
A Strategy with Novel Evolutionary Features for the Iterated Prisoner's Dilemma

Jiawei Li

School of Computer Science, University of Nottingham, Nottingham, NG8 1BB, UK

jwl@cs.nott.ac.uk

Graham Kendall

School of Computer Science, University of Nottingham, Nottingham, NG8 1BB, UK

gkx@cs.nott.ac.uk

Abstract

In recent iterated prisoner's dilemma tournaments, the most successful strategies were those that had identification mechanisms. By playing a predetermined sequence of moves and learning from their opponents' responses, these strategies managed to identify their opponents. We believe that these identification mechanisms may be very useful in evolutionary games. In this paper one such strategy, which we call collective strategy, is analyzed. Collective strategies apply a simple but efficient identification mechanism (that just distinguishes themselves from other strategies), and this mechanism allows them to only cooperate with their group members and defect against any others. In this way, collective strategies are able to maintain a stable population in evolutionary iterated prisoner's dilemma. By means of an invasion barrier, this strategy is compared with other strategies in evolutionary dynamics in order to demonstrate its evolutionary features. We also find that this collective behavior assists the evolution of cooperation in specific evolutionary environments.

Keywords

Iterated prisoner's dilemma, invasion barrier, identification mechanism, evolutionarily stable.

1 Introduction

Since the iterated prisoner's dilemma (IPD) tournaments of Robert Axelrod and *The Evolution of Cooperation*, the IPD has heavily been studied in game theory, machine learning, and evolutionary computation communities (Axelrod, 1980a, 1980b, 1984; Fogel, 1993; Hoffmann, 2000). Methodologies of evolutionary computation have been applied to IPD, for example, evolution strategies (Fogel, 1993; Thibert-Plante and Charbonneau, 2007), evolutionary programming (Fogel, 1995), and co-evolution (Darwen and Yao, 2002). Another topic of interest is the interaction of learning and evolution. It shows that adaptive players, learning from the games in which they are involved, are more likely to survive than non-adaptive players in evolutionary IPD. An effective adaptation method is to identify and exploit your opponents. In this paper, we investigate how a rule-based identification mechanism can be applied by *collective strategies* to identify the opponents and how collective behaviors assist in improving the fitness of a group of strategies.

In the 2004 and 2005 IPD tournaments (Kendall, Yao, and Chong, 2007), some strategies with identification mechanisms performed extremely well and defeated well-known strategies in round-robin competitions. These mechanisms identified an

opponent according to its response to a certain sequence of cooperate and defect. In the 2004 tournament, a team from the University of Southampton (Rogers et al., 2007) introduced a group of strategies that were designed to recognize each other through a known series of 5 to 10 moves at the start of each game. Once two Southampton players recognized each other, they would act as a master or a slave. A master will always defect while a slave will always cooperate in order for the master to win the maximum payoff. If a Southampton player recognized that the opponent was not a partner, it would immediately defect to minimize the score of the opponents. The Southampton group strategies succeeded in defeating any non-grouped strategies and won the top three positions in the competition. However, the master/slave approach is not always possible, for example, when the rules of the competition only allow one entry per person. Competition 4 of the 2005 tournament mirrored the original competition of Axelrod, in which only one entry was permitted for each participant. The winner, Adaptive Pavlov, also had an identification mechanism. By analyzing the moves it made, and then adopting the corresponding optimal strategy in its interaction, the strategy was able to identify most of the opponents and thus to respond almost optimally in all games in which it participated (Li, 2007).

Strategies with identification mechanisms behave differently from others not only in round-robin tournaments but also in evolutionary dynamics. We found that some of these strategies were especially strong in maintaining a homogeneous population. One such strategy, which we call *collective strategy*, will be analyzed in this paper.

1.1 Previous Research

Computing the best response to an unknown opponent has long been studied by researchers of multi-agent systems. With respect to IPD, several approaches have been developed to solve the problem of learning optimal responses to a deterministic or mixed strategy (Carmel and Markovitch, 1998; Darwen and Yao, 2002). However, the problem is usually intractable because of the computational complexity, and finding the best response for an arbitrary strategy can be non-computable (Nachbar and Zame, 1996). This is one of the reasons why there have not been many strategies that have identification mechanisms submitted to IPD tournaments.

A strategy with a simple identification mechanism, which was named handshake by Robson (1990), first appeared in evolutionary IPD (see Chong et al., 2007, for a review of evolutionary IPD). This strategy defects in the first move and cooperates in the second move. If the opponent behaves the same as handshake does, it will keep cooperating in subsequent moves. Otherwise, it will always defect. This “initial defect then cooperate” can be seen as a password. Any strategy that knows this password (or behaves like it by chance) may evoke handshake’s cooperation while others trigger defection. This demonstrates that handshakes have an effect on promoting cooperation in evolutionary IPD. This strategy has also been analyzed in the context of IPD with choice and refusal options (Stanley, Ashlock, and Tesfatsion, 1994; Ashlock et al., 1996).

Just like handshakes, recently developed identification mechanisms identify the opponents in a simple way. They were designed to identify common strategies, rather than every possible strategy. This greatly reduces the amount of computation required and makes identifying an opponent in limited rounds possible as long as the strategy is one of the common strategies. These identification mechanisms played a determined sequence of cooperate and defect and categorized the opponents from their response (Li, 2007).

A fundamental problem in evolutionary game theory is to explain how cooperation can emerge and persist in a population of self-interested individuals. Many psychological and economic experiments (Cooper et al., 1996; Kahn and Murnighan, 1993) have shown that even in finite IPDs, subjects would not necessarily apply a strategy such as always defects (AllD). Game theorists (Harsanyi, 1967; Kreps et al., 1982; Simon, 1990) explained these experimental results in terms of incomplete information, reputation, and bounded rationality. It has been proven that information exchange aids in cooperation between players in games with incomplete information (Farrel, 1995; Wiseman and Yilankaya, 2001). However, in games like IPD where communication between players is not allowed, how cooperation emerges is still unclear. An identification mechanism is an approach for players to derive information about the opponent from the history of interactions. When two players with an identification mechanism meet, their identification mechanisms play a role of communication between them. Therefore, identification mechanisms can also aid cooperation between players.

