J. theor. Biol. (1998) **195,** 53–67
Article No. jt980780

Evolution of Strategies in the three-person Iterated Prisoner's Dilemma Game

MASANAO MATSUSHIMA AND TAKASHI IKEGAMI*

Institute of Physics, The Graduate School of Arts and Sciences, University of Tokyo,
3-8-1 *Komaba, Meguro-ku, Tokyo* 153, *Japan*

(*Received on* 26 *November* 1997, *Accepted in revised form on* 6 *July* 1998)

A three-person iterated prisoner's dilemma (3p-IPD) game is studied. The present simulation uses a lattice model of finite state automata. In particular, the similarity between a noisy 2p- and a noiseless 3p-IPD game is discussed. It is known that, due to the noise, Tit for Tat loses its robustness and is taken over by more complex strategies in a noiseless IPD game. But in the 3p-IPD game, even without noise, Tit for Tat loses it robustness and is also taken over by more complex strategies. It is found that similar strategies take over Tit for Tat in both situations. We thus remark that the role of noise in the two-person game is replaced by the third player in the three-person game. As a result, the strategies diversify in both the noisy 2p- and the quiet 3p-IPD game. It is also found that game strategies in an automaton form can be understood as a combination of defensive and offensive substructures. A recognition of these substructures enables us to study the mechanism of robustness in the strategies of the 3p-IPD game.

1. Introduction

Mutual cooperation can be sustained when players use nice and retaliatory strategies. Axelrod's early computer tournaments (Axelrod, 1984) have revealed that a Tit for Tat strategy (TFT) is such a strategy. Although this conclusion is a rather hypothetical one, TFT is nevertheless found to be robust and non-evolutionary.

We, however, conversely address the question as to how the game strategies can be diversified. In nature, we see a variety of strategies in a wide range of societies, from insects to animals. In these societies, we observe mutual cooperation among species but we do not expect to see commonly used simple strategies. Instead, we see a variety of strategies which are rather tactical and sophisticated.

From recent studies on dynamical systems approaches to game worlds, we have at least three explanations for the mechanism of diversity. The first possible mechanism is noise, which causes erroneous actions in games. Individual players will make mistakes due to the noise, being compelled to play probabilistic strategy instead of pure strategies. A noisy game makes it possible that every choice will be taken with a positive probability. This fact is used to show that noisy game can have evolutionary stable strategies (Boyd, 1989). But at the same

*Author to whom correspondence should be addressed; E-mail: ikeg@sacral.c.u-tokyo.ac.jp

0022–5193/98/021053 + 15 $30.00/0

time it can enhance the variety of strategies. The role of noise in the iterated prisoner's dilemma game is to destabilize mutual cooperating states by creating a high level of distrust between the players. For example, two TFTs will alternatively defect each other. Because of the resulting distrust, TFT will be replaced by strategies which are more tolerant of defections (Molander, 1985; Muller, 1987; Bendor *et al.*, 1991; Nowak & Sigmund, 1993), but at the same time more complex strategies will evolve (Lindgren, 1992, 1994; Ikegami, 1994).

As the second possible mechanism, we propose that diversity is already embedded in the nature of some dilemma games. Without introducing noise into a game, we can construct a certain set of strategies which can exploit TFT (Boyd & Lorberbaum, 1987). It has already been stated in Axelrod (1984) that no optimal strategy exists when the expected future iteration is very long. Further it has been shown that there are no strategies that constitute Evolutionary Stable Strategy (ESS) in the repeated IPD game (Lorberbaum, 1994). In other words, when a strategy dominates a population, there is always a chance that it will be taken over or be subverted by an alternative strategy. Absence of ESS in noiseless game potentially leads to a variety of strategies. But, as we said, noisy game can effectively generate a variety of strategies. The effectiveness of a large variety of strategies is dependent on the third mechanism.

The third possible mechanism is a restriction of the evolutionary paths of strategies. If we assume that new strategies can emerge from other strategies, a runaway or a so-called open-ended evolution (Lindgren, 1992; Ikegami, 1994) can be frequently observed in simulations. In other words, evolution does not lead to ESS, and a diversity of strategies will appear, instead.

In game theory, it is known that not every equilibrium is rationalizable and not every rationalizable outcome is reasonable (van Damme, 1987). The evolutionary dynamics gives a basis for such concept. When we study games with evolving strategies, we have to interpret a player's action more in terms of evolutionary dynamical context. Namely, not every rationalizable outcome is evolutionarily possible. So-called bounded rationality is practically caused by evolution of the possible strategies. On the other hand, we see that two-player analysis cannot be simply applied to the multi-player game. Nice strategies found in a two-person game generally do not work in more than three-person games. Rationalizable outcomes between two players can be destabilized by the rest of the players in many-person games.

In fact, we observe qualitatively different phenomena in more than a three-person game (Akimov & Soutchanski, 1994; Rapoport, 1994; Molander, 1992; Dugatkin, 1990; Alexrod, 1986). Social contracts such as abuse and coalition begin to appear when we have at least three persons. What makes a many-person game complex is that the effectiveness of retaliatory action is not assured. That is, players cannot defect directly against the right person. Even if players can recognize the defector, they still cannot effectively punish the right person because of simultaneous playing. Therefore, the usual TFT is no longer effective here. Molander (1992) analysed that new equilibrium states emerge with a mix of pure defection with a conditionally cooperative strategy. When a player punishes one player, he may have to punish other nice or cooperative players at the same time. We will show in this paper that this helpless situation will, in turn, evolve into a diverse variety of strategies.

2. Modeling

2.1. Three-person prisoner's dilemma

The original two-person prisoner's dilemma game consists of the following payoff matrix (Table 1) where C and D correspond to "Cooperative" and "Defective" actions, respectively. Each matrix element, (S_1, S_2) corresponds to the scores of players 1 and 2, respectively, which satisfies two dilemma inequalities, $T > R > P > S$ and $2R > T + S$.

In the three-person prisoner's dilemma, each player has to play simultaneously with the other two players. The payoffs of this game are obtained by simply superimposing each of three combinations of moves. That is, if the three players' moves are C, C, and D, the assigned payoffs are

$$\frac{R+S}{2}, \frac{R+S}{2} \text{ and } T$$

to each of three players, respectively. We can define the payoffs of the N-person prisoner's dilemma in a similar way. Scores for a player's move (C,D) against $N-1$ other players are given by

$$\frac{(n-1)R+(N-n)S}{N-1}$$

and

$$\frac{nT+(N-n-1)P}{N-1},$$

respectively. Here we assume n cooperative and $N-n$ defective actions. This natural extension of two-person prisoner's dilemma to N-person is also studied by Molander (1992) and others (see e.g. Hamburger, 1973).

