Robustness against Longer Memory Strategies in Evolutionary Games.

Eizo Akiyama *

1 Players as finite state automata

In our daily life, we have to make our decisions with our restricted abilities (bounded rationality), which may be developed through learning and evolution that utilize our *past experiences*. Recently, such evolutionary phenomena have been studied by many researchers using evolutionary game settings, where *repeated games* are played by finite state automata (players), especially after the analysis and computer experiment on the repeated prisoners dilemma by [?].

In this work, we discuss the *effective memory size* of finite automata to remember past experiences, and show that if fthere exist a Neutral Stable Strategy of the effective memory size that is also in Pareto optimal state, then it cannot be invaded by any players with longer memory size. We also discuss the possibility of open-ended evolution when there isn't any effective memory size for a game.

2 Memory, strategy, game structure and social phenomena

Let us consider a game whose payoff matrix is shown in Table **??**. (The Avatamsaka Game [?] In this game, a player's point depends only on the other's behavior. If player 2 chooses behavior D (defect), the point player 1 can gain is only 1.0 no mater what behavior he or she chooses. If player 2 chooses behavior C (cooperation), the point for player 1 is always 2.0. Therefore, the game theoretical conclusion is that there is no motive for each player to change his/her current behavior in this game. In other words, any choice of any player is best response to the other's choice in every situation, and a set of the mixed strategies of all the players should always be a Nash Equilibrium. However, if we consider the case where palyers' have memories of past behaviors, the result will be quite different one line the prisoner's dilemma case, which will be discussed later.

*Institute of Policy and Planning Sciences, University of Tsukuba, Tel +81-298-54-9455, Fax +81-298-53-3848, eizo@sk.tsukuba.ac.jp

Table 1: Payoff matrix for the "Avatamsaka game": Two matrices below show the points gained by Player 1 (left) and Player 2 (right). The columns indicate the actions Player 1 would choose and the rows Player 2. As we only deal with symmetric games (i.e. the payoff structures from the viewpoint of Player 1 and of Player 2 are identical), such as the game in this table, we will show the payoff matrices only for Player 1 in the rest of this paper.

The payoff for player 1

	D	C
D	1	2
C	1	2

The payoff for player 2

	D	C
D	1	1
C	2	2

Besides the Avatamsaka game, we introduce typical four 2x2 symmetric dilemma games, Prisoner's dilemma, Chicken, Leader's and Hero's dilemma games (following the classification of [**?**]) to see the effect of game structures on the evolutionary dynamics. The payoff matrices are given in Table **??**. While Avatamsaka game has a non-generic[1]payoff matrix, Rapoport classified generic and symmetric 2x2 games and found out only four games are non-trivial.

Let us explain briefly about the games (c) and (d). (We assume (a) Prisoner's dilemma and (b) Chicken game are well known.)

(c) In the Leader's dilemma game, D represents "to act positively and to wish to be a leader," while C represents "to act passively and to follow the leader." If both can successfully differentiate their roles into the *leader* and the *follower*, the productivity of the entire society will be improved. However, it should be noted that the leader can get more profit than the follower can in this game. On the other hand, if the both insist on becoming the leader, it turns out to be the worst result. The Leader's dilemma game can be applied as a model for power struggles.

(d) Also in the Hero's dilemma game, D is "to act positively" and C is "to act passively." However, the follower (passive) is more advantageous than the leader (positive) in this game. Here, action D represents "to contribute actively to the society with self-sacrifice," which should be called the *heroic behavior*. However, if both of them try to act heroically, it turns out to be a waste. The Hero's dilemma game may be applied as a model for the president election in a residents' association, in which nobody is willing to be elected, but somebody has to.

3 The model

We deal with *multi-stage (or repeated, iterated) games* of the above five 2x2 games (Avatamsaka + 4 dilemma games) in this paper, and we call one play

[1]There exist payoff ties.

