择<sup>[7-9]</sup>等机制都可以促使自私的个体放弃自己的利益而采取合作. 由Hamilton原理<sup>[7]</sup>, 基于基因相似性的“亲缘”可以支持合作<sup>[8-9]</sup>. 在生物界, 一些特征可

收稿日期: 2018-03-11; 录用日期: 2018-08-18.

†通信作者. E-mail: drui@163.com; Tel.: +86 373-3040087.

本文责任编辑: 王龙.

国家自然科学基金项目(61773156), 河南省高校重点科研计划项目(19A120006)资助.

Supported by the National Natural Science Foundation of China (61773156) and the Key Scientific Research Project in Universities of Henan Province (19A120006).

以反映亲缘关系,有些弱的甚至具有潜在欺骗性的亲缘关系以“标签”形式呈现,如“绿胡须”(可继承、可观察、最初可能具有任意性的某种特征),当个体更倾向于跟自己的某些特征相似的个体合作时,“基于标签捐赠”或“绿胡须效应”也能支持合作的进化<sup>[10]</sup>. Dawkins演示了“基于标签捐赠”在没有家族亲缘关系的个体间是如何促进合作的<sup>[11]</sup>. 近几年,“基于标签捐赠”依然是研究热点<sup>[12-14]</sup>. 王龙等研究了“群内有爱,群外仇恨”的区域利他主义的形成及其对合作的促进<sup>[15]</sup>.

个体之间发生作用的频率依赖于其策略的现象称为个体间的一种非均匀连接现象<sup>[16]</sup>,文献[16-17]建立了囚徒困境演化博弈的非均匀连接模型,并解析分析了非均匀连接对支持合作的作用.“亲缘选择”机制和“基于标签捐赠”机制都是通过增加合作与合作之间相互作用的概率来促进合作的演化.在重复囚徒困境博弈中,存在支持合作演化的“直接互惠”机制<sup>[3-4]</sup>.与非均匀连接相比,“直接互惠”机制是通过个体学习实现合作促进的,条件是博弈以较大的概率重复进行.但是,在人类社会的诸多博弈中,两个特定个体的接触是转瞬即逝的,再次相遇并进行博弈的概率很低.当两个个体的博弈行为可以被群体中的其他个体观察到并在种群中进行宣传时,“声誉”才可能通过非直接互惠机制促进合作<sup>[5-6]</sup>.与声誉机制相比,“标签”机制可以起到与“声誉”机制相似的促进合作演化的作用,需要的认知要简单的多.在生物群体中,这种“标签”可能是颜色、形状、气味等等;对于人类社会,“标签”可能是文化、信仰、语言、肤色等.

将群体中的个体视为智能体,把多智能体系统通过局部信息交换、个体自主决策,最终协同完成任务的过程可视为一个博弈的过程,各参与者的策略随时间演变可以达到均衡,实现合作<sup>[18-20]</sup>.由于难以获取个体和环境的全部信息,现有文献较少考虑群体之外的因素对博弈的反馈和控制.事实上,对于社会组织中的囚徒困境博弈,政府、行业协会等机构可以成为合作的推动者<sup>[21]</sup>.文献[22-23]对囚徒困境博弈给出了基于奖励-惩罚设计的促进合作的控制方法.

本文以经济社会中的某些群体(比如物流园区或双边平台)中的个体(企业或个人)间博弈为背景研究合作演化的促进机制与控制问题.结合基于标签捐赠机制、直接互惠机制及间接互惠机制的思想,研究囚徒困境博弈中促进合作演化的控制问题.在群体形成初期,由于彼此间缺乏了解、没有信任关系,个体决策时很少选择合作,如何保护和壮大有限的合作个体是群体初创期的核心问题;当群体合作率大幅提升之后,少量的不合作个体对整个群体的运行会造成极为严重的负面影响,如何驱逐不合作个体,尽早进入合作稳定状态是各个理性的利害相关者极为关注的问题.

本文以政府、行业协会、集群或平台的运营管理方等独立的第三方为控制者,通过对个体合作情况的监察,形成对群体信息的反馈,协助个体选择策略从而促进合作.

## 2 “贴标签”控制对博弈的影响

考察有限大群体 $\Omega$ 中的对称性两个体博弈,设每个个体的纯策略集为 $\Delta = \{C, D\}$ ,一次博弈的收益矩阵如下:

$$\begin{array}{c} C \\ D \end{array} \begin{array}{cc} C & D \\ \left( \begin{array}{cc} R, R & S, T \\ T, S & P, P \end{array} \right) \end{array} \quad (1)$$

如果式(1)中的收益满足如下条件:

$$T > R > P > S, \quad (2)$$

则称该博弈是囚徒困境.这里: $C$ 表示合作, $D$ 表示背叛.

本文假设博弈满足如下的条件:

$$2R > T + S. \quad (3)$$

在重复博弈中,如果式(3)不满足,则单纯合作的社会效率低于双方参与者约定轮流选择合作和背叛,促进合作演化没有意义.显然,背叛是囚徒困境博弈的占优策略,没有特殊机制作用时,合作将演化失败,合作既无法产生(合作无法小概率入侵背叛群体)也无法存续(小概率背叛策略可以入侵合作策略).

### 2.1 直接互惠机制下的博弈

在非重复博弈中,收益矩阵(1)描述的博弈有唯一的均衡( $D, D$ ),导致的收益是低效率的( $P, P$ ).在无限重复囚徒困境中,Trim-tigger(触发机制)中的个体的纯策略集为 $\Delta_1 = \{\text{GRIM}, \text{ALLD}\}$ ,其中:ALLD表示永远背叛,GRIM为:第1回合选择合作,而后只要对方不背叛就继续选择合作,一旦对方背叛,那么他将永远选择背叛.设 $\delta$ 为囚徒困境中参与人的贴现因子( $\delta$ 是下次博弈中一个单位收益的当前值),则当 $\delta > (T - R)/(T - P)$ 时,GRIM是子博弈精炼纳什均衡,合作可以成功演化<sup>[24]</sup>.在 $m$ 次重复囚徒困境博弈中,GRIM不再是子博弈精炼纳什均衡,但是当 $m > (T - P)/(R - P)$ 时,博弈具有2个ESS:GRIM和ALLD.从演化动力学角度分析,低频率的ALLD不能成功入侵GRIM,即这样的机制可以实现合作成功演化.Axelrod<sup>[3]</sup>指出,在重复囚徒困境中,一报还一报(tit for tat, TFT)可以促进合作成功演化.设每个个体的纯策略集为 $\Delta_2 = \{\text{TFT}, \text{ALLD}\}$ ,在无限重复博弈中,当博弈双方有一定的耐心时,为了避免后面的博弈会遭到对方的惩罚,博弈双方可能总是选择合作.TFT虽然不是子博弈精炼纳什均衡,但可以成为某些机制下的ESS.

