基于名声机制的重复囚徒困境合作博弈分析

李 栋[1] 蒋军利[2] 唐晓嘉[1]

(西南大学逻辑与智能研究中心 重庆 400715)[1] (西南大学经济管理学院 重庆 400716)[2]

摘 要 处于重复囚徒困境下的博弈者存在相互合作的可能性。在此基础之上,探讨了一类具有两种状态的名声机制,并发现其中只有3个马尔可夫策略是高效的强健完美纳什均衡。研究表明,跟好名声者合作和背叛坏名声者的策略是最具吸引力的一个策略,此合作可最终成功实现并且持续下去。

关键词 名声机制,重复囚徒困境,强健完美纳什均衡

中图法分类号 B819 **文献标识码** A

Analysis of Cooperative Game in Repeated Prisoners' Dilemma Based on Reputation Mechanisms

LI Dong[1] JIANG Jun-li[2] TANG Xiao-jia[1]

(Institute of Logic and Intelligence, Southwest University, Chongqing 400715, China)[1]

(College of Economics and Management, Southwest University, Chongqing 400716, China)[2]

Abstract The players under the repeated Prisoners' Dilemma have the possibility of mutual cooperation. On this basis, We examined a class of two-state reputation mechanisms, and found that only three reputation mechanisms are efficient robust perfect Nash equilibrium in Markov strategies. The research shows that the strategy to cooperate with good opponents and defect against bad opponents is a global attractor, and hence cooperation is successfully achieved and sustained in the long run.

Keywords Reputation mechanisms, Repeated prisoners' dilemma, Robust perfect Nash equilibrium

1 引言

人类社会与其他动物群体的一个重要区别在于,人与人之间可以通过运用个人理性而达到某种形式的合作。人与人之间的合作,是人类文明社会的基础[1]。关于合作如何形成的研究具有重大的理论意义和实践价值,已经成为社会科学、自然科学乃至计算机科学的研究热点。在每个人都具有自私动机的情况下,人们怎样才能通过社会博弈而自发地产生合作?囚徒困境(Prisoners' Dilemma)[2]所揭示的正是这样的问题。特里弗斯(Trivers)提出过他自己的见解,并新造了"互惠利他主义"(reciprocal altruism)这个专用术语[3],弗里德曼(Friedman)提供了一个严格的证明:如果博弈者比较重视自己未来的收益,那么合作就能够在不定次的重复囚徒困境(indefinitely repeated Prisoners' Dilemma)中自我实现(self-enforcing)[4]。但是,这种直接互惠机制存在着巨大的局限——个体之间必须具有很大的重复交往机会。在当今社会中,一次性的快捷的交往方式逐渐成为主流,个体很难与同一对手交往多次,因而直接互惠机制由于缺乏现实合理性而限制了这个理论的应用范围。因此,间接互惠开始受到研究者们的重视。美国密歇根大学的 Alexander 于 1986 提出了"间接互惠"这个概念,并且认为间接互惠是人类道德、伦理和法律体系的基础[5]。

后来,又有学者提出"间接收益(indirect reciprocity)"这个概念,即一个人现在选择合作的未来收益,不是来自当前对局中的收益预期,而是间接来自基于当前合作得到的名声的未来对局。名声机制在间接互惠的研究中处于核心地位,诺瓦克(Nowak)与西格蒙德(Sigmund)设计出了一个模型,在这个模型中,群体中的所有成员都携带着一种基于过去行为的名声标记[6]。但是,这个间接收益模型存在一个很大的困难:它不能对博弈者因为背叛一个好名声者而采取的背叛策略与博弈者为了惩罚一个坏名声者而采取的背叛策略做出区分。要做出这样的区分,名声转换函数(从当前的名声转换成将来的名声)需要参考个体当前的行动和对手的名声。Kandori[7]与 Okuno-Fujiwara 和 Postlewaite[8]认识到了这一问题,并提出一种"标签装置"作为解决方案。