Table 2: 2x2 dilemma games (based on the classification of [?])

(a)Prisoner's dilemma

	D	C
D	1	5
C	0	3

(b)Chicken game

	D	C
D	0	5
C	1	3

(c)Leader's dilemma

	D	C
D	0	5
C	3	1

(d)Hero's dilemma

	D	C
D	0	3
C	5	1

in the repetition *a stage game*. Furthermore, we refer to C or D in a stage game as *an action*, while a complete plan in a multi-stage game for a player to decide his/her action based on the past information as *a strategy*. We investigate the evolutionary phenomena in these five games using an evolutionary model presented by [?].

3.1 Dynamics of the frequencies of the strategies

Suppose that there are infinitely many agents in a population of a game world. Two of the Agents living in the game world are chosen randomly from the population to play an infinitely repeated 2x2 game. They decide their actions according to their own strategies during the course of the repetition. Their strategies can be classified into several categories in the game world. Suppose that the total population of the agents in the game world is 1.0 and the fraction of the agents having strategy $i\,(i = 0, 1, \ldots, n-1)$ in a specific generation t is x_i. The fitness f_i of strategy i acquired in the game world (the expected value of the payoffs in all the games in which strategy i is involved), of course, depends on the strategy distribution $\boldsymbol{x} = (x_0, x_1, \ldots, x_{n-1})$. (e.g. The fitness of a strategy would be decreased if many unfavorable strategies should dominate the game world.)

Assume that the population share of a strategy increases or decreases according to the fitness of the strategy and that a small fraction of newborn agents appear as mutants of other strategies. Considering the fact that the average fitness of all agents is $\frac{x_i f_i}{\Sigma_i x_i f_i}$, the population share x_i' of strategy i in the next generation would be:

$$x_i' = \left(\frac{f_i}{\Sigma_i x_i f_i} x_i + u\right) / (1 + nu),\ i = 0, \ldots, n-1.$$

where u is the constant term to represent the effect of the uniform and time-independent mutation (i.e. mutation rate). In this paper, we choose two cases $u = 0$ (no mutation) and $u = 0.0004$ in order to see the effect of the mutation. (The change of u does not affect the results essentially as long as $0 < u << 1$ in most cases.

3.2 Iterated game with action-noise and the strategies with past memory

In the game world, agents play iterated 2x2 games where infinitely many rounds of stage games are repeated. Each agent has the strategy to decide the actions in the next stage game based on the memory of the actions that he and the opponent did in previous rounds. We investigate the following three cases: (1) Agents decide their actions without memory. (This condition is denoted as "$m = 0$" case.) (2) Agents decide the action referring to the previous action of the opponent ("$m = 1$") (3) Agents decide the action referring to the actions of himself and of the opponent ("$m = 2$").

The number of possible pure strategies in this repeated game depends on the agents' memory size. In the case of no memory ($m = 0$), there are only two possible pure strategies: "Always D (which we denote as "AIID" or "S_0^0")" and "Always C (AIIC= S_1^0)." [2] In the case of $m = 1$, we denote a strategy of an agent as p_0p_1 ($p_0, p_1 \in [0, 1]$) where $p_0(p_1)$ is the probability to play "C" when the opponent's last move was "C" ("D"). For example, the strategy "00" (i.e. $p_0 = 0$, $p_1 = 0$) always play "D." There are four possible pure strategies for $m = 1$ such as 00, 01, 10, 11, which we denote as S_0^1 S_1^1 S_2^1, S_3^1 respectively. Note that 00=S_0^1, 11=S_3^1 are same as AIID and AIIC. 10=S_2^1 means that the agents act the same as the opponent did previously, which is so-called "Tit-For-Tat (TFT)". 01=S_1^1 means the opposite action from the opponent's previous action, and this is "Anti-TFT (ATFT)".