在 $m$ 次重复囚徒困境中, TFT和ALLD博弈的收益矩阵与GRIM和ALLD博弈的收益矩阵相同, 由此, 本文不区分TFT和GRIM, 统称为 $C$ , 称ALLD为 $D$ . 相应地, 对任何时间 $t$ , 设合作者和背叛者在群体中的占比分别为 $x(t)$ 和 $y(t)$ , 相应的群体状态定义为 $z(t) = (x(t), y(t))$ . 由于 $x(t) + y(t) = 1$ , 下面用合作者占群体的比率 $x(t)$ 表示群体状态.

在大群体中, 如果个体之间的相遇是等概率随机的, 则博弈在两个固定个体之间重复进行的概率极低, 从而合作者几乎总是选择合作, 这时本文定义的合作者几乎等同于ALLC(总是合作). 在与 $D$ 的博弈中, ALLC将演化失败<sup>[3]</sup>. 本文设计贴标签控制, 以实现2个目的: 1) 促进合作的成功演化, 使得合作成为ESS; 2) 在控制者对群体合作状态不满意时, 提升合作者在群体中的比率.

## 2.2 贴标签控制下的博弈

在经济社会的某些企业集群中发生的囚徒困境博弈, 由于群体中的个体数量庞大, 个体之间的相互作用不固定, 中小企业的信誉形成与传播也很难自发进行. 在这种情况下, 集群中公正的第3方可以扮演企业行为信息的搜集、发布等角色, 以一种相对高效率的方式促进成员企业的相互识别. 比如物流园区中的园区经营者和平台推广者、政府派出的园区管委会或园区中的行业协会等都有可能发挥类似作用. 鉴于此, 本文假设集群中存在控制方, 在外部集中控制背景下分析集群成员合作演化过程, 并评价控制实施效果.

**定义1** 控制方: 指独立于博弈群体之外, 能够识别群体中的部分或全部个体在已经发生的博弈中的策略选择, 并通过将识别结果传递给群体中的所有个体影响其在后续博弈中的策略选择的人或组织.

**定义2** 贴标签控制: 控制者定期监察一次博弈中的个体策略选择并将监察结果作为个体分类标签贴给个体, 这种通过贴标签行为决定个体策略的控制称为贴标签控制.

这种对个体定期贴标签的控制造成博弈中的个体策略选择按阶段发生变化, 称这样一个阶段的博弈为一轮博弈, 下面给出“一轮博弈”的定义.

**定义3** 一轮博弈: 称从某次监察开始至本次监察结束的所有博弈为一轮博弈.

为了更严谨地分析控制策略对囚徒困境中合作演化所起的作用, 本文做以下假设.

**假设1** 博弈轮数假设: 博弈轮数无限, 每一轮博弈中的博弈次数为 $m + 1 (m \in N)$ 次.

**假设2** 复制动态假设: 群体状态的演化按复制子动态进行, 个体复制发生在当前一轮博弈后, 下一轮博弈开始前, 个体收益指一轮博弈的总收益, 不考

虑个体的迁入和迁出.

**假设3** 均匀连接假设: 每一次博弈中个体相遇是均匀的, 即博弈双方在群体中等概率随机配对, 两个个体连续相遇两次的概率为0.

**假设4** 贴标签假设: 在每一轮博弈中, 控制者对第1次博弈的过程和结果以一定的概率识别, 一个合作者被识别的概率为 $\alpha (0 \leq \alpha \leq 1)$ , 控制者在第2次博弈开始前给识别到的合作者贴上标签, 这种标签携带到本轮博弈的后 $m$ 次博弈中, 背叛者不会被贴上标签.

在一轮博弈中, 贴标签后群体 $\Omega$ 由2个子群组成: 贴上标签的个体组成的子群 $\Omega_1$ 和未贴上标签的个体组成的子群 $\Omega_2$ ,  $\Omega_1 \cup \Omega_2 = \Omega$ ,  $\Omega_1 \cap \Omega_2 = \emptyset$ , 这里 $\emptyset$ 表示空集. 一轮博弈中, 贴标签之后至本轮博弈结束,  $\Omega_i (i = 1, 2)$ 不变. 个体复制和下一轮博弈的贴标签行为导致 $\Omega_i (i = 1, 2)$ 发生改变. 按照复制规则, 合作者复制的子代依然是合作者, 背叛者复制的依然是背叛者, 与是否被识别无关.

用“ $Y$ ”表示“贴标签”, “ $N$ ”表示不“贴标签”. 贴标签完成了从博弈群体 $\Omega$ 到 $\{Y, N\}$ 的映射 $\sigma: \Omega \rightarrow \{Y, N\}$ . 当贴标签的概率 $\alpha = 1$ 时,  $\sigma(C) = Y$ ,  $\sigma(D) = N$ ; 当 $0 \leq \alpha < 1$ 时, 有

$$\begin{cases} p\{\sigma(C) = Y\} = \alpha, \\ p\{\sigma(C) = N\} = 1 - \alpha, \\ p\{\sigma(D) = N\} = 1, \end{cases} \quad (4)$$

这里 $p\{\cdot\}$ 表示事件 $\{\cdot\}$ 发生的概率.

**假设5** 贴标签下的策略选择: 在每轮博弈的第1次博弈中, 合作者选择合作, 背叛者选择背叛. 其后的 $m$ 次博弈中, 个体可以观察到自己是否被贴上标签, 也可以观察到其他个体是否被贴上标签, 当且仅当2个个体都来自子群 $\Omega_1$ 时, 他们选择合作, 其他情况下2个个体都选择背叛.

由假设3, 贴标签控制不影响2个个体发生博弈的概率, 但影响个体的策略选择, 具体分析见本文第3部分. 假设4和5是基于合作识别的贴标签规则, 本文第5部分将讨论基于背叛识别的贴标签规则.