One of the main motivations for studying IPD strategies is to find an evolutionarily stable strategy (ESS). According to Maynard Smith (1982), an ESS is a strategy such that, if all the members of a population adopt it, then no mutant strategy could invade the population under the influence of natural selection. However, no known strategy in the IPD is an ESS when players maintain long-term relationships (Boyd and Lorberbaum, 1987; Lorberbaum, 1994). A defecting strategy that can stand invasions from cooperative strategies cannot prevent other defecting strategies from successful invasions. Similarly, a cooperative strategy that can stand invasions from defecting strategies cannot resist invasions from other cooperative strategies. It is difficult to estimate the direction of evolution in an evolutionary game where no ESS exists. For example, in an evolutionary process where four strategies of always cooperates (AllC), always defects (AllD), tit for tat (TFT), and generous tit for tat (GTFT; like TFT but occasionally cooperates after a defection) are involved (Kahn and Murnighan, 1993), evolution starts off toward defection but then veers toward cooperation. TFT strategies play a key role in invading the population of defectors. GTFT strategies, and then the more generous AllCs gradually become dominant once cooperation is widely established, and this provides an opportunity for AllD to invade again.

The definition of an ESS is not concerned with the parameters of the population dynamics. It ensures that an ESS can be stable in a varied environment of IPD and a population of ESSs becomes a static system where evolution stops (Lorberbaum, 1994). Any currently known strategy is excluded under this definition. However, when specific evolutionary dynamics are concerned, there possibly exists a certain strategy that can be evolutionarily stable. This ESS is dynamic-dependent. We find that collective strategies are able to keep a dynamically stable population in a specific environment where there are periodic invasions of small clusters of AllDs.

1.2 Collective Strategy

Collective strategy (CS) is designed to make its group members acquire higher fitness than others. To achieve this objective, a CS cooperates with its kin members and defects against others. Its identification mechanism is designed to distinguish itself from most known strategies.

A CS behaves in this way: It always cooperates in the first move and defects in the second move. If the opponent also cooperates in the first move and defects in the second move, CS will cooperate until the opponent defects. Otherwise, CS will always defect.

Table 1: General formulation of the prisoner’s dilemma payoff matrix.

		Player 2		R = reward for mutual cooperation S = sucker’s payoff
		Cooperate	Defect	
Player 1	Cooperate	(R, R)	(S, T)	T = temptation to defect P = punishment for mutual defection
	Defect	(T, S)	(P, P)	

Thus, a CS will only cooperate with its kin strategies and defect against any others. When two CSs meet, they both play a sequence of “CDCCC.” Few other strategies behave in this manner when interacting with a CS. A cooperative strategy seldom defects in the second move after cooperating in the first round. Few defecting strategies choose to cooperate in the second move after defecting and receiving the maximum payoff in the first round. Therefore, most known strategies can be distinguished from CS after two rounds in an IPD. In this simple way, CS manages to cooperate with its kin strategies and defect against others.

CS is different from other strategies with an identification mechanism, for example, handshakes and Southampton’s group strategies. Firstly, because few known strategies behave as “CDCCC” in interacting with CS, CS can distinguish itself from most known strategies in a short period (2–3 rounds), which makes CS almost the most efficient strategy that is able to distinguish itself from others. Secondly, CS conditionally cooperates with its kin members, which prevents it from being exploited. Unlike handshakes, CS will always defect once the opponent that is considered to be a kin member defects.

CS is not designed to win IPD tournaments, but to maintain a stable population in the evolutionary IPD. A single CS performs poorly in IPD game. However, a group of CSs is powerful in evolution. A population of CSs keeps cooperation among group members and defects against those outside the group. The fitness of the group increases in the population as the size of the group increases.

It will be more difficult to invade a population of CS than to invade any other well-known strategy. One of the reasons why a strategy such as TFT cannot prevent another strategy such as ALLC from invading is that they behave the same when interacting with one another. If a population cannot distinguish between itself and an invader, then to stop an invasion becomes impractical. However, strategies with identification mechanisms could distinguish one type of strategy from another. Thus, these strategies have the potential to be evolutionarily stable in IPD games. In this paper, CS will be analyzed and compared with both cooperative and defecting strategies in order to show its novel evolutionary features.

2 Iterated Prisoner’s Dilemma

The general form of the PD is shown in Table 1, where $R, S, T,$ and P denote, respectively, reward for mutual cooperation, sucker’s payoff, temptation to defect, and punishment for mutual defection, and $T > R > P > S$ and $R > \frac{1}{2}(S + T)$. The two constraints motivate each player to play noncooperatively and prevent any incentive to alternate between cooperation and defection. In an IPD game, two players have to choose their mutual strategy repeatedly and have the option to retain a memory of the previous behaviors of both players.

Our objective is to analyze CS in evolutionary IPDs where a population of strategies plays IPD with one another and each strategy has the chance to produce offspring proportional to its fitness (payoff) (Kraines and Kraines, 2000). At each time step, the

strategies with higher fitness reproduce and replace those with lower fitness. In this way, the population converges to strategies with higher fitness. In order to compare between different IPD strategies, it needs a measure of evolutionary stability.

Suppose that there are two types of strategies in the population, A and B . Let $E(B, A)$ denote the payoff strategy B receives in interacting with strategy A . The strategy A is evolutionarily stable if

$$E(A, A) \geq E(B, A), E(A, B) > E(B, B) \tag{1}$$

for all $A \neq B$. This form of definition is taken from Thomas (1985).

The concept of an ESS considers those situations when a single mutant invades an infinite population of homogeneous strategies. It is not concerned with the structure of the population, the selection scheme, and other parameters of evolutionary dynamics. If an ESS exists for an evolutionary game, the evolutionary process is, in theory, likely to converge to the state where ESSs are common. However, if no ESS exists, it is difficult for game theorists to forecast the result of the evolution. Unfortunately, no known strategy is evolutionarily stable in IPD. We have to seek another measure to evaluate the evolutionary stability of IPD strategies.

Nowak and Sigmund (1992) considered the size of the cluster that is needed for an invader to invade a finite population of a particular strategy. The minimal cluster size for one strategy to invade another can be treated as an invasion barrier. If the invasion barrier for a strategy is low, it means that a small cluster of invaders can successfully invade and it is difficult to maintain a homogeneous population. On the contrary, if the invasion barrier for a strategy is high, successful invasion requires a large cluster of invaders. Therefore, different strategies can be compared by means of their invasion barriers. Strategies with a higher invasion barrier are evolutionarily stronger than those with a lower invasion barrier. By means of replacing the quantity of each strategy in the population with the ratio of the quantity to the size of population, an invasion barrier can be used in evolutionary IPDs with both finite and infinite populations.

