The Effect of Memory Size on the Evolutionary Stability of Strategies in Iterated Prisoner's Dilemma

Jiawei Li, *Member, IEEE,* and Graham Kendall, *Senior Member, IEEE*

***Abstract*—The iterated prisoner's dilemma is an ideal model for the evolution of cooperation among payoff-maximizing individuals. It has attracted wide interest in the development of novel strategies since the success of tit-for-tat in Axelrod's iterated prisoner's dilemma competitions. Every strategy for iterated prisoner's dilemma utilizes a certain length of historical interactions with the opponent, which is regarded as the size of the memory, in making its choices. Intuitively, longer memory strategies must have an advantage over shorter memory strategies. In practice, however, most of the well known strategies are short memory strategies that utilize only the recent history of previous interactions. In this paper, the effect of the memory size of strategies on their evolutionary stability in both infinite length and indefinite length n-person iterated prisoner's dilemma is studied. Based on the concept of a counter strategy, we develop a theoretical methodology for evaluating the evolutionary stability of strategies and prove that longer memory strategies outperform shorter memory strategies statistically in the sense of evolutionary stability. We also give an example of a memory-two strategy to show how the theoretical study of evolutionary stability assists in developing novel strategies.**

***Index Terms*—Evolutionary stability, game theory, iterated prisoner's dilemma, strategies.**

I. Introduction

The PRISONER'S dilemma is a non-zero-sum game in which two players try to maximize their payoff by cooperating with, or betraying the other player [34], [35]. The payoff matrix of the game is shown in Fig. 1.

In the payoff matrix, R, S, T, and P denote reward for mutual cooperation, Sucker's payoff, temptation to defect, and punishment for mutual defection, respectively, and $T > R > P > S$. The constraint motivates each player to play noncooperatively.

When both players are rational and they make their choice independently, the theoretical outcome of the game is a Nash equilibrium, in which both players choose to defect and each receives a punishment for mutual defection. It is worse for each player than the outcome they would have received if they had cooperated.

		Prisoner 2	
		Cooperate	*Defect*
Prisoner 1	*Cooperate*	(R,R)	(S,T)
	Defect	(T,S)	(P,P)

Fig. 1. Payoff matrix of the Prisoner's Dilemma.

In the iterated prisoner's dilemma (IPD) game, two players have to choose their mutual strategy repeatedly, and they also have memory of their previous behaviors and the behaviors of the opponents. $R > \frac{1}{2}(S+T)$ is set to prevent any incentive to alternate between cooperation and defection. IPD is considered an ideal test bed for the evolution of cooperation among selfish individuals and it has attracted wide interest since Robert Axelrod's IPD tournaments and the Evolution of Cooperation [4], [5], [11], [21], [27], [28].

The winner of Axelrod's tournaments was the tit-for-tat (TFT) strategy. TFT starts with cooperation, and then copies the opponent's previous move. Axelrod attributes the success of TFT to its properties of nice, forgiving, retaliating and simple. However, later research has shown some weaknesses of TFT, such as vulnerability to noise and being unable to exploit unconditional cooperators [28], [39]. Since then, researchers have attempted to develop novel strategies that can outperform TFT in either round-robin IPD competitions or evolutionary dynamics and some IPD strategies have been developed, for example, win-stay lose shift [29], generous-TFT [33], gradual [6], and very recently group strategies [22], [24], [36], and zero-determinant [31], [37]. Also, evolution of strategies in replicator dynamics with infinite or finite population, spatial, and noisy environments has been studied in [1], [2], [8], [9], [15]–[17], [19], [20], [26] and [32].

Different IPD strategies may use different lengths of historic interactions. TFT, for example, is a memory-one strategy, only making use of information from the previous stage of interaction in making its next decision. Tit-for-two-tats is a memory-two strategy. Most well known IPD strategies are memory-one and memory-two strategies and only a limited number of strategies use a memory size greater than three.

Intuitively, the players with longer memories can perform at least as well as those with shorter memories. However, it

Manuscript received December 28, 2012; revised October 5, 2013; accepted October 9, 2013. Date of publication October 21, 2013; date of current version November 26, 2014. This work was supported by the Engineering and Physical Sciences Research Council, U.K., under Grant EP/H000968/1.

J. Li is with the School of Computer Science, ASAP Research Group, University of Nottingham, Nottingham NG8 1BB, U.K. (e-mail: jiawei.li@nottingham.ac.uk).

G. Kendall is with the School of Computer Science, ASAP Research Group, University of Nottingham, U.K. and with the University of Nottingham, Jalan Broga, Semenyih 43500, Malaysia (e-mail: graham.kendall@nottingham.ac.uk; graham.kendall@nottingham.edu.my).

Digital Object Identifier 10.1109/TEVC.2013.2286492

is still not clear whether longer memory strategies outperform shorter memory strategies. Few IPD strategies developed in either experiments or tournaments use a very long memory. A nontrivial question is whether or not a longer memory grants a strategy advantage in IPD?

Experiments have revealed that memory helps learning and cooperation in evolution [13], [14], [41]. Posch [30] studied win-stay, lose-shift with diverse memory size and showed that win-stay, lose-shift with longer memory performed better than those with shorter memory in computer simulations. Ashlock and Roger [3] showed that several strategies utilizing long term memory outperformed control strategies in evolutionary IPD. Press and Dyson [31] proved that the shortest memory strategy sets the rule of two-player IPD games. That is, longer memory strategies do not have an advantage over shorter memory strategies in two-player IPD. However, this result cannot be extended to n-player ($n > 2$) IPD cases and the memory-one zero-determinant strategies are not evolutionarily stronger than those known IPD strategies [37], [42].