由假设5, 在每轮博弈中, 从第2次博弈开始, 策略对 $(C, C)$ 和 $(D, D)$ 出现的频率分别为 $(\alpha x)^2$ 和 $1 - (\alpha x)^2$ , 策略对 $(C, D)$ 和 $(D, C)$ 不再出现. 由上面的讨论可知, 第 $i$ 轮第 $j$ 次博弈中各策略收益与总体状态 $x(t)$ 、贴标签后的 $\Omega_i (i = 1, 2)$ 及博弈次数 $j$ 有关, 与博弈轮数 $i$ 没有关系. 记一轮博弈中合作的总收益为 $w_C(x)$ , 背叛的总收益为 $w_D(x)$ , 则

$$w_C(x) = [R + \alpha^2 m(R - P)]x + mP + S(1 - x), \quad (5)$$

$$w_D(x) = Tx + mP + P(1 - x). \quad (6)$$

基于以上假定的贴标签控制原理如图1所示.

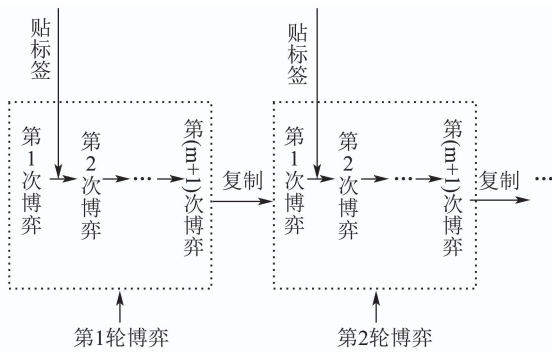


图 1 贴标签控制规则

Fig. 1 The principle of tag-based control

由假设3, 贴标签控制下的个体收益(5)–(6)相当于如下新一轮博弈的收益矩阵:

$$\begin{matrix} & C & D \\ \begin{matrix} C \\ D \end{matrix} & \begin{pmatrix} m\alpha^2(R - P) + R + mP & S + mP \\ T + mP & (m + 1)P \end{pmatrix} & \end{matrix} \quad (7)$$

实际上, 一轮博弈的后  $m$  次博弈中个体选择服从假设5规定的策略选择规则, 收益矩阵(7)中的收益  $a_{ij}$  不再是策略对  $(i, j)$  下的  $i$  策略收益<sup>[16–17]</sup>. 使用收益矩阵(7)的理由如下: 1) 合作与背叛在一轮博弈中的收益  $w_C(x)$  和  $w_D(x)$  对状态  $x(t)$  线性; 2) 在贴标签控制下, 一个合作个体如果进入一个全是合作者或者是背叛者的群体中, 其在一轮博弈中的期望收益为  $m\alpha^2(R - P) + R + mP$  或者  $S + mP$ . 收益  $T + mP$  和  $(m + 1)P$  可以做类似理解.

**命题 1** 若个体在每轮博弈中的带标签博弈次数  $m$ 、贴标签概率  $\alpha$  和博弈收益参数之间满足如下关系:

$$m\alpha^2 > \frac{T - R}{R - P}, \quad (8)$$

则贴标签控制下, 合作是 ESS.

**证** 记  $F(i, j)$  表示一个  $i$  个体进入一个全是  $j$  个体的群体中, 其在一轮博弈中的期望收益 ( $i, j = C, D$ ). 由收益矩阵(7),  $F(C, C) = m\alpha^2(R - P) + R + mP$ ,  $F(D, C) = T + mP$ . 条件(8)满足时,

$$F(C, C) > F(D, C).$$

由文献[25], 策略  $C$  是正 ESS, 从而是 ESS.

与重复博弈相比, 贴标签控制不要求博弈是两个个体固定配对重复进行, 而是指贴标签后的  $\Omega_i (i = 1, 2)$  不发生变化, 这里的  $\Omega_i (i = 1, 2)$  在一轮博弈中不发生变化包含3层含义: 1) 群体中的个体不发生死亡和复制; 2) 个体属性不发生变化, 从而总体状态不发生变化; 3) 群体的2个子群不发生变化. 所以, 本文考虑的贴标签下的重复博弈满足假设3.

### 3 贴标签控制对合作演化的促进途径

在假设3下, 无贴标签控制时, 大群体囚徒困境博弈中几乎不存在重复博弈, 从而合作者总是选择合作. 贴标签的概率为  $\alpha$  时, 一轮博弈的第1次博弈中, 策略对  $(C, C)$  和  $(D, D)$  出现的概率分别为  $x^2$  和  $(1 - x)^2$ , 策略对  $(C, D)$  和  $(D, C)$  出现的概率均为  $x(1 - x)$ . 从第2次博弈开始的后  $m$  次博弈中, 策略对  $(C, C)$ ,  $(D, D)$  出现的概率分别  $(\alpha x)^2$  和  $1 - (\alpha x)^2$ , 策略对  $(C, D)$  和  $(D, C)$  不再出现. 记  $t$  时刻合作对占总策略对的比率为  $c(t)$ , 则

$$c(t) = \frac{x^2(t) + m[\alpha x(t)]^2}{m + 1}. \quad (9)$$

由假设4,  $0 \leq \alpha < 1$  时,  $c(t) < x^2(t)$ , 从而贴标签控制下博弈中呈现的合作对占总策略对的比率下降, 从这个角度看, 贴标签不是促进合作, 而是抑制了合作. 下面从两个角度分析贴标签控制是如何通过影响博弈中不同策略的收益从而促进合作的.

记无贴标签控制时在后  $m$  次博弈中合作与背叛的收益分别为  $v'_{jC}(x)$  和  $v'_{jD}(x)$ .

**命题 2** 如果假设1–5成立, 则对于任意的贴标签概率  $\alpha (0 \leq \alpha \leq 1)$  和任意的总体状态  $x(t) \in (0, 1)$ , 在每轮博弈的后  $m$  次博弈中, 与无贴标签相比, 背叛的收益减小, 并且背叛的收益不高于合作的收益, 即  $v_{jD}(x) < v'_{jD}(x)$ ,  $v_{jC}(x) \geq v_{jD}(x)$ ,  $j = 2, 3, \dots, m + 1$ .

**证** 由贴标签作用前的博弈收益矩阵(1)和假设3知

$$\begin{cases} v'_{jC}(x) = xR + (1 - x)S, \\ v'_{jD}(x) = xT + (1 - x)P, \end{cases} \quad (10)$$

$$\begin{cases} v_{jC}(x) = \alpha[\alpha xR + (1 - \alpha x)P] + (1 - \alpha)P, \\ v_{jD}(x) = P. \end{cases} \quad (11)$$

无论贴标签的概率  $\alpha$  的大小,  $v_{jD}(x) < v'_{jD}(x)$  和  $v_{jC}(x) \geq v_{jD}(x)$  (其中等号当且仅当  $\alpha = 0$  时成立)恒成立, 贴标签控制总是使得背叛的收益相对于贴标签之前减少, 这种贴标签控制能抑制背叛的投机收益并通过这种抑制实现促进合作.