Consider a population consisting of two types of strategies, A and B . p_A and p_B are the frequencies of A and B in the population respectively. The IPD between two strategies is denoted by the payoff matrix:

$$\begin{matrix} & A & B \\ A & \begin{pmatrix} a & b \end{pmatrix} \\ B & \begin{pmatrix} c & d \end{pmatrix} \end{matrix} \tag{2}$$

Then the fitness of two strategies, E_A and E_B , can be expressed as:

$$E_A = ap_A + bp_B, \quad E_B = cp_A + dp_B. \tag{3}$$

The condition for A to invade B is $E_A > E_B$, or $(a - c)p_A > (d - b)p_B$. $(a - c)p_A < (d - b)p_B$, on the other hand, is the condition for B to invade A . Therefore, $(a - c)p_A = (d - b)p_B$ is the transition point for the evolution. Without loss of generality, suppose that $a > c$ and let $\lambda = p_A/p_B$. The invasion barrier can be expressed as

$$\lambda_0 = \frac{d - b}{a - c}. \tag{4}$$

In the case of $\lambda = \lambda_0$, two strategies will coexist and maintain their current frequencies in the population. $\lambda > \lambda_0$ and $\lambda < \lambda_0$ indicate contrary directions of the evolution. In one case, A tends to become dominant and B tends to die out; in the other case, A dies out and B becomes dominant.

Equation (1) is equivalent to $a \geq c$ and $b > d$. This means that $E_A > E_B$ holds for any frequency of A . λ_0 does not exist in this situation, or in other words, there is no invasion barrier for an ESS to invade another strategy.

By using the index of invasion barrier, one strategy can be compared with another. $\lambda_0 = 1$ denotes that two strategies are equally evolutionarily strong because they will coexist if their frequencies in the population are equal. $\lambda_0 < 1$ means that A is evolutionarily stronger than B ; and B is evolutionarily stronger than A if $\lambda_0 > 1$.

CS players only cooperate with their group members and defect against others. This group behavior makes them especially strong in evolution. CS will be analyzed in evolutionary IPDs where cooperative strategies and defecting strategies are involved. In this paper, cooperative strategies denote those strategies that will not defect before their opponent does, and defecting strategies indicate those that always defect in the first round. Our definition of cooperative and defecting strategies does not include strategies with identification mechanisms.

The set of cooperative and defecting strategies includes most of the well-known IPD strategies. IPD strategies can be categorized according to whether they respond to their opponents. Strategies that do not respond to the opponent either play a predetermined sequence of C and D or just play randomly. It is obvious that AllD is the optimal strategy to deal with them. Strategies that respond to the opponent behave differently when interacting with different opponents and they may also play randomly occasionally. The way a strategy responds to the opponent can be either simple or complicated. However, almost all of the successful strategies in IPD have adopted simple ways of responding. This is the reason why simple identification mechanisms could work well.

3 Evolutionary Features of CS

CS differs from any well-known strategy in terms of not only the identification mechanism but also its evolutionary features. In an evolutionary IPD where defecting strategies, cooperative strategies, and CSs are involved, one of the three types of strategies will eventually dominate the population and the others die out. In this section, CS is compared with defecting strategies, cooperative strategies, and a mixture of these two strategies in order to show its novel evolutionary features.

3.1 CS vs. Defecting Strategies

AllD is chosen as the representative for defecting strategies. Against AllD, CS is playing in the worst position against defecting strategies because any other defecting strategy will not receive higher payoffs than AllD in interacting with CS, and CS will not receive lower payoffs in interacting with any other defecting strategy than it does with AllD. Most defecting strategies aim to exploit the opponent. If a defecting strategy is designed to seek cooperation with the opponent, it is essentially a cooperative strategy except defecting in the first move. Strategies like this can be dealt with as cooperative strategies and they are not taken into consideration in this subsection.

Consider an evolutionary IPD with the payoff matrix shown in Table 1. Let r denote the number of rounds for each IPD (some of the literature applies a constant w to denote

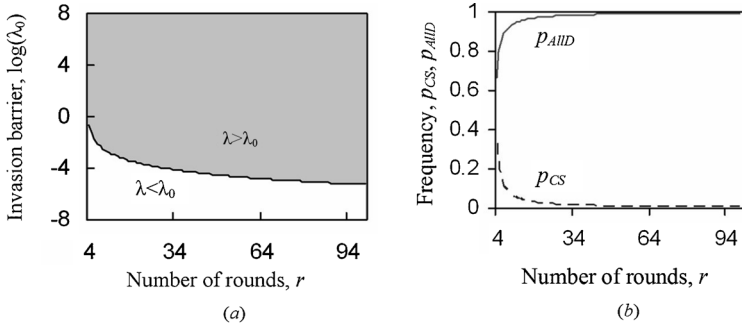


Figure 1: Invasion barrier between CS and AllID shows at what cluster size each strategy can invade the other. (a) $\log(\lambda_0)$ is shown as a function of r . The solid line indicates the invasion barrier between CS and AllID, and the gray-shaded region shows the region where evolution favors CS replacing AllID. (b) The solid line indicates the minimal frequency of AllID that is needed to invade CS, and the dashed line the minimal frequency of CS to invade AllID.

the probability for another round ($0 < w < 1$). Then, the average number of rounds for each IPD is expressed as $r = 1/(1 - w)$. If strategies A and B denote CS and AllID, respectively (refer to Equation (2)), then we have $a = (r - 1)R + P$, $b = (r - 1)P + S$, $c = (r - 1)P + T$, and $d = rP$. There is $a > c > d > b$ provided $r - 1 > \frac{T-P}{R-P}$. According to Equation (4), there is

$$\lambda_0 = \frac{P - S}{(r - 1)(R - P) + P - T}.$$

In Figure 1, the invasion barrier λ_0 is shown as a function of r provided $T = 5$, $R = 3$, $P = 1$, and $S = 0$. It shows that a small cluster of CSs can invade a population of AllID. However, a large cluster of AllIDs cannot invade CS. For example, when $r = 10$, a cluster of CSs greater than 6.7% in the population can invade AllID. However, the necessary condition for AllID to invade CS is that the number of AllID exceeds 93.4% in the population. It is more difficult for AllID to maintain a homogeneous population than CS. Therefore, CS is evolutionarily stronger than AllID.

Spatial games show the effect of the structure of the population on evolutionary dynamics. In spatial IPD, each strategy has its position and only plays IPD with its immediate neighbors. After interactions between neighbors, each site is occupied either by its original owner or by one of the neighbors, depending on who scores the highest total payoff in interacting with its neighbors (Nowak and May, 1993; Lindgren and Nordhal, 1995; Ishibuchi and Namikawa, 2005). CS competes with AllID in a 200×200 square lattice. Each strategy occupies a cell and plays rounds of PD with its neighbors. In the next generation, an individual cell is occupied with the strategy that received the highest payoff among the eight immediate neighbors and the cell itself.

Let us consider the extreme situation where a CS is surrounded by 24 AllIDs (Figure 2(a)). When there is self-interaction, the payoffs for CS and its each neighbor, AllID, are $(r - 1)R + P + 8((r - 1)P + S)$ and $(r - 1)P + T + 8rP$ respectively.