In this paper, we investigate the effect of memory size on the evolutionary stability of IPD strategies. In order to evaluate the evolutionary stability of IPD strategies, we develop the concept of a counter strategy. A strategy is a counter strategy against another strategy if it receives no less payoff than any strategy in interacting with the opponent. Based on this concept, a number of theorems, which show the relationship between the length of the history a strategy uses and its evolutionary stability, are proven. The contributions of this paper include following:

1) We propose a theoretical methodology for evaluating the evolutionary stability of IPD strategies. Based on the concept of a counter strategy, the evolutionary stability of a strategy is evaluated based on whether the strategy is a counter strategy against itself and the probability that the strategy is a counter strategy against an arbitrary strategy.
2) The effect of the memory size of IPD strategies on their evolutionary stability is analyzed. We prove that longer memory strategies have an advantage over shorter memory strategies in both infinite length and indefinite length n-IPD. A longer memory strategy has a higher probability of winning against an arbitrary strategy than a shorter memory strategy.

The rest of paper is structured as follows. Section II introduces the concept of counter strategy and how the evolutionary stability of IPD strategies can be evaluated by means of a counter strategy. Section III presents three theorems and their proofs. The theorems show the effect of memory size of strategies on their evolutionary stability. Finally, Section IV has concluding remarks.

II. Evolutionary Stability of IPD Strategies

An indefinite length IPD has a discount rate ω ($1 \geq \omega > 0$). The game continues with probability ω and the expected number of iterations of the game are $1/(1-\omega)$. If $\omega = 1$, the game is infinite. In this paper, we focus on n-player infinite length and indefinite length IPDs and ignore the finite length IPD. Let S_i ($i = 1, \ldots n$) denote the strategy of the ith player and $E(S_i, S_j)$ the payoff of S_i playing against strategy S_j. Then, the payoff of S_i playing against all n strategies (including playing against itself) can be expressed as $E(S_i) = \sum_{j=1}^{n} E(S_i, S_j)$.

TABLE I
Some Well Known IPD Strategies With Different Size of Memories

Memory-zero	Memory-one
Always-cooperate Always-defect Random	Tit-for-tat, GRIM triger Generous TFT, Contrite TFT Win-stay-lose-shift, Gradual Zero-determinant
Memory-two	**Memory-$L(L > 2)$**
Tit-for-two-tats Two-tits-for-tat	Fortress Prober Group strategies

A. Counter Strategy

Let H_L denote a history of L moves of interactions between two players before the current move. A strategy that makes use of H_L is called a memory-L strategy. A player should have at least L length of memory in order to adopt a memory-L strategy.

The strategies in IPD can be categorized according to the length of memory they use. Some well known IPD strategies are shown in Table I. Descriptions of all strategies used in this paper can be found in [23]. For those strategies that have variable memory length, the memory length is the longest memory that they use. Let $f(H_L)$ denote a memory-L strategy. The strategy space of a length-L IPD is determined by $\{f, H_L\}$ and the strategy space of an infinite or indefinite length IPD is determined by $\{f, H_\infty\}$.

In order to evaluate the evolutionary stability of IPD strategies, we introduce the concept of a counter strategy. A strategy S is a counter strategy (CS) against another strategy S_1 if, for any strategy S'

$$E(S, S_1) \geq E(S', S_1). \quad (1)$$

A CS receives the highest payoff in playing 2-IPD against another strategy. In an indefinite length 2-IPD, for example, always-defect (AllD) is a CS against always-cooperate; TFT is a CS against TFT. For an arbitrary strategy, there must be at least one CS against it.

Specially, strategy S is a CS against itself (CSI) if, for any strategy S'

$$E(S, S) \geq E(S', S). \quad (2)$$

A strategy S is a CS against a population of strategies $\{S_i\}$ ($i = 1, \ldots, n$) if, for any S' and S_i, there is

$$E(S, S_i) \geq E(S', S_i). \quad (3)$$

If a strategy is a CS against a population, it is a CS against any member of the population and thus, it receives the highest payoff in playing n-IPD against the population.

There is an equivalence between a CS against a mixed strategy and a CS against a population. Assume that S is a CS against a population of n strategies $\{S_i\}$. Consider the mixed strategy $\bar{S}$ that assigns probabilities $p_i = 1/n$ to S_i

$$\bar{S} = \begin{cases} S_1, & p_1 \\ S_2, & p_2 \\ \cdots \\ S_n, & p_n. \end{cases}$$

The expected payoff for S playing against $\bar{S}$ is

$$E(S, \bar{S}) = p_1 E(S, S_1) + p_2 E(S, S_2) + \cdots + p_n E(S, S_n) = \frac{1}{n}\sum_{i=1}^{n} E(S, S_i).$$

According to the definition of a CS, we have for any S'

$$E(S, \bar{S}) \geq \frac{1}{n}\sum_{i=1}^{n} E(S', S_i) = E(S', \bar{S}).$$

Thus, S is also a CS against $\bar{S}$. On the other hand, if a strategy is a CS against a mixed strategy, it must be a CS against a population that contains all pure strategies of the mixed strategy. As shown in the following subsection, the concepts of CSs can be used to evaluate the evolutionary stability of IPD strategies.

