Evolution of Reciprocal Cooperation in the Avatamsaka Game

Eizo Akiyama[1] and Yuji Aruka[2]

[1] Institute of Policy and Planning Sciences, University of Tsukuba, Tel
  +81-298-54-9455, Fax +81-298-53-3848 `eizo@sk.tsukuba.ac.jp`
[2] Faculty of Commerce, Chuo University, `aruka@tamacc.chuo-u.ac.jp`

Summary. The Avatamsaka game is investigated both analytically and using computer simulations. The Avatamsaka game is a *dependent game* in which each agent's payoff depends completely not on her own decision but on the other players'. Consequently, any combination of mixed strategies is a Nash equilibrium.

Analysis and evolutionary simulations show that the socially optimal state becomes evolutionarily stable by a Pavlovian strategy in the repeated Avatamsaka game, and also in any kind of dependent game. The mechanism of the evolutionary process is investigated from the viewpoint of the agent's memory and mutation of strategies.

1 The Avatamsaka Game as a dependent game

1.1 Robinson Crusoe economy and multiple-person games

In our daily life, we must sometimes make decisions to improve our current situation. If utilities for the results depend only upon our own decision, the problem is usually called an optimization problem, which a solitary person, such as Robinson Crusoe (or a person whose payoff is unrelated to the decisions of other people) meets ("Robinson Crusoe economy" [Neumann and Morgenstern 1944]). What matters first in this kind of situation is the possibility of constructing the function itself. Once the function is constructed, solving for the optimal behavior is a computational problem to determine the strategy that optimizes the function. As von Neumann and Morgenstern pointed out in their book "Theory of Games and Economic Behavior," the difficulty in solving for the optimal strategy in one-person games, such as the Traveling Salesman Problem, is a *technical* rather than *conceptual* one.

In multiple person problems, however, the result for each one will depend in general not merely upon his decision but on those of the others as well. Moreover, all participants desire maxima simultaneously. Consequently, it is

usually not possible for a player to decide his best behavior without considering others' decision-making processes. Such a situation in general is called a *game*[1].

1.2 independent / dependent game

Let us refer to the Robinson Crusoe problem, in which decisions of others do not have any effect on his utility, as *an independent game*; Robinson Crusoe's effort can be termed *independent optimization*. (In the context of the evolutionary model, it will be called an *independent adaptation*.) In this paper, we consider a class of games that is opposite of the independent game: *dependent games.*

In a dependent game, the result for each agent will depend merely on the decisions of others, but not at all on his own decision. Every player's effort to increase his profit in a dependent game is called a *dependent optimization*. (In the context of the evolutionary model, it will be called a *dependent adaptation*.) Any multiple-person problem is representable as a formulation that lies between the independent game and the dependent game, in principle.

By definition, any change of a player's action in a dependent game does not directly improve her utility. However, if the game is repeated there is a chance her current action might serve as a signal to induce beneficial behaviors from others in the future.

For an example of dependent games, let us consider internet auction sites. Many internet auction sites adopt a "user rating system," with which each buyer rates his experiences with sellers, in order to prevent dishonesty among sellers. In this case, rating sellers high or low does not affect the current utility of the buyers at all, which exemplifies the dependent game situation. Note that not a few members of auction sites *repeatedly* participate in auctions there. As a result, the rating behaviors are repeated for the future, which usually motivates the buyers to give honest ratings. If their participation were not repeated, it would not matter for a player, for example, to rate a good seller "low" without any consideration (and forever leave the auction site).

1.3 The Avatamsaka game

As the simplest example of dependent games, let us consider the *Avatamsaka Game*[2][Aruka 2001]. Its payoff matrix is shown in Table 1. In this game, a

---

[1]In the theory of games, von Neumann and Morgenstern had an insight into the fact that this problem is not an ordinary maximization problem and that the formalization of this type of problem should take the form of a *matrix*. As a result, they succeeded in formalizing the problem as one in which every individual *"can determine the variables which describe his own actions, but not those of the others."* In addition, *"those alien variables cannot, from his point of view, be described by statistical assumptions."*

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 that Player 1 would choose and the rows indicate those for Player 2. Because we only deal with symmetric games (i.e. the payoff structures from the viewpoints 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 remainder of this paper.