**命题 3** 如果在每轮博弈中贴标签的概率  $\alpha$  满足如下条件:

$$\alpha > \sqrt{\frac{x(R - S) + S - P}{x(R - P)}}, \quad (12)$$

则在该轮博弈的后  $m$  次博弈中, 合作的收益比贴标签之前增大, 即  $v_{jC}(x) > v'_{jC}(x)$ .

**证** 当式(12)满足时, 由收益(10)–(11)可知  $v_{jC}(x) > v'_{jC}(x)$  成立.

**推论 1** 条件(12)不满足时,合作的收益与背叛的收益都因贴标签控制而减小,但仍然满足 $v_{jC}(x) > v_{jD}(x)$ .

条件(12)不满足时,虽然贴标签机制仍然能促进合作的演化,但是博弈中每个个体的收益都是降低的.此时,贴标签控制可以视为合作者有代价地对背叛行为进行惩罚,背叛的损失高于合作者付出的代价并且实现了 $v_{jC}(x) > v_{jD}(x)$ .

贴标签控制对一轮博弈中个体总收益的影响不仅与贴标签的概率 $\alpha$ 有关,还与总体状态 $x(t)$ 有关.记无贴标签控制时,合作与背叛在一轮博弈中的总收益分别为 $w'_C(x), w'_D(x)$ ,下面的2个命题说明贴标签控制与总体状态对个体收益影响的交互作用.

**命题 4** 对于任意的贴标签概率 $\alpha \in [0, 1)$ 和任意的总体状态 $x \in (0, 1)$ ,与无贴标签控制相比,贴标签控制降低了背叛在一轮博弈中的总收益,即 $w_D(x) < w'_D(x)$ .

**命题 5** 当设置贴标签控制但 $\alpha = 0$ 时,若 $x > (P - S)/(R - S)$ ,与无贴标签控制相比,贴标签控制降低了合作的总收益;否则,贴标签控制增加了合作的总收益.

命题5说明,当群体的合作率较低时,贴标签控制本身可以保护合作者,与博弈重复次数及贴标签概率无关.群体合作率较低时,贴标签可以减少合作被背叛欺诈的概率,实现有效保护合作.当合作率较大时,如果管理者设置了贴标签控制,但是执行力太低,即 $\alpha$ 太小,则这样的贴标签控制在阻止背叛的同时,也牺牲了合作以较大比例获得双赢的机会.

## 4 控制策略设计与促合作效果分析

### 4.1 带有控制的复制子动态

这里,作者考虑较长一段时间内(或者无限时间跨度)的博弈行为,通过对 $\alpha(t)$ 的调整来影响状态变量 $x(t)$ ,从而协调控制目标的实现情况.

为简化公式,记 $\pi = T + S - P - R$ ,  $\pi' = T + S - 2P$ ,则假设1-5下,合作的复制动态方程为

$$\dot{x}(t) = (-\pi x + S - P)(1 - x)x + mt(R - P)(1 - x)x^2\alpha^2(t). \quad (13)$$

考虑博弈起始阶段合作与背叛混合存在,即在 $t = 0$ 时刻,系统初值为

$$x(0) = x_0, 0 < x_0 < 1. \quad (14)$$

**注 1** 由文献[25],贴标签控制下囚徒困境的ESS与合作演化动态(13)的稳定平衡点具有一致性:合作或者背叛是ESS,对应 $x = 1$ 或者 $x = 0$ 是动态(13)的稳定平衡点;合作与背叛均为ESS,对应动态(13)是双稳定系统.

**注 2** 博弈的ESS的吸引域与合作演化动态(13)的稳

定点的吸引域一致.

**注 3** 虽然博弈各方在无贴标签控制和 $\alpha(t) = 0$ 的贴标签控制下的收益不同,但是在忽略时间刻度的情况下,由式(13),二者具有相同的演化动态.所以,本文对无贴标签控制时和 $\alpha(t) = 0$ 时的合作演化动态不加区别.

### 4.2 促合作控制设计

在每轮博弈中,控制者观察第1次博弈中的合作行为并将合作者贴上标签,通过控制贴标签概率 $\alpha$ ,使得在其后的 $m$ 次博弈中,一个贴有标签的个体经历合作对 $(C, C)$ 的次数均值为 $\max$ .显然, $\alpha$ 越大, $c(t)$ 就越高.但是,由于 $\alpha(t)$ 受控制者识别能力和控制成本的限制,致使控制量 $\alpha(t)$ 的上界受限,这里假设 $\alpha(t)$ 的最大值为 $M$ ,则容许控制集合为

$$\mathcal{U} = \{\alpha(t) | 0 \leq \alpha(t) \leq M \leq 1\}. \quad (15)$$

**定理 1** 对于任意的 $\alpha(t) \in \mathcal{U}$ ,背叛恒为ESS.

**证** 一个合作个体进入背叛群体时,其收益为 $F(C, D)$ ,群体中的背叛个体的收益为 $F(D, D)$ ,由收益矩阵(7), $F(C, D) = S + mP$ , $F(D, D) = (m + 1)P$ .由 $S < P$ ,得 $F(C, C) > F(D, C)$ .由文献[25],小群体的合作者不能入侵背叛群体,从而背叛恒为ESS.

虽然引入控制 $\alpha(t)$ 不能改变背叛是囚徒困境的ESS这个结果,但是满足一定条件的 $\alpha(t)$ 可以使得合作也是ESS,并且可以通过增大 $\alpha(t)$ 来增加合作在双稳定系统的吸引域.

**定理 2** 如果容许控制(15)的上界 $M$ 和每轮博弈的博弈次数 $m + 1$ 满足如下条件:

$$M > M_1 = \sqrt{\frac{T - R}{m(R - P)}}, \quad (16a)$$

$$m + 1 > m_1 = \frac{T - P}{R - P}, \quad (16b)$$

则存在控制律 $\alpha(t)$ ,使得合作是博弈的ESS.

**证** 设计控制律

$$\alpha(t) = \sqrt{\frac{1}{2}(M^2 + M_1^2)}. \quad (17)$$

由式(16b), $M_1 < 1$ ,所以存在 $M < 1$ 使得式(16a)成立.显然,控制律(17)满足 $\alpha(t) \in \mathcal{U}$ .在控制律(17)下,博弈中的个体收益满足 $F(C, C) > F(D, C)$ ,所以控制律(17)下合作是博弈的ESS.

