Proceedings of the 2007 IEEE Symposium on
Computational Intelligence and Games (CIG 2007)

Information Sharing in the Iterated Prisoner's Dilemma Game

Ayman Ghoneim, Hussein Abbass, Michael Barlow
Defence and Security Applications Research Center
University of New South Wales
Australia
{a.ghoneim,h.abbass,m.barlow}@adfa.edu.au

***Abstract*— In the Iterated Prisoner's Dilemma (IPD) game, players normally have access to their own history, without being able to communicate global information. In this paper, we introduce information sharing among players of the IPD game. During the co-evolutionary process, players obtain access, through information sharing, to the common strategy adopted by the majority of the population in the previous generation. An extra bit is added to the history portion in the strategy chromosome. This extra bit holds a value of 0 if the decisions to cooperate were greater than the decisions to defect in the last generation and 1 if otherwise. We show that information sharing alters the dynamics of the IPD game.**

I. Introduction

Real life situations exhibit complex behaviors that affect the decisions of all parties involved. Simple games with rich dynamics have been used to understand emergent behaviors in complex situations. The Prisoner's Dilemma (PD) game, despite its simplicity, has been used extensively in modeling several complex real-world problems such as in international politics, economics and social systems [3].

In the age of globalization, whether to share information or not is becoming an important, sometimes controversial, issue. Regardless of the different views on the topic, with the existence of media, internet, and laws such as freedom of information, information sharing is becoming an inevitable concept that we cannot avoid.

Information sharing has the potential to create a shared understanding of the world [12]. The process of information sharing includes both the provision of information as well as the confirmation of the validity of previously received information [12].

The Iterated Prisoner's Dilemma (IPD) game has been used as a platform for understanding the dynamics of many complex situations. Oh [11] modeled internet searching agents as players in the IPD game, where these players seek to acquire a certain piece of information from the internet. If players act selfishly, each will send a lot of queries to different sites and this will increase the traffic and waste common resources (sites). However, if these players cooperate and share information (not just the information found from search results but information also about the strategies used for search), the resources (sites) will be wisely used.

In the IPD game, there has always been an assumption that players have access to their own history, without being able to communicate global information. In this paper, we introduce information sharing among players of the IPD game. During the co-evolutionary process, players obtain access, through information sharing, to the common strategy adopted by the majority of the population in the previous generation. An extra bit is added to the history portion in the strategy chromosome. This extra bit holds a value of 0 if the total decisions to cooperate was greater in number than the decisions to defect in the last generation, and 1 if otherwise. This extra bit represents the information shared by all players and has the effect of doubling the length of the chromosome. The information shared between players remains constant for a complete generation. We investigate the effect of information sharing with different history steps and different temptation levels. The effect of noise in the shared information is also studied.

The rest of the paper is organized as follows: in the following section, we introduce the IPD game. Experiments, results and discussions then follow.

II. Iterated Prisoner's Dilemma

The Prisoner's Dilemma (PD) game is a non-zero sum and non-cooperative game. The basic form of the PD game is a two-player game where there is a single available choice to each player: to cooperate or defect. The payoff matrix of the PD game (Figure 1) must satisfy two conditions related to the payoff values for different actions that may be taken by each player [8], [14]:

- $T > R > P > S$
- $2 \times R > (S + T)$

	Cooperate	Defect
Cooperate	R R	T S
Defect	S T	P P

Fig. 1. The payoff matrix of the 2-player PD Game.

The PD game models the conflict between self interest (being selfish) and the group interest; hence the dilemma. An individual rationality alone leads to a poor outcome because of the existence of a Pareto optimal solution if both actors cooperate. Iterated Prisoner's Dilemma (IPD) is a series of repeated rounds of the PD game. This feature makes the PD game more capable of modeling complex situations where future interaction between the actors is influenced by their history during playing the game [3], [1], [2]. For sufficiently large weight (discount factor) for future interactions, cooperation can emerge spontaneously. This is a very interesting characteristic to observe how cooperation may evolve among a group of potentially selfish players [14]. In many real-world situations, the evolution of cooperation is considered the best solution for the long run because it represents the maximum benefit for the group or society. As such numerous studies have been conducted of the dynamics of the IPD game in order to discover under what conditions cooperation evolves.

