Memory boosts turn taking in evolutionary dilemma games

Tao Wang a,⁎, Zhigang Chen b, Lei Yang a, You Zou c, Juan Luo a

a College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China
b School of Software, Central South University, Changsha 410083, China
*c High Performance Computing Center, Central South University, Changsha 410083, China

A R T I C L E   I N F O

Article history:
Received 20 September 2014
Received in revised form 30 March 2015
Accepted 30 March 2015
Available online 1 April 2015

Keywords:
Cooperation
R reciprocity
ST reciprocity
P reciprocity
TFT
W5LS

A B S T R A C T

Spontaneous turn taking phenomenon can be observed in many self-organized systems, and the mechanism is unclear. This paper tries to model it by evolutionary dilemma games with memory mechanism. Prisoner’s dilemma, Snowdrift (including Leader and Hero) and Stag-hunt games are unified on an extended S–T plane. Agents play game with all the others and make decision by the last game histories. The experiments find that when agents remember last 2-step histories or more, a kind of cooperative turn taking (CAD) bursts at the area of Snowdrift game with restriction of S+T<2R and S≠T, while the consistent strategy (DorC) gathers on the line of S+T>2R and S=T. We also find that the system’s fitness ratio greatly improved with 2-step memory.

© 2015 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Spontaneous turn taking can be seen as an action of rotation or oscillation in populations (Jiang et al., 2006). This phenomenon has been found in many different systems such as the density oscillator (Steinbock et al., 1998), ticking hour glass (Wu et al., 1993), RNA Polymerase traffic on DNA (Sneppen et al., 2005) and traffic of ants (Dussutour et al., 2005). Turn taking can also solve the problem of pedestrians passing a bottleneck (Burstedde et al., 2001; Helbing and Molnár, 1995; Helbing et al., 2000, 2005).

In this paper we introduced evolutionary game theory to model the action of turn taking in self-organized populations (Gintis, 2000; Hofbauer and Sigmund, 1998; Maynard Smith, 1982; Nowak, 2006a). It may also be able to illustrate the systems mentioned above. Evolutionary game theory (EGT) is the application of game theory to evolving populations of life forms in biology. EGT is useful in this context by defining a framework of contests, strategies, and analytics into which Darwinian competition can be modeled (Maynard-Smith and Price, 1973).

The payoff matrix of a 2-people-2-strategy game is shown in Table 1. There are three kinds of relationships between the players: player cooperates with each other and gets payoff (R (R reciprocity); player defects with each other gets payoff P (P reciprocity); player cooperates with defector gets S and the opponent gets T (ST reciprocity). There exists social metaphor for dilemma situations (Rand and Nowak, 2013) such as Prisoner’s dilemma game (PD, T>R, P>S), Snowdrift game (SD, T>R, S>P), Stag hunt game (SH, R>T, P>S).

In Hawk dove game (also named as snowdrift game, T>R, S>P) (Maynard-Smith and Price, 1973), when facing a new territory, if one player choose cooperation (dove strategy, to avoid combat and may lose the territory), the best strategy for the other is defection (hawk strategy, to combat in any case and may acquire the territory), and vice versa. However, the cooperator usually get lower payoff than the defector which will injure the cooperators' enthusiasm if the game is played repeatedly. Is there any way to help the 2 player to reach a state of "turn taking" in ST reciprocity: the one who cooperate will defect next time; and the one who defect will cooperate next time, and going on?

Some studies revealed the relationship between turn taking and evolutionary games. There are some dilemma games that related to turn taking, including leader game, Hero game and common-pool resources game (Browning and Colman, 2004; Lau and Mui, 2012; Rapoport, 1967). These games all belong to the subtype of Snowdrift game, and the game matrix meets the conditions of T+S>2R.

