The Eighth International Symposium on Operations Research and Its Applications (ISORA'09)
Zhangjiajie, China, September 20–22, 2009

Social Behavior in the Simulation of Iterated Prisoner's Dilemma

Hong-Wei Zhang Kuan-Kuan Zhou Neng-Bing Hu

School of Mathematics, Changsha University of Science and Technology,
ChangSha 410004, China

Abstract Most of economic action can regarded as a special case of game theory. This article through analysis of computer simulation for the characteristics of Iterated Prisoner's Dilemma, a repeated game model based on the evaluation of the cumulative aspiration tense is built and the result of modulation is analyzed. The results show that it can provide dynamic description for the behavior of social group along with time prolong, and it could provide aids for repeated game research.

Keywords Iterated Dilemma; The Cumulative Aspiration Strategy; Lonely Strategy; Simulation Society

With the development of decision support system, the simulation of iterated has attracted substantial interest in recent years. According to member's rulers, we create a simulation model, and from the analysis of the result of simulation, we can provide useful supporting to the problems [3]. Form the Von Neumann model of cellular automata in this paper, based on the members' rational behavior; we establish an iterated prisoner's dilemma model in a simulation society. We join loner strategy in this model to avoid frozen state, because of the members might deceive each other. And, we introduce the mechanism of social movement. Finally, we carry out the computer simulation to demonstrate the influence that the strategy of focusing on the evaluation, mechanisms and the strength of the flow has made on social behavior.

1 Iterated prisoner's dilemma

Prisoner's dilemma (PD) is a classical problem of game theory [2]. In this game, there are two players; they have two strategies, cooperation or defection, denote them by C or D, respectively. But the Players are reluctant to take risks or no hope of participating in social activities frequently [1]. Relying on their own efforts, they only obtain small income but it is safe at least. So, we introduce loner strategy; denote it by L [4]. C and D participates in the game, and L does not to participate in it, only stable but less pay. Then, we can come to the introduction of voluntary participation in the prisoner's dilemma strategy payment matrix (Table 1).

Table 1: Payment matrix

	C	D	L
C	4	0	2
D	5	1	2
L	2	2	2

2 Simulation of the social environment

Model simulation is conducted in the social structure of the population size in 20×20, by adopting Von Neumann neighbors' structure, C, D and L were placed on the lattice 20×20 randomly, as illustrated in the payment matrix of Table 1. Concerning the population mobility in the physical social environment, on the condition of maintaining social stability of the structure, we use a simple and low mobility structure in this model, where the flowing process is described as follows: Swap positions of randomly selected m players with those of randomly selected neighbors.

3 Evaluation strategy

We use the cumulative aspiration strategy [5]. Given the player's individual rational, at the same time, considering the 4 neighbors' strategy and payment in full, and analyzing the players' behavior. The rules can better reflect the actual psychological characteristics of players.

3.1 The level of expectation

Definition: in the s phase of game,E_i^s is the expectation level of player i,p_i^s is the actually paid of player i.Expection diversity is the difference between aspiration levels and actually paid,denote it by ε_i^s,where

$$\varepsilon_i^s = E_i^s - p_i^s.$$

3.2 Model description

We create the repeated game model with N players.Let the players's set is $A = \{1, 2, \cdots, N\}$,the set of all strategy is $O = C, D, L$,$N(i)$ is the set of all player i's neighbors,t_i is player i's strategy,then player i's limited portfolio strategy is $T_i = \{s, (t_i, t_j) | t \in O, i \in A, j \in N(i)\}$.Stage of the payment function is $p_i : T \longrightarrow R$,it is the stage payment of player i,when the portfolio strategy T is achieved.

Now se assume the game in $s = 1, 2, ..., n$ phase for repeat.Player i choice strategy $t = (t_i^1, t_i^1, \cdots t_i^n), t_i^j \in O$ independently,and then get pay $p_i(t)$.We denote game history sets by h:

$$h = \{(t^1, t^2, \cdots, t^s) | s \geq 0, t^{\xi} \in O, \xi = 1, 2, \cdots, s\}.$$

In this model, players will determine the next phase of the strategy through relying on their own game history and pay, and the game history (only the latest three times) and pay of their neighbors.