B. Evolutionary Stability

An evolutionarily stable stategy (ESS) is a strategy such that, if all the members of a population adopt it, then no mutant strategy can invade the population under the influence of natural selection. According to [25], the condition for a strategy S to be ESS is that for any S'

$$E(S, S) \geq E(S', S). \tag{4}$$

This condition is actually equivalent to a Nash equilibrium in a two-player IPD. In order for a homogeneous population to resist invasion of mutant strategies, a more restrictive condition is defined. According to [38], the condition for a strategy S to be ESS is that for any S'

$$\begin{aligned} E(S, S) &\geq E(S', S) \\ E(S, S') &> E(S', S'). \end{aligned} \tag{5}$$

This condition guarantees that an ESS always outperforms mutant strategies so that a homogeneous population can be maintained in evolutionary dynamics. However, most evolutionary algorithms for IPD do not satisfy the hypothesis of ESS and it has been proven that no strategy is ESS in infinite length or indefinite length n-player IPD [7], [40]. Except in specific situations, the condition of ESS cannot be used to analyze the evolutionary stability of IPD strategies.

There is a relationship between CS and ESS. A strategy is ESS if it is the only CS against all IPD strategies. Assume that S is a CS against all strategies. For any strategy S', we have

$$E(S, S) \geq E(S', S)$$

and

$$E(S, S') \geq E(S', S').$$

Comparing the inequalities with (5), we know that S is ESS if it is the unique CS against all strategies. Since the condition of ESS is too strict, we need another criterion to measure the evolutionary stability of IPD strategies.

It is easy to verify that any CSI is Maynard Smith's definition of ESS. If strategy S is a CSI, (4) always holds for any strategy S'. There is an infinite set of CSIs in n-IPD. Some well known strategies, for example, AllD, TFT, and tit-for-two-tats (TFTT) are such strategies. A CSI is an equilibrium choice for a player and thus it is superior to any non-CSI in maintaining a homogeneous population.

In order to evaluate the evolutionary stability of CSIs, we need another criterion. Given a strategy S, let $p(S)$ denote the probability that S is a CS against an arbitrary IPD strategy. A strategy with a high value of $p(S)$ is more likely to win in an n-IPD competition or evolution. We call $p(S)$ the evolutionary stability value of S. For any strategy, there is $1 \geq p(S) \geq 0$. A strategy is ESS if it is the only strategy satisfying $p(S) = 1$.

Here, we give an example of computing the evolutionary stability value of a random strategy. Let S denote a random strategy that chooses between C and D in every move, and S' an arbitrary strategy. In an indefinite length n-IPD with discount rate ω, the expected length of interaction between S and S' is $1/(1-\omega)$. It needs to make the correct choice in every move for S to be a CS against S'. Because S is a random strategy, the probability that it makes the correct choice in each move is 0.5. Therefore, the probability that S is a CS against S' is $0.5^{\frac{1}{1-\omega}}$.

Different IPD strategies can be compared according to whether or not they are CSI and their evolutionary stability values. A CSI outperforms any non-CSI and a CSI with a higher evolutionary stability value outperforms another CSI with a lower evolutionary stability value. In the following section, several theorems are given to show that longer memory strategies outperform shorter memory strategies.

III. Effect of Memory Length on Evolutionary Stability

Different IPD strategies use different history lengths of interactions among players to determine their choices. In a finite length IPD that iterates exactly L rounds, the history length a strategy can access is at most L. In an infinite length IPD, however, the history length is infinite. A nontrivial question is whether the full history length is useful for an individual player to interact with others optimally?

It has been proven that every finite history length is possible to occur in an infinite length n-IPD. This is expressed as the following theorem [40].

Theorem 1: In the infinite length n-IPD where the probability of further interaction is sufficiently high, every finite history of interactions among the n players occurs with positive probability in any evolutionarily stable mixture of pure strategies.

It is easy to verify that Theorem 1 also holds in an indefinite length n-IPD. In an indefinite length IPD, the probability that the game continues to any limited number of stages is positive. Thus, every finite history occurs with positive probability.

A direct conclusion from Theorem 1 is that every finite history may be useful for a strategy to interact with an arbitrary strategy optimally. As to the relationship between the history length a strategy uses and its evolutionary stability, we have the following theorems.

Theorem 2: For any strategy that uses a limited history length, there always exist some strategies with longer memory against which the strategy cannot be a counter strategy.

Proof: Consider a memory-L strategy S. Let H_L denote a specific L length history of interactions between S and an arbitrary strategy S_1.

a) If S is a pure strategy, without loss of generality, assume that S responds D (or C) to H_L. Let H'_{L+1} denote a $L+1$ length history in which the first L moves are the same as H_L and the opponent plays D (or C) at the last move. Another strategy S_2 is defined as below

$$S_2 = \begin{cases} S_1 \text{ in the first L moves} \\ always-defect \text{ if } H'_{L+1} \\ cooperate \text{ otherwise.} \end{cases} \quad (6)$$

S_2 is a memory-$(L+1)$ strategy. We now prove that S is not a CS against S_2. Consider a strategy S' defined as below

$$S' = \begin{cases} S \text{ in the first L moves} \\ C \text{ if } H_L \\ S \text{ otherwise.} \end{cases} \quad (7)$$

In the 2-IPD between S' and S_2, S_2 plays cooperate after $L+1$ moves, while in the 2-IPD between S and S_2, S_2 plays always-defect after $L+1$ moves. It is easy to verify that $E(S, S_2) < E(S', S_2)$.

b) If S is a mixed strategy, without loss of generality, assume that S responds D to H_L with probability $p > 0$ (responds C with probability $1 - p$). Let H'_{L+1} denote a $L+1$ length history in which the first L moves are the same as H_L and the opponent plays D at the last move. Consider two strategies S_2 and S' as defined in (6) and (7). There is $E(S, S_2) < E(S', S_2)$.