本文将首先分析囚徒困境中实现合作的潜在可能性,然后探讨如何通过名声机制(Reputation Mechanisms, RM)实现群体内的合作,最后做出简要的总结。

2 囚徒困境中实现合作的可能性

假设在一个二人博弈中,每个博弈者拥有两种可能的选择:合作(Cooperate, C)和背叛(Defect, D),每种可能选择的预期收益如下表所列。

到稿日期:2012-06-03 返修日期:2012-10-10 本文受教育部人文社会科学规划项目(10YJA72040003)资助。

李 栋(1981—),男,博士生,主要研究方向为现代逻辑与人工智能,E-mail:swnuld@sina.com;蒋军利(1980—),女,博士后,主要研究方向为博弈与经济管理;唐晓嘉(1954—),女,教授,博士生导师,主要研究方向为现代逻辑与人工智能,E-mail:tangxj@swu.edu.cn(通信作者)。

		列博弈者 C	列博弈者 D
行博弈者	C	R=1,R=1	S=−c,T=b
	D	T=b,S=−c	P=0,P=0

注:行博弈者的收益列于前面;b>1,c>0,且 2>b−c>0。

作为博弈一方的"行"可以选择合作或者背叛;同时另一方"列"也在合作或者背叛中进行选择,这些选择的收益组合形成如上表所列的 4 种可能结果。在这个博弈中,如果双方都选择合作策略 C,那么他们都将获得较高收益(1,1),这是对合作的奖励(reward),用 R 表示;如果他们都选择背叛策略 D,那么他们都将获得比较低的收益(0,0),遭受对背叛的惩罚(punishment),用 P 表示;如果任何一方在对手选择合作策略时选择背叛策略,那么背叛者将得到最高的收益 b,这一高收益成为背叛的诱惑(temptation),用 T 表示,而此时合作者却得到很低的收益 $-c$,意味着傻瓜(sucker)的收益,用 S 表示。

在这种收益结构下,选择背叛策略 D 将是占优策略(dominant strategy)。因为,如果对手选择合作策略 C,那么此时你选择背叛策略将比选择合作策略获取更大的收益,即 $T>R$。如果对手选择背叛策略 D,那么此时你选择背叛策略仍比选择合作策略获益更多,即 $P>S$。因此,在囚徒困境的情况下,无论对方怎么做,自己选择背叛都将是最好的策略。同样的逻辑对另一个人也同样适用,即无论你如何选择,对方也一定会选择背叛。这样,将会出现双方相互背叛的情形,双方都只能得到较少的收益 $P=0$,这比双方合作所能得到的收益 $R=1$ 差很多。个体追求最大利益的理性却导致双方实际得到的收益比可能得到的收益少很多,这就是所谓的"困境"。"囚徒困境"是对一些非常普遍的情形的简单抽象。在这些情形中,从个人的角度考虑,背叛是最好的选择,但双方背叛又会导致很不理想的结果。"囚徒困境"的定义要求这 4 个可能的结果之间保持一定的关系:(1)$T>R>P>S$;(2)$R>(T+S)/2$,即博弈者不能通过轮流背叛对方来摆脱困境。

如果互相之间都没有对方的任何信息,对于背叛不存在有效的惩罚机制,任何一次轻信都将可能导致损失,那么无论是一次博弈还是多次重复博弈,背叛都是最好的选择。但是,如果群体中的所有成员都携带着一种基于过去行为的名声标记,通过选择合作策略而获得好名声,具有好名声的个体在未来的对局中将会获得更大的收益。即一个人现在选择合作策略的好处不是直接来自当前对局中的收益预期,而是间接来自基于当前合作得到的名声的未来对局。在这种情况下博弈者能够预期到背叛的后果,如果他们关心自己未来的收益,那么就会出现合作的可能性。