在社会经济系统中,控制律是控制者所实施的控制力相对于时间的函数.对于集群或双边平台的管理者来说,控制律是其在主观管理能力范围内实现控制目标所要求的控制力水平在时间上的反应,它跟控制能力、控制成本和控制目标紧密相关.

由定理1、定理2知,当容许控制(15)的上界 $M$ 和每轮博弈的博弈次数 $m + 1$ 满足(16)时,存在控制

律 $\alpha(t)$ , 使得囚徒困境具有2个纯策略ESS. 下面对双稳定系统(13)的不稳定内点进行讨论, 定义并找到入侵屏障 $x=r$ : 设合作的吸引域为 $\{x \mid r < x \leq 1\}$ , 则 $r \leq 1/2$ 时称合作为风险占优策略, 此时合作的吸引域大于背叛的吸引域;  $r \leq 1/3$ 时称合作为优胜策略, 合作的吸引域大于背叛的吸引域的2倍.

**定理3** 如果容许控制(15)的上界 $M$ 和每轮博弈的博弈次数 $m+1$ 分别满足如下不等式:

$$M > M_2 = \sqrt{\frac{T+P-R-S}{m(R-P)}}, \quad (18a)$$

$$m+1 > m_2 = \frac{T-S}{R-P}, \quad (18b)$$

则存在控制律 $\alpha(t)$ , 使得合作是博弈的风险占优策略.

**证** 由 $S < P$ 得 $M_2 > M_1$ . 设计控制律

$$\alpha(t) = \sqrt{\frac{1}{2}(M^2 + M_2^2)}. \quad (19)$$

由定理2, 条件(18)满足时,  $\alpha(t) \in \mathcal{U}$ . 控制(19)下, 合作是ESS. 由定理1, 博弈有2个ESS, 所以演化动态式(13)是双稳定系统, 当 $x=1/2$ 时, 个体收益满足

$$F(C, x) - F(D, x) = 0,$$

从而博弈的不稳定内点均衡为 $x^* = 1/2$ . 所以控制(19)下合作的吸引域为 $\{x \mid 1/2 < x \leq 1\}$ , 由注2, 合作是博弈的风险占优策略.

**定理4** 如果容许控制(15)的上界 $M$ 和每轮博弈的博弈次数 $m+1$ 分别满足如下不等式:

$$M > M_3 = \sqrt{\frac{\pi + 3(P-S)}{m(R-P)}}, \quad (20a)$$

$$m+1 > m_3 = \frac{\pi'}{R-P}, \quad (20b)$$

则存在控制律 $\alpha(t)$ , 使得合作是博弈的优胜策略.

**证** 当条件(20)满足时, 设计控制律

$$\alpha(t) = \sqrt{\frac{1}{2}(M^2 + M_3^2)}, \quad (21)$$

则 $M_3 < \alpha(t) < M$ ,  $\alpha(t) \in \mathcal{U}$ . 控制(21)下, 合作是ESS, 演化动态(13)是双稳定系统且博弈的不稳定内点均衡为 $x^* = 1/3$ . 所以控制(21)下合作的吸引域为 $\{x \mid 1/3 < x \leq 1\}$ , 由注2, 合作是博弈的优胜策略.

由定理2-4,  $m_3 > m_2 > m_1$ 且 $M_3 > M_2 > M_1$ , 即对合作的吸引域要求越高, 则对带标签博弈次数 $m$ 和贴标签控制 $\alpha(t)$ 的要求就越高.

**定理5** 如果容许控制(15)的上界 $M$ 和每轮博弈的博弈次数 $m+1$ 满足如下不等式:

$$M > M_4 = \sqrt{\frac{\pi x_0 + P - S}{m(R-P)x_0}}, \quad (22a)$$

$$m+1 > m_4 = \frac{\pi'x_0 + P - S}{(R-P)x_0}, \quad (22b)$$

则存在控制律 $\alpha(t)$ , 使得 $x=1$ 是合作演化动态(13)在任意初值下的稳定平衡点, 合作的吸引域是 $(x_0, 1]$ .

**证** 当式(22)成立时, 设计控制律

$$\alpha^2(t) = \beta M^2 + (1-\beta) \frac{\pi x(t) + P - S}{m(R-P)x(t)}, \quad (23)$$

这里 $0 < \beta < 1$ 是控制者设计的常量, 则 $M_4 < \alpha(t) < M$ . 控制律(23)下合作演化动态如下:

$$\dot{x}(t) > \beta m(R-P)(1-x)^2. \quad (24)$$

令 $x' = x - 1$ , 式(24)等价于

$$\dot{x}'(t) < -\beta m(R-P)x'(1-x')^2. \quad (25)$$

基于动态系统稳定性理论,  $x=1$ 是合作演化动态(13)的稳定平衡点. 由注2, 控制(23)下合作的吸引域为 $(x_0, 1]$ .

**注4** 控制律(23)设计的 $\alpha(t)$ 随设计参数 $\beta$ 增大而增大; 同时, 随参数 $\beta$ 的增大, 受控系统(13)的状态变化速度加快, 即合作率提升速度加快.

**注5** 本文定理2-5证明中设计的控制律 $\alpha(t)$ 只是满足定理的一种控制形式.

**注6** 由于本文假设每一轮博弈的博弈次数是固定的, 所以相对于定理2-5, 定理条件中对博弈次数的要求是对应命题成立的充要条件.

定理5看上去完全解决了囚徒困境博弈中的合作的生存与合作率提升问题, 但是,  $M_4$ 和 $m_4$ 随 $x_0$ 的减小而增加, 并且有 $\lim_{x_0 \rightarrow 0^+} m_4 = +\infty$ . 所以, 当 $x_0$ 很小时, 条件(22)对 $m$ 和 $\alpha(t)$ 要求很高, 这些要求的实现是困难的.

条件(22)说明促进合作率提升的控制律的存在性取决于系统初值, 演化起始阶段合作者占比群体总量的比例 $x_0$ 越小, 就要求合作有越大的吸引域, 对控制 $\alpha(t)$ 的要求也就越高. 所以, 这种合作识别的贴标签控制比较容易实现群体合作率较高时的合作积极性保持, 而不太容易实现群体合作率低时的状态改善.

换个思路, 考虑识别群体中的背叛行为, 讨论基于背叛识别的贴标签控制与基于合作识别的贴标签控制的异同.

## 5 基于背叛识别的贴标签控制

现在考虑控制者对每轮博弈的第1次博弈中的个体背叛行为进行识别, 通过对背叛者贴标签来保护合作. 假设控制者对背叛识别的概率为 $\alpha'$  ( $0 \leq \alpha' \leq 1$ ), 当合作者与一个未被贴上标签的个体相遇时, 选择合作, 否则选择背叛; 背叛者总是选择背叛.