2.2. REPRESENTATION OF STRATEGY BY FINITE AUTOMATON

We assume that each player chooses his move according to his current state and the moves of the opponents. A finite-state automaton (FSA) is used as a model for individual strategy to stipulate the individual's next move (Rubinstein, 1986; Miller, 1996; Lindgren, 1996). Usually the input to FSA for N-person game is given the choice of the majority of the other players (see e.g. Akimov & Soutchanski, 1994). We here use the full combination of the other players action pattern. In the case of $N=3$, they are CC, CD and DD. An example of FSA we study in this paper is shown in Fig. 1. In the diagram, states

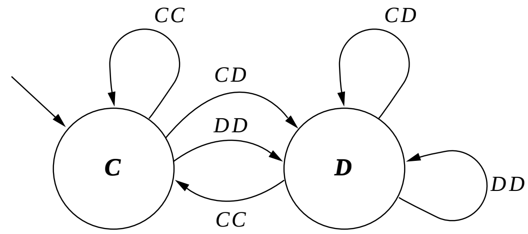


FIG. 1. The strategy as a finite automaton. The symbols in the state circules denote the next action. On the transition arcs, two symbols correspond to the actions performed by the other two players. The initial action of the strategy is denoted by the arc without symbols.

are shown by nodes which are connected by arcs. A current node represents the current state of the player and the specified action indicated within the node represents the player's next move. The initial move is indicated by an arc without a label. Updating of the current state is specified by arcs named by the other two players' moves. For example, the strategy represented by this diagram behaves as follows.

(1) It initially cooperates; (2) at the C node, it cooperates as long as at least either one of two players cooperate. Otherwise it changes the node from C to D and defects; (3) at the D node, it defects as long as other two players both defect. Otherwise it changes the node from D to C and cooperates.

TABLE 1
The payoff matrix of the two-person prisoner's dilemma game

		Player 2	
		C	D
Player 1	C	(R, R)	(S, T)
	D	(T, S)	(P, P)

2.3. SIMULATION SCHEMA

We have four basic processes for the present simulation.

(1) First, place an initial set of strategies on the two-dimensional plane. We have concentrated on evolutions from initial strategies with two nodes. The six initial strategies are shown in Fig. 2. Initially, each lattice site is occupied randomly with one of the six strategies. We use these strategies in particular as we can easily interpret their strategic meanings. However, the following observations are robust to the initial choice of strategies.

(2) Strategies change their automaton structures by the following genetic operations. There are countless numbers of possible FSA structures, and we cannot have every FSA at one time. We thus tentatively introduce some

"genetic" operators to generate new strategies by making an analogy between FSA and genes:

- flipping a label of a randomly selected node ($C \leftrightarrow D$);
- changing the initial move by reputting the unlabeled arc to other node;
- reassigning arcs;
- adding a new node with a label and three outgoing arcs;
- removing a randomly selected node. Arcs to the node will be reassigned.

By sequentially combining above operators, we can reach almost all possible FSA structures from any arbitrary structures. This is the criterion on which we adopt those operators. By using more global operators (such as recombination or fusion), we can have a different evolutionary process. But as for the first step, we restrict it to the above local operators.

(3) Each player takes part in 12 3p-IPD games with different combinations of players. A technical problem with the N-person game is the necessary increase in computational time for simulating a round-robin game. We here study a spatial game for the sake of simulating a large set of strategies. A square plaquette having four corners contains a single player. It plays the IPD game with those players sharing the same corner. In Fig. 3, individual A shares the left upper corner with three other players. Therefore, the possible combinations among A, B, C, and D are, (A,B,C), (A,B,D), (A,C,B), and (B,C,D). Since one plaquette (one player) has four corners, each individual takes part in 12 ($=4 \times 3$) different games. We use 50×50 plaquettes so that the total number of games played in one generation is 10 000, where each game constitutes 100 iterations. In order to avoid the effect of boundaries of the spatial structure, we adopt a periodic boundary here.

(4) Strategies are updated by the results of the game. In each generation, a strategy at each plaquette is replaced by the best strategy of the neighboring plaquettes. If two or more strategies are equally effective, the strategy with the least number of nodes is replaced so as to prevent the automata from meaninglessly increasing the complexity of the game. If the automata have the same number of nodes, one is selected randomly. The replacement occurs simultaneously at every plaquette for each generation.

The above processes from (2) to (4) define one generation of the present game dynamics. By recursively iterating the processes (2)–(4), we analyse the evolutionary processes generated from the initial set of strategies (1) in the following sections.

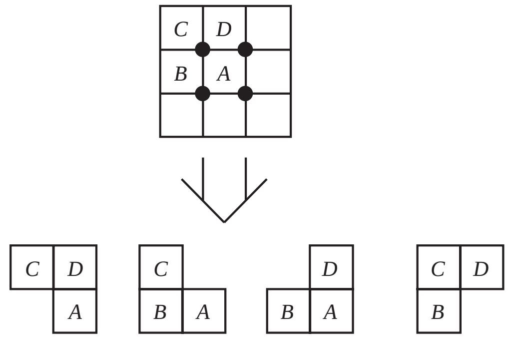


FIG. 3. Players sharing the same corner participate in the 3p-IPD games. For each generatioin of a game, each player takes part in 12 3p-IPD games with different combinations of players.

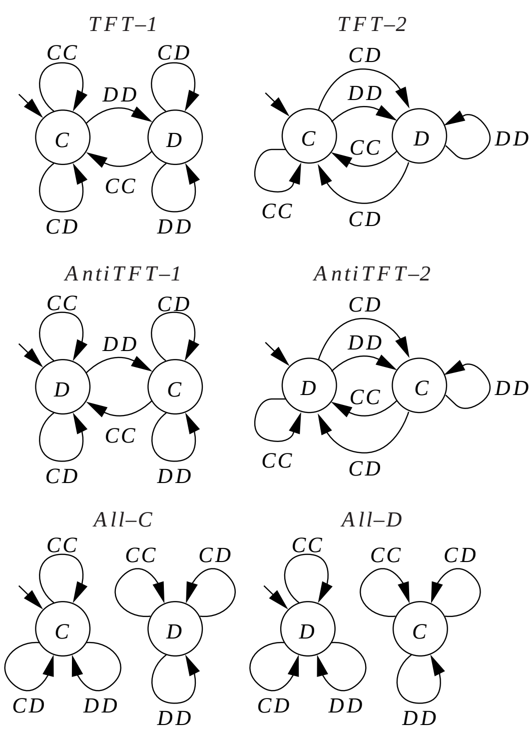


FIG. 2. Initial set of strategies of finite automata consist of two nodes. They are corresponding to the memory length one strategies. Different from the 2p-IPD game, we have different types of TFTs and anti-TFTs for the 3p-IPD game.

3. Simulation Results

3.1. A TYPICAL EVOLUTIONARY PATHWAY

Through the simulations, we fix the payoff matrix to $(R,S,T,P) = (3,0,5,1)$, which was adopted in Axelrod's original contest. The mutation rates of adding or removing nodes are set to 0.003, whereas the other mutation rates are set to 0.01 per each strategy.