3 名声机制的基本概念

在诺瓦克和西格蒙德模型最简单的形式中,如果一个博弈者最近选择的是合作策略(Cooperate,C),那么他将得到一个好名声(Good reputation,G);如果他最近选择的是背叛策略(Defect,D),那么他将得到一个坏名声(Bad reputation,B)。群体中个体的基本变量是拥有好名声或者拥有坏名声这两种状态。假设群体中的个体是成对对局的,并且个体在每次对局之后的名声取决于他所做出的行动和他与对手在对局之前的名声。也就是说,一个名声机制的输入变量包括这个个体现在的行为、他的名声以及他对手的名声。由于有好名声 G 与坏名声 B 之分,而且是成对对局的,因此组合起来就有 4 种名声有序对:{GG,GB,BG,BB}。每个名声有序对都具有特定的含义,前一个字母代表对手的名声,后一个字母代表自己的名声。比如,名声有序对 GB 就意味着:对手拥有一个好名声,且自己拥有一个坏名声。每一个个体都有合作 C 与背叛 D 这两种可能的行为,所以一个名声机制 RM 就有 8 种不同的输入变量:{GGC,GGD,GBC,GBD,BGC,BGD,BBC,BBD}。对于每一个输入,RM 输出 G 或 B,即个体都可能获得一个好名声 G 或者获得一个坏名声 B,因此,一共就有 $2^8=256$ 种可能的名声机制 RMs,每一种名声机制都可以通过一个表格表示出来,如表 1 所列。

表 1 RM1

当前的名声		自己当前的行为	
对方	自己	合作 C	背叛 D
G	G	G	B
G	B	G	B
B	G	G	G
B	B	B	B

表 1 显示出了 RM1 的名声组合情况,前两列给出了对局双方当前的名声,后两列给出了自己当前采取一定行动之后所得到的名声。当一个个体与一个坏名声者(B)对局时,如表 1 的后两行所示,这时不用考虑他的行为,其名声将保持不变;当一个个体与一个好名声者(G)对局时,如果他选择合作策略 C(第 3 列),那么他的名声将会变好;如果他选择背叛策略 D(最后一列),那么他的名声会变坏。需要注意的是,当一个好名声者为了惩罚一个坏名声者而选择背叛策略时,他不会因此被惩罚而获得一个坏名声,如表 1 的第 3 行最后一列;一个坏名声者也不会因为他对另一个坏名声者使用合作策略而获得一个好名声,如表 1 的最后一行第 3 列。这里,名声机制充分考虑了参与者行为的动机,区分了公正与自私的背叛行为[9]。

马尔可夫策略是将{GG,GB,BG,BB}中的每一个名声有序对映射到可能的行为{C,D}所得到的策略组合。令这 4 个名声有序对的顺序固定不变,一个马尔可夫策略可以被描述为形如 DCDD 的四元组,它的意思是:当名声有序对是 GB 时,选择合作策略 C,否则选择背叛策略 D。CCCC 表示总是选择合作策略;DDDD 表示总是选择背叛策略;CCDD 表示当对方拥有好名声时选择合作,当对方拥有坏名声时选择背叛。在 GB 状况下的合作恢复自己好名声的手段;在 BG 状况下的背叛不会导致自己丧失好名声而给予对手惩罚。这里一共有 $2^4=16$ 种不同的马尔可夫策略。

4 名声机制合作博弈的理论模型

4.1 强健均衡

均衡是博弈论的核心概念,是指博弈达到的一种稳定状态,任何一方都不愿意单独改变自己的策略,最重要也是最著名的均衡是纳什均衡(Nash Equilibrium)。假设有 n 个人参与博弈,在给定其他人策略的条件下,每个参与人的选择自己的最优策略,从而使自己利益最大化,所有参与人的策略构成一个策略组合。纳什均衡指的是这样一种策略组合,这种策略组合由所有参与人的最优策略组成,即在给定别人策略的情况下,没有人有足够理由打破这种均衡。如果一个策略组合在所有可能博弈路径上都能达到均衡,那么这个策略组合