The strategy in the case of $m = 2$ can also be described as p_0 p_1 p_2 p_3 ($p_i \in [0, 1]$). p_0, p_1, p_2 and p_3 represent the probability to play "C" when, in the previous round, (the agent's move, the opponent's move) = (C, C), (C, D), (D, C) and (D, D), respectively. There are 16 possible pure strategies such as 0000, ..., 1111 = $S_0^2 \cdots .S_{15}^2$. Thus, the number of possible pure strategies in total is 2^m for the memory size m of an agent.

To make it simple, we assume that each game would be repeated *infinitely* between two agents chosen from the population. Moreover, agents would make mistakes with probability p during the repetition. (The agent with the strategy to play "C" may sometimes play "D" by mistake against his/her will.) This means that the uncertainty (*action-noise*) exists in agents' cognitive functioning. Taking into account the fact that the transition of the probability distribution for the 2^m states above is a Markov process, the transition matrix can be defined uniquely depending on the two agents' strategies. The probability distribution in the steady state corresponds to the eigenvector with the eigenvalue 1 of the transition matrix as long as $p > 0$. The average payoff in an iterated infinitely game can be given from the probability distribution of the steady state and the stage game payoff matrix. A slight change of p does not essentially affect the analysis result introduced in the next section as long as $p << 1$. (The theoretical values found in the next section such as the equilibria of game dynamics are the values at $p \to 0$ limit. We use $p = 0.01$ for the computer simulations.)

[2] S_n^m means the nth strategy with memory size of m.

As a short note for strategies at $m = 2$, S^2_0=0000 and S^2_{15}=1111 mean AllD, AllC and S^2_{10}=1010 corresponds to TFT, S^2_5=0101 to ATFT. Moreover, S^2_8=1000 is "GRIM" (that play "C" only when the previous choices of the both players were "C"), S^2_9=1001 is so-called "PAVROV" [**?**]. The last two strategies can be formed only if $m \geq 2$. Note that the nomenclature used in the above is the one usually used in the "Prisoners' dilemma" researches, and so, it might be sometimes misleading to use the above strategy names in *other* games than Prisoner's Dilemma, although we do basically use the above names in this paper to avoid confusion.

4 Brief summary of the results

Let us briefly sum up the results.

In case of Avatamsaka game and of Chicken game, the effective memory size is $m = 2$. That is, $m = 2$ strategy PAVROV is a neutrally stable strategy (NSS) and does not allow the prevalence of *ANY* strategy including those with longer memory size. In case of Leader's or Hero's dilemma game with action noise, the effective memory size is $m = 4$. That is, there is a $m = 4$ strategy that is NSS and does not allow the prevalence of any strategy including those with longer memory size.

It can be shown that if there exists, for some memory size, a NSS that is in a Pareto optimal state, then it should be NSS for any strategy with any memory size. This result implies that further evolutions that increase memory size are meaningless, and evolutionary phenomena for such games always have some end-points at NSSs.

On the other hand, such NSS cannot be found for prisoner's dilemma with action noise, and actually we can observe open-ended evolution with the increase of the memory size in the computer simulation of this game.

References

[Aruka 2001] Aruka, Y., "Avatamsaka Game Structure and Experiment on the Web," in Aruka, Y.(ed), Evolutionary controversies in Economics, Srpinger, Tokyo, 2001, pp. 115-132.

[Axelrod 1984] Axelrod, R., "The Evolution Of cooperation." 1984: Basic Books.

[Rapoport and Guyer 1966] Rapoport, A. and M. Guyer (1966) "A Taxonomy of 2×2 games" General Systems 11:203-214.

[Nowak and Sigmund 1993a] Nowak, M. A. and Sigmund, K. (1993), "Chaos and the evolution of cooperation." Proc. Natl. Acad. Sci. USA 90: 5091-5094.

[Nowak and Sigmund 1993b] Nowak, M. A. and Sigmund, K. (1993), "A strategy of win-stay lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game." Nature 364: 56-58.