Thus, S is not a CS against S_2. ■

Theorem 3: Any strategy that uses a limited history length cannot be an ESS in an infinite length or indefinite length n-IPD.

Proof: It is a direct conclusion from Theorem 2 that any limited memory strategy cannot be ESS in an infinite length n-IPD. Consider a memory-L strategy S in an indefinite length n-IPD with discount rate $\omega > 0$. Since the probability that the game continues more than L stages is positive, there is always a positive probability that S meets a memory-$(L+1)$ strategy against which S is not a CS. Thus, S cannot be an ESS. ■

Lemma 1: If a strategy is a CS against a memory-L strategy, this strategy will eventually play a periodic sequence in the 2-IPD against the memory-L strategy and the period will be $\leq L$.

Proof: Consider an arbitrary memory-L strategy S_1. S_2 is a CS against S_1. If $L = 0$, S_2 must be always-defect and Lemma 1 holds. We only need to consider the case of $L > 0$.

a) Assume that S_1 is a pure strategy. Let h_1 and h_2 denote two continuous length-L history of interactions between S_1 and S_2, s_1 and s_2 denote two sequences of moves played by S_1 in h_1 and h_2, respectively, and l_1 and l_2 denote two sequences of moves played by S_2 in h_1 and h_2, respectively, as shown in Fig. 2.

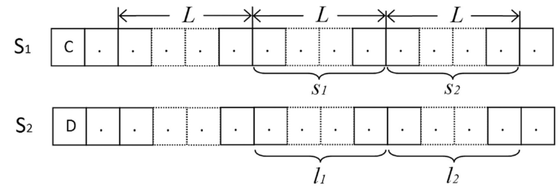

Fig. 2. Two continuous length-L sequences of moves in the interaction between S_1 and S_2.

Because S_1 is a memory-L pure strategy, the first move of s_2 is determined by s_1 and l_1. Since S_2 is CS against S_1, l_1 must be the optimal sequence of moves such that S_2 receives the highest payoff in playing against both s_1 and the first move of s_2. Given s_1, l_1 and the first move of s_2, the second move of s_2 is determined, etc. Given s_1 and l_1, each move of s_2 is determined and thus l_2 is determined. Since l_1 is the payoff-maximizing sequence, there must be $l_2 = l_1$.

b) If S_1 is a mixed strategy, without loss of generality, assume that S_1 assigns probabilities over m pure strategies S'_i, each of which is memory-L or less

$$S_1 = \begin{cases} S'_1, & q_1 \\ S'_2, & q_2 \\ \dots \\ S'_m, & q_m. \end{cases}$$

S_2 is a CS against any pure strategy S'_i $(i = 1, \dots m)$. According to the proof in section **a**, S_2 will eventually play a periodic sequence against any S'_i and the period will be $\leq L$. ■

Theorem 4: There are always longer memory strategies that have a higher probability of being a CS against an arbitrary strategy than a shorter memory strategy in an infinite length or indefinite length n-IPD.

Proof: We first consider the case of infinite length n-IPD. Let S' be an arbitrary strategy. If S' is a memory-L strategy, a CS against S' must play a periodic sequence with period $\leq L$ according to Lemma 1. Without loss of generality, let q_i $(i = 1, \dots L)$ denote the probability that the CS plays a sequence whose period is equivalent to i. There is $q_i \geq 0$ and $\sum_{i=1}^{L} q_i = 1$.

Let $p_L(S)$ denote the probability of S being a CS against an arbitrary memory-L strategy. For a memory-zero strategy S_0, there is $p_L(S_0) = \frac{1}{2}q_1$ if S_0 plays a sequence of pure C (or D). If S_0 plays a periodic sequence whose period is equivalent to two, we have $p_L(S_0) = \frac{1}{4}q_2$, and so on

$$p_L(S_0) = \begin{cases} \frac{1}{2}q_1, \text{ if } S_0 \text{ plays a period-1 sequence} \\ \frac{1}{4}q_2, \text{ if } S_0 \text{ plays a period-2 sequence} \\ \dots \\ \frac{1}{2^L}q_L, \text{ if } S_0 \text{ plays a period-L sequence.} \end{cases}$$

Thus, the highest value of $p_L()$ for a memory-zero strategy is

$$p_L(S_0) = max\left(\frac{1}{2}q_1, \frac{1}{4}q_2, \dots, \frac{1}{2^L}q_L\right).$$

A memory-one strategy can shift between a determined sequence and a period-2 sequence. The highest value of $p_L()$ for a memory-one strategy is

$$p_L(S_1) = max(\frac{1}{2}q_1 + \frac{1}{4}q_2, \frac{1}{8}q_3 + \frac{1}{4}q_2, \ldots, \frac{1}{2^L}q_L + \frac{1}{4}q_2).$$

Similarly, there are $p_L(S_L) = \frac{1}{2}q_1 + \frac{1}{4}q_2 + \ldots + \frac{1}{2^L}q_L$, where S_L is the memory-L strategy that has the highest value of $p_L()$.