Crowley et al. (1998) found turn taking phenomenon in evolutionary games. They studied iterated complementarity dilemma game, and observed two special cooperation mechanism: cooperation alternating with defection (CAD), which means that both players taking alternative strategies: (cooperate, defect), (defect, cooperate), ...; the other is defection or cooperation (DorC),

⁎ Corresponding author.
E-mail address: wangtao1999@foxmail.com (T. Wang).
which means that both players taking one strategy and do not change: (cooperate, defect), (cooperate, defect), .... These two mechanisms both belong to ST reciprocity and CAD is a kind of turn taking. We call CAD as cooperative turn taking for it can help agents overcome the dilemma of games like Snowdrift.

Browning and Colman (Browning and Colman, 2004; Colman and Browning, 2009b) studied iterated Prisoner’s dilemma, Snowdrift game, Hero game and Leader game. They made agents remember the last 3-step histories and adopt genetic algorithm as strategy evolutionary rule, and they also observed turn taking strategies like CAD.

Tanimoto and Sagara (2007) defined a special 2-dimensional Dg–Dr game plane, which including Prisoner’s dilemma, Snowdrift, Stag Hunt, Leader and Hero (Fig 1(b)). Wakiyama and Tanimoto (2011) studied the system when agents gamed and updated strategies in different groups on Dg–Dr plane. Agents had one-step memory and CAD strategy was found in the area of harmony game.

An important application of EGT is in the researches of cooperation mechanism (Axelrod and Hamilton, 1981; Nowak, 2006b). Cooperation is the process of groups of organisms working or acting together for their common benefit, as opposed to working in competition for selfish benefit (Kohn, 1992). Since Nowak found that cooperation can exist in evolutionary Prisoner’s dilemma on lattice. In 1992 (Nowak and Robert, 1992), many scientists were interested in the issue of how cooperation exist in rational individuals (Hauert and Szabo, 2005; Nowak et al., 2010; Perc and Szolnoki, 2010; Roca et al., 2009; Szabó and Fáth, 2007).

In the study of turn taking in evolutionary game, memory is an important mechanism. The origin of the study should start from Axelrod’s tournaments in 1980s (Axelrod, 1984, 1987; Axelrod and Hamilton, 1981). He conducted computer tournaments to find the strategy that performs best in a population playing PD game. TFT (Tit-for-tat) won both of the tournaments. However, Nowak found that WSLS (win-stay, lose-shift) outperforms TFT in 1993 (Nowak and Sigmund, 1993). Both TFT and WSLS are based on memory. There are many studies about memory mechanism and strategies (Alonso-Sanz, 2009; Alonso-Sanz and Martin, 2006; Deng et al., 2010; Imhof et al., 2007; Liu et al., 2010; Posch, 1999; Qin et al., 2008; Tanimoto and Sagara, 2007; Wang et al., 2006).

ST reciprocity is a subtype of cooperation. People usually think that bilateral cooperation is the key to overcome the dilemma of 2-person game. However, in the studies of Chen et al. (2013) and Wang et al. (2014), the system’s cooperation ratio is composed by R

---

Table 1
Payoff matrix for a 2 × 2 game.

<table>
<thead>
<tr>
<th>Ego</th>
<th>Cooperation (C)</th>
<th>Defection (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>S</td>
<td>P</td>
</tr>
<tr>
<td>T</td>
<td>P</td>
<td>R</td>
</tr>
</tbody>
</table>

---

Fig. 1. Extended S–T plane. (a) general S–T plane; (b) Dg–Dr plane defined by Tanimoto; (c) unified extended S–T plane. PD is Prisoner’s dilemma game, SD is Snowdrift game, SH is Stag hunt game and HG is Harmony game (without dilemma). In (b) and (c), PD game includes area of $S + T > 2R$, which is not the common definition, we call it extend PD game.
reciprocity and ST reciprocity. Wang et al. (2014) showed that memory promotes ST reciprocity while denotes R reciprocity in Snowdrift game, which cause the decrease of cooperation and the increase of the system's fitness.