Payoff for player 1

	D	C
D	1	2
C	1	2

Payoff for player 2

	D	C
D	1	1
C	2	2

player's points depend only on the other agent's behavior. If player 2 chooses behavior D (defect), the point player 1 can gain is only 1.0, no matter what behavior player 1 chooses. If player 2 chooses behavior C (cooperation), player 1 is bound to receive 2.0.

In this game, any combination of both players' mixed strategies is a *Nash equilibrium* because neither player has an incentive to change their strategy. In addition, any mixed-strategy is a *neutrally stable strategy* from the viewpoint of evolutionary game theory.

1.4 Repeated interactions in a society, players' cognitive abilities, and mistakes

Let us imagine that we are actually under circumstances similar to the Avatamsaka game. In this case, various thoughts might come to mind such as "Should I be satisfied with his defection, even though we are in a Nash equilibrium state?" "Can I somehow induce his cooperation?" "Can I somehow use my behavior as a signal to him?" etc. The reason we may not be satisfied with our current game play is that, if we actually encounter a situation like

---

[2]Avatamsaka is a well-known Mahayana Buddhist Sutra. The late Shigeo Kamata, Professor Emeritus, University of Tokyo, who was working under the field of Buddhist philosophy, skillfully illustrated the situation of Heaven and Hell in view of Avatamsaka. (See Kamata [Kamata 1988], pp.167-168.)

Suppose that two men are seated across from each other at a table. They are bound with ropes except for one arm. Each is given an overly long spoon. Thereby, they cannot serve themselves because of the awkwardness of the long spoon, but there is sufficient food for both of them on the table. If they cooperate in concert to provide each other with a meal, they can both be happy. This defines the Heaven, or Paradise. Alternatively, one will be kind enough to provide the other with a meal, but the other might not have the same feeling of cooperation. This case pays the other, which must give rise to a feeling of hate for the inconsiderate opponent. This describes a situation denoted as Hell. (See [Aruka 2001] for the philosophical and sociological significance of this game.)

this game, something other than the activity of "analyzing the Table 1 in the way of traditional game theory" appears in our mind, which might cause mutual cooperation under Avatamsaka game interactions.

For example, we usually interact not only with a certain individual, but also with other community members in the real world. Therefore, it is necessary for us to consider, comprehensively, relationships with various types of individuals (*various interactions in a society*). In addition, we sometimes, or often, consider not only current, but also future relationships with the person (*repeated interactions*). In this case, although my current action has no meaning for me now, but it might have some meaning as a "signal" to derive cooperation of the opponent in the future. Whether or not repeated interaction works as a signal to derive mutual cooperation depends on the perceptions and cognition of both players. Note that, in the course of repeated interactions, we must also consider the fact that we are beings who cannot be free from mistakes in the real world. And so, cognitive abilities to deal with players' mistakes are crucial for formation of cooperation (*players' cognitive abilities and mistakes*). Furthermore, we do not fix our strategy. We sometimes change our strategy to adapt to our current situation (*mutation / modification of the strategies*).

This paper investigates the following points to study the evolution of cooperation under a dependent game: the Avatamsaka game.

1. How can we derive cooperative behavior of the opponents in Avatamsaka situations, when we live in a society in which we interact repeatedly with various players who might sometimes make mistakes? Players surely need cognitive abilities to induce another's cooperation. To what extent are such abilities required for deriving cooperation?
2. How do social phenomena change with players' cognitive abilities? For example, an agent who remembers only the previous actions of the opponents, such as so-called "Tit-For-Tat strategy" in the Prisoner's Dilemma game, may bring about completely different social phenomena from agents who remember the previous behaviors of themselves and of their opponents.
3. Do capricious changes of strategies (mutation) have a fundamental effect on the evolutionary phenomena?

2 The model

This study addresses a *multi-stage (or repeated, iterated)* Avatamsaka game; we refer to one iteration in the game as 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 action based on the past information as a *strategy*. We investigate evolutionary phenomena in the Avatamsaka game using an evolutionary model whose formulation was presented by [Nowak and Sigmund 1993a] for the study of the Prisoner's Dilemma game.