For any memory-K ($K \geq L$) strategy, the highest value of $p_L(S_K)$ is

$$p_L(S_K) = \frac{1}{2}q_1 + \frac{1}{4}q_2 + \ldots + \frac{1}{2^L}q_L.$$

Thus, we have

$$\begin{cases} p_L(S_L) > p_L(S_{L-1}) > \ldots > p_L(S_1) > p_L(S_0) \\ p_L(S_K) = p_L(S_L) \quad \text{(for any } K \geq L). \end{cases} \tag{8}$$

Because S' is an arbitrary strategy, without loss of generality, assume that the probability that S' is a memory-i strategy is q'_i. There is $q'_i \geq 0$ and $\sum_{i=0}^{\infty} q'_i = 1$

$$p(S) = \sum_{i=0}^{\infty} q'_i p_i(S).$$

Let S_L be the memory-L strategy that has the highest value of $p()$ in the set of all memory-L strategies and S_{L+1} be the memory-($L+1$) strategy that has the highest value of $p()$ in the set of all memory-($L+1$) strategies. There is

$$p(S_L) = \sum_{i=0}^{\infty} q'_i p_i(S_L)$$

$$p(S_{L+1}) = \sum_{i=0}^{\infty} q'_i p_i(S_{L+1}).$$

According to (8), we have

$$\begin{cases} p_i(S_L) < p_i(S_{L+1}) \quad \text{(if } i > L) \\ p_i(S_L) = p_i(S_{L+1}) \quad \text{(if } i \leq L). \end{cases}$$

Thus, there must be $p(S_{L+1}) > p(S_L)$ for any limited number L in an infinite length n-IPD.

In an indefinite length n-IPD, the probability that the game continues more than any limited number of stages is positive. There is always a positive probability that the game continues i stages such that $p_i(S_L) < p_i(S_{L+1})$ and thus $p(S_{L+1}) > p(S_L)$ holds in indefinite length n-IPD. ■

IV. Example of Memory-2 Strategy

In this section, we use a memory-2 (MEM2) strategy to show how the theoretical analysis in the previous section assists in developing IPD strategies. MEM2 behaves like TFT in the first two moves and then it shifts among three strategies, AllD, TFT and TFTT, according to the interactions with the opponent in the last two moves. The logic for MEM2 to choose which strategy to play based on the following rules:

A. If the payoff in two moves is $2R$ (two mutual cooperations), then play TFT in the following two moves.

B. If the payoff is $T+S$, then play TFTT in the following two moves.

C. In all other cases, play AllD in the following two moves.

D. If AllD has been chosen twice, always play AllD.

MEM2 will cooperate with cooperative strategies according to Rule A, and it can restore cooperation from an occasional defection because of Rule B. Rule D makes sure that MEM2 defects against those periodic or random strategies. It is easy to verify that MEM2 is a CSI strategy.

TABLE II

Result of a Round-Robin IPD Competition With 19 Strategies

Strategy	Score	Strategy	Score
MEM2	3.045	PAVLOV	2.362
GRIM	2.799	HM	2.314
GRADUAL	2.749	AllC	2.283
TFT	2.668	STFT	2.246
FBF	2.608	PCD	2.243
GTFT	2.598	AllD	2.221
CTFT	2.587	RAND	2.149
TFTT	2.529	RTFT	2.132
RP	2.399	NEG	2.081
NP	2.364		

We run a round-robin IPD competition with 19 strategies, which include MEM2, AllD, TFT, TFTT, and some strategies that have appeared in research papers. The strategies play an IPD with each other and the discount rate of IPD is 0.98, which means that the average length of the IPD is 50 moves. The scores (average payoff per move) are listed in Table II. MEM2 receives the highest payoff, significantly higher than other CSI strategies.

We also run a series of evolutionary IPD simulations. The initial population contains $x= 5, 6 \ldots 10$ strategies randomly chosen from the 19 strategies in Table II. Each strategy has 20 identical copies. Stochastic universal sampling is used to select parents for the next generation. The parents simply copy their strategies to produce offspring and no mutation is carried out. An evolutionary IPD is run for 100 generations. As the outcome of any single evolutionary IPD is affected by chance, we repeat each evolutionary IPD with the same x value for 50 000 times, and gather statistics on the outcomes. Two measures, average fitness and average frequency in the population, are used to measure the performance of the strategies. They are average values over the results of 50 000 evolutionary IPDs.

The results of simulations are shown in Figs. 3–6. The average fitness and frequencies of all 19 strategies in the population after 100 generations are shown in Figs. 3 and 4. MEM2 outperforms other strategies in all settings of x. The average fitness and frequency of four strategies, MEM2, AllD, TFT, and TFTT as functions of generation are given in Figs. 5 and 6. It shows that the fitness of MEM2 is significantly higher than other strategies at the beginning of simulations, which leads to its higher frequency in the population. In most of the simulations, defective strategies became extinct after 50 generations and only cooperative strategies remained in the population. This was the reason why the fitness of some cooperative strategies tended to be equal at the end of evolution.