Understanding the properties of successful strategies in IPD is vital to our understanding of the dynamics of the game. Axelrod [3], [1] attempted to discover the properties of successful strategies in 2-players PD game through the formation of a computer tournament of 14 strategies that were submitted by different researchers. The tournament was held in a round robin form (each strategy plays with each other strategy including itself and the RANDOM strategy). Axelrod discovered that properties like “to be nice” (not to be the first to defect), “to be forgiving” (have propensity to cooperate after others defection avoiding defection echo that will lead to unending mutual punishment) and “to be provocative” (not to be exploited) existed in the top ranked strategies. The winner TIT FOR TAT (TFT) strategy (start by cooperation and then do whatever the other player does) depends on reciprocity. Axelrod held a second tournament [2] after announcing the results and analysis of the first one, 62 strategies participated and the winner was again TFT. The results of the second tournament was very surprising because all participants knew the results of the first tournament but no one could get a better performing strategy than TFT.

A more sophisticated way was needed to investigate the conditions of cooperation. Axelrod [1], [5] proposed the idea of using genetic algorithms to evolve more complex strategies. These strategies co-evolve in a population of competitive strategies.

Lindgren [10] started with very simple strategies and used Genetic Algorithms (GA) to evolve them to more complex ones. Axelrod [4], [5] used GA for evolving strategies where the strategy representation contains a history portion which is used in remembering the players’ actions for the previous l history steps. If there were 3 players and two history steps, then the history portion will consist of 6 bits (2 bits for each player indicating her own previous actions and 4 bits indicating the other players’ actions). The rest of the strategy representation will be a lookup table of size 2^{nl} where n is the number of players. Each possible combination of a history has a corresponding action.

Yao and Darwen [14] proposed another representation that is more space-effective than Axelrod’s representation in n player games. In their representation, the history portion in the strategy representation will hold the players own history and the number of players who cooperated in each of the considered historical steps. This representation overcomes the drawbacks of Axelrod’s representation like keeping unnecessary information about each player’s action and the chromosome length that is significantly affected by the number of players [14]. The rest of the strategy chromosome is also a lookup table.

Different ways for evaluating the fitness of the evolving strategies were suggested. Axelrod [4], [5] used 8 representative strategies from his second tournament, similarly in [9], six fixed strategies (ALLC, ALLD, TFT, TFTT, PAVLOV and RANDOM) were used in evaluating the fitness, where these six strategies provide a good mix of cooperators, defectors and strategies utilizing memory. Darwen and Yao [7], [14], [8] used co-evolution for evaluating the fitness, where each strategy in the population plays against every other strategy in the population, causing the environment to continuously evolve.

Darwen and Yao [7] used a GA to investigate the time needed for the population to converge and the effect of seeding the initial population with well known strategies such as TFT. Also in [14] the effect of the number of players and the number of history steps taken into account on the evolution of cooperation were discussed. Yao [13] studied evolutionary stable strategies (Collective Stability [3]), where strategies are called stable if they can’t be invaded by other strategies.

Vital features were neglected in the PD abstraction formulation like the possibility of communication between players and uncertainty about the other players’ previous actions [3]. Introducing new features to the PD game and considering different scenarios for the game were very helpful to move the PD game closer to modeling complex real world situations. Introducing different levels of cooperation in the PD game and investigating their influence on the emergence of full mutual cooperation was investigated in [8]. The introduction of multiple levels of cooperation into IPD helps in studying the dynamics of real-world situations that offer intermediate responses between full cooperation and full defection. Chang and Yao [6] introduced noise to the IPD game, investigated the effect of different (low and high) noise levels and how modeling mistakes in the players’ decisions influence the evolution of cooperation and the behavioral diversity in the multiple levels of cooperation (how different the played choices are in the game). Also studied was the effect of reputation on the dynamics of the game [15] where information about players’ past actions are available for future opponents.

III. Experiments

We used GA for investigating the interplay of information sharing and temptation level on the dynamics of IPD. A population of 101 strategies is initialized randomly. Axelrod’s representation [4], [5] is used (sufficient for 2-players PD game). Each player is evaluated by playing against each other player in the population; hence, each player plays $(N-1)$ 2-players IPD game. Each game lasts for a certain pre-defined number of rounds equal to 100 in our experiments.

After all players finish playing against each other, each player is awarded a cumulative payoff from the played games. The fitness is calculated by dividing a strategy’s cumulative payoff by the number of games it participated in multiplied by the number of rounds in each game to obtain the average payoff per round for this strategy.