2.1 Evolutionary dynamics with mutational effect

Suppose that there are infinitely many agents in a population (a game world). All agents in the population play a round-robin tournament with each other in a generation. In each match, an infinitely iterated Avatamsaka game is played between two agents. Suppose that the total population of the agents in the game world is 1.0 and that the fraction of the agents having strategy i ($i = 0, 1, \ldots, n-1$) in a generation t is x_i, where $\sum_i x_i = 1.0$.

The fitness f_i of strategy i is defined as the expected value of the payoffs in all games in which strategy i is involved in the population. Because the expected payoff of strategy i depends on the probability of meeting with strategies against whom he can gain high (or low) payoffs, f_i depends on the strategy distribution $\boldsymbol{x} = (x_0, x_1, \ldots, x_{n-1}) \in \Delta$ over the population, where $\Delta = \{\boldsymbol{x} : \forall i\, x_i \geq 0 \text{ and } \sum_i x_i = 1.0\}$

Assume that the population share of a strategy either 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 $\frac{x_i f_i}{\sum_i x_i f_i}$, the population share x_i' of strategy i in the next generation $t + 1$ would be the following.

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

where u is the constant term representing the effect of the uniform and time-independent mutation: the mutation rate. For computer simulations, we choose two cases $u = 0$ (no mutation) and $u = 0.0004$ (with mutation) to illustrate the effects of mutation. The change of u does not substantially affect the results if $0 < u << 1$.

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

In the repeated Avatamsaka game, each agent uses a strategy to determine her action in the next stage game based on the memory of actions that she and the opponent made in previous rounds. We investigate the following three cases in this paper: (1) Each agent continues the initial action without any consideration of past information. This condition is denoted herein as the "$m = 0$ (no memory)" case. (2) Each agent chooses her action referring to the previous action of the opponent ("$m = 1$") (3) Each agent determines her action referring to the previous actions of herself and of the opponent ("$m = 2$"). As shown below, the number of possible pure strategies in this repeated game depends on the agents' memory size. Let us denote the nth strategy with memory size of m by S_n^m.

In the case of no memory ($m = 0$), only two possible pure strategies exist: "Always D (which we denote as AllD or S_0^0)" and "Always C (AllC or S_1^0)."

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 plays "D." Four possible pure strategies exist for $m = 1$. They are 00, 01, 10, 11, which we denote respectively as S_0^1, S_1^1, S_2^1, and S_3^1. Note that $00 = S_0^1$ and $11 = S_3^1$ are equivalent to AllD and AllC. $10 = S_2^1$ means that the agent acts in the same manner as the opponent did previously: the so-called Tit-For-Tat (TFT) strategy. $01 = S_1^1$ means that the agent chooses the opposite action of the opponent's previous action: the Anti-TFT (ATFT) strategy.

The strategy in the case of $m = 2$ can also be described as $p_0p_1p_2p_3$ ($p_i \in [0, 1]$), where p_0, p_1, p_2 and p_3 represent the probabilities 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 $0000, \ldots, 1111 = S_0^2 \ldots S_{15}^2$. Similarly, the total number of possible pure strategies is 2^{2^m} for the memory size of m.

For simplicity, we assume that each game would be repeated *ad infinitum* between two agents chosen from the population. Moreover, agents are assumed to make *mistakes* with probability p during the repetition. (e.g. The agent with the strategy to play "C" may sometimes play "D" by mistake against his will.) This means that 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 normalized eigenvector with eigenvalue 1 of the transition matrix as long as $p > 0$. The average payoff in an infinitely iterated game can be given from the probability distribution of the steady state and the stage game payoff matrix. (See [Nowak and Sigmund 1993a].) A slight change of p does not substantially affect the analysis result introduced in the next section as long as $p << 1$. (Basically we assume $0 < p << 1$. Theoretical values found in the next section such as the equilibria of game dynamics are the values gained at $p \to 0$ limit. For computer simulations, we use $p = 0.01$.)