The researches above showed that turn taking is possible to evolve in evolutionary games. However, there still remain challenges. The first is that memory's effect to turn taking is not clear in evolutionary games. Some studies fix the memory length to three (Browning and Colman, 2004; Colman and Browning, 2009a) and some others fix it to one (Tanimoto and Sagara, 2007), so different memory length's effect to turn taking is not clear. The second is that distribution of turn taking on general game parameters is not clear. Some papers studied fixed payoff matrix, and some others studied self-defined payoff plane, so how turn taking distribute on the general S–T plane is not clear.

This paper will study the system based evolution system on an extended S–T plane with different memory length. The study contents include CAD and DorC's distribution, the system's cooperation and fitness, and R, ST, P reciprocities' distribution.

2. The model

2.1. Game matrix

In order to reduce parameter, the game matrix \[
\begin{bmatrix}
R & S \\
T & P
\end{bmatrix}
\] usually reduces to \[
\begin{bmatrix}
1 & S \\
T & 0
\end{bmatrix}.
\]
To Prisoner's dilemma (PD), \(T > 1, S < 0\); to Snowdrift(SD) game, \(T > 1, S > 0\); to Stag hunt (SH) game, \(T < 1, S < 0\); to harmony game (HG) without dilemma, \(T < 1, S > 0\) (Fig. 1(a)).

In order to define dilemma, Tanimoto defined two indicators (Tanimoto and Sagara, 2007):

\[
\text{Dg} = T - R
\]
\[
\text{Dr} = P - S
\]

If one of the indicators is greater than zero, the game is a dilemma game. To the definition, PD (\(Dg > 0, Dr > 0\)), SD (\(Dg > 0\), including leader and hero) and SH (\(Dr > 0\)) are all dilemma games. Letting \(R = 1, P = 0\), he got the payoff matrix \[
\begin{bmatrix}
1 & -Dr \\
1 + Dg & 0
\end{bmatrix}.
\]
Based on the matrix, Tanimoto defined a 2-dimensional plane. It is obvious that Dg = T - 1, Dr = -S, so the S–T plane and Dg–Dr plane are actually the same.

This paper unified the general S–T plane and the Dg–Dr plane to extended S–T plane (Fig. 1). As shown in Fig. 1(c), dilemma games including Prisoner's dilemma, Snowdrift (including leader and hero) and Stag hunt are all on the plane.

2.2. Memory mechanism

Following previous work (Chen et al., 2013; Wang et al., 2014), we design a memory mechanism to code the game history and strategies. One-step memory means that agents can remember history for the last game. Cooperation is coded as “1” and defection as “0”. Thus, there are four possible historical interactions for one-step memory (Table 2): 00 (the focused player defects and the opponent defects), 01 (the focused player defects and the opponent cooperates), 10, and 11. Depending on the history, agents choose to cooperate or defect. Thus, the strategy can be coded as “***”, where * represents 0 or 1. The position of * indicates a particular history. Thus, this 4-bit encoding can represent 16 different strategies. For example, 1001 is WSLS strategy and 1010 is TFT (high bit on the left). Table 2 shows the codes for always cooperates (ALLC), always defects (ALLD), TFT, and WSLS strategies. When agents can remember more than one-step, the relationship between memory length and strategies is shown in Table 3.

2.3. Strategy update rule

There are many strategy update rule such as replicator rule, unconditional imitation rule, etc. (Perc and Szolnoki, 2010). In the studies of repeated dilemma game with memory, genetic algorithm is a common update rule (Axelrod, 1987; Browning and Colman, 2004; Colman and Browning, 2009b; Crowley, 2001; Tanimoto and Sagara, 2007).

In this paper we adopt genetic algorithm as the update rule. There are three steps of the algorithm: selection, crossover and mutation. The first step is selection. Every agent randomly selected a neighbor. If the neighbor's payoff was higher than the focal agent, then the following steps are executed; else the focal agent keeps the strategy unchanged and go to the third step. The second step is crossover. The focal agent and the neighbor's strategies are coded into binary strings. The position of crossover is at the middle of the strategy. One of the two children strings is randomly selected as the strategy of next generation. When memory length = 0, and the strategy is only one bit, then it had 0.5 ratio to change to the neighbor's strategy. The third step is mutation. Every bit of the new strategy has a mutation rate of 0.01. A sketch map of crossover and mutation is in Fig. 2. We discussed the algorithm in Section 4.