MEM2 shows an example of integrating several different strategies to form a new CSI strategy. Each strategy has its advantages and disadvantages. AllD is a CS against all

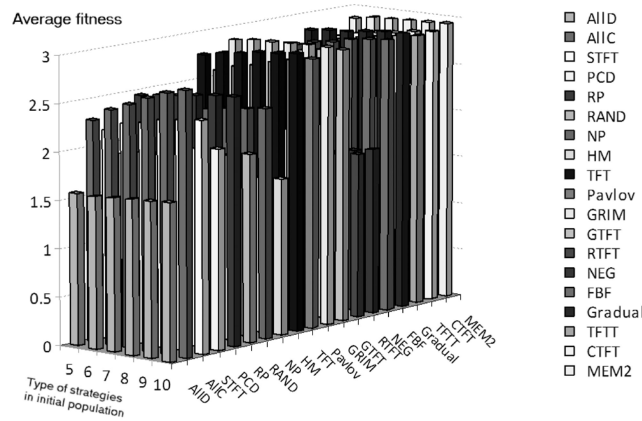

Fig. 3. Average fitness of strategies at generation 100.

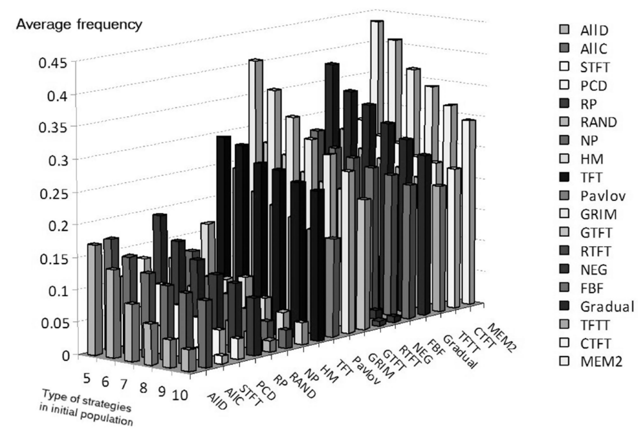

Fig. 4. Average frequencies of strategies at generation 100.

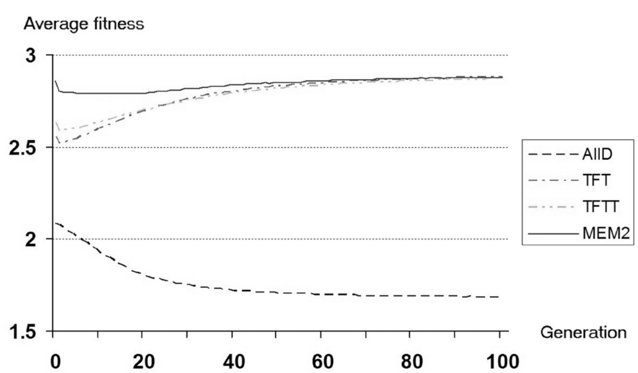

Fig. 5. Fitness of strategies MEM2, AllD, TFT, and TFTT (x = 10).

memory-zero strategies and it receives low payoffs in interacting with most of the memory-nonzero strategies. TFT is a CS against those strategies that cooperate with the opponent conditionally. A situation that TFT cannot handle well is a long series of mutual retaliations evoked by a single defection. TFTT performs well in this situation by playing one more cooperation. However, it can be exploited by the strategies that alternatively play C and D. MEM2 inherits the advantages of the three strategies and thus outperforms them in evolutionary IPD. The idea of combining different strategies can be used to develop longer memory strategies.

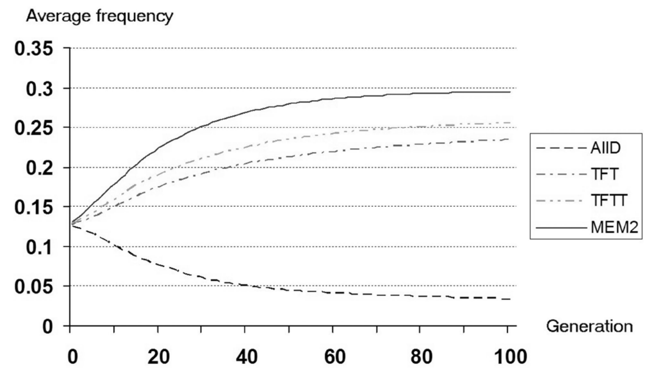

Fig. 6. Frequency of strategies MEM2, AllD, TFT, and TFTT (x = 10).

V. Conclusion

The condition of ESS is so strict that no strategy can be ESS in infinite length and indefinite length n-IPD. With the absence of an ESS, a criterion is needed to evaluate the evolutionary stability of IPD strategies. Based on the concept of a counter strategy, we have proposed a theoretical methodology in which the evolutionary stability of a strategy is evaluated by whether it is CSI and the probability it is a CS against an arbitrary strategy. Different strategies can be compared with each other. The effect of memory size on evolutionary stability is studied by means of this methodology.

The memory length used by a strategy has a significant influence on its evolutionary stability in n-IPD. We have proved that longer memory strategies outperform shorter memory strategies in the sense of evolutionary stability. A well-designed strategy that uses a longer memory statistically receives higher payoffs than shorter memory strategies in interacting with an arbitrary opponent and thus it is more likely to be dominant in evolution.

It may be difficult to theoretically check whether a strategy is a CS against another strategy, especially when both strategies use a long memory, which makes it difficult to compute the evolutionary stability value of an arbitrary strategy. In practice, it is possible to compute approximate evolutionary stability values of IPD strategies by means of statistical methodologies. For example, the performance of strategies in evolution can be measured by running a series of races [18]. The idea is further developed to generalization, a measure for learning performance in co-evolution by using a set of randomly chosen test strategies. It has been proven that an estimated value will approach the true value as the size of the set of unbiased test strategies increases [10], [12]. As the application of generalization, a statistical methodology that takes into account outcomes across varying n-IPD competitions has been used to evaluate the performance of IPD strategies [23]. With this methodology, a series of n-IPD competitions are run and the strategies in each competition are randomly chosen from a set of representative strategies. The performance of a strategy is evaluated according to its win rate, which is the frequency of achieving the highest payoff in a single competition. The win rate of a strategy is considered an approximation of its evolutionary stability value. In this way, the evolutionary stability value can be computed with greatly reduced computational complexity.