Let $\eta \in (0,1)$ is the update coefficient of player i's expectation level,$\alpha \in [0,1]$ is the degree what the player emphasize particularly on the greatest history pay or the current payment.Then player i's expectation level on the Pre-s stage as follows:

(1) $s=0, E_i^s(h)=E_i^0$. (E_i^0 is the initial expectation)

$$(2)s=1, M_i(h)=\alpha \cdot \max_{\xi \le s} p_i(t^{\xi})+(1-\alpha)\cdot p_i, M_k=\max_{\substack{s-2\le\xi\le s\\ k\in N(i,h)}} p_k(t^{\xi}),$$

$$E_i^{s+1}(h)=\eta(\eta E_i^s+(1-\eta)M_i)+(1-\eta)(\eta E_k^s+(1-\eta)M_k).$$

After the stage game, it is a repeated game. All players may understand the neighbor's strategy and payment information fully, and thus, player can get the optimal strategy according to actual strategy from the payment which player gets from game completely [3]. Next, player chooses the next stage strategy from it. We establish strategy evaluation function as follows.

$$Ev_i=\begin{cases} 0, s=0 \\ \sum_{\xi=1}^{s} I_{t_i^{\xi}-t_i}\cdot \varepsilon_i^s, s>0 \end{cases}.$$

where,ε_i^s is the difference of expectation,$\varepsilon_i^s=E_i^s-p_i^s, t_i\in O$,

$$I_{t_i^{\xi}=t_i}=\begin{cases} 0, player\ \ i\ \ dosen't\ \ choose\ \ t_i\ \ on\ \ \xi\ \ stage \\ 1, player\ \ i\ \ choose\ \ t_i\ \ on\ \ \xi\ \ stage \end{cases}.$$

Players almost select the best strategy on the next phase of the game, which is the strategy of the lowest cumulative aspiration difference [3]. Here, we join the mobility mechanism in simulation society, which flows as followsčžSwap positions of randomly selected player with those of randomly selected neighbors.

4 Analysis of simulation results

4.1 Analysis of indicators

1. The numbers of players with the strategies (C, D, L) in the simulation society, which is gathered for statistics, when every game is completed. With the times of repeated game be sequence, draw curves for the number of players who adopt three different strategies in various stages.

2. The average scores of the players with the strategy (C, D, L) in the simulation society. Upon completion of each game, the pay matrix of Table 1 is used. Gather of after each player scores each game, and according to the different strategies the player used calculate the average scores of the strategy. Draw curve for average scores (which is completed by using MATLAB) with the times of repeated game being time sequence.

4.2 No punishment mechanism

Figure 1 shows the curve of average scores which players get in all stage of game, and the curve of players' strategy number, in terms of scoring or on the number, D strategy

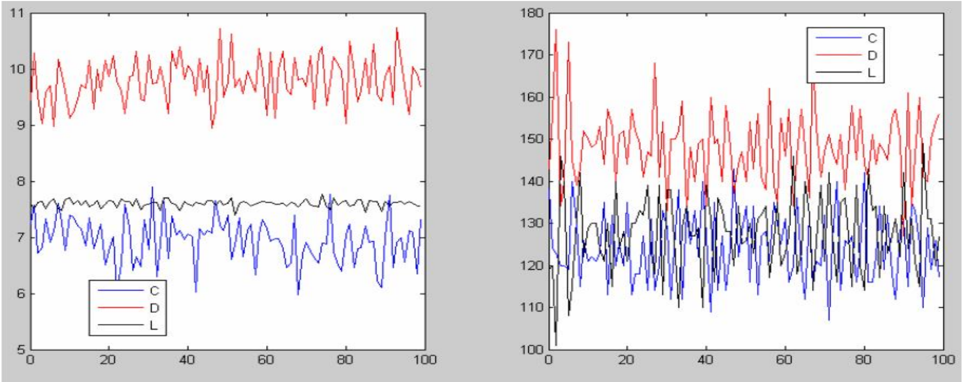


Figure 1: Initial Expectation:$E_i^0 \in [8,16], E_i^0 \in Z$,Update Coefficient:$\eta = 0.8$,Degree:$\alpha = 1$(Player entirely on the history of the greatest value to pay),Repeated Times:100.

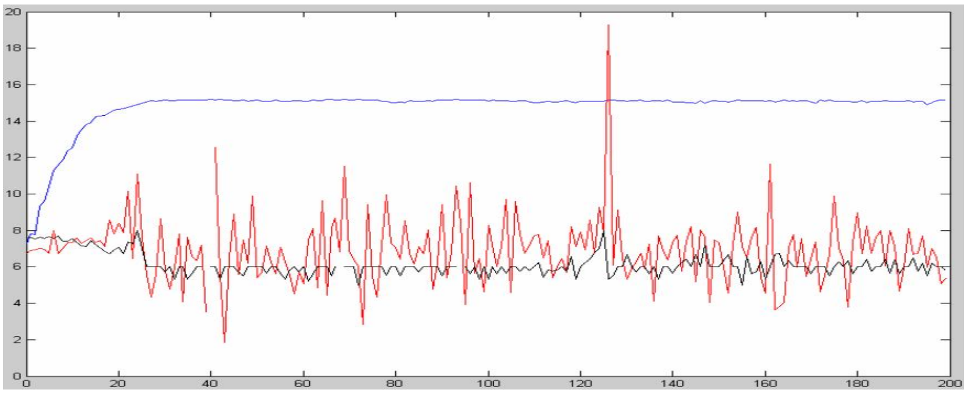


Figure 2: Initial Expectation: $E_i^0 \in [8,16], E_i^0 \in Z$,Update Coefficient: $\eta = 0.8$,Degree: $\alpha = 1$(Player entirely on the history of the greatest value to pay),Repeated Times:200.