Linear ranking selection is used where a new fitness is calculated for each strategy according to its position in the population, with the smallest average payoff ranked at the first position and the highest average payoff at the last position. Equation 1 is used, with SP denoting the selection pressure. We set SP to 1.25 in our experiments. We then apply a one-point crossover and bit mutation for generating new offspring with probabilities of 0.6 and 0.01 respectively. These parameters settings were used by Yao and Darwen [14].

$$Fitness(Pos) = 2-SP+2*(SP-1)*\frac{(Pos-1)}{(PopSize-1)} \quad (1)$$

We experimented with three different history lengths of 0, 1, and 2 for different temptation levels. To decide on an appropriate temptation level, we need to maintain the conditions of the PD game. If we substitute $R+t$ for T, we can get a bound on t as $0 < t < R-S$. Therefore, if we use $R=3$, $S=0$ and $P=1$, the temptation t is $0 < t < 3$. We experimented with 6 different values of t as follows: 0.1, 0.5, 1, 1.5, 2 and 2.5.

In the case of information sharing, there are 1010000 decisions taken per generation. If the total number of decisions to cooperate was greater than the number of decisions to defect, the extra bit will hold 0, otherwise 1. As this is a 2-player game, there is no way to have an odd number to avoid ties; hence our arbitrary choice of assuming that a 50-50 split in number of decisions is interpreted as a global choice of defect.

IV. Results and Discussions

We used a zero-history with different temptation levels to validate the code. The results shown in Figure 2 indicated, as expected, that defection is the Nash-equilibrium in this setup. The temptation level, regardless of how small it is, had no impact on the level of defection.

We then tested a history length of one and two respectively, with different temptation levels. Once more these experiments were designed to validate our code and to establish the basis of comparison against information sharing. The results were consistent with the literature. Cooperation becomes the dominant strategy with low temptation levels. As the temptation level increases, the percentage of the population cooperating (as being indicated by the average payoff) decreases. Interestingly, at a high level of temptation (2.5), cooperation is still the dominant strategy in the population despite the minority who are trying to exploit the dominant cooperative behavior.

The effect of different temptation levels become clear in one and two history step(s), where, as the temptation level increases, the player's temptation to defect becomes higher, this makes the average payoff drop as the level of cooperation decreases (Figures 3, 4). The decrease in the average payoff in early generations is a normal behavior reported in the literature, as at the first generations the defective strategies exploit the cooperative ones until all the existing strategies in the population become defective ones and no more exploitable strategies exist, then cooperation starts to evolve as players find that cooperating pays more than defecting.

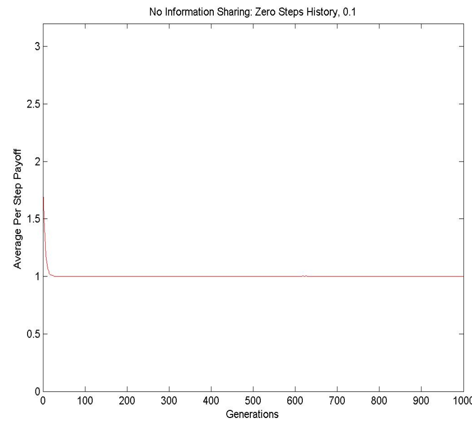

Fig. 2. The average payoff in 30 runs for Zero history steps and temptation level 0.1

By introducing information sharing, all players have access to a shared view for what was the majority type of decision in the last generation. This shared information represents the public belief for a certain issue. In the real-world this belief can be affected by several things like media or news coverage. Our investigation of the effect of information sharing in PD raises an important question about the effect of this on the evolution of cooperation. We observed that the addition of information sharing led the average payoff per-step to decrease as shown in Figure 5 compared to Figure 3 and also in Figure 4 compared to Figure 6.

Several studies have been conducted to investigate the effect of noise, whether this noise represents mistakes made by players or the wrong (opposite) implementation for the players' chosen action. Here we considered noise acting on the information shared among the players, the noise level (probability) is evaluated per player. This mimics disinformation or mistakes in media coverage of events, affecting the general public's view of a certain issue. Investigating the impact of noise on public opinion, and the results that stem from such opinions, and in particular how the degree of noise in information impacts decisions is an important area.

TABLE I

A sample of the best strategy for different temptation levels and history of 1.