It is worth comparing 3p-IPD strategies with 2p-IPD strategies (see Fig. 2). For example, Tit for Tat-like strategies can also be generated in this three-person game. A *TFT*-1 is more generous than *TFT*-2, in the sense that it defects only when two players defect. However, once the game has reached the defecting state, it is harder to recover a state of mutual cooperation than with *TFT*-2. In other words, *TFT*-2 is a more forgiving TFT than *TFT*-1. *TFT*-1 cannot spread in the same way as *TFT*-2. The reason for this will be clarified later.

At the early stage of evolution runs, a defective strategy (*All*-*D*) always spreads rapidly, but it lasts for only a few generations. Instead, Tit for Tat of type 2 (*TFT*-2) grows rapidly, forming spreading islands. Then all strategies are replaced by *TFT*-2. This is because all three of the *All*-*D* games are mutually defective, but one of the three *TFT*-2s is mutually cooperative, so that a cluster of *TFT*-2s can gradually increase in size.

After *TFT*-2 has spread, the tempo of evolution becomes much slower (e.g. on the order of thousands of generations). Maintaining a cooperative society, *TFT*-2 gradually decreases its population size. We then enter the third stage of evolution. Several mutants with two nodes take over the population. The two most powerful strategies within this group are depicted in Fig. 4 as *X* and *Y*. These mutants can frequently dominate half of a population.

These two strategies coexist by temporally sharing a large population. But this coexisting state does not last permanently. Suddenly, strategies which have three nodes invade. They spread rapidly by exploiting *Strategy*-*X*, and occupying a large portion of the population. The *Strategy*-*Y* is strong against the three-node strategies, however, and it finally repels them. The potion of the population using *Strategy*-*X* begins to increase again, and the former coexisting state between *Strategy*-*X* and *Strategy*-*Y* recovers. These dynamic courses will be repeated again. The typical population dynamics are shown in Fig. 6.

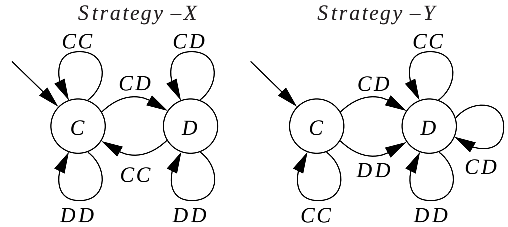


FIG. 4. The two strategies are found to be effective in the present simulation. A *Strategy*-*Y* is a natural generalization of the Friedman [taken from Axelrod (1984)] to 3p-IPD game. This strategy never cooperates if it is once defected.

All the three-node strategies that can invade are quite similar in their structure and their performance. Such as:

1. they start with the *D* node;
2. they never stay in a (D,D,D) state;
3. they play *C* after two successive (D,D,D) states;
4. they can stay in a (C,C,C) state.

Two typical structures of three-node strategies are shown in Fig. 5.

4. Conditions of Robust Strategies

4.1. THE CONDITION OF COOPERATIVENESS

We examined what strategies can be "robust" or evolutionarily stable in the three-person IPD game. Unless a strategy is evolutionarily stable or at least robust, it can easily be invaded or overrun by mutant strategies. Tit for Tat is known to be robust in the two-person IPD game.

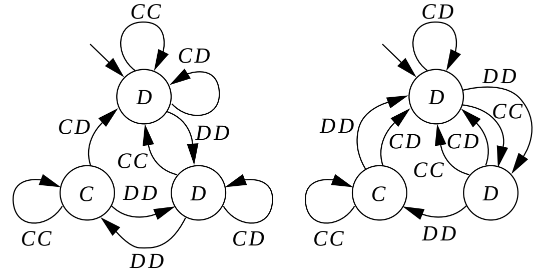


FIG. 5. The three-node strategies which can invade the population of *Strategy*-*X* and -*Y*.

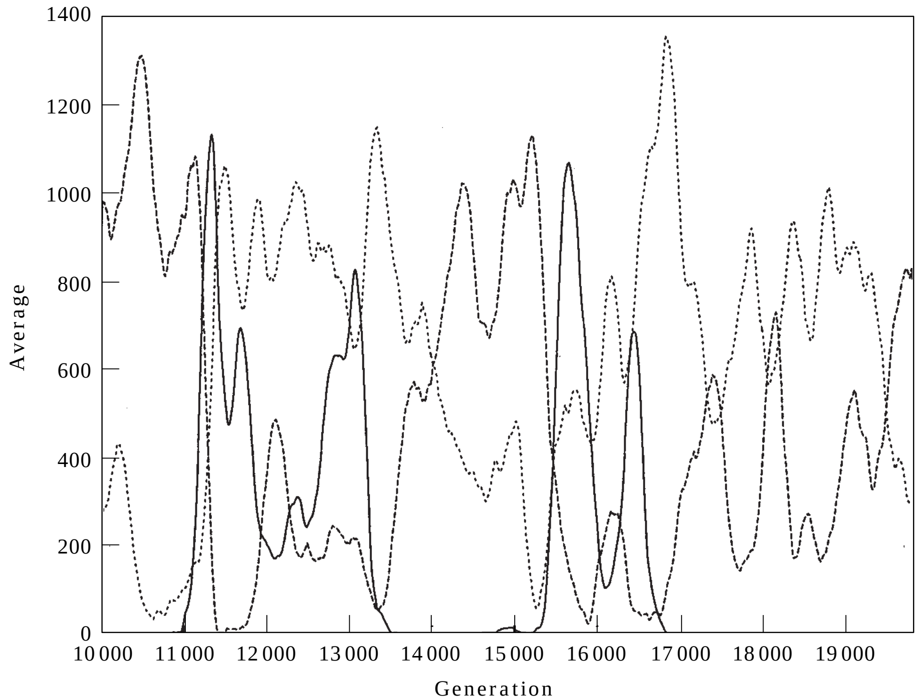


FIG. 6. Time evolution of population for three dominating classes of strategies (X, Y and three-node strategy). The class of three node strategies (——) are intermittently invading the population of X (– – – –) and Y (· · · · ·).

It is, nevertheless, not evolutionarily stable. In the present situation, ESS is not relevant for discussing the stability of strategies. The important stability criterion for a strategy is if it is stable when faced with their direct mutants.

It is not necessary that robust strategies be mutually cooperative with one another. In the two-person IPD game, playing all-D against ill strategies can be very robust. But they cannot keep mutual cooperation themselves. The problem there is how to recognize which strategy to cooperate with. We consider a strategy to be mutually cooperative when it meets "the condition of cooperativeness". By this term, we mean that a mutually cooperating state (C,C,C) is sustained by given strategies. In terms of automata structures, a CC labeled arc from a C-labeled node will be connected to the C-labeled node.