If $r > \frac{T+R+6P-8S}{R-P}$, CS will outperform its neighbors and occupy the cells of those neighbors in the next generation. Because an AllID receives a lower payoff when there

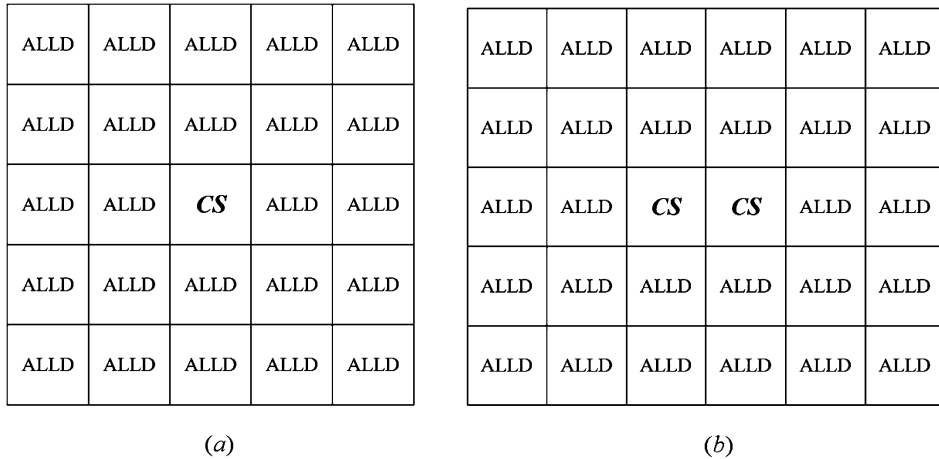


Figure 2: CS competes with AllD in two-dimensional spatial IPD game. (a) A CS surrounded by AllDs. With self-interaction, a single CS outperforms its neighbors if $r > 7$. (b) Two joint CSs surrounded by AllDs. Without self-interaction, two joint CSs will outperform the neighbors and develop to a larger group if $r > 8.5$.

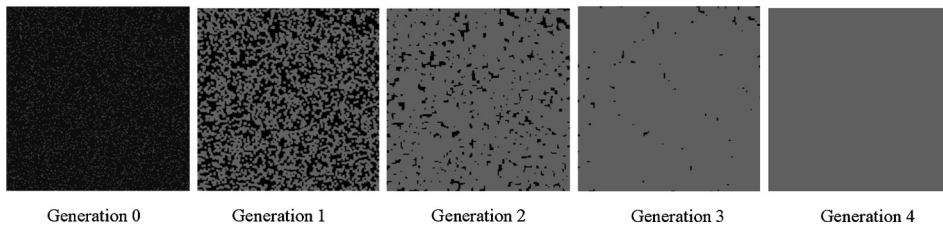


Figure 3: CS competes with AllD. Black represents an AllD and gray represents a CS. The game starts with a population of 10% CS (and 90% AllDs). Conditions: 200×200 square lattice, fixed boundaries, eight neighbors plus self-interaction, random initial configurations, initial frequency of CS $p_{CS} = 0.1$.

are more AllD among its neighbors and a CS receives a higher payoff when there are more CSs among its neighbors, a CS will expel all AllDs if the above condition holds. For example, $r > 7$ is the sufficient condition for a single CS to expel AllDs provided $T = 5, R = 3, P = 1$, and $S = 0$.

When there is no self-interaction, a single CS surrounded by AllDs will always receive a lower payoff than its neighbors and be replaced in the next generation. However, two joint CSs can survive. Consider the situation of two joint CSs surrounded by AllDs (Figure 2(b)). Without self-interaction, the payoff each CS receives is $(r - 1)R + P + 7((r - 1)P + S)$, and payoff for AllD is $2(r - 1)P + 2T + 6rP$ or $(r - 1)P + T + 7rP$.

If $r > \frac{2T+R+4P-7S}{R-P}$, CSs will receive a higher payoff than their neighbors. For example, $r > 8.5$ is a sufficient condition for two joint CSs to expel AllDs provided $T = 5, R = 3, P = 1$, and $S = 0$.

Figures 3 and 4 show typical evolutionary processes of CS vs. AllD. In Figure 3, 10% CS and 90% AllD strategies are randomly placed in a 200×200 square lattice. Each

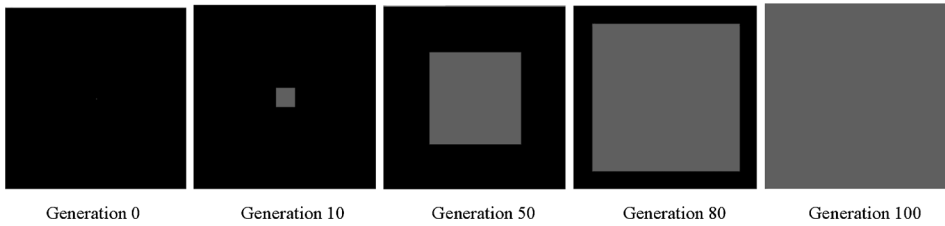


Figure 4: A single CS invades a population of AllD. Black represents an AllD and gray represents a CS. The game starts with the state where a CS lies in a population of AllDs. Conditions: 200×200 square lattice, fixed boundaries, eight neighbors plus self-interaction, initial frequency of CS $p_{CS} = 0.000025$.

strategy plays a 10-round IPD with its eight neighbors, with self-interaction. The results show that AllD dies out after four generations. In Figure 4, AllD strategies occupy a 200×200 square lattice except for a single CS in the middle of the lattice. Other conditions are the same as that of Figure 3. It shows how a single CS is able to invade a population of AllD.

3.2 CS vs. Cooperative Strategies

Again, a representative strategy, GRIM, is chosen to compete against CS, by which CS is playing in the worst position against cooperative strategies. GRIM starts with cooperate but plays always defect after one defection. Any other cooperative strategy will not receive higher payoffs than GRIM in interacting with CS, and CS will not receive lower payoffs in interacting with any other cooperative strategy than it does with GRIM. Therefore, if CS outperforms GRIM it defeats all cooperative strategies in evolutionary IPD.

If strategies A and B denote CS and GRIM respectively (refer to Equation (2)), then we have $a = (r - 1)R + P$, $b = (r - 2)P + R + T$, $c = (r - 2)P + R + S$, and $d = rR$. There is $d > b$ and $a > c$ provided $r - 1 > \frac{T-P}{R-P}$. According to Equation (4), we have

$$\lambda_0 = 1 - \frac{T - R + P - S}{(r - 1)(R - P) + 3P - 2R - S}.$$

In Figure 5, the invasion barrier λ_0 is shown as a function of r provided $T = 5$, $R = 3$, $P = 1$, and $S = 0$. CS is evolutionarily stronger than GRIM because the minimal cluster size for CS to invade the other strategy is relatively smaller than that for GRIM. For example, when $r = 10$, a cluster of CSs more than 45.2% in the population can invade GRIM. However, it needs a cluster of GRIMs more than 54.9% in the population to outperform CS.