### Table 2
Examples of the coding of strategies when \(m = 1\).

<table>
<thead>
<tr>
<th>Game history of two players</th>
<th>Payoff of the first one</th>
<th>The next step of the first one</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ALLC</td>
</tr>
<tr>
<td>00 (defect, defect)</td>
<td>P</td>
<td>1</td>
</tr>
<tr>
<td>01 (defect, cooperate)</td>
<td>T</td>
<td>1</td>
</tr>
<tr>
<td>10 (cooperate, defect)</td>
<td>S</td>
<td>1</td>
</tr>
<tr>
<td>11 (cooperate, cooperate)</td>
<td>R</td>
<td>1</td>
</tr>
</tbody>
</table>

### Table 3
The relationship between memory length and coding of strategies (bit).

<table>
<thead>
<tr>
<th>Memory length</th>
<th>Coding length of game history</th>
<th>Strategies’ length</th>
<th>Number of strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>(2^1)</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>4</td>
<td>(2^4)</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>16</td>
<td>(2^{16})</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>64</td>
<td>(2^{64})</td>
</tr>
</tbody>
</table>
The process of the experiment is as follows:

1. For different memory length from 0 to 3, the following steps ((2)–(4)) is done.
2. We sampled the extended S–T plane every 0.5 unit, and the total samples are 15 × 15 = 225. Every sample point corresponding to a specific payoff matrix, and at each sample point, the following steps ((3)–(4)) was done.
3. N = 2000 agents are well-mixed (to exclude the network structure’s effect), and each focal agent picks up her opponent randomly. All the agents gained Lr = 1000 generations (step (4));
4. Each agent randomly selected 100 neighbors and gamed 100 times with them. Then genetic algorithm was executed to generation new generation.
5. All the following statistic data were based on 100 samples (steps (1)–(4)).

### 3. Experiments and analysis

#### 3.1. Turn taking strategy (CAD) and consistent strategy (DorC)

Turn taking is the focus of this paper. Previous researches concentrate on 1-step memory and 3-step memory, and the relationship between memory length and spontaneous turn taking is not clear. We studied the evolutionary system with memory length from 0 to 3. When system run to a stable state (1000 generations), we analyzed the last 6-step history. If the history looks like 011001100110 or 100110011001, then it is CAD (Table 4); if the game history looks like 0101010101 or 101010101010, it is DorC. It is not accurate to study CAD and DorC just by strategies. The ratio of CAD and DorC are defined as follows:

\[ r_{CAD} = \frac{n_{CAD}}{n_{total}} \]  \hfill (1)

\[ r_{DorC} = \frac{n_{DorC}}{n_{total}} \]  \hfill (2)

where \( r_{CAD}, r_{DorC} \) are the ratio of CAD and DorC, \( n_{CAD}, n_{DorC} \) are the number of histories like CAD, DorC, and \( n_{total} \) is the total number of histories between all the nodes.

We show the distribution of turn taking (CAD) in Fig. 3. We found that when agents remember 2-step history, spontaneous turn taking spring up on the S–T plane. When memory length \( m = 0 \), CAD does not exist; When \( m = 1 \), CAD exist in a small area in SD game. When \( m = 2 \), the distributions change drastically. CAD expands rapidly to the area of \( S + T > 2R \). When \( m = 3 \), the result is similar to \( m = 2 \).

We also show the distribution of consistent strategy (DorC) in Fig. 4. When memory increases from 0 to 1, DorC increases at the area of SD game; while when \( m = 2 \) and 3, DorC strategies gathers on the line of \( S = T \).

Some results need further analyses. In Fig 4(c) and (d), DorC is on some discrete points near the axis. In fact, in the experiment, the sampling density on the S–T plane is every 0.5 unit one point. So these discrete points in Fig 4 is the sampling points, and DorC actually distribute strictly on the line \( S = T \) in SD game area.