A game played against players with the same strategy is defined as a "self"-game. We will see that robust strategies in the three-person IPD game can implicitly "recognize" whether they are playing a self-game or not, and at the same time they meet the condition of cooperativeness. In the original two-person IPD game, such recognition is not expected. A TFT also fails to recognize itself.

4.2. INVASION AND DEFENSE

Suppose a situation where a type-B strategy is surrounded by type-A strategies (Fig. 7). A defender (as type A) has to play eight self-games and four games against the same A and an invader B. An invader (as type-B) has to play all 12 games against two As. We define a game in which one plays against one player employing the same strategy and another employing a different strategy as a "two–one" game (i.e. A–A–B), whereas a game played against two other players using the same strategy as a "one–two" game (i.e. A–B–A). It is necessary to analyse the two–one game for an analysis of

A	A	A	A	A
A	A	A	A	A
A	A	B	A	A
A	A	A	A	A
A	A	A	A	A

FIG. 7. An assumed situation where an individual type-B appears in the population of type-A.

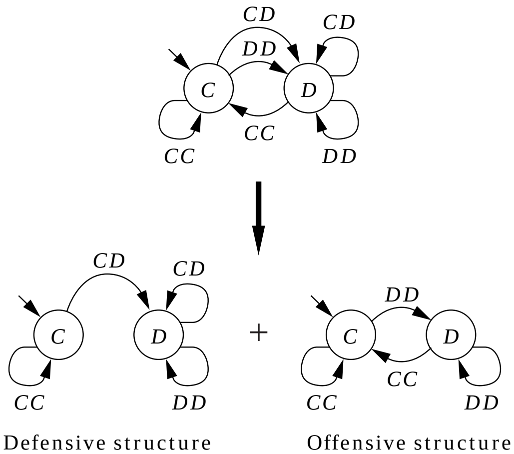
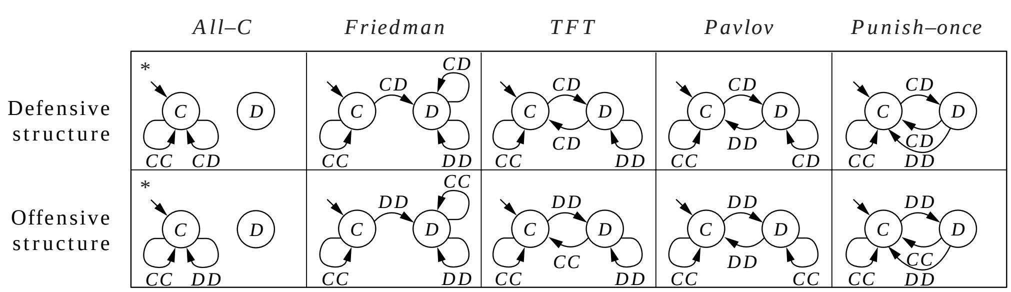


FIG. 8. A strategy of finite automaton is represented by two substructures with respect to the strategic content: the defensive and offensive structures.

defensive ability and the one–two game to analyse offensive ability. In order to see the most severe conditions, we assume that all strategies satisfy the conditions of cooperativeness.

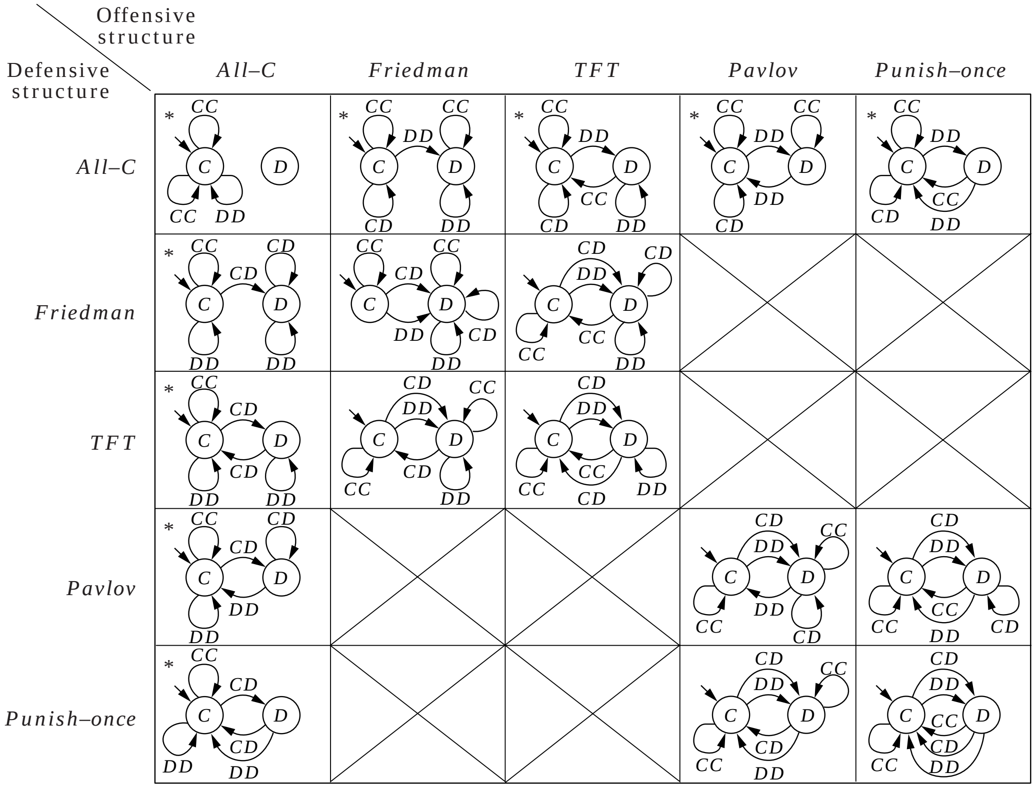
For the systematic analysis, we divide each strategy into two substructures; one is used for the defensive situation and the other for the offensive situation. Figure 8 provides an example. The original structure of the strategy is drawn as a combination of the defensive and offensive structures. In the offensive situation, the one–two game will be played. Therefore, a structure with DD-labeled arcs from C-labeled nodes and CC-labeled arcs from D-labeled nodes has to be analysed. In the defensive situation the substructures without these arcs should be analysed.