As a short note for strategies with $m = 2$, $S_0^2 = 0000$, and $S_{15}^2 = 1111$ mean AllD, AllC and $S_{10}^2 = 1010$ correspond to TFT, and $S_5^2 = 0101$ to ATFT. Moreover, $S_8^2 = 1000$ is "GRIM" (that plays "C" only when the previous choices of the both players were "C"), $S_9^2 = 1001$ is so-called "PAVLOV" [Nowak and Sigmund 1993b]. The last two strategies can be formed only if $m = 2$. Note that the nomenclature used in the above is the one usually used in "Prisoner's Dilemma" studies. For that reason, it might be sometimes misleading to use the above strategy names in *other* games, but we use those names in this paper for simplicity.

3 Analysis and simulation results

In this section, we show both the analytical results and the results of computer simulations of the evolutionary Avatamsaka game. To capture the nature of evolutionary phenomena, we especially investigate (i) the dynamics of the population share of the strategies, and (ii) the dynamics of agents' actions in each iterated Avatamsaka games.

### 3.1 Agents without any memory of past actions ($m = 0$)

When agents have no memory of past actions ($m = 0$), only two possible pure strategies exist: AllD (S_0^0) and AllC (S_1^0). The payoff matrix for $m = 0$ case is given in Table 2. Obviously, any combination of both players' mixed strategies is a Nash equilibrium. Therefore, the subset of strategies $\mathbf{x} \in \Delta$ that are in Nash equilibrium with themselves, $\Delta^{NE} = \{\mathbf{x} \in \Delta : (x, x)$ is a Nash equilibrium$\}$, is equivalent with the whole mixed-strategy space Δ.

Furthermore, any mixed strategy can be a Neutrally Stable Strategy (NSS), but there is no Evolutionary Stable Strategy (ESS) in this game. (The set of NSS, $\Delta^{NSS} = \Delta$, and the set of ESS, $\Delta^{ESS} = \emptyset$.)

Table 2. Payoff matrix of the iterated Avatamsaka game for $m = 0$ strategies: The procedure to derive this matrix is given in Section 2.2. Because $m = 0$ matrix is simple, we can understand its intuitive meaning. If the opponent is AllD, a player's payoff is $(1 - p) + 2p = 1 + p$ because the opponent's action is "D" with probability $1 - p$ and "C" with probability p, where p is the probability of agent's making mistake.

	AllD	AllC
AllD (S_0^0)	1 + p	2 - p
AllC (S_1^0)	1 + p	2 - p

An example of population dynamics in the case where the agents have no memory ($m = 0$) and no mutation occurs ($u = 0$) is shown in Fig. 1(a). Frequencies of the strategies remain exactly the same as in the initial generation. In the Avatamsaka game, an agent always obtains an identical payoff whether he chooses C or D unless the other acts differently. Because we cannot change the future actions of the strategies AllD and AllC (S_0^0 and S_1^0) that play "D" and "C" forever by their definition, the fitness of either AllD or AllC is determined only by the distribution of AllD and AllC over the population. Therefore, the respective fitness values of AllD and AllC are identical: $f_0 = f_1$. Consequently, the population share of AllC or of AllD does not change with generations. This population dynamics corresponds to the fact that Δ^{NSS} of this game is identical to Δ.
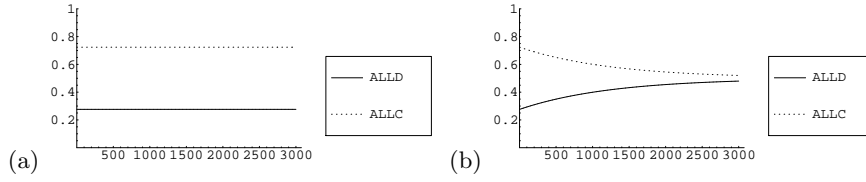
Fig. 1. Population dynamics in the Avatamsaka game world in which agents have no memory of past actions, $m = 0$: The fraction of the strategies in the population is shown as a function of generation. (a) When the mutation rate $u = 0$, the population share of each of the two strategies does not change with generations. (b) When $u = 0.0004$, both converge to 0.5.