This experiment shows that when agents have 2-step memory, turn taking (CAD) will emerge on S–T plane with the condition: \( S + T > 2R, S \neq T \). This area includes most of SD game, all the leader and hero game, and even includes part of PD and HG.

#### 3.2. Cooperation ratio and system’s fitness

In the study of dilemma game, system’s cooperation ratio and average fitness are important factors. Cooperative turn taking is a component of system’s cooperation. The system’s cooperation

![Fig. 2. Sketch map of strategy crossover and mutation.](image)
**Fig. 3.** Distribution of spontaneous turn-taking (CAD) on the extended S–T plane with memory length \( m = 0, 1, 2, 3 \).

**Fig. 4.** Distribution of consistent strategy (DoC) on the extended S–T plane with memory length \( m = 1, 2, 3 \).
The distribution of system’s cooperation ratio and fitness with different memory length on the extended S–T plane. The first column is cooperation ratio and the second column is fitness. Line 1–4 display memory length from 0 to 3, respectively.

ratio $r_c$ and average fitness $\pi_{\text{ave}}$ are defined as follows:

$$r_c = \frac{n_c}{(n_c + n_d)}$$

$$\pi_i = \frac{p_{fi}}{(k \times \ell)}$$
Fig. 6. Theoretical max fitness $F_{\text{max}}$ on extended S–T plane. We got it by Formula (9).

$$\pi_{\text{ave}} = \frac{\sum_{i=1}^{N} \pi_i}{N}$$

(8)

where $n_c$ ($n_d$) is number of cooperation (defection) strategies adopted by the agents in a game round. $\pi_i$ is the fitness of the agent $i$, $\pi_{\text{ave}}$ is the system’s fitness, where $P_{fi}$ is the total payoffs agent $i$ get in a round, $k$ is agent $i$’s degree and $t$ is the repeated game times in a round. $N$ is the total number of agents.

We studied the system with memory from 0 to 3 and got Fig. 5. We can see that when memory increases step by step, cooperation ratio increases rapidly at the area between PD and SH (area B); while between SD and HG (area A), cooperation is decreased. That means memory not necessarily promotes cooperation in dilemma games (Wang et al., 2014). We can see that increasing memory promotes the system’s fitness.

We make further study to see how much memory affects system’s fitness. Firstly we get the theoretical max fitness $F_{\text{max}}$ by the definition:

$$F_{\text{max}} = \left\{ \begin{array}{ll} R, & S + T \leq 2R \\ \frac{S + T}{2}, & S + T > 2R \end{array} \right.$$

(9)

Then we get Fig. 6. Next we divided the system’s fitness by $F_{\text{max}}$ and got fitness ratio (Fig. 7). We can see that when agents remember no less than 2-step history, the system get high fitness ratio in the whole game area. It is interesting that fitness ratio adjacent to the line of $S+T=2R$ is lower than other area. It means that agents are in chaos of whether to cooperate or turn taking with others at these areas. We can also see that cooperation ratio and fitness ratio are not synchronously increasing with memory’s increase.

3.3. Three kinds of reciprocity R, ST, P

As we have introduced that turn taking is a subtype of ST reciprocity, studying the distribution of R, ST and P ratio laid a foundation for the further study of turn taking. It can also help us understand the composition of cooperation.

Fig. 7. Fitness ratio on the extended S–T plane with different memory length. We divided the fitness in Fig. 5 by $F_{\text{max}}$ of Fig. 6 and got this figure.
In 2-people-2-strategy game, players have a choice to cooperate, C, or to defect, D. Then there are three kinds of reciprocity between two players: R reciprocity (players both chose C, they get payoff R), P reciprocity (players both chose D and get payoff P) and ST reciprocity (one player chose C and the other chose D, they get S and T). When we talk about cooperation ratio, it actually includes 2 parts: all the R reciprocity and half of the ST reciprocity. These reciprocities have been studied by some researchers (Browning and Colman, 2004; Crowley, 2001; Tanimoto and Sagara, 2007).