First we decompose all possible two-node substrategies into substructures in Fig. 9. These strategies meet the conditions of cooperativeness in the sense that they can keep a mutually cooperating state (C,C,C) as mentioned. *All-C* is the always-cooperating strategy as in the 2p-IPD game. Friedman never defects unless at least one of the opponents defects. If it is defected, it will defect in all later rounds. This Friedman is also a natural extension of Friedman in 2p-IPD game (Axelrod, 1984). TFT starts with cooperation and repeats the opponent's last move. Strategy of Pavlov [taken from Nowak & Sigmund (1993)] and *Punish-once* is explained from that of TFT. Pavlov does the opposite to TFT after his defection. Namely, TFT repeats the invader's last move but Pavlov does the inverse action. On the other hand, *Punish-once* is a more generous strategy than TFT. When the *Punish-once* punishes once, he will cooperate in the next regardless of the other players' action. As will be discussed in the next section, the invader who has different strategy from the dominant strategy is automatically detected in this 3p-IPD game. In a defensive mode, at most one player is the invader. Hence the CD labeled arc from a C-labeled node means the invader's defection. On the other hand, the CD labeled arc from a D-labeled node means the invader's cooperation. But in the offensive mode, there are no CD arcs. Here the CC labeled arc from any node means the defender's cooperation and the DD labeled arc from any node means the defender's defection.



The automation in the cells labeled with an asterisk seems to lack two arcs. We have omitted them, for they may point to any node.

FIG. 9. All substrategies which can be represented by finite automata with two nodes are presented. The substructure of each automaton is highlighted with respect to the defensive and offensive properties. For example, TFT as a defensive structure is defined to have only CC and CD arcs in its structure. On the other hand, it is defined to have only CC and DD arcs as an offensive structure.



The automation in the cells labeled with an asterisk seems to lack two arcs. We have omitted them, for they may point to any node.

FIG. 10. The classification of any finite automata with two nodes by using the substructures defined in Fig. 9.

In Fig. 10 we inversely compose automata from those substrategies in Fig. 9. We again show the strategies which maintain the condition of cooperativeness in an initial move C. These composing/decomposing studies of automata will be a strong tool for the following analyses.

4.3. STABILITY ANALYSIS

Scores of two–one games determine the stability of a cooperative society, which consists of strategies maintaining the condition of cooperativeness. We again study the situation of Fig. 7. A strategy B, as an invader is assumed to choose the best move against the defender strategy A. As are mutually cooperative in their self-games. If the strategy A chooses C actions, it scores exactly the same amount as the B score. Hence B prevails in the population only neutrally. In order to prevail in more than a neutral mode, B should at least take D action somewhere. Suppose that an invader takes D action in a certain round. We will show the progression of events after the first action of B's D action.

In Table 2 we summarize our analysis of the stability of the five defence structures shown in Fig. 9. Concerning the situation of Fig. 7, we assume that the corresponding invader can choose the best move against the defender strategies. It should be noticed that such ideal invaders may not be modeled by the two node FSA. We just analysed the stability by assuming the worst case.

We call strategies "stable", "quasi-stable", and "unstable" when scores of the A player against B are clearly higher, nearly equal, or clearly lower in comparison with those of B, respectively. A stable strategy can certainly repel any possible invaders. When it is quasi-stable, it repels invaders or allows them to coexist neutrally. When it is unstable, it allows some invaders to enter. In the worst case, invaders spread rapidly through the society with an unstable strategy.

TABLE 2

Stability of strategies against invasion can be analysed with respect to their defensive structures. When an invader takes the optimal actions, the expected move patterns of strategies are depicted in the second row. The corresponding averaged and total scores are depicted in the third and forth row. If the total score is larger than its invader, it is called stable. If the scores are equal, it is called neutral. Otherwise, it is called unstable

Defensive structure	*All-C*		*Friedman*		*TFT*		*Pavlov*		*Punish-once*	
	Defenders	Invader	Defenders	Invader	Defenders	Invader	Defenders	Invader	Defenders	Invader
Moves	*C C*	*D*	*C C*	*D*	*C C*	*D*	*C C*	*D*	*C C*	*D*
	C C	*D*	*D D*	*D*	*D D*	*C*	*D D*	*D*	*D D*	*D*
	C C	*D*	*D D*	*D*	*C C*	*C*	*C C*	*D*	*C C*	*D*
					C C	*C*	*D D*	*D*	*D D*	*D*
Averaged score (in the game)	1.5	5	1	1	3	3	1.25	3	1.25	3
Total score (in the generation)	30	60	28	12	36	36	29	36	29	36
Stability	Unstable		Stable		Quasi-stable		Unstable		Unstable	

An example of a stable defense structure with two nodes is called *Friedman*, as shown in Table 2. We show the time evolution of population size with respect to the sub-strategies in Fig. 11. All strategies are summed up with respect to their defensive forms and are overlaid. The dominance of *Friedman*-type is clearly seen.

5. Stability Analysis of Strategies from Evolutionary Dynamics

5.1. The Case of Axelrod's Payoff Matrix

We have seen that stable two-node strategies have a *Friedman*-type defensive structure. In this section we analyse our simulation results in Section 3 with the same substructure decomposition. We discuss why, out of all the strategies with the *Friedman*-type defensive structure, *Strategy*-*X* and *Strategy*-*Y* are selected as the predominant strategies and why three-node strategies emerge and spread but finally disappear.

A *Strategy*-*X* consists of the *Friedman*-type defensive structure and the *All*-*C*-type offensive structure. In short, this strategy prevails because the defensive and offensive structures are combined most effectively. We have checked the stability for *Strategy*-*X* under the Fig. 7 environment, where the invader *B* does the best against *A*. A *Strategy*-*X* exploits its mutant, scoring almost as high as a mutual cooperation state (Table 3). In general, the scores of defenders decrease with the appearance of mutant strategies. However, it is often the case that an invader appears a mutant from defenders in the evolutionary simulations. Some mutant strategies are only different from the parental strategy by the initial node. If the mutants act as prey to the parental strategies, the parent strategy becomes stable and *Strategy*-*X* can take a great advantage over its mutant (Table 3).

Table 3

Stability of Strategy-X *against the initial state-charged is analysed. Four out of* 12 *three-person games include one mutant player. The third row shows the averaged score of the four games which include mutants. Since* Strategy-X *always cooperates against the own* Strategy-X, *the rest of eight games give three points to* Strategy-X. *As the result, the total score of a* Strategy-X *in one generation is computed as* $36(=3\times 8+3\times 4)$ *point*

Strategy	*Strategy*-*X*	
	Defenders	Mutant
Moves	*C* *C*	*D*
	D *D*	*C*
	D *D*	*C*
Averaged score (in the game)	3	0
Total score (in the generation)	36	0

Table 4

The patterns of moves expected among one Strategy-X *and two three-node strategies*

Strategy	Three-node strategies	*Strategy*-*X*
Moves	*D* *D*	*C*
	D *D*	*C*
	D *D*	*C*
Averaged score (in the game)	3	0

The degree of stability is very much dependent on the organization of the payoff matrix. Here we extend the same payoff matrix as the one used by Axelrod. With some other matrices, *Strategy*-*X* can result in a score even higher than those achieved by its mutants. In such cases the excellence of the *Strategy*-*X* becomes more pronounced. Having a weak part in the offensive structure will provide this advantage over the mutants. In the case of *Strategy*-*X*, most mutants will become prey to the *Strategy*-*X*.