As r increases, then $\log(\lambda_0) \rightarrow 0$, which means that a population with an almost equal number of CS and GRIM remains stable when the number of iterations is infinite. If the quantity of CSs is equal to, or more than, one half of the population, it will expel GRIM.

CS and GRIM are almost equally evolutionarily strong in spatial IPD. Figure 6 shows a spatial IPD game with the same parameters as Figure 3. It starts with a population of 50% CS and 50% GRIM. After 10 generations, the game enters a stable state in which neither CS nor GRIM can invade the regions occupied by the other strategy. In

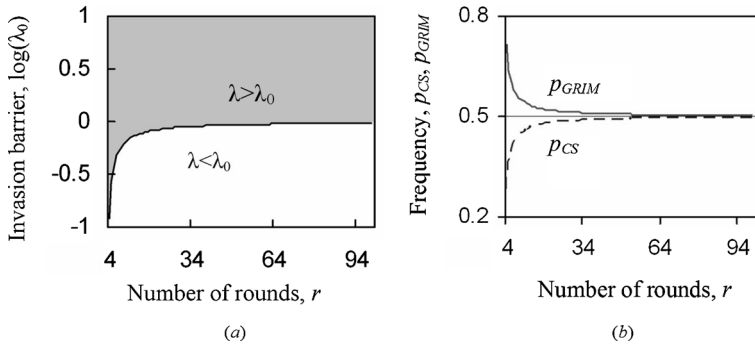


Figure 5: Invasion barrier between CS and GRIM is shown. (a) $\log(\lambda_0)$ is shown as a function of r . The solid line indicates the invasion barrier between CS and GRIM, and the gray-shaded region shows the parameter region where evolution favors CS replacing GRIM. (b) The solid line indicates the minimal frequency of GRIM that is needed to invade CS, and the dashed line the minimal frequency of CS to invade GRIM. The plot has the opposite effect of Figure 1; specifically, increasing r reduces the advantage of CS compared to GRIM.

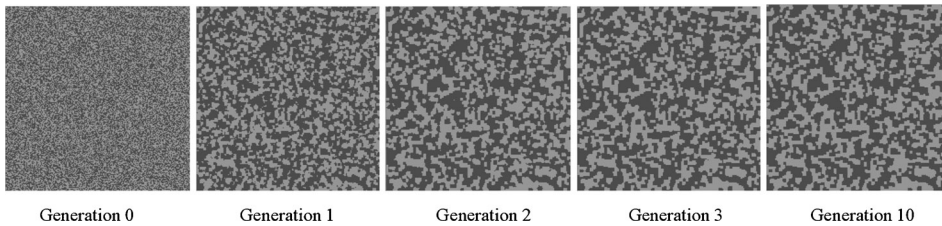


Figure 6: CS competes with GRIM in spatial IPD. Light gray represents GRIM and dark gray represents CS. The game starts with 50% CS (and 50% GRIM). Conditions: 200×200 square lattice, fixed boundaries, eight neighbors plus self-interaction, random initial configurations.

comparison with the fact that CS expels GRIM in an infinite population with equal numbers of both strategies, Figure 6 shows the border effect of spatial IPD.

CS outperforms GRIM in term of resisting invasions by other strategies. Most cooperative strategies, such as TFT and AllC, do as well as GRIM in a cooperative environment. They coexist with GRIM in a population so that defecting strategies are given a chance to invade. For example, AllD can invade a population of GRIM indirectly. Because GRIM cannot expel AllC, a single AllD will successfully invade if AllC has occupied a significant percentage of the population. CS does not coexist with cooperative strategies. In fact, CS does not coexist with any strategy that is identified to not be a kin member. Thus, the chance for a strategy to indirectly invade a population of CS is small.

3.3 CS vs. a Mixture of Defecting and Cooperative Strategies

TFT cannot expel other cooperative strategies such as AllC and GRIM under an evolutionary selection scheme; AllD cannot expel other defecting strategies such as suspicious TFT (STFT). CS differs from most other strategies in the way that it stands in the

population. In an evolutionary IPD that involves CS, CS will either die out or occupy the whole population.

Consider a population that consists of three strategies: CS, TFT, and AllD. TFT behaves the same as GRIM when interacting with CS and AllD. Let p_A , p_B , and p_C denote the frequencies of CS, TFT, and AllD, respectively. The fitness of each strategy can be computed as:

$$\begin{aligned}
 F_{CS} &= p_A((r-1)R + P) + p_B((r-2)P + R + T) + p_C((r-1)P + S) \\
 F_{TFT} &= p_A((r-2)P + R + S) + p_B r R + p_C((r-1)P + S) \\
 F_{AllD} &= p_A((r-1)P + T) + p_B((r-1)P + T) + p_C r P.
 \end{aligned}$$

Comparing F_{CS} with F_{TFT} and F_{AllD} , the conditions for CS outperforming the other strategies are:

$$\left\{ \begin{array}{l} \frac{p_A}{p_B} > 1 - \frac{T - R + P - S}{(r-1)(R-P) + 3P - 2R - S} \\ \frac{p_A}{p_C} > \frac{R - S - \frac{1}{p_C}(R-P)}{(r-1)(R-P) + 2P - R - T} \\ r - 1 > \frac{T - P}{R - P} \end{array} \right. \quad (5)$$

According to Equation (5), if the number of CSs is not less than the number of TFTs and the number of AllDs is not large enough (e.g., not more than 93.4% providing $r = 10$), CS will expel all other strategies. In this situation, evolution will converge to the state of a homogenous CS population.

When interacting with CS, no cooperative strategy receives a higher payoff than TFT and no defecting strategy receives a higher payoff than AllD. On the other hand, CS will not receive a lower payoff if we substitute TFT by other cooperative strategies or substitute AllD by other defecting strategies. Therefore, Equation (5) gives sufficient conditions for CS to outperform cooperative and defecting strategies.

When the strategies involved in the evolution are not limited to just cooperative and defecting strategies, it will be difficult to find a general form of the conditions as Equation (5). With the concept of an invasion barrier, we can estimate what the theoretical evolutionary strongest strategies should be.

Let us assume that a group of strategies can identify their opponents at no cost and they just cooperate with group members and defect against others. Suppose that strategy A in Equation (2) is such a strategy. For any B , there must be $a \geq d$ and $b \geq c$, and then $E_A \geq E_B$. This means that A must be among the evolutionarily strongest strategies. This strategy is only possible in theory because it requires that there is communication between strategies so that they know each other before the game. In IPD games, however, strategies have to identify the opponents in the interaction processes. It can be verified that at least two rounds of interaction are required to distinguish between some of the well-known strategies. The identification mechanism of CS is among the most efficient. Therefore, CS must perform well if it can successfully distinguish other strategies from its kin members.