就被称作完美纳什均衡(perfect Nash Equilibrium)。

假设在一个群体中存在着若干个个体,个体之间存在无穷的潜在对局轮回。在每个轮回中任意一个个体都将与群体中其他的个体进行对局,而且它们在每一个对局轮回之后的名声都会根据名声机制 RM 进行更新。这里个体之间所进行的重复博弈并非在两个特定的个体之间展开,而是在这个群体内的任意两个个体之间进行。另外,在每一个对局轮回之后,个体不再进行对局的概率为 $1-\delta(1>\delta>0)$。因此,就平均状况来看,每个个体都有 $1/(1-\delta)$次对局。

我们首先考虑仅由一个马尔可夫策略组成的群体的稳定状态。如果在每一次对局之后,一个马尔可夫策略 m 做出了一个对其自身的最优回应(best-response),那么这个马尔可夫策略 m 就是一个完美纳什均衡。

定理 1 对任意名声机制 RM,DDDD 总是一个完美纳什均衡,且群体的长期收益为 0。

证明:如果一个个体的对手总是选择背叛策略 D,那么他的收益是 $P=0$ 或者 $S=-c$,因此,这个个体未来的收益折算成当前价值 V 最多为 0。总是选择策略 D 将会达到这一最大值 0,如果当前选择策略 C,则未来收益的价值为 $-c+\delta V<0$。因此,选择 DDDD 策略是对 DDDD 策略的最优回应。

定理 2 对于名声机制 RM1(见表 1),如果 $\delta>\max\{b-1,c\}/(1+c)$,那么 CCDD 是一个完美纳什均衡,且群体的长期收益为 $1/(1-\delta)$。

证明:相对于对方的策略 CCDD,对于名声有序对 BG 和 BB 来说,同在定理 1 中的证明,背叛策略 D 是个体的占优策略,其预期收益为 $b+\delta V_G$ 和 δV_B。对于名声有序对 GG,当且仅当 $1+\delta V_G\geqslant b+\delta V_B$,即 $\delta(V_G-V_B)\geqslant b-1$ 时,如果后者选择合作策略 C,那么前者也选择合作策略 C 是最优的。对于名声有序对 GB,当且仅当 $-c+\delta V_G\geqslant\delta V_B$,即 $\delta(V_G-V_B)\geqslant c$ 时,如果后者选择背叛策略 D,那么前者选择合作策略 C 是最优的。总的来看,我们需要 $\delta(V_G-V_B)\geqslant\max\{b-1,c\}$。令 ψ 代表群体中拥有坏名声的个体所占的比例,则

$$V_G=(1-\psi)(1+\delta V_G)+\psi(b+\delta V_G),\lambda_{GB}=0$$

$$V_B=(1-\psi)(-c+\delta V_G)+\psi\delta V_B,\lambda_{BG}=0$$

注:λ_{GB}为由一个好名声变成一个坏名声的概率,λ_{BG}为由一个坏名声变成一个好名声的概率。

以上两个公式相减得到:$(V_G-V_B)=(1-\psi)(1+c)+\psi b+\psi\delta(V_G-V_B)$。这里唯一的可稳定状态是 $\psi=0$,所以根据上式得到$(V_G-V_B)=(1+c)$。因此,条件 $\delta(V_G-V_B)\geqslant\max\{b-1,c\}$可以被转变为 $\delta>\max\{b-1,c\}/(1+c)$。所以,当 $\delta>\max\{b-1,c\}/(1+c)$时,CCDD 是一个纳什均衡。

对于一个完美纳什均衡的完全刻画需要一个马尔可夫策略和一个与马尔可夫策略相一致的名声的稳定分布。任意给定一个马尔可夫策略作为候选均衡,可以得到一个 2×2 的矩阵,在这个矩阵中由一个好名声变成一个坏名声的概率为 λ_{GB},仍保持其好名声的概率为$(1-\lambda_{GB})$;由一个坏名声变成一个好名声的概率为 λ_{BG},仍保持其坏名声的概率为$(1-\lambda_{BG})$。