However, such an advantage over the mutants can also turn out to be a disadvantage. We have analysed the stability of strategies with respect to the two–one game, but sometimes the results of a few one–two games cannot be neglected. Because *Strategy*-*X* has an explicit "security" hole in its offensive substrategy, it is inversely exploited in one–two games by the three-node strategies (Table 4). Three-node strategies can attack the weak part of *Strategy*-*X*, which is stable in our analysis in situations such as that presented in Fig. 7. In such cases, the effect of

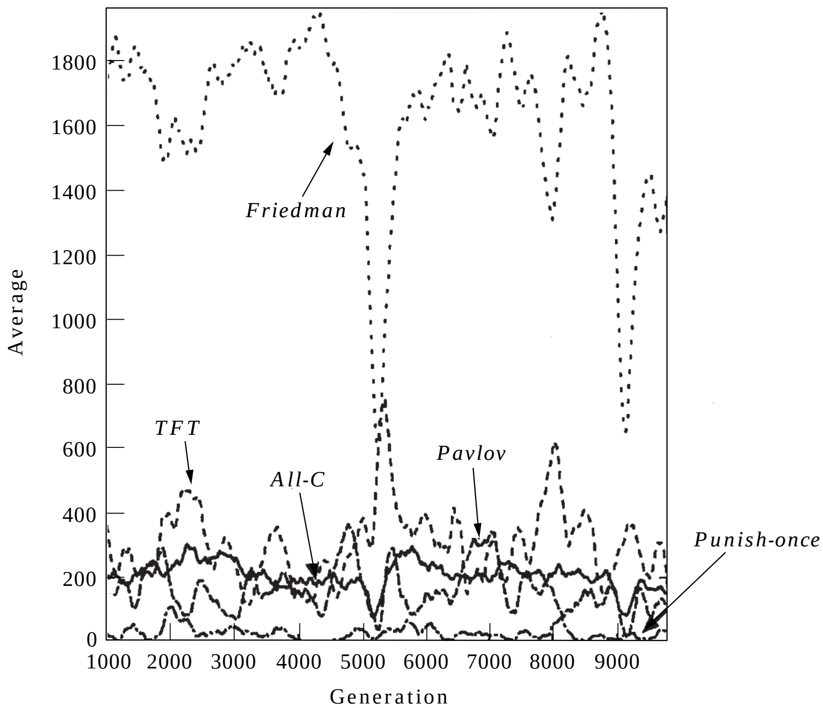


FIG. 11. Time evolution of population size of strategies with respect to defensive substructures. The spikes of *Friedman*-type substructures are due to the invasion of three-node strategies.

one–two games must not be neglected. This is why the population of *Strategy-X* suddenly crashes even with a three-node strategy. But no stable two-node strategy can outperform *Strategy-X*.

What about the stability of more than node three strategies? Testing their stability in the Fig. 7 situation, we have found that node three strategies are quasi-stable due to the cyclic defection of period three (Table 5). Namely, a node three strategy can be periodically exploited by the ideal invader. However, because a node three strategy satisfies the condition of cooperativeness, it scores higher than the invader on average. The pattern of action which this three-node strategy takes, which involves never cooperating until being defected against twice successively, is known as a robust strategy in the two-person IPD game (Lindgren, 1992). The pattern of action is also useful when a strategy invades *Pavlov*- or *Punish-once*-type strategies whose starting node is the *D* state. *Pavlov* and *Punish-once* happen to be the parent of such node-three strategy which shows such useful action of patterns.

In the same way, we can predict the appearance of four-node strategies which can spread within the three-node strategies. The ideal four-node strategies are more stable because they cause a defective cyclic defection of period four against the three-node strategies. Unfortunately,

TABLE 5
Stability of the three-node strategies against the optimal mutant. Averaged and total score are computed as well as in Table 3

Strategy	Three-node strategies		
	Defenders		Invader
Moves	*D*	*D*	*D*
	D	*D*	*D*
	C	*C*	*D*
	D	*D*	*D*
	D	*D*	*D*
	C	*C*	*D*
Averaged score (in the game)	1.17		2.33
Total score (in the generation)	28.7		28.0
Stability	Quasi-stable		

the chance in practice of having an ideal four-node strategy is remote. It is true, however, that the three-person IPD game can be open-ended with respect to the size of the automaton.

The other main strategy we have found in our simulation (see Fig. 4) is a *Strategy-Y*. In this strategy, both the defensive and offensive structures are *Friedman*-type. In other words, it plays *Friedman* both offensively and defensively. We see that the behavior resulting from the *Strategy-Y* is a natural extension of the *Friedman* strategy in the two-person IPD game. *Strategy-Y* never cooperates once it detects the presence of another strategy. Therefore, it is strong against any invader including the three-node strategies. After the three-node strategies predominate, *Strategy-Y* appears to repel the three-node strategies. *Strategy-Y* coexists with *X*, however *Strategy-X* is more generous than *Y* due to the *All-C* type offensive form. Thus *Strategy-X* can again spread by pushing back *Strategy-Y* after it repels the three-node strategies. In brief, *Strategy-Y* is good at defense, but it does not spread easily.

5.2. PERTURBATION OF THE PAYOFF MATRIX

In the prisoner's dilemma, the payoff matrix must follow two inequalities, $T > R > P > S$ and $2R > T + S$ (subsection 2.1). So far we have fixed these variables as $(R,S,T,P) = (3,0,5,1)$. We now consider variations of these values. By fixing R and S, we give a payoff matrix of $(R,S,T,P) = (3,0,p,q)$, with $3 < p < 6$ and $0 < q < 3$.

The phase diagram in Fig. 12 summarizes the effects of variations in the payoff matrix*. The predominance of *Strategy-Y* and *All-D* strategy roughly change along the depicted arrows.

Transitions in strategies along the diagonal line are shown in Fig. 13. To the right side of this line, strategies can exploit the ideal invaders. *All-D* is such a strategy. On the other hand, the strategies containing *Pavlov*- or *Punish-once*-type strategies are being exploited. The three-node strategy and *Strategy-X* can either exploit or be exploitable depending on the phase.

*Refer to the Appendix for the calculation of the phase boundaries.