If the mutation process is introduced ($u = 0.0004$) in this $m = 0$ Avatamsaka game world, the observed phenomenon is altered completely. In this case, the frequencies (x_0, x_1) of the strategies (AllD, AllC) converge to the equal population share, $(1/2, 1/2)$, with generations (Fig. 1(b)), and the entropy of the configuration of the population is maximized.

### 3.2 Considering the opponent's previous action ($m = 1$)

Payoff structure

When agents have a memory of the opponent's previous action ($m = 1$), they can implement four kinds of pure strategies, AllD, ATFT, TFT and AllC, which are denoted respectively as S_0^1, S_1^1, S_2^1, and S_3^1. The payoff matrix for $m = 1$ case is given in Table 3.

In this case, $\Delta^{NE} = \{(u(1-v), v/2, v/2, (1-u)(1-v)) : u, v \in [0,1]\}$ because the expected payoff for any mixed strategy $x \in \Delta$ against $\mathbf{x}^* \in \Delta^{NE}$ is always $3v/2 + (1-v)(1-u-p)$, which is independent of the value of x. Furthermore, the set of NSS, $\Delta^{NSS} = \{((1-v)/2, v/2, v/2, (1-v)/2) : v \in [0,1]\}$. However, there is no ESS.

Table 3. Payoff matrix of the iterated Avatamsaka game for $m = 1$ strategies: The procedure to derive this matrix is given in Section 2.2.

	AllD	ATFT	TFT	AllC
AllD (S_0^1)	$1+p$	$2-2p-2p^2$	$1+2p-2p^2$	$2-p$
ATFT (S_1^1)	$1+p$	$3/2$	$3/2$	$2-p$
TFT (S_2^1)	$1+p$	$3/2$	$3/2$	$2-p$
AllC (S_3^1)	$1+p$	$1+2p-2p^2$	$2-2p-2p^2$	$2-p$
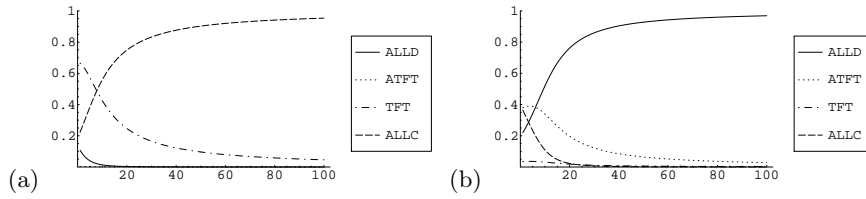
Fig. 2. Avatamsaka game: $m = 1$, mutation rate $= 0$

Population dynamics

The initial strategy distribution has no influence on the future population dynamics in the case of $m = 0$, but it does in the case of $m = 1$ with no strategic mutation, in most cases.

In this game world, where each agent may have a memory of the previous move of the other, the population share $x = (x_0, x_1, x_2, x_3)$ converges either to $(1, 0, 0, 0)$ or $(0, 0, 0, 1)$ (the dominance by AllD (S_0^1) or by AllC (S_3^1)) with generations in most cases other than the knife-edge cases, for which x is exactly in Δ^{NE} in the initial generation. ($x \in \Delta^{NE}$, of course, does not change with generations.) Convergence depends on the initial distributions of ATFT(S_1^1) and TFT(S_2^1) (Fig. 2). AllC continues to increase when the frequency of TFT is greater than that of ATFT ($x_2 > x_1$); the AllD population continues to increase when $x_1 > x_2$.

Because AllD always chooses D regardless of past actions, the average payoff of *any* opponent of AllD is always 1 in the Avatamsaka game, while that of the opponent of AllC is always 2. Consequently, the effect of the strategy distribution of AllD/AllC on the fitness of any strategy is always the same: it does not affect the population dynamics.

It is ATFT or TFT that may affect the population dynamics. When meeting with AllC, TFT always chooses "C" in response to AllC's previous "C," while ATFT always chooses "D" (except for the change of the action by noise.) That is, TFT benefits AllC and, in the same way, $ATFT$ benefits AllC.