我们用 ψ 表示某一轮回群体中具有坏名声的个体所占的百分比,在下一次对局中,这个比例将会由 ψ 变成$\lambda_{GB}-\psi(\lambda_{GB}+\lambda_{BG})$。显然,$\psi^*$ 处于稳定状态,当且仅当 $\psi^*(\lambda_{GB}+\lambda_{BG})=\lambda_{GB}$。$\psi^*$ 将被称作是可稳定的,当且仅当对所有 $\psi<\psi^*$,$[\lambda_{GB}-\psi(\lambda_{GB}+\lambda_{BG})]>0$ 或者对所有 $\psi>\psi^*$,$[\lambda_{GB}-\psi(\lambda_{GB}+\lambda_{BG})]<0$。

如果(1)对所有 $\delta\in(\delta^*,1]$,$\delta^*<1$,这是一个完美纳什均衡;(2)ψ 的稳定状态是一个稳定的策略,在 δ 的值不是强健的情况下,这是一个完美的均衡;那么把一个马尔可夫策略和一个与之相应的名声稳定分布叫作一个强健的完美均衡。在一个策略的完美均衡中的最大可能收益是 $1/(1-\delta)$,我们把任何一个长期收益为 $1/(1-\delta)$的强健完美均衡称作是高效的。

定理 3 在这 16 个名声机制中,有 13 个具有唯一的强健完美纳什均衡 DDDD,并且每个个体在可稳定状态都有一个坏名声。其余 3 个名声机制具有高效性的强健完美均衡,并且每个个体在可稳定状态都有一个好名声。

这 3 个具有高效性的强健均衡的名声机制包括 RM1,另外两个是 RM2 和 RM3,分别在表 2 与表 3 中表示出来。

表 2 RM2

当前的名声		自己当前的行为	
对方	自己	合作 C	背叛 D
G	G	G	B
G	B	G	B
B	G	G	G
B	B	G	B

表 3 RM3

当前的名声		自己当前的行为	
对方	自己	合作 C	背叛 D
G	G	G	B
G	B	B	B
B	G	G	G
B	B	G	B

RM2 与 RM1 的区别仅体现在最后一行,在 RM1 中两个坏名声者对局时其名声将保持不变,而在 RM2 中坏名声者如果现在选择合作策略 C,那么他将会获得一个好名声。在 RM2 中,对于 $\delta>\max\{b-1,c\}/(1+c)$,CCDC 是一个强健的完美纳什均衡,并且群体的长期收益为 $1/(1-\delta)$,这与 RM1 具有相同的长期收益。

RM3 与 RM2 的区别仅在于第二行。在 RM2 中,当坏名声者与好名声者对局时,如果他选择合作策略,那么他将恢复其好名声。而在 RM3 中,当坏名声者与好名声者对局时,即使他选择合作策略,也不能得到一个好名声。在 RM3 中,如果 $\delta>(b-1)/b$,则 CDDC 是一个强健的完美纳什均衡,并且这个群体的长期收益为 $1/(1-\delta)$。

这 3 个高效的名声机制具有以下两个基本特征:(1)当两个好名声者对局时,他们所获得的名声取决于他们对行为的选择:如果选择合作策略 C,那么就获得好名声 G;如果选择背叛策略 D,就获得坏名声 B。(2)当一个好名声者与一个坏名声者对局时,这个好名声者可以惩罚那个坏名声者而选择背叛策略 D,而且好名声者不会因此而失去其好名声。

4.2 基于名声机制 RM1 的强健均衡求解

我们首先把时间进行分段,并编上号 $t=1,2,\cdots$,而且每一个时间段都拥有无穷的对局轮回,也编上号 $n=1,2,\cdots$。

由于一个个体在 RM1 下与一个坏名声者对局时,不论他选择什么行为,其名声都将保持不变,因此对他来说背叛策略 D 是最佳选择。也就是说,对于任意 $ab\in\{C,D\}^2$,马尔可夫策略 abDD 都是相对于马尔可夫策略 abCD,abDC 和 abCC