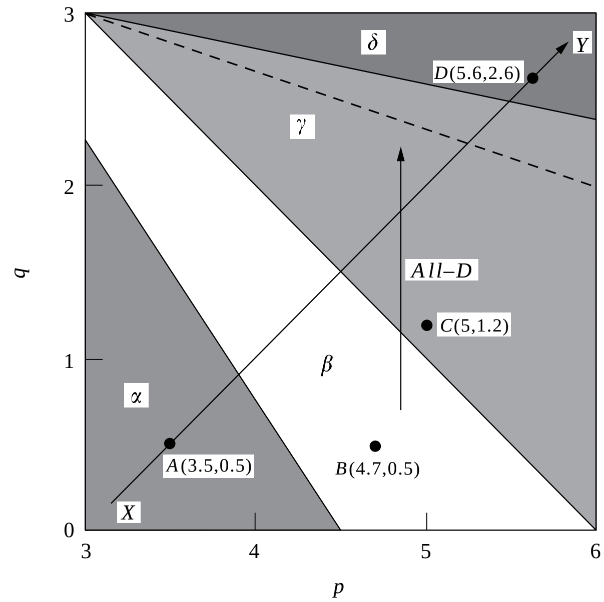


FIG. 12. Approximate phase diagram for the 3p-IPD game when the payoff matrix $(R,S,T,P) = (3,0,p,q)$ is varied. Typical population dynamics for each phase is drawn in Fig. 13.

On the other hand, strategies along the vertical line (i.e. a constant p value) become defections. A large q value implies that the defection is profitable as well as being mutually cooperative. Thus *All-D* becomes dominant on the large q region. Bad mutants easily emerge to exploit cooperative strategies. For example, the three-node strategies are weak against *All-D*. The number of strategies which contain *All-C*-, *Pavlov*-, or *Punish-once*-type defense structures generally decreases with the appearance of *All-D*.

In Phase-α the strategies which contain *Pavlov*- or *Punish-once*- defense structure are quasi-stable. They can cooperate with their *initial-state changed* mutants. We emphasize again the effect of invasion by the initial-state-changed mutants. A *Strategy-X* can gain against such mutants. However, mutual cooperation is more beneficial than exploiting the mutants. Therefore, strategies which can cooperate with the initial-state-changed mutants can outperform others. For example, strategies which contain a *Pavlov*-defensive structure can surpass those which contain a *Friedman*-defensive structure. The three-node strategies which can exploit the

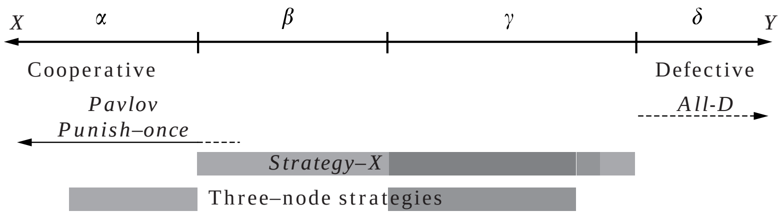


FIG. 13. Transitions of dominant strategies along the diagonal line in Fig. 12. If we increase (decrease) the value of p and q, more defective (cooperative) actions are observed. *Pavlov* and *Punish-once* strategies are getting more stable when the value of p and q is small. On the other hand, *All-D* becomes stable at large p and q values. The *Strategy-X* and three-node strategies become dominant in the middle of the range. The darker line implies that the population size becomes larger.

strategies containing *Pavlov*- or *Punish-once*-defensive structures can emerge, but they cannot spread widely in the population.

In Phase-β, the constituent strategy is similar to those we have seen with the score $(R,S,T,P) = (3,0,5,1)$. We only note that *Strategy-X* becomes a little more common strategy. As in Phase-α, mutual cooperation is better than exploiting mutants. Hence the three-node strategies seldom emerge in this phase.

But in Phase-γ, exploiting mutants becomes more beneficial than mutual cooperation. Hence *Strategy-X* can emerge more frequently. But at the same time, the three-node strategies also appear more easily. Therefore, a trio-relationship among *Strategy-X*, *Strategy-Y*, and the three-node strategies, is frequently observed here. Also, we have to note that *All-D* becomes especially strong in the large q region. It is worth noting that the evolutionary path of three-node strategies changes. They are the off–offsprings of *All-D* rather than those containing *Pavlov*- or *Punish-once*-defensive structures (Fig. 14). There is a broken line in the boundary around Phase-δ. *Strategy-X* becomes unstable in the above area of the line.

As we have discussed, *Strategy-X* is stable in two–one games but is severely exploited in one–two games by strategies such as *All-D*. In Phase-δ, that loss of the one–two game becomes effective so that *All-D* can invade and spread the population of *Strategy-X**. *Strategy-Y* and the strategies which contain *Friedman* defensive structure and *TFT* offensive structures dominate the population.

*Our stability analysis says nothing about the spread of invaders.

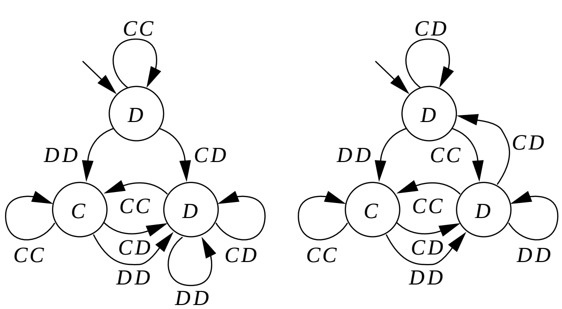


FIG. 14. The structures of the other three-node strategies found in Phase-γ of Fig. 12.

6. Discussion

In the three-person IPD game without noise, we have shown that a certain set of strategies coexist instead of one dominant strategy. It is true that introducing spatial structure drastically changes the repertoire of the strategies. But the preference of coexistence rather than dominancy of strategies is already found in the many person IPD game without spatial structure. It is reported that the stable equilibrium states of the $(N > 2)$p-IPD games contain both (mostly conditional) cooperators and defectors (Molander, 1992; Dugatkin, 1990; Boyd & Richerson, 1988).

Most of the studies are restricted to the fixed size of strategies due to the computational resource. But as we have discussed in the introduction, evolutionary dynamics does control the diversity of strategies. Though our simulation is limited to $N = 3$ case, we have not fixed the size of strategies. Thus the stability of strategies are discussed in the light of open-ended evolution (e.g. the environments with evolving strategies.) Then the practical examples of robust strategies in the 3p-IPD game have been revealed. Those strategies are not simple generalization of Tit for Tat strategies, and they induce mutant strategies which are similar to those in the noisy 2p-IPD game. The equilibrium states cannot form stable fixed points but form dynamically unstable oscillations.

In the two-person noisy IPD game, TFT becomes weak against defections. TFT again cannot work effectively in the 3p-IPD game without noise. In this way, the role of noise is replaced by the third player. As a result, strategies are diversified in both simulations. In the noisy 2p-IPD game, if a strategy proceeds so that "if one player is accidentally defected, it will resume cooperation only when both defect twice successively" is known to be robust (Lindgren, 1992). In a noiseless 3p-IPD game, one of the primary three-node strategies resumes co-operation when all three players defect twice successively, similar to the 2p-IPD case. In the noisy two-person IPD game, the cause of the noise itself has not been discussed. The present simulation provides one possible source for such noise generation. That is, the third person can act as a noise source for the other players.