A situation that CS cannot handle well is the misidentification of a strategy. Some strategies may behave the same as CS in interacting with a CS but behave differently in

interacting with other strategies. An example is the deterministic strategy that automatically plays the sequence “CDCCC...”. AllD and GRIM can (almost) optimally exploit this strategy but CS cannot. If there are lots of these strategies in the population, CS will not perform as well as it does in other circumstances. A homogeneous CS population cannot resist invasions from these types of strategies. This leaves the possibility for other strategies to indirectly invade the population.

A solution for the problem of misidentification is to improve the identification mechanism so that CS can identify these fake kin strategies. For example, if CS plays an extra defect in the 20th round when the opponent is supposed to be CS, many of these fake CS strategies can be identified. This improvement is achieved at the cost of increasing computational complexity and a loss of fitness. However, there is a more efficient approach for a population of CS to get rid of fake kin strategies. By means of introducing a small group of invaders that are not strong enough to expel CS, CS can successfully maintain a stable population. This will be discussed in the following section.

4 Toward Evolutionary Stability in Heterogeneous IPDs

The definition of an ESS is not concerned with the parameters of the population dynamics. Any currently known strategy is excluded under this definition if the number of iterations of IPD is large enough. One way to make some strategy to be ESS is to restrict the possible strategies (Lorberbaum et al., 2002). For example, if there are only AllC and AllD, AllD becomes an ESS. However, if any strategy is permitted for an evolutionary IPD, some restrictions are needed for a homogeneous population to remain stable in the evolution. We find that some strategies may be evolutionarily stable when specific evolutionary dynamics are defined. A dynamics-dependent ESS can maintain a stable population if the evolutionary dynamics remains unchanged. CS is able to maintain a dynamically stable population by introducing periodic invaders of AllD.

First of all, we are going to prove that any strategy that can be distinguished by CS's identification mechanism cannot invade a homogeneous CS population.

In a homogeneous population, the fitness of CS is $(r - 1)R + P$. The necessary condition for an invader to successfully invade is that it receives an equal or higher payoff than $(r - 1)R + P$ in interacting with CS. Once CS identifies that the opponent is different from itself, it will always defect, and then its opponent will receive a payoff no more than P per round. Before the invader is identified, it cannot defect against CS more than once. Therefore, the invader must receive a lower payoff than CS provided $r - 1 > T - P/R - P$.

Possible invaders for a homogeneous population of CS are those strategies that cannot be effectively identified. When the number of those strategies in the population is sufficiently large, there is the possibility for other strategies to invade. Consider an evolutionary IPD where stochastic selection is involved. At each time step, one individual is chosen for reproduction, in proportion to its fitness, and its offspring replaces a randomly chosen individual. Suppose the selection rate is a constant p , which is the probability of one individual being replaced at each time step. Let CD+AllC denote the deterministic strategy that plays C in first round and D in second round and then behaves as AllC from third round to the end. CD+AllC belongs to those strategies that CS cannot distinguish from itself, and each CD+AllC can invade a population of CS with probability p . If CD+AllC strategies appear continuously, it will definitely invade the population. After CD+AllC invades the population, some other strategies,

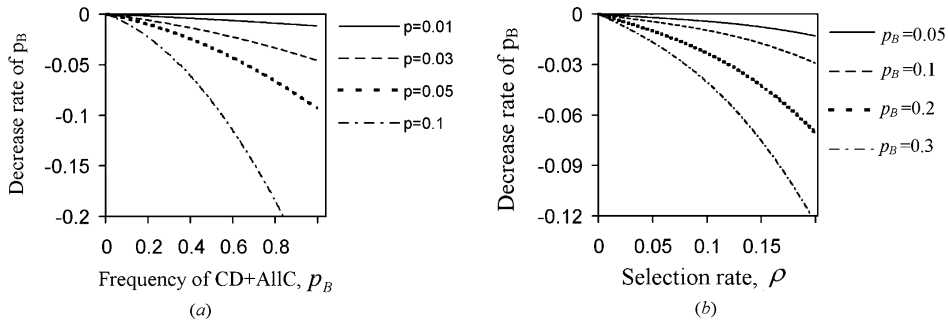


Figure 7: Invasion of AllD makes the frequency of CD+AllC decrease provided $T = 5$, $R = 3$, $P = 1$, $S = 0$, and $r = 10$. (a) Decrease rate of CD+AllC is shown as the function of the initial frequency of CD+AllC. (b) Decrease rate of CD+AllC is shown as the function of the selection rate ρ . It shows that both high initial frequency and high selection rate lead to a high decrease rate of the frequency of CD+AllC.

for example TFT, will make the ratio of CD+AllC increase because CD+AllC receives higher payoffs than CS by cooperating with TFT. Thus, continuous invasions of TFT will help CD+AllC expel CS. After CD+AllCs have occupied a sufficient percentage of the population, TFT is then going to take over.

However, CS can maintain dominance by introducing periodic invasions of AllD, which will prevent the number of fake CSs from increasing.

Consider a population that consists of CS and CD+AllC. The frequencies of CS and CD+AllC are p_A and p_B , respectively ($p_A + p_B = 1$). The selection rate is ρ . After an invasion of a group of AllDs, the fitness of each strategy can be expressed as

$$\begin{aligned}
 F_{CS} &= (1 - \rho)((r - 1)R + P) + \rho((r - 1)P + S) \\
 F_{CD+AllC} &= (1 - \rho)((r - 1)R + P) + \rho((r - 1)S + P) \\
 F_{AllD} &= (1 - \rho)p_A((r - 1)P + T) + (1 - \rho)p_B((r - 1)T + P) + \rho r P.
 \end{aligned}$$

Because there is always $F_{CS} > F_{CD+AllC}$, the ratio of the number of CD+AllC to that of CS, p_B/p_A , will decrease. Figure 7 shows the decrease rate of p_B as functions of the selection rate ρ and the initial frequency of CD+AllC provided $T = 5$, $R = 3$, $P = 1$, $S = 0$, and $r = 10$.

Continuous invasions of AllD and a long period of selection will be needed to remove those CD+AllCs, especially when CD+AllCs only have a small presence in the population. However, the frequency of CD+AllC will decrease quickly, to a low level, because a high frequency of CD+AllC will greatly increase the fitness of AllD, which means that the fitness of CD+AllC will be far below the average fitness. Therefore, the mutations of AllD are sufficient for CS to be dominant in the population.

Among those strategies that play C in the first round and D in the second round, CS is the optimal strategy in interacting with AllD. Thus, AllD will be able to restrain most of those fake CSs from flourishing. For example, Figure 8 shows that the invasion of a small cluster of AllD is efficient in restraining the frequency of CDC+TFT. CDC+TFT denotes the strategy that plays CDC in the first three rounds and then behaves as TFT.