总体状态为 $x(t)$ 时, 带标签博弈中策略对 $(C, C)$ ,  $(D, D)$ 出现的概率分别为 $x^2(t)$ 和 $1-x^2(t)-2(1-\alpha')$

$(1-x)x$ ,  $(C, D)$ 和 $(D, C)$ 出现的概率均为 $(1-\alpha') \cdot (1-x)x$ . 与无贴标签控制相比, 基于背叛识别的贴标签控制不改变合作对 $(C, C)$ 出现的频率, 但是,  $(C, D)$ 对和 $(D, C)$ 出现的频率为由 $(1-x)x$ 降为 $(1-\alpha')(1-x)x$ , 降低了合作被欺诈的概率, 并且,  $\alpha'$ 越大, 合作被背叛欺诈的可能性越低. 与基于合作的贴标签控制相比, 基于背叛的贴标签控制保持了合作对占总策略对的比率 $c(t)$ , 但是 $\alpha' < 1$ 时不能消除合作被背叛欺诈的可能性.

假定 $\alpha'(t)$ 的容许控制集合为如下定义的 $\mathcal{U}'$ :

$$\mathcal{U}' = \{\alpha'(t) | 0 \leq \alpha'(t) \leq M' \leq 1\}. \quad (26)$$

基于背叛识别的贴标签控制下的合作演化动态如下:

$$\begin{aligned} \dot{x}(t) = & -(m+1)(T-R)(1-x)x + \\ & m(\pi'x + P - S)(1-x)x\alpha'. \end{aligned} \quad (27)$$

对应定理1-5, 下面不加证明地给出基于背叛识别的贴标签控制的相应结果.

**定理 1'** 对于任意的 $\alpha'(t) \in \mathcal{U}'$ , 背叛是ESS.

由定理1、定理1', 无论是对合作者贴标签, 还是对背叛者贴标签, 都不能让合作成功入侵背叛群体.

**定理 2'** 如果容许控制(26)的上界 $M'$ 满足如下不等式:

$$M' > M'_1 = \frac{(m+1)(T-R)}{m(T-P)}, \quad (28)$$

并且每轮博弈的次数 $m+1$ 满足(16b), 则存在控制律 $\alpha'(t)$ , 使得合作是ESS.

**定理 3'** 如果容许控制(26)的上界 $M'$ 满足如下条件(29):

$$M' > M'_2 = \frac{(m+1)(T+P-R-S)}{m(T-S)}, \quad (29)$$

并且每轮博弈的次数 $m+1$ 满足式(18b), 则存在控制律 $\alpha'(t)$ , 使得合作是风险占优策略.

**定理 4'** 如果容许控制(26)的上界 $M'$ 满足如下不等式:

$$M' > M'_3 = \frac{(m+1)[\pi + 3(P-S)]}{m\pi'}, \quad (30)$$

并且每轮博弈的次数 $m+1$ 满足式(20b), 则存在控制律 $\alpha'(t)$ , 使得合作是优胜策略.

**定理 5'** 如果容许控制(26)的上界 $M'$ 和每轮博弈的博弈回合数 $m+1$ 满足如下不等式:

$$M > M'_4 = \frac{(m+1)(\pi x_0 + P - S)}{m[\pi'x_0 + P - S]}, \quad (31)$$

并且每轮博弈的次数 $m+1$ 满足式(22b), 则存在控制律 $\alpha'(t)$ , 使得 $x=1$ 是合作演化动态(27)在任意的初值下的稳定平衡点, 合作的吸引域是 $(x_0, 1]$ .

比较定理 $i$ 与定理 $i'$ ,  $i, i' = 2, 3, 4, 5$ , 为了促使合

作的成功演化, 基于合作识别和基于背叛识别对带标签博弈次数 $m$ 的要求一致, 但是对贴标签控制的控制上限的要求不同. 当带标签博弈次数 $m$ 比较大时, 基于合作识别的贴标签控制对控制上限的要求较低. 下面讨论这两种贴标签控制的实现难度.

考虑如下关于 $x_0$ 的一元二次方程:

$$\begin{aligned} & [(m+1)^2(R-P)\pi - m\pi'^2]x_0^2 + \\ & [(m+1)^2(R-P) - 2m\pi'](P-S)x_0 - \\ & m(P-S)^2 = 0. \end{aligned} \quad (32)$$

设 $x_0^L$ 和 $x_0^H$ 分别表示式(32)的较小和较大的根, 则 $x_0^L x_0^H < 0$ . 对于给定的博弈收益矩阵(1)和带标签博弈次数 $m$ , 如果如下不等式成立:

$$(R-P)(T-R)(m+1)^2 > \pi'(T-P)m, \quad (33)$$

则 $x_0^H \in (0, 1)$ . 由 $M'_4$ 与 $M_4$ 的定义, 当 $x_0 < x_0^H$ 时,  $M'_4 < M_4$ ; 当 $x_0 > x_0^H$ 时,  $M'_4 > M_4$ . 所以, 在控制的初始时刻, 如果群体合作率较低, 则基于背叛的贴标签控制比较易于实现合作率提升, 反之, 基于合作识别的贴标签控制更易于实现.

当条件(33)不满足时, 带标签博弈次数 $m$ 不满足式(16b), 即不存在满足条件(33)的带标签博弈次数 $m$ 和贴标签控制律 $\alpha'(t)$ , 使得合作是ESS. 从而, 这两种控制的可实现程度由群体状态决定. 但是, 两种控制都不能改变在群体合作率较低时, 为了促进合作而对带标签博弈次数要求较高的约束.

## 6 数值仿真

在博弈次数足够大且控制无上限约束(即贴标签的概率可以是1)时, 合作可以在任意小的初值下成功演化, 即合作的吸引域可以是 $[\varepsilon, 1]$ , 这里 $\varepsilon > 0$ 是常量. 但是, 由演化动态(13)知, 控制项 $\alpha^2(t)$ 的系数是 $m(a-d)(1-x)x^2$ , 当 $x(t)$ 接近于0或者1时, 控制对合作演化产生的影响非常有限, 即当合作者在群体中所占的比例比较小时, 其成功演化要求较大的控制 $\alpha(t)$ 和较大的带标签博弈次数 $m$ . 当合作者在群体中所占的比例比较大时, 对控制 $\alpha(t)$ 和 $m$ 要求降低, 但是控制的变化对演化速度影响也很小. 下面举例说明控制律设计的多样性和不同控制律之下演化均衡点及演化速度的差异.