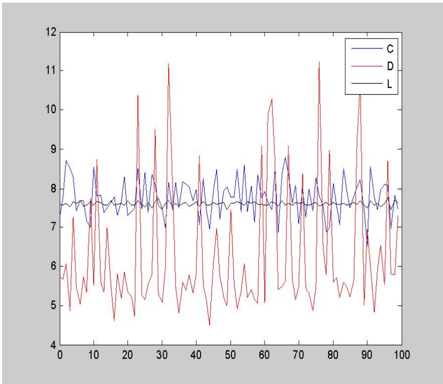
is better than the C strategy in the who simulation process, which is due to that there exits in the prisoner's dilemma game a Nash equilibrium (From Table 1). Because of the introduction of the strategy is lonely, in the whole simulation process, neither side has not completely suppressed the other side. It allows the three to form a strategy similar to the "scissors - a burden - a hammer," the cycle of comparative advantage, in order to effectively avoid the common deception of the "frozen" (A disappeared state that two strategies co-exist, and the system proceeds to an absorbed state of pure strategy).

4.3 Set punishment mechanism

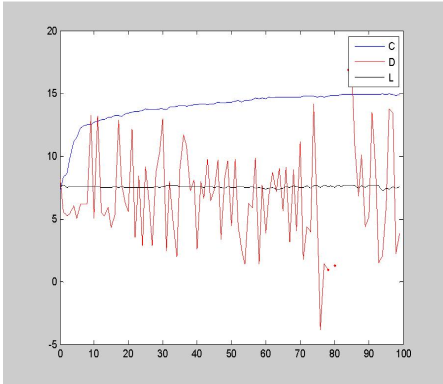
Punishment weight: $w = count\{(t_i = D) AND (t_j = C) | j \in N(i), (t_i, t_j) \in T_i\}$.

Punishment Coefficient: $f = 3$.Update the current player of the actual payment after the punishment to,$p_i^s = p_i^s - w \cdot f \cdot rand, rand \in (0, 2.5)$.

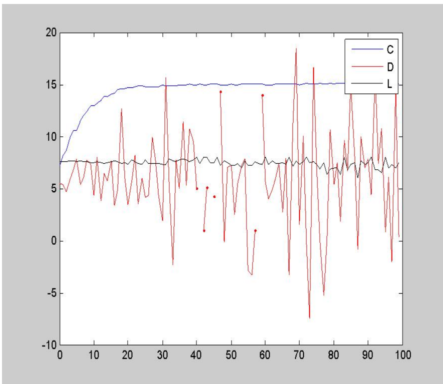
As for C strategy player, his average income is much larger than that of D strategy



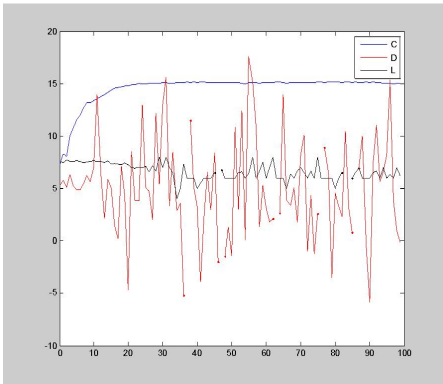
(a)



(b)

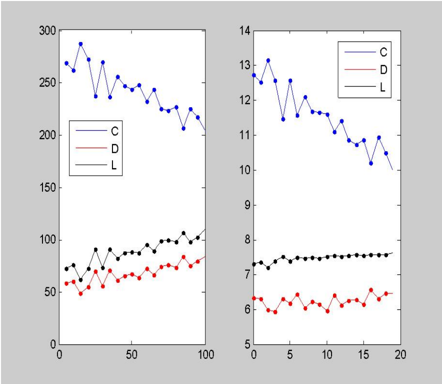


(c)

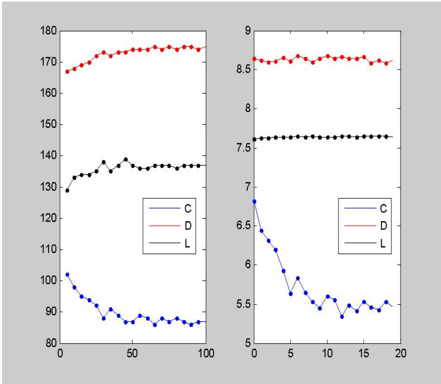


(d)