But we should be careful that the similarity between noiseless 3p-IPD game and noisy 2p-IPD depends on spatial structures and the given payoff matrix. In the last section, three node strategies appearing in Phase-γ have no similar structures to noisy 2p-game. It is worth noting that strong strategies with four nodes will appear at the lower q (a score for mutual defection) with $p + q > 6$ regions, where mutual defection leads to the lower score. Amplitudes of population number also become chaotic in this region. It has been reported that the two-dimensional noisy IPD game (Lindgren, 1994) accelerates the memory elevation and thus a variety of strategies in the higher values of p and q regions. We have shown here that at least three-person is necessary to enhance a variety of strategies when there exists no noise. The effects of spatial structure should be further investigated.

A new strategy analysis we have shown in this paper, on the other hand, does not depend on the payoff-matrix or spatial structure. That is, strategies in the three-person game can be categorized by deconstructing the subautomaton into defensive and offensive structures. For example, the key *Strategy X* is understood as a strategy with a *Friedman* defensive form and an *All-C* offensive form. That we can make this kind of classification means that at the same time a strategy can recognize a strategy identical to its own. Because each player has to play with the other two players simultaneously, it is often difficult to punish only one of the two players. A player who behaves differently from the other two will be identified as a stranger. This is similar to social insects such as the bumble bee which can identify their enemies only by comparing their behavior with others.

This work is partially supported by the Grant-in-Aids (No. 07243102) for Scientific Research. Of priority are "System Theory of Function Emergence" from the Japanese Ministry of Education, Science and Culture.

REFERENCES

AKIMOV, V. & SOUTCHANSKI, M. (1994). Automata simulation of *N*-person social dilemma games. *J. Conflict Resolution* **38,** 138–148.

AXELROD, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.

AXELROD, R. (1986). An evolutionary approch to norms. *Amer. Poli. Sci. Rev.* **80,** 1095–1111.

BENDOR, J., KRAMER, R. M. & STOUT, S. (1991). When in doubt: cooperation in a noisy prisoner's dilemma. *J. Conflict Resolution* **35,** 691–719.

BOYD, R. (1989). Mistakes allow evolutionary stability in the repeated prisoner' dilemma game. *J. theor. Biol.* **136,** 47–56.

BOYD, R. & LORBERBAUM, J. P. (1987). No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. *Nature* **327,** 58–59.

BOYD, R. & RICHERSON, P. J. (1988). The evolution of reciprocity in sizable groups. *J. theor. Biol.* **132,** 337–356.

DUGATKIN, L. A. (1990). *N* person games and the evolution of cooperation: a model based on predator inspection in fish. *J. theor. Biol.* **142,** 123–135.

HAMBURGER, H. (1973). *N* person prisoner's dilemma. *J. Math. Soc.* **3,** 27–48.

IKEGAMI, T. (1994). From genetic evolution to emergence of game strategies. *Physica D* **75,** 310–327.

LINDGREN, K. (1992). Evolutionary phenomena in simple dynamics. In: *Artificial Life II* (Langton, C. G. *et al*., eds) pp. 295–312. Redwood City, CA: Addison-Wesley.

LINDGREN, K. & NORDAHL, M. G. (1994). Evolutionary dynamics of spatial games. *Physica D* **75,** 292–309.

LINDGREN, K. (1996). Evolutionary dynamics in game-theory models. In: *The Economy as an Evolving Complex Systems II* (Arthur, W. B., Durlauf, S. & Lane, D., eds). pp. 337–367. Reading, MA: Addison-Wesley.

LORBERBAUM, J. (1994). No strategy is evolutionarily stable in repeated prisoner's dilemma. *J. theor. Biol.* **168,** 117–130.

MILLER, J. H. (1996). The coevolution of automata in the repeated prisoner's dilemma. *J. Econ. Behav. Org.* **29,** 87–112.

MOLANDER, P. (1985). The optimal level of generosity in a selfish, uncertain environment. *J. Conflict Resolution* **29,** 611–618.

MOLANDER, P. (1992). The prevalence of free riding. *J. Conflict Resolution* **36,** 756–771.

MULLER, U. (1987). Optimal retaliation for optimal cooperation. *J. Conflict Resolution* **31,** 692.

NOWAK, M. & 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.

RAPOPORT, A. (1994). Editorial comments. *J. Conflict Resolution* **38,** 149–151.

RUBINSTEIN, A. (1986). Finite automata play the repeated prisoner's dilemma. *J. Econ. Theory* **39,** 83–96.

VAN DAMME, E. (1987). *Stability and Perfection of Nash Equilibria*. Berlin: Springer-Verlag.

APPENDIX A

The Boundary between Phase-α and β

When the payoff matrix is given by $(R,S,T,P) = (3,0,p,q)$, the average scores for the strategy containing the *Pavlov*- or *Punish-once*-type defensive structures, and the invader in Table 2 are

$$\frac{1.5 + q}{2}$$

and

$$\frac{p + q}{2},$$

respectively. Therefore, the total score is:

$$\begin{aligned} \textit{defender} \ldots\ & 3 \times 8 + \frac{1.5 + q}{2} \times 4 \\ &= 27 + 2q \\ \textit{invader} \ldots\ & \frac{p + q}{2} \times 12 \\ &= 6p + 6q \end{aligned}$$

In order to be quasi-stable: $[27 + 2q > 6p + 6q]$

APPENDIX B

The Broken Line across Phase-γ

Consider the situation shown in Fig. B1, where A denotes *Strategy-X* and B represents the best invader against A. Each B plays four two–one games and eight one–two games, but the constitution of the games that A plays depends on the site. That is, the one above or below B, named A_1, plays seven self, four two–one, and one one–two games, and the one on the right or left side of B, referred to as A_2, plays eight self and four two–one games. According to Tables 2 and 4, when the payoff matrix is given by $(R,S,T,P) = (3,0,p,q)$, the total score is:

$$\begin{aligned} A_1 \ldots\ & 3 \times 7 + q \times 4 + 0 \times 1 \\ &= 21 + 4q \\ A_2 \ldots\ & 3 \times 8 + q \times 4 \\ &= 24 + 4q \\ B \ldots\ & \frac{p + q}{2} \times 4 + q \times 8 \\ &= 2p + 10q \end{aligned}$$

Considering (the score of A_1) < (the score of A_2), in order for *Strategy-X* not to be stable: $24 + 4q < 2p + 10q$.

APPENDIX C

The Boundary between Phase-γ and δ

Also consider the situation shown in Fig. B1. In order for the invaders to spread, the total score needs to be above that in a situation of mutual cooperation, that is: $2p + 10q > 36$.

A	*A*	*A*	*A*	*A*	*A*
A	*A*	*A*	*A*	*A*	*A*
A	*A*	*B*	*B*	*A*	*A*
A	*A*	*A*	*A*	*A*	*A*
A	*A*	*A*	*A*	*A*	*A*

FIG. B1. Two individuals of type-*B* within the population of type-*A*.