As shown above, the appearance of strategies with $m = 1$ gives rise to population dynamics that are dependent on the strategy distribution in the initial generation. Note that the relation $x_2 > x_1$ or $x_1 > x_2$ will not change with generations because the expected payoff of TFT and that of $ATFT$ are identical, which can be known by Table 3. Consequently, if $x_2 > x_1$ in the initial generation, the fraction of AllC continues to increase while that of AllD continues to decrease.

Effect of mutation

The power of the convergence to the AllD/AllC society observed in the above is not so strong because it results solely from the difference between the relative fraction of TFT to ATFT, not on the interaction *within* AllD (AllC)
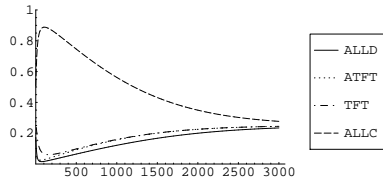
Fig. 3. Avatamsaka game: $m = 1$, mutation rate $u = 0.0004$

agents. When the mutation is introduced to the evolutionary process, $\mathbf{x}$ converges to $(1/4, 1/4, 1/4, 1/4)$ with generations (Fig. 3). What happens here is the triumph of the mutational effect that increases the entropy of the populational configuration over the adaptation power, which is also observed in the $m = 0$ case.

The world of $m = 1$ shares the same nature with that of $m = 0$ in the sense that all existing strategies cannot conduct independent optimization because the increase of the fitness of AllC (AllD) depends completely on TFT(ATFT). Consequently, the fitness of AllC also approaches that of AllD with generations because the fraction of TFT, whose fitness is always identical to ATFT's, approaches that of ATFT by mutational effect.

The mutation mechanism in principle has the capability of increasing the entropy of the population structure. On the other hand, there usually exists in a game the power that comes from players' adaptation to the payoff matrix. We can say that the Avatamsaka game at $m = 0$ or $m = 1$ is one of the games where the effect of the adaptation is overshadowed by the mutational effect that increases the entropy. In other words, dependent adaptation (See section 1.2) by itself is insufficient to overcome the mutational effect.

3.3 Considering the player's and the opponent's actions (m=2)

Let us consider the case of $m = 2$, in which an agent can refer to both her own and the opponent's previous action to decide the next action. The salient difference is that, while the fitness of each of $m = 0, 1$ strategies depends completely on the population shares of the other strategies, several $m = 2$ strategies exist whose fitness depends on the shares of all strategies including that of the strategies themselves.

The payoff matrix for $m = 2$ strategies is a little complicated function of p; because we have insufficient space for describing the entire matrix (and doing it does not seem so informative), we visualize the payoff structure in Fig. 4 by computing payoffs for $p = 0.01$ and by drawing them as a figure and a graph. It can be confirmed analytically that the "PAVLOV[3] ($1001=S_9^2$)" is the ESS of this game. That fact can also be confirmed with Fig. 4.

---

[3]As in the study of the iterated Prisoner's Dilemma game [Nowak and Sigmund 1993b], PAVLOV is a good strategy to form and main-

Population dynamics of the $m = 2$ Avatamsaka game without mutation are shown in Fig. 5. In this setting, the game world converges to the exclusive winning of PAVLOV. Note that the result does not substantially differ even if there is mutation.

Dominance by PAVLOV is realized for any initial condition as long as the initial population share corresponds to the interior of the mixed strategy space Δ. In other words, "independent adaptation" through the use of a $m = 2$ strategy, PAVLOV, overshadows the effect of mutation.

Let us consider why PAVLOV is strong in this game. The PAVLOV policies are as follows: (1) If I feel good (the opponent's action is C), keep doing the current action, whether C or D. (2) Never comply with the situation if I feel uneasy (the opponent's action is D) by sending a signal to the opponent by changing my action[4]. With these two features, PAVLOV can maintain a reciprocal relationship once it is formed because a Pavlovian player continues C as long as he feels good.

Furthermore, two PAVLOVs interacting with each other can quickly recover from betrayal relationships caused by action noise. Assume that two PAVLOVs play a repeated Avatamsaka game, as shown below. PAVLOV's action should be C if both cooperated (CC) in the previous round.

```
PAVLOV  |  C C D* D C C C ...
PAVLOV  |  C C C  D C C C ...
```

In this figure, the mark "*" denotes the betrayal of the upper player by mistake (noise) that is supposed to happen with probability p in this study. While the upper player continues D, the lower Pavlovian player, who feels uneasy, sends a signal to the upper player by changing her action into D in the next round (DD), thereby engendering the recovery of mutual cooperation CC two rounds later.