在囚徒困境博弈的收益矩阵(1)中, 假设 $T=5, R=3, P=0, S=-1, t=0$ 时的合作率分别设为 $x_0=0.9, 0.5, 0.34, 0.1$ . 由定理5,

$$m \geq 5, M > \sqrt{\frac{x_0+1}{3mx_0}}$$

时, 在各个初值下都存在贴标签控制实现合作率提升, 设博弈的带标签博弈次数 $m=6$ . 本例中, 当 $x_0 > 0.0588$ 时, 基于合作识别的贴标签控制对控制上界要

求较低. 由于本例  $x_0 \geq 0.1$ , 所以只设计基于合作识别的带标签控制.

为了观察控制变化对合作率变化的影响, 分别取如下 4 种贴标签的概率:

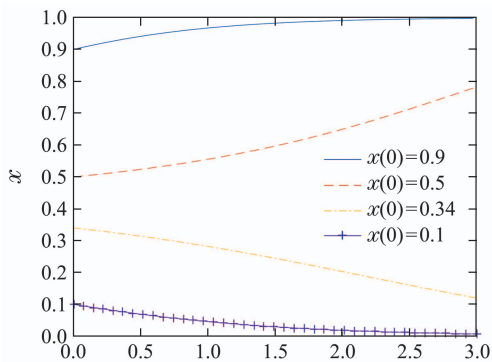
$$\alpha_1 = 0.42 + 0.01, \alpha_2 = 0.6 + 0.01,$$

$$\alpha_3 = \sqrt{\frac{x+1}{18x}} + 0.01,$$

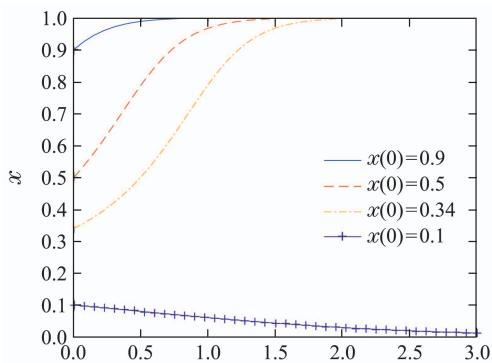
$$\alpha_4 = \begin{cases} \frac{2}{9}\sqrt{3} - 0.2, & x \geq 0.6, \\ \frac{2}{9}\sqrt{3} + 0.2, & x < 0.6, \end{cases}$$

得到合作动态的演化轨迹如图.

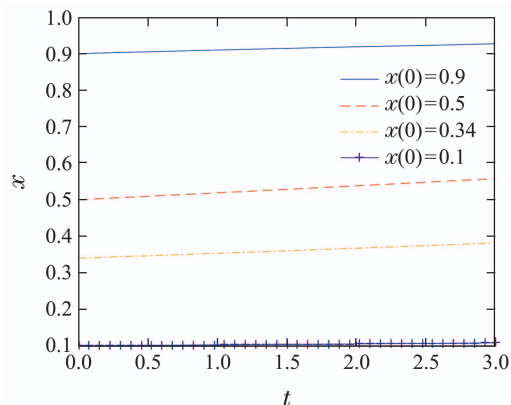
由图 2 可以看出, 当状态初值为 0.9 和 0.5 时, 控制  $\alpha_1$  和  $\alpha_2$  都能使状态  $x(t)$  随时间  $t$  单调递增, 即能使得合作率提升, 但是当状态初值为 0.34 时, 较大的  $\alpha_2$  可以提升合作率, 较小的  $\alpha_1$  却不能. 与控制  $\alpha_1$  相比较, 较大的控制  $\alpha_2$  使得系统初值为 0.9 和 0.5 时的状态轨迹更快地收敛于 1, 即合作所占的比率以更快的速度增加直至充满整个群体. 当系统初值为 0.1 时, 控制  $\alpha_1, \alpha_2$  虽然不能改变演化结果, 但是可以延缓合作衰减的速度.  $\alpha_3$  使得演化在这种控制作用下可以始终基本维持目前状态(由于控制  $c$  中包含不随状态变化的微小常量, 演化中合作率呈现非常缓慢的增加). 这种形式的控制要求同样是由系统初值决定的, 相对于  $x_0$  的 4 种取值 0.9, 0.5, 0.34 和 0.1, 对应的控制 0.1 分别为 0.352, 0.409, 0.478 和 0.791.



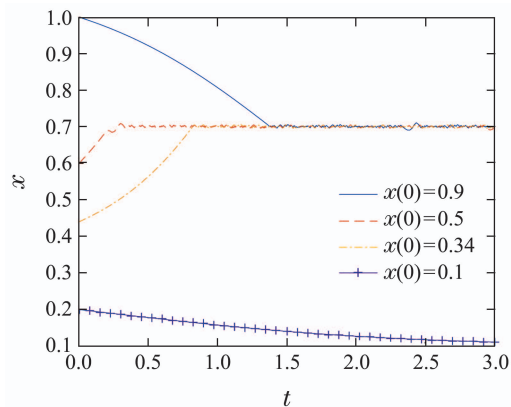
(a)  $\alpha(t) = \alpha_1$



(b)  $\alpha(t) = \alpha_2$



(c)  $\alpha(t) = \alpha_3$



(d)  $\alpha(t) = \alpha_4$

图 2 控制  $\alpha_i$  下的状态轨迹

Fig. 2 The state's trajectories under control law  $\alpha_i$  ( $i = 1, 2, 3, 4$ )

控制者并不总是将控制目的设定为合作者占领整个种群, 有时, 控制目的是实现合作者在群体内占比长期稳定在某一满意值即可, 比如, 希望  $x(t)$  稳定在 0.6. 控制  $\alpha_4$  可以实现这一目的. 由于  $\alpha_4$  最大值为 0.586, 控制  $\alpha_4$  在系统初值为 0.9, 0.5, 0.34 下可以实现控制目的, 但是, 当初值为 0.1 时, 控制不足以驱动状态正向演化, 合作演化失败. 同时, 可以通过控制律  $\alpha_4$  实现合作演化动态具有内点稳定点的结论.

### 7 结论与展望

本文讨论了囚徒困境博弈的促进合作演化问题, 引入了控制者的概念, 设计了贴标签控制规则, 指出了贴标签控制对促进合作的作用并分析促合作的实现途径. 本文引入动态反馈控制的概念, 建立合作演化动态模型并分析了可通过贴标签控制促进合作演化的条件, 讨论了合作的吸引域的大小与控制大小及博弈的其他参数之间的关系. 给出了囚徒困境的内点稳定的概念, 讨论了使得博弈在给定内点实现稳定的控制条件. 文章设计了基于合作和基于背叛的 2 种控制设计, 比较了两种控制的促合作途径、促合作效果和设计困难.