Theorems 2–4 do not necessarily hold in an IPD with noise where the players are assumed to make mistakes occasionally. The effect of noise on the performance of IPD strategies varies although noise generally has a negative effect on the persistence of cooperation. Some strategies are more robust than others in a noisy environment. The effect of memory size in IPDs with noise will be one topic of our future research.

REFERENCES

[1] W. Ashlock and D. Ashlock, "Changes in prisoner's dilemma strategies over evolutionary time with different population sizes," in *Proc. IEEE Congr. Evol. Comput.*, Jul. 2006, pp. 297–304.

[2] D. Ashlock and E-Y. Kim, "Fingerprinting: Visualization and automatic analysis of prisoner's dilemma strategies," *IEEE Trans. Evol. Comput.*, vol. 12, no. 5, pp. 647–659, Oct. 2008.

[3] D. Ashlock and N. Roger, "The impact of long-term memory in the iterated prisoner's dilemma," in *Intelligent Engineering Systems through Artificial Neural Networks*, C. Dagli, E. Akay, P. Chen, B. Fernandez, and J. Ghosh. Eds. New York, NY, USA: Amer. Soc. Mech, Eng., 2009, ch. 31.

[4] R. Axelrod, "More effective choice in the prisoner's dilemma," *J. Conflict Resolution*, vol. 24, no. 3, pp. 379–403, 1980.

[5] R. Axelrod, *The Evolution of Cooperation*. New York, NY, USA: Basic Books, 1984.

[6] B. Beaufils, J. Delahaye, and P. Mathieu, "Our meeting with gradual: A good strategy for the iterated prisoner's dilemma," in *Artificial Life V: Proceedings of the 5th International Workshop on the Synthesis and Simulation of Living Systems.* Cambridge: MIT Press, 1997, pp. 202–209.

[7] R. Boyd and J. Lorberbaum, "No pure strategy is evolutionarily stable in the repeated prisoner's Dilemma game," *Nature*, vol. 327, pp. 58–59, May 1987.

[8] R. Chiong and M. Kirley, "Effects of iterated interactions in multiplayer spatial evolutionary games," *IEEE Trans. Evol. Comput.*, vol. 16, no. 4, pp. 537–555, Aug. 2012.

[9] S. Chong and X. Yao, "Behavioral diversity, choices and noise in the iterated prisoner's dilemma," *IEEE Trans. Evol. Comput.*, vol. 9, no. 6, pp. 540–551, Dec. 2005.

[10] S. Chong, P. Tino, and X. Yao, "Measuring generalization performance in coevolutionary learning," *IEEE Trans. Evol. Comput.*, vol. 12, no. 4, pp. 479–505, Aug. 2008.

[11] S. Chong, J. Humble, G. Kendall, J. Li, and X. Yao, "Iterated prisoner's dilemma and evolutionary game theory," in *The Iterated Prisoners' Dilemma: 20 Years On*, G. Kendall, X. Yao, and S. Chong, Eds. Singapore: World Scientific, 2007, 23–62.

[12] S. Chong, P. Tino, D. Ku, and X. Yao, "Improving generalization performance in co-evolutionary learning," *IEEE Trans. Evol. Comput.*, vol. 16, no. 1, pp. 70–85, Feb. 2012.

[13] P. Darwen, "Co-evolutionary learning by automatic modularisation with speciation," Ph.D. dissertation, Australian Defence Force Academy, School of Computer Science, Australia, 1996.

[14] P. Darwen and X. Yao, "On evolving robust strategies for iterated prisoner's dilemma," in *Proc. Lecture Notes Artif. Intell., Progress Evol. Comput.*, 1995, vol. 956, pp. 276–292.

[15] D. Fogel, "The evolution of intelligent decision-making in gaming," *Cybern. Syst.*, vol. 22, no. 2, pp. 223–236, 1991.

[16] D. Fogel, "Evolving behaviors in the iterated prisoner's dilemma," *Evol. Comput.*, vol. 1, no. 1, pp. 77–97, 1993.

[17] D. Fogel, "On the relationship between the duration of an encounter and the evolution of cooperation in the iterated prisoner's dilemma," *Evol. Comput.*, vol. 3, no. 3, pp. 349–363, 1995.

[18] D. Fogel and A. Fraser, "Running races with Fraser's recombination," in *Proc. Congr. Evol. Comput.*, 2000, pp. 1217–1222.

[19] H. Ishibuchi, K. Hoshino, and Y. Nojima, "Evolution of strategies in a spatial IPD game with a number of different representation schemes," in *Proc. IEEE Congr. Evol. Comput.*, Jun. 2012, pp. 1–8.

[20] H. Ishibuchi and N. Namikawa, "Evolution of iterated prisoner's dilemma game strategies in structured demes under random pairing in game playing," *IEEE Trans. Evol. Comput.*, vol. 9, no. 6, pp. 552–561, Dec. 2005.