The relationship between cooperation ratio and the ratio of R, ST and P can be described as follows:

\[ r_c = \frac{r_R + r_{ST}}{2} \]  \hspace{1cm} (9)

\[ r_d = \frac{r_P + r_{ST}}{2} \]  \hspace{1cm} (10)

\[ r_{ST} = r_S + r_T = r_S \times 2 = r_T \times 2 \]  \hspace{1cm} (11)

\( r_c, r_d, r_R, r_S, r_T, r_P, r_{ST} \) means the ratio of cooperation, defection, R, S, T, P and ST reciprocity. The result is in Fig. 8.

We can see from the figure that when memory increases, ST reciprocity increases dramatically. When memory length \( m \geq 2 \), ST reciprocity mainly distributes at the area of \( S + T > 2R \). Memory’s effect to R reciprocity is complex; at the area between SD and HG, R reciprocity decreases with memory’s increase; at the area between

Fig. 8. Distribution of R, ST, P ratio with different memory length on extended S–T plane. Column 1, 2, 3 are the distribution of R, ST and P, respectively; line 1, 2, 3, 4 display memory length from 0 to 3, respectively.
PD and SH, R reciprocity increases with memory’s increase. P reciprocity monotonically decreases with memory’s increase. When there is no memory, we found that even when add the ratio of CAD and DorC together, the sum is less than the ratio of ST reciprocity. It means that there exist other erratically ST reciprocity (for example, the game history looks like 010110, 011010, 101001, etc., in 3-step history) when agents have no memory.

4. Conclusions and discussion

We studied evolutionary dilemma game on extended S–T plane with memory mechanism. There are mainly two findings in this work:

1. When agents can remember 2-step history or more, turn taking (CAD) occupies S + T > 2R and S ≠ T on the extended S–T plane (which even including part of extended Prisoner’s dilemma game and Harmony game).

2. When agents can remember 2-step history or more, consistent strategy (DorC) occupies S + T > 2R and S = T on the extended S–T plane.

3. When agents can remember 2-step history or more, the system gets high fitness ratio (>0.6) in all the dilemma games. The result shows that cooperation ratio and fitness are two different factors to an evolutionary game system.

Number of strategies is relatively large vis-à-vis the number of agents N = 2000. For instance, in case of 4-step memory, the number of strategies is $2^4 \times 2^4 = 256$. That means many good strategies may not join the system’s evolution. However, two factors help the system test more strategies than the number of agents: the high mutation rate ($r = 0.01$ at each bit of the strategy) when agents update their rule and all the results are based on 100 samples.

We do not select the agents proportional to their fitness as the common genetic algorithm did, but the agents are randomly selected. Because in well mixed population, some good strategies will die earlier if they are selected proportional to the fitness. We have also do experiments with the common genetic algorithm, and the results are not so clear to support the conclusions. It seems that 2-step memory can overcome the dilemma of Snowdrift game and help agents get high payoff. The result maybe can illustrate the turn taking phenomenon in biological systems; also it may give hints to manmade self-organized system, such as self-organized network (Liu et al., 2013, 2014), multi-robot system, etc. As Snowdrift game models the dilemma of resources competition and conflict, the study maybe could shed light to the regional conflicts and wars all around the world today.

Acknowledgements

We thank advices given by the two anonymous reviewers, and the editor’s kind help for polishing the paper. This work is supported by “Fundamental Research Funds for the Central Universities”, partly supported by the National Natural Science Foundation of China under Grant 61379110, 61379058, 61272062, 61202289, the Priority Development Areas Project 20120162130008 of Chinese Ministry of Education and the fundamental research funds for the Central Universities of China.

References


Wakiyama, M., Tanimoto, J., 2011. Reciprocity phase in various $2 \times 2$ games by agents equipped with two-memory length strategy encouraged by grouping for interaction and adaptation. Biosystems 103, 93–104.