Figure 3: (a) Initial Expectation: $E_i^0 \in [8,16], E_i^0 \in Z$,Update Coefficient: $\eta = 0.8$,$P = 0.2$,Degree: $\alpha = 1$(Player relies entirely on the history of the greatest value to pay). (b) Initial Expectation:$E_i^0 \in [8,16], E_i^0 \in Z$,Update Coefficient:$\eta = 0.8$,$P = 0.2$,Degree:$\alpha = 0$(Player is totally dependant on the current value of the payment).(c) Initial Expectation: $E_i^0 \in [8,16], E_i^0 \in Z$,Update Coefficient: $\eta = 0.8$,$P = 0.2$,Degree:$\alpha_D = 1, \alpha_{CL} = 0$(D strategy for the player's use is totally dependent on the largest value of payment,and the player who uses C, L strategy is totally dependent on the current payment).(d) Initial Expectation: $E_i^0 \in [8,16], E_i^0 \in Z$,Update Coefficient: $\eta = 0.8$,$P = 0.2$,Degree:$\alpha_D = 1, \alpha_{CL} = 0.15$(D strategy for the player's use is totally dependent on the history of the largest value of the payment,and the player who uses play of the strategy C,L focuses on the current payment).



(a)



(b)

Figure 4:

player, which indicates that the penalty mechanism effectively curbs the emergence of a betrayal, and brings the arrival of co-operation. But in the process of simulation, obviously, we can see that in the first game, the average income of D strategy player gets a sudden increase which is more than that of C strategy player; it is because the punishment failure appears in the process of simulation, causing the betrayal to get a great chance of big earnings.

4.4 Punishment disturbance

If the punishment comes to failure to accidentally or the punishment efforts could't be ensured,we set the possibly P as penalty failure.If the penalty failure then punishment factor will be update as $f = f - rand, rand \in (1,2)$.(The following simulation repeated 100 times)

From the test in terms of the factor , we can see that, under the premise of the accidental penalty failure or the uninsured penalty efforts, the player with the strategy D focus more on the largest history payment. In other words, the utmost payment shall be used to upgrade the expectation value in the next stage, thus forcing the player with the strategy C to focus more on the utmost history payment when upgrading the expectation value in the next stage. Hence, the player with the D strategy is likely to acquire the opportunities for his development, while the player with the strategy C can not proceed with greed, step by step, but use the current payment as the upgraded standard for the expectation value in the next stage. Only in this way can be suppressing the player with the strategy D entirely (see Figure 3(a), (b), (c) and (d)). Although the player with the strategy D may gain more earnings under the extremely accidental circumstances than the player with the strategy C, but in the whole process of simulation there emerges the disappearance of the player with the strategy D (the intermittent part in the figures showing that no players adopt this strategy). Therefore, if the punishment strategy comes to instability, the best way for the player with the strategy C shall not advance blindly, but set the current earnings as the standard for upgrading the expectation valued in the following stage.

4.5 Simulation of Social Mobility

Flowing process is as follows: Swap the positions randomly selected m players with randomly selected neighbors.Let $m \in [5, 100]$,5 grade for a man-made changes, for a total of 20 simulation, set the average numbers of changes in average earnings in the record punished mechanism (Figure 4(a)) and no punishment mechanism (Figure 4(b)). According to the changes in the number of mobile, the analysis indicators and punishment mechanism shown in the section 4.1 draw Figure 4(a), Figure 4(b).

It is not difficult to see that, with the increasing number of mobile players, the average numbers of players and the average scores of players with strategy C is on a downward trend, while those of the player with strategy D is on slowly rising trend. The results show that the flow is not conducive to the emergence of cooperation.

5 Conclusion

From the above simulation, we can see that, it can not get fully success to reach groups cooperation, only by player's individual rational behavior (expectations). While adding external interference (set punishment mechanism), however, we can achieve the effect of groups cooperation successfully, and bring L strategy to avoid the appearance of the "frozen". Of course, there will be punishment disturbance. In this paper we also give the results of the analysis of the disturbance. Finally, we find that, social mobility, especially, the frequent and radical mobility is not conducive to the emergence of cooperation.

References

[1] Gert-Jan de Vreede, DanielT T Van Eijck. Modeling and simulating organizational coordination [J]. Simulation& Gaming, 1998, 29(1): 60-87.

[2] Zhang Weiying. Game Theory and Information Economics M. Shanghai ShanLian Publishing, 1996

[3] Zhang Tao Sun Liyan. A Study of Simulation in Repeated Game Based on the Evaluation of the Cumulative Aspiration Tense Systems Engineering 2002, 111(3):87-89

[4] Li Yun. The Evolution of Cooperation in the Prisoner's Dilemma DaLian University of Technology 2006:11-16

[5] Zhang Fa. Social Cooperation in the Context of Iterated Prisoner's Dilemma: A Simulation Approach Systems Engineering Theory Methodology Applications 2004,11(2):112-119