[21] G. Kendall, X. Yao, and S. Chong, "The iterated prisoners' dilemma 20 years on," in *Advances in Natural Computation 4*. Singapore: World Scientific, 2007.

[22] J. Li, "How to design a strategy to win an IPD tournament," in *The Iterated Prisoners Dilemma: 20 Years On*, G. Kendall, X. Yao, and S. Chong Eds. Singapore: World Scientific, 2007, ch. 4.

[23] J. Li, P. Hingston, and G. Kendall, "Engineering design of strategies for winning iterated prisoner's dilemma competitions," *IEEE Trans. Comput. Intell. AI Games*, vol. 3, no. 4, pp. 348–360, Dec. 2011.

[24] J. Li and G. Kendall, "A strategy with novel evolutionary features for iterated prisoner's dilemma," *Evol. Comput.*, vol. 17, no. 2, pp. 257–274, 2009.

[25] J. M. Smith, *Evolution and the Theory of Games*. Cambridge, U.K.: Cambridge Univ. Press, 1982.

[26] S. Mittal and K. Deb, "Optimal strategies of the iterated prisoner's dilemma problem for multiple conflicting objectives," *IEEE Trans. Evol. Comput.*, vol. 13, no. 3, pp. 554–565, Jun. 2009.

[27] M. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.

[28] M. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg, "Emergence of cooperation and evolutionary stability in finite populations," *Nature*, vol. 428, no. 6983, pp. 646–650, 2004.

[29] M. Nowak and K. Sigmund, "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game," *Nature*, vol. 364, no. 6432, pp. 56–58, 1993.

[30] M. Posch, "Win-stay, lose-shift strategies for repeated games–memory length, aspiration levels, and noise," *J. Theoretical Biol.*, vol. 198, no. 2, pp. 183–195, 1999.

[31] W. Press and F. Dyson, "Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent," *PNAS*, vol. 109, no. 26, pp. 10409–10413, 2012.

[32] H. Quek, K. Tan, C. Goh, and H. Abbass, "Evolution and incremental learning in the iterated prisoner's dilemma," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 303–320, Apr. 2009.

[33] D. Rand, H. Ohtsuki, and M. Nowak, 'Direct reciprocity with costly punishment: Generous tit-for-tat prevails," *J. Theoretical Biol.*, vol. 256, no. 1, pp. 45–57, 2009.

[34] A. Rapoport, "Optimal policies for the prisoner's dilemma," Psychometric Lab., Univ. North California, Berkeley, CA, USA, Tech. Rep. 50, MH-10006, 1966.

[35] A. Rapoport, *Two-person Game Theory*. New York, NY, USA: Dover, 1999.

[36] A. Rogers, R. K. Dash, S. D. Ramchurn, P. Vytelingum, and N. R. Jenning, "Error-correcting codes for team coordination within a noisy iterated prisoner's dilemma tournament," in *The Iterated Prisoners Dilemma: 20 Years On*, G. Kendall, X. Yao, and S. Chong Eds. Singapore: World Scientific, 2007, ch. 9.

[37] A. Stewart and J. Plotkin, "Extortion and cooperation in the prisoner's dilemma," *PNAS*, vol. 109, no. 26, pp. 10134–10135, 2012.

[38] B. Thomas, "On evolutionarily stable sets," *J. Math. Biol.*, vol. 22, no. 1, pp. 105–115, 1985.

[39] J. Wu and R. Axelrod, "How to cope with noise in the iterated prisoner's dilemma," *J. Conflict Resolution*, vol. 39, no. 1, pp. 183–189, 1995.

[40] X. Yao, "Evolutionary stability in the *n*-person iterated prisoner's dilemma," *Biosystems*, vol. 37, no. 3, pp. 189–197, 1996.

[41] X. Yao and P. Darwen, "An experimental study of N-person iterated prisoner's dilemma games," *Informatica*, vol. 18, no. 4, pp. 435–450, 1994.

[42] C. Hilbe, M. A. Nowak, and K. Sigmund, "The evolution of extortion in iterated Prisoners dilemma games," in *Proc. Nat. Acad. Sci.*, vol. 110, no. 17, pp. 6913–6918, 2013.

Jiawei Li (M'12) received the B.Sc. degree in ship engineering and the Ph.D. degree in fluid mechanics from Harbin Engineering University, Harbin, China, in 1992 and 1998, respectively.

He is currently a Research Fellow with the School of Computer Science, University of Nottingham, Nottingham, U.K. His research interests include hyper-heuristic, adaptive learning, and evolutionary game theory.

Graham Kendall (M'03–SM'10) received the B.S. degree in computation (first class, honors) from the Institute of Science and Technology, University of Manchester, Manchester, U.K., in 1997 and the Ph.D. degree in computer science from the University of Nottingham, Nottingham, U.K., in 2001.

His previous experience includes over 20 years in the information technology industry, where he held both technical and managerial positions. He is currently a Professor of computer science with the University of Nottingham and is currently based at their Malaysia campus, where he holds the position of Vice-Provost (Research and Knowledge Transfer). He is a Director of two Nottingham, U.K.-based companies and the CEO of two Malyasia-based companies.

Dr. Kendall is a Fellow of the Operational Research Society. He is an Associate Editor of nine international journals, including IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION and IEEE TRANSACTIONS ON COMPUTATIONAL INTELLIGENCE AND AI IN GAMES. He chaired the Multidisciplinary International Conference on Scheduling: Theory and Applications in 2003, 2005, 2007, 2009, and 2011, and has chaired several other international conferences.