文章分析的模型简单, 很多情况没有考虑, 如贴标签的准确性, 本文只考虑漏贴 ( $\alpha < 1$  或者  $\alpha' < 1$ ), 而



没有考虑误贴(将背叛策略者贴上合作标志或者将合作者贴上背叛标志)。另外,考虑标签信息如何影响个体的学习方法是进一步的研究方向。

### 参考文献:

- [1] KENNEDY D, NORMAN C. What don't we know? *Science*, 2005, 309(5731): 75 – 102.
- [2] NOWAK M A. Five Rules for the evolution of cooperation. *Science*, 2006, 314(5805): 1560 – 1563.
- [3] MAY R M. The evolution of cooperation. *Bioscience*, 1981, 292(5821): 291 – 292.
- [4] ZHENG X D, LI C, YU J R, et al. A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the prisoner's dilemma game. *Journal of Theoretical Biology*, 2017, 420(3): 12 – 17.
- [5] NOWAK M A, SIGMUND K. Evolution of indirect reciprocity. *Nature*, 2005, 437(7063): 1291 – 1298.
- [6] WEDEKIND C, MILINSKI M. Cooperation through image scoring in humans. *Science*, 2000, 288(5467): 850 – 852.
- [7] HAMILTON W D. The genetical evolution of social behavior. *Journal of Theoretical Biology*, 1964, 7(1): 17 – 52.
- [8] MAYNARD S J, PRICE G R. The logic of animal conflict. *Nature*, 1973, 246(11): 15 – 18.
- [9] KASIEWICZ A, SZUBERTKRUSZYNSKA A, RADWAN J. Kin selection promotes female productivity and cooperation between the sexes. *Science Advances*, 2017, 3(3): e1602262.
- [10] RIOLO R L, COHEN M D, AXELROD R. Evolution of cooperation without reciprocity. *Nature*, 2001, 414(6862): 441 – 443.
- [11] DAWKINS R. The selfish gene. *Quarterly Review of Biology*, 2006, 110(466): 781 – 804.
- [12] HELLER J, ZHAO J, ROSENFELD G, et al. Characterization of greenbeard genes involved in long-distance kind discrimination in a microbial eukaryote. *Plos Biology*, 2016, 14(4): e1002431.
- [13] GRUENHEIT N, PARKINSONK, STEWART B, et al. A polychromatic 'green beard' locus determines patterns of cooperation in a social amoeba. *Nature Communications*, 2017, 8: 14171.
- [14] MCDONALD P G, ROLLINS L A, GODFREY S. The relative importance of spatial proximity, kin selection and potential 'green beard' signals on provisioning behavior among helpers in a cooperative bird. *Behavioral Ecology & Sociobiology*, 2016, 70(1): 133 – 143.
- [15] WANG Long, WU Te, ZHANG Yanling. Feedback mechanism in coevolutionary games. *Control Theory & Applications*, 2014, 31(7): 823 – 836.  
(王龙, 吴特, 张艳玲. 共演化博弈中的反馈机制. 控制理论与应用, 2014, 31(7): 823 – 836.)
- [16] TAYLOR C, NOWAK M A. Evolutionary game dynamics with non-uniform interaction rates. *Theoretical Population Biology*, 2006, 69(1): 243 – 252.
- [17] DONG Rui, WANG Xianjia, CHEN Lin. The controller design for promoting the evolution of cooperation in the prisoner's dilemma based on the non-uniform interaction rates. *Systems Engineering – Theory & Practice*, 2017, 37(10): 2582 – 2591.  
(董瑞, 王先甲, 陈琳. 囚徒困境中基于非均匀连接的促合作控制设计. 系统工程理论与实践, 2017, 37(10): 2582 – 2591.)
- [18] WANG Long, DU Jinming. Evolutionary game theoretic approach to coordinated control of multi-agent systems. *Journal of Systems Science & Mathematical Sciences*, 2016, 36(3): 302 – 318.  
(王龙, 杜金铭. 多智能体协调控制的演化博弈方法. 系统科学与数学, 2016, 36(3): 302 – 318.)
- [19] GUAN Yonqiang, JI Zhijian, ZHANG Lin, et al. Recent developments on controllability of multi-agent systems. *Control Theory & Applications*, 2015, 32(4): 421 – 431.  
(关永强, 纪志坚, 张霖, 等. 多智能体系统能控性研究进展. 控制理论与应用, 2015, 32(4): 421 – 431.)
- [20] XUE Lei, WANG Qingling, SUN Changyin, et al. Game theoretical approach for the leader selection of the second-order multi-agent system. *Control Theory & Applications*, 2016, 33(12): 1593 – 1602.  
(薛磊, 王庆领, 孙长银. 博弈论框架下的二阶多智能体系统领导者选择算法. 控制理论与应用, 2016, 33(12): 1593 – 1602.)
- [21] PAN Feng, XI Bao, WANG Lin. Analysis on environmental regulation strategy of local government based on evolutionary game theory. *System Engineering – Theory & Practice*, 2015, 35(6): 1393 – 1404.  
(潘峰, 西宝, 王琳. 基于演化博弈的地方政府环境规制策略分析. 系统工程理论与实践, 2015, 35(6): 1393 – 1404.)
- [22] WANG X J, DONG R, CHEN L. The optimal control for promoting the cooperation in evolution game generated by prisoner's dilemma. *Acta Mathematica Scientia*, 2018, 38(1): 73 – 92.
- [23] CHEN X, SASAKI T, BRANNSTROM Å, et al. First carrot, then stick: how the adaptive hybridization of, incentives promotes cooperation. *Journal of the Royal Society Interface*, 2015, 12(102): 20140935.
- [24] MILLER F P, VANDOME A F, MCBREWATER J, et al. *Grim Trigger*. Saarbrücken, Germany: Alphascript Publishing, 2010.
- [25] TAYLOR P D, JONKER L B. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 1978, 40(2): 145 – 156.

### 作者简介:

董瑞 副教授, 博士, 研究方向为系统控制、博弈论, E-mail: drui@163.com;

陈琳 博士研究生, 研究方向为系统控制、博弈论, E-mail: 293312573@qq.com.

王先甲 教授, 博士, 研究方向为博弈论、决策分析, E-mail: wangxj@whu.edu.cn.