的弱占优策略。为了使模型简单，将忽略劣势策略，假定群体仅由这 4 种策略组成：{DDDD, CDDD, DCDD, CCDD}，而且可以把它进一步简化为 $M^* \equiv \{DD, CD, DC, CC\}$。

用 X_m 表示按程序运用策略 $m \in M^*$ 的博弈者所占的人数比例，则 $\sum_{m \in M^*} X_m = 1$。对于 M^* 中的任一个策略 m，用 ψ_m 表示个体拥有一个坏名声的概率，所以就用 $(1-\psi_m)$ 表示拥有一个好名声的比例。用 X_{mB} 表示运用策略 m 的坏名声者所占的人数比例，则 $X_{mB} = X_m \psi_m$。

所有的博弈者在每个阶段的第一个轮回都会进行对局。在每一次对局之后，博弈者被剔除不再对局的概率为 $1-\delta$，因此，在一个时间段中博弈者参与对局的预期数量为 $1/(1-\delta)$。在一个轮回中，当博弈者与一个好名声者对局时，他的名声能够发生改变。用 $\rho_m(t,n)$ 表示对局库中在时间段 t 的第 n 个轮回的开始使用策略 m 的坏名声者所占的比例，需要指出的是，$\rho_m(t,1) = \psi_m(t)$。最后，使 $\bar{\rho}(t,n) \equiv \sum_{m \in M^*} \rho_m(t,n)$，即对局库中坏名声者所占的总比例。

用 $\pi_{mG}(t)$ 和 $\pi_{mB}(t)$ 表示采用策略 m 在时间段 t 的所有对局轮回的累积收益：

$$\pi_{mG}(t) = k_t[V_{mG}(t) - A(t)]$$

$$\pi_{mB}(t) = k_t[V_{mB}(t) - A(t)]$$

这里 $V_{mG}(t)$ 和 $V_{mB}(t)$ 是策略 m 在时间段 t 的起点分别拥有一个好名声和一个坏名声的条件下的累积收益。

$$A(t) = \sum_{m \in M} x_m(t)[(1-\psi_m(t))V_{mG}(t) + \psi_m(t)V_{mB}(t)]$$

这是时间段 t 的所有对局轮回的群体平均收益。策略 m 在时间段 t 的平均增长率为：

$$\pi_m(t) \equiv [1-\psi_m(t)]\,\pi_{mG}(t) + \psi_m(t)V_{mB}(t)$$

虽然增长率是基于每一时间段开始的名声定义的，但是名声会在一个时间段内发生变化。用 $\mu_{mB}(t)$ 表示在时间段 t 有一个坏名声的条件下运用策略 m，它将会在时间段 t 的终点也获得一个坏名声的概率；用 $\mu_{mG}(t)$ 表示在时间段 t 的起点有一个好名声的条件下运用策略 m，它将会在时间段 t 的终点获得一个坏名声的概率。因此，策略 m 在时间段 t 的终点拥有一个坏名声的比例为：

$$x_{mB}(t^+) = x_m(t)\{[1-\psi_m(t)][1+\pi_{mG}(t)]\mu_{mG}(t) + \psi_m(t)[1+\pi_{mB}(t)]\mu_{mB}(t)\}$$