Analysis of the $m = 2$ payoff matrix shows that PAVLOV is the ESS for $m \leq 2$. Furthermore, some additional consideration allows us to know that PAVLOV is also *evolutionarily stable in the whole pure-strategy space* including the longer-memory case if $w < 1$, in which w is a constant probability to another round of repetition[5].

---

tain mutual cooperation under the circumstances with action noise. However, PAVLOV is not an evolutionary stable strategy in the Prisoner's Dilemma with noise because the best response to PAVLOV is AllD in the Prisoner's Dilemma.

[4]Choosing C or D functions only as a "signal" in the Avatamsaka game because the choice does not affect the player's own payoff.

[5]The only thing players in the Avatamsaka game should do is to derive the other's C as frequently as possible. In so doing, the best response to PAVLOV's D that happens by mistake should be D in the next step, which will stop the PAVLOV's continuing D *as quickly as possible*. Therefore, the best response should have a code ((a) $CD \rightarrow D$). The best-response strategy that has fallen into mutual defection with PAVLOV would recover from the tragic state toward the mutual cooperation in the next step ((b) $DD \rightarrow C$). The best response strategy, of course, would retain
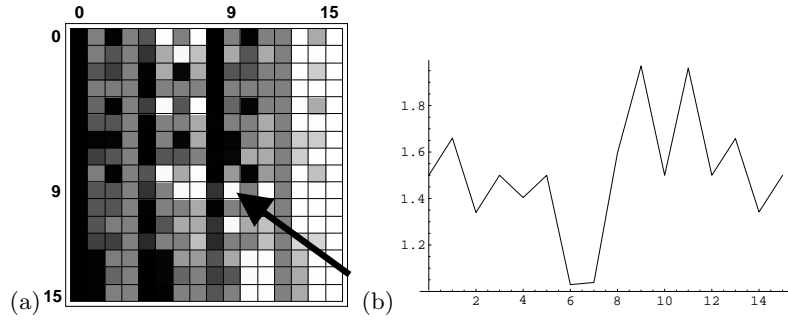
Fig. 4. (a) An illustration of the payoff matrix of the $m = 2$ Avatamsaka game: The vertical axis shows the pure strategy of player 1 and the horizontal axis shows that of player 2. The figures 0, 9, 15 on the vertical and horizontal axes correspond to pure strategies S_0^2, S_9^2, S_{15}^2, respectively. The shade on the tiles represents the payoff for the iterated Avatamsaka game with $p = 0.01$. The higher the payoff, the whiter the color on the correspondent tile. We can visually know the best-response relations among $m = 2$ pure strategies in this figure. (b) The payoff when facing PAVLOV (S_9^2); $p = 0.01$ is plotted as a function of the strategy number. Both (b) and the arrow in (a) indicate that the only best response for PAVLOV is PAVLOV itself. (The payoff for PAVLOV against PAVLOV is $2 - 3p + 6p^2 - 4p^3$.)
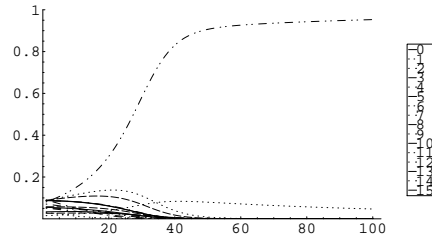
Fig. 5. Population dynamics with $m = 2$ and $u = 0$.

4 Discussion

This paper shows evolutionary phenomena in a dependent game, the Avatamsaka game. Let us sum up the results.