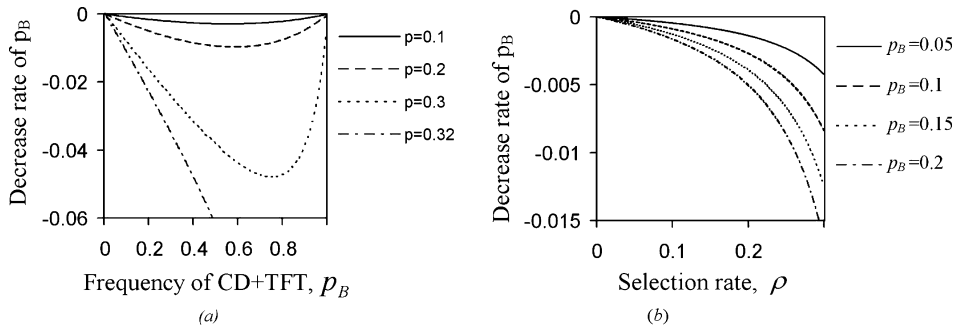


Figure 8: Invasion of AllD is efficient in restraining the frequency of CDC+TFT provided $T = 5, R = 3, P = 1, S = 0,$ and $r = 10$. (a) Decrease rate of CDC+TFT is shown as the function of the initial frequency of CDC+TFT. (b) Decrease rate of CDC+TFT is shown as the function of the selection rate ρ .

According to Section 3.1, it is difficult for AllD to expel CS. The necessary condition for AllD to expel CS is that AllD occupies the majority of the population, for example, more than 93.4% when $r = 10$. In a large range of selection rates, CS can always expel the invaders.

We carried out a simulation of evolutionary IPD with a replicator dynamic. Evolution begins with a population of CS. Until 3,000 generations, 10% of the population is replaced by a cluster of invading strategies every 200 generations. Each type of strategy in the population will then decrease by 10% except the same strategy as the invader. Each invader is randomly selected from the following strategies: AllD, AllC, TFT, GRIM, Pavlov, STFT, CD+AllC, CDC+TFT, Handshake, Fortress4. Pavlov, also known as win-stay loss-shift, cooperates when both sides have cooperated or defected on the previous move, and defects otherwise. STFT behaves like TFT but defects on the first move. Fortress4 (Ashlock and Ashlock, 2006) initially defects and keeps defecting until the opponent has defected three times, and after the third defection, cooperates once. If the opponent does likewise, Fortress4 keeps cooperating until the opponent defects. If the opponent cooperates in response to either of the first three defections or if it defects in response to any cooperation, Fortress4 will defect and return to the initial state. Besides these invaders, 2% of the population will be replaced by AllD every 30 generations. The replicator equation is

$$\dot{p}_i = p_i(F_i - \bar{F}) \tag{6}$$

where p_i denotes the frequency of i th strategy in the population, F_i denotes its fitness, and \bar{F} denotes the average fitness of all strategies (Nowak and May, 1993). The size of the population remains constant in the evolution.

The results of the simulation show that CS is able to keep a dynamically stable population when there are periodic invasions of AllD. AllD helps CS remove fake kin members and to maintain a single strategy population. Then, CS can efficiently resist the invasions of other strategies that can be identified by its identification mechanism. Figure 9 shows the frequency of CS as a function of generations. It is worthwhile noting that the invasions of CDC+TFT and CD+AllC (as indicated by the dotted line in Figure 9)

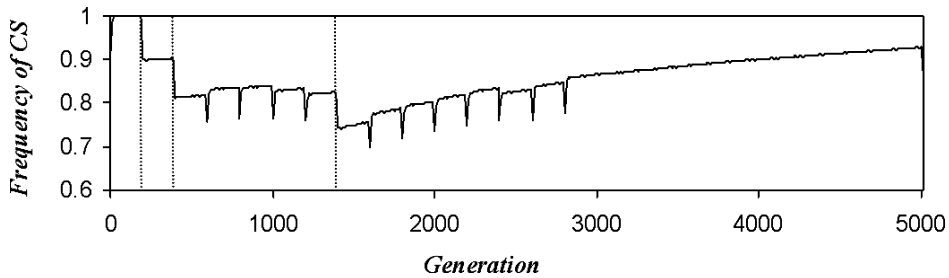


Figure 9: CS remains dynamically stable in an evolutionary IPD where the population is frequently invaded by different strategies. Except for the invasions of AllID every 30 generations, there were totally 15 invasions within 3,000 generations, and the invaders were, in time sequence, Fortress4, CDC+TFT, CDC+TFT, AllC, Handshake, STFT, TFT, CD+AllC, AllC, AllID, Handshake, GRIM, STFT, Fortress4, and Pavlov. Dotted lines indicate the invasions of CDC+TFT and CD+AllC.

lead to an effective reduction of the frequency of CS, and it takes a relatively long time for CS to remove these fake kin members.

With periodic invasions of AllID, those strategies that cannot be identified by CSs' identification mechanism can invade but will be kept at low frequencies in the population. Although the frequency of CS fluctuates when invasions occur, CS always expels most of the invaders.

5 Conclusions

Collective strategy is evolutionarily strong in maintaining a stable population. Because of its identification mechanism, cooperation is achieved between CS members. On the other hand, invaders are distinguished and are then defected against. This active action to expel invaders makes CS especially stable in evolution. By means of invasion barrier, we have shown that CS is evolutionarily stronger than most known strategies.

CS is a conditional defector: it does not defect against its group members. The advantage of this behavior becomes obvious especially when different groups compete for survival. Losers in the competition will suffer the risk of dying off. Thus, cooperation between individuals is crucial for the group to survive. The emergence of cooperation is then inevitable between the group members.

The collective behavior of CS shows another way of payoff maximization. CS is designed to maximize the fitness of its group in evolutionary games. When a group of CSs predominates in the evolution (occupies more resources or produces more offspring than other individuals and groups), each CS does, in fact, maximize its own payoff. This collective behavior shows the power of cooperation since a group of CS easily overwhelms any singleton strategy.

When evolutionary dynamics are considered, it is possible to find some strategies that can be evolutionarily stable in specific circumstances. A population of CSs can remain dynamically stable in the evolutionary IPD by introducing periodic invaders of AllID. Just like the predators that hunt weak animals, invasions of AllID prevent the population of CS from degenerating. Then, the population of CSs could maintain stability against invaders. This aids in explaining how cooperation persists in the

evolution. Cooperation between individual strategies can remain only if both sides profit from the long-term payoffs. The success of CS in evolutionary IPD shows evidence that collective behavior maximizes the fitness of the group and then each individual's long-term payoff is maximized.