不考虑名声，采用策略 m 的人数比例为：

$$x_m(t^+) = x_m(t)\{[1-\psi_m(t)][1+\pi_{mG}(t)] + \psi_m(t)[1+\pi_{mB}(t)]\}$$

$$= x_m(t)[1+\pi_m(t)]$$

4.3 基于派系的名声机制刻画

名声机制的刻画是以群体为基础的，但是群体之间是存在着一定差异的，比如在一些群体内部存在不同的派系，本节就这种情况作专门的讨论。一个派系表示共同分享一个名声机制的一个小群体，从个体的角度来看，一个派系的另一种解释是：一个个体的行为能够被他自己所处的派系的其他成员观察到，并且这个个体也能观察到这个派系内其他成员的行为，所以，个体与处于同一派系内的其他个体对局所得到的名声，将会影响到自己今后在派系内对局时的收益。但是，个体与派系外的个体所进行的对局，却不会影响其在自己的派系内的名声。因此，个体对自己的派系之外的对手使用背叛策略 D 是一个弱占优策略，我们可以通过在所有自己派系外的对局中使用背叛策略 D 来扩展前面提到的马尔可夫策略。

用 r_k 表示派系 k 中的一个个体与派系 k 中的另一个个体随机的对局的概率，我们就得到了定理 2 的一个关于派系的扩展。

定理 4 对于名声机制 RM1，如果 $\delta > \max\{b-1, c\}/[r_k(1+c) + (1-r_k)\max\{b-1, c\}]$，那么 CCDD 是一个完美纳什均衡，且群体的长期收益为 $r_k/(1-\delta)$。

显然，当 $r_k = 1$ 时，定理 4 就与定理 2 完全相同了。如果派系之间正在为珍稀资源而竞争，那么具有最大 r_k 值的派系具有最大的生存潜力，因此，对局过程是很关键的。如果 r_k 值在派系的大小上是单调的，那么其中最大的派系将最终占据支配地位。如果 r_k 值对于一个规模有限的群体是优化的，那么几个派系能够共存。

结束语 如何实现合作是合作博弈研究中的一个难点，并具有重要的现实意义。我们通过分析发现，处于重复囚徒困境下的博弈者存在着相互合作的潜在可能性，但是，这种直接互惠机制存在着巨大的局限——个体之间必须具有很大的重复交往机会。在当今社会中，一次性的快捷的交往方式逐渐成为主流，个体很难与同一对手交往多次，因而直接互惠机制由于缺乏现实合理性，限制了这个理论的应用范围。而名声机制则是突破困境实现合作的一条有效途径。研究表明，跟好名声者合作和背叛坏名声者的策略是一个最具吸引力的策略，合作可因此最终成功实现并且持续下去。此外，由于派系的存在，有关名声的信息不能在不同派系顺利流通，这大大影响了名声机制的应用。如何使名声机制的功能突破派系的约束，在更大的群体中发挥作用？这需要有效的信息流通模式，将如名声这样的重要信息在各个派系间实现共享。本文的分析方法给我们提供了一种研究思路，对我们拓展开来研究其他问题具有重要的启发意义。

参 考 文 献

[1] 罗伯特·阿克塞尔罗德. 合作的复杂性——基于参与者竞争与合作的模型[M]. 梁捷，高笑梅，译. 上海：上海世纪出版集团，2008

[2] 罗伯特·阿克塞尔罗德. 合作的进化[M]. 吴坚忠，译. 上海：上海世纪出版集团，2007

[3] Trivers R. The evolution of reciprocal altruism [J]. The Quarterly Review of Biology, 1971, 46: 35-57

[4] Friedman J. A non-cooperative equilibrium for supergames [J]. Review of Economic Studies, 1971, 38: 1-12

[5] Alexander R D. Ostracism and indirect reciprocity: The reproductive significance of humor [J]. Ethology and sociobiology, 1986, 7: 253-270

[6] Nowak M, Sigmund K. Evolution of indirect reciprocity by image scoring [J]. Nature, 1988, 393: 573-577

[7] Kandori M. Social norms and community enforcement [J]. Review of Economic Studies, 1992, 59: 63-80

[8] Okuno-Fujiwara M, Postlewaite A. Social norms and random matching games [J]. Games and Economic Behavior, 1995, 9: 79-109

[9] Brandt H. Sigmund K. The good, the bad and the discriminator [J]. Errors in direct and indirect reciprocity, 2006, 239: 183-194