─────────

mutual cooperation ((c) $CC \rightarrow C$) If the best-response strategy has produced a mistake, it would defect again ((d) $DC \rightarrow D$) to revert quickly to mutual cooperation with PAVLOV. (Otherwise, PAVLOV would continue D.) (a)–(d) as a whole shows that the only strategy that can derive most frequent cooperation from PAVLOV is PAVLOV itself, provided $w < 1$. (If $w = 1$ (infinite repetition), a strategy that differs only in his first, say, 100 rounds from PAVLOV can have exactly the same payoff as PAVLOV, on average.) Consequently, if all players' strategies are PAVLOV, they are in a strong perfect equilibrium [Boyd 1989], and so PAVLOV is the ESS.

4.1 Memory size, mutation and social phenomena

First, the memory size generally has a large effect on evolutionary dynamics. For a memory size of $m = 0$, strategic interactions do not have any effect on the population dynamics. Consequently, the population shares of the strategies do not change from the initial generation. When $m = 1$, a strategic interaction becomes more influential than in the case of $m = 0$. As a result, the population converges to either AllC or AllD state, but the convergence depends strongly on the initial configuration of population. If the population consists of agents who consider the previous actions of both themselves and the opponent ($m = 2$), the cooperative society is self-organized in the population by PAVLOV for any initial distribution of strategies that corresponds to a mixed strategy in the interior of Δ. However, further extension of memory size ($m \geq 2$) has no substantial effect on evolutionary phenomena: PAVLOV or its analog always dominates the population. Particularly, if $w < 1$, PAVLOV alone dominates the population. In this sense, an *effective memory-size* of 2 exists in the Avatamsaka game.

Second, the effect of mutation is remarkable in cases of the Avatamsaka game with $m = 0, 1$, in which only dependent adaptation is possible. However, in the case of $m = 2$, the mutational effect is overshadowed by the independent adaptation by PAVLOV, so that the introduction of mutation cannot have a strong effect on evolutionary phenomena. Summing up, investigation of the Avatamsaka game shows clearly that the entropy of a populational configuration is maximized by mutation if there is only dependent adaptation, while the entropy can be decreased when there is an independent adaptation for which a certain size of memory is required for agents in the Avatamsaka game.

4.2 Efficiency in the dependent game, and the origin of inefficiency

Third, PAVLOV is the only ESS in the Avatamsaka game with $m \leq 2$. In addition, PAVLOV is the only ESS in the Avatamsaka game for any m if $w < 1$, where w is a constant probability of another repetition.

Note that, in *any kind of dependent game* with any number of players and with any number of strategies, a *socially optimal state* in which everybody mutually cooperates is achievable as an evolutionarily stable state through PAVLOV-like strategies. Let us consider a Pavlovian strategy used by an agent who continually changes his action until all others each chooses the most beneficial action for him; that is, the Pavlovian strategy keeps sending signals by changing his action to the others until he feels comfortable. It can be easily understood that all players' mutual cooperation is the only stable state. Furthermore, it is the Pavlovian strategy that derive most frequent cooperation of opponent Pavlovians, which means that the best response to Pavlovian strategies is the Pavlovian strategy itself, and that Pavlovian strategy is ESS (See footnote 5 in Section 3.3)

Summing up, the world of the *dependent game* will eventually engender a stable state that is socially optimal. The world of the *independent game*, in which players each independently confront an optimization problem, will also lead to a socially optimal state wherein the only thing that agents should do is to choose their own optimal strategies, if we disregard computational complexity of the problem. Although players in dependent games should have memories of previous actions and players in independent games do not, both the dependent and the independent games should engender socially optimal states through adaptation.

Seen from another perspective, the origin of social inefficiency that we observe in reality arises from both technical and conceptual difficulties: Technical difficulty lies in computational problems for players to calculate optima, while conceptual difficulty comes from the existence of the *entanglement* of dependent and independent game structures (e.g. the Prisoner's Dilemma) but neither from the structure of completely dependent games nor from that of completely independent games.

Acknowledgment

References

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

[Boyd 1989] Boyd, R. (1989), "Mistakes allow evolutionary stability in the repeated prisoner's dilemma game," *Journal of Theoretical Biology* **136**,47–56.

[Neumann and Morgenstern 1944] von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton University Press, NJ, USA

[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.

[Kamata 1988] Kamata, S. (1988), *Kegon no Shiso (The Thought of Avatamsaka)*, in Japanese, Kodan-sha, Tokyo.