The identification mechanisms provide a means of information exchange between IPD players since direct negotiation is not permitted. With identification mechanisms, IPD players are able to "learn" the opponents and then collective behavior is possible. Unlike those unconditional cooperators, CS defects against any other strategies although cooperation would be achieved if it cooperated. This is the reason why a population of CSs can effectively expel invaders. Therefore, it is more likely that cooperation emerges among group members and then spreads to the rest of the population.

Acknowledgments

We would like to thank the three anonymous reviewers for their constructive comments, which have led to a much improved paper. The work of the first author was partly supported by visiting scholar's funding provided by the Chinese Scholarship Council and Harbin Institute of Technology. The work of the second author was supported by an EPSRC (Engineering and Physical Sciences Council) grant (Ref: GR/S63465/1).

References

- Ashlock, W., and Ashlock, D. (2006). Changes in prisoner's dilemma strategies over evolutionary time with different population sizes. *Proceedings of the 2006 Congress on Evolutionary Computation*, pp. 297–304.
- Ashlock, D., Smucker, M., Stanley, E., and Tesfatsion, L. (1996). Preferential partner selection in an evolutionary study of prisoner's dilemma. *BioSystems*, 37:99–125.
- Axelrod, R. (1980a). Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, 24:3–25
- Axelrod, R. (1980b). More effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, 24:379–403.
- Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.
- Boyd, R., and Lorberbaum, J. (1987). No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature*, 327:58–59.
- Carmel, D., and Markovitch, S. (1998). How to explore your opponent's strategy (almost) optimally. *Proceedings of the International Conference on Multi Agent Systems*, pp. 64–71.
- Chong, S., Humble, J., Kendall, G., Li, J., and Yao, X. (2007). Iterated prisoner's dilemma and evolutionary game theory. In G. Kendall, X. Yao, and S. Y. Chong, (Eds.), *The Iterated Prisoners Dilemma: 20 Years On*, chapter 2. Singapore: World Scientific.
- Cooper, R., Jong, D., Forsythe, R., and Ross, T. (1996). Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, 12(2):187–218.
- Darwen, P., and Yao, X. (2002). Coevolution in iterated prisoner's dilemma with intermediate levels of cooperation: Application to missile defence. *International Journal of Computational Intelligence and Applications*, 2(1):83–107.
- Farrell, J. (1995). Talk is cheap. *The American Economic Review*, 85(2):186–190.

- Fogel, D. (1993). Evolving behaviors in the iterated prisoner's dilemma. *Evolutionary Computation*, 1(1):77–97.
- Fogel, D. (1995). On the relationship between the duration of an encounter and the evolution of cooperation in the iterated prisoner's dilemma. *Evolutionary Computation*, 3(3):349–363.
- Harsanyi, J. (1967). Games with incomplete information played by “Bayesian” players. *Management Science*, 14(3):159–182.
- Hoffmann, R. (2000). Twenty years on: The evolution of cooperation revisited. *Journal of Artificial Societies and Social Simulation*, 3(2). Accessed 4/08/09 at <<http://www.soc.surrey.ac.uk/JASSS/3/2/forum/1.html>>.
- Ishibuchi, H., and Namikawa, N. (2005). Evolution of iterated prisoner's dilemma game strategies in structured demes under random pairing in game playing. *IEEE Transaction on Evolutionary Computation*, 9(6):552–561.
- Kahn, L., and Murnighan, J. (1993). Conjecture, uncertainty, and cooperation in prisoners' dilemma games: Some experimental evidence. *Journal of Economic Behavior and Organisms*, 22:91–117.
- Kendall, G., Yao, X., and Chong, S.Y. (Eds.). (2007). *The Iterated Prisoner's Dilemma: 20 Years On*. Singapore: World Scientific.
- Kraines, D., and Kraines, V. (2000). Natural selection of memory—One strategy for the iterated prisoner's dilemma. *Journal of Theoretical Biology*, 203:335–355.
- Kreps, D., Milgrom, P., Roberts, J., and Wilson, R. (1982). Rational cooperation in the finitely repeated prisoner's dilemma. *Journal of Economic Theory*, 27:245–252.
- Li, J. (2007). How to design a strategy to win an IPD tournament. In G. Kendall, X. Yao, and S. Y. Chong, (Eds.), *The Iterated Prisoner's Dilemma: 20 Years On*, chapter 4. Singapore: World Scientific.
- Lindgren, K., and Nordhal, M. (1995). Cooperation and community structure in artificial ecosystems. In C. G. Langton, (Ed.), *Artificial Life, An Overview* (pp. 15–37). Cambridge, MA: MIT Press.
- Lorberbaum, J. (1994). No strategy is evolutionarily stable in the repeated prisoner's dilemma. *Journal of Theoretical Biology*, 168:117–130.
- Lorberbaum, J., Bohning, D., Shastri, A., and Sine, L. (2002). Are there really no evolutionarily stable strategies in the iterated prisoner's dilemma? *Journal of Theoretical Biology*, 214:155–169.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge, UK: Cambridge University Press.
- Nachbar, J., and Zame, W. (1996). Non-computable strategies and discounted repeated games. *Economic Theory*, 8:103–122.
- Nowak, M., and May, R. (1993). The spatial dilemmas of evolution. *International Journal of Bifurcation and Chaos*, 3:35–78.
- Nowak, M., and Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, 355:250–252.
- Robson, A. (1990). Efficiency in evolutionary games: Darwin, Nash, and the secret handshake. *Journal of Theoretical Biology*, 144:379–396.
- Rogers, A., Dash, R. K., Ramchurn, S. D., Vytelingum, P., and Jennings, N. R. (2007). Error-correcting codes for team coordination within a noisy iterated prisoner's dilemma tournament. In G. Kendall, X. Yao, and S. Y. Chong, (Eds.), *The Iterated Prisoner's Dilemma: 20 Years On*, chapter 9. Singapore: World Scientific.

- Simon, H. (1990). A mechanism for social selection and successful altruism. *Science*, 250(4988):1665–1668.
- Stanley, E., Ashlock, D., and Tesfatsion, L. (1994). Iterated prisoner’s dilemma with choice and refusal of partners. In C. Langton, (Ed.), *Artificial Life III*, pp. 131–175. Reading, MA: Addison Wesley.
- Thibert-Plante, X., and Charbonneau, P. (2007). Crossover and evolutionary stability in the prisoner’s dilemma. *Evolutionary Computation*, 15(3):321–344.
- Thomas, B. (1985). On evolutionarily stable sets. *Journal of Mathematical Biology*, 22:105–115.
- Wiseman, T., and Yilankaya, O. (2001). Cooperation, secret handshakes, and imitation in the prisoner’s dilemma. *Games and Economic Behavior*, 37:216–242.