Temptation level	Strategy
0.1	c d d d
1.5	c d c d
2.5	c d d d

In order to investigate what type of strategies evolved in

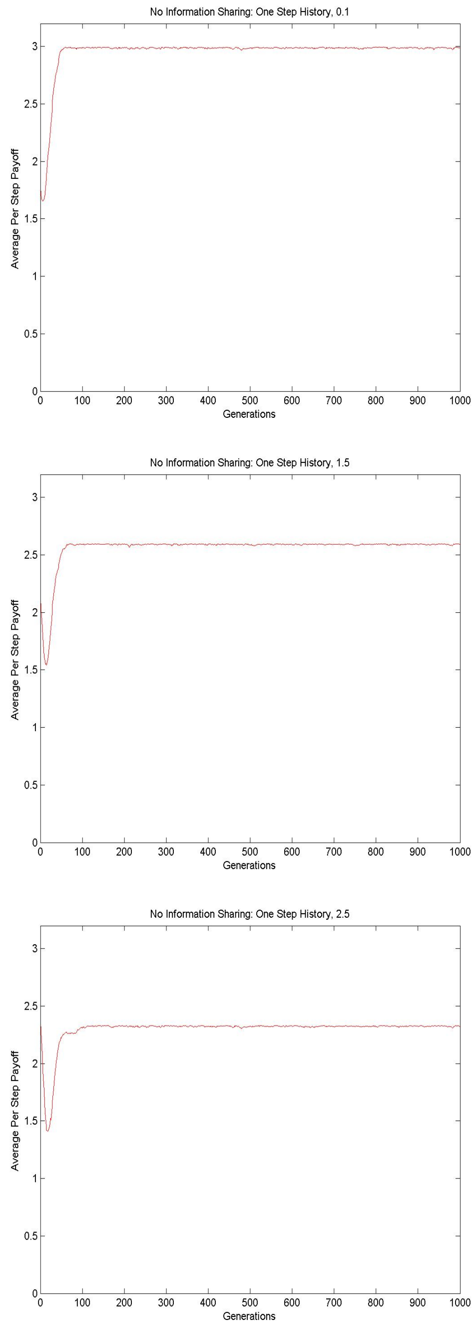

Fig. 3. The average payoff in 30 runs for One history step and different temptation levels: 0.1, 1.5 & 2.5

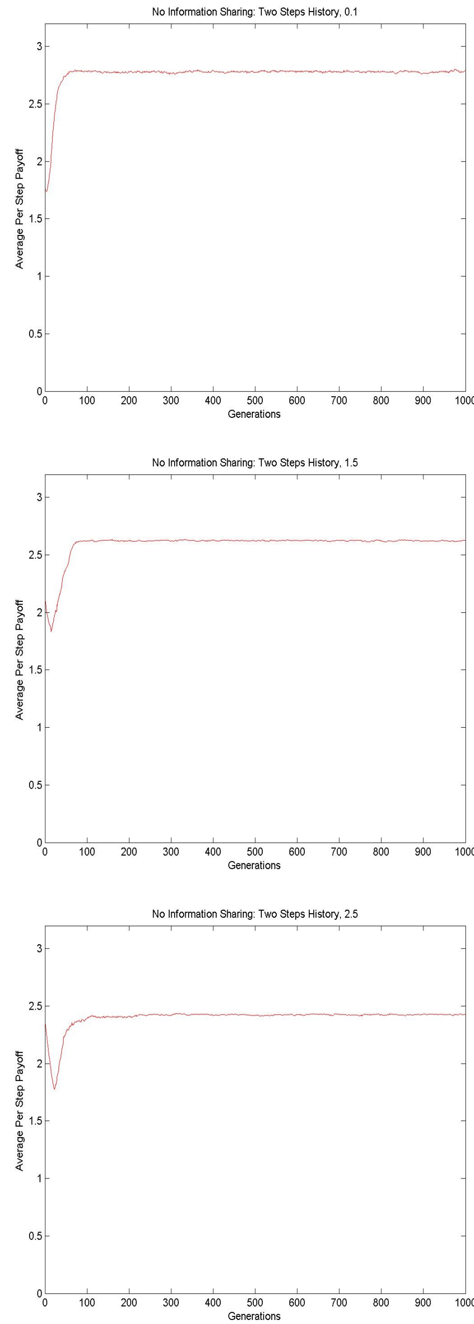

Fig. 4. The average payoff in 30 runs for Two history steps and different temptation levels: 0.1, 1.5 & 2.5

different scenarios, we analyzed the best strategy evolved in a sample run. Table I shows the strategies' structure neglecting the history portion when information sharing is not used. As we considered one of the best strategies (higher payoff), this strategy is intending to defect to exploit the cooperative strategies in the population. Although at temptation level 0.1 in one history a higher cooperation level is obtained, the best strategy just cooperates if the history was mutual cooperation otherwise it defects. This means that the strategy is not a forgiving one and will defect forever once the opponent defects.

Adding information sharing (Table II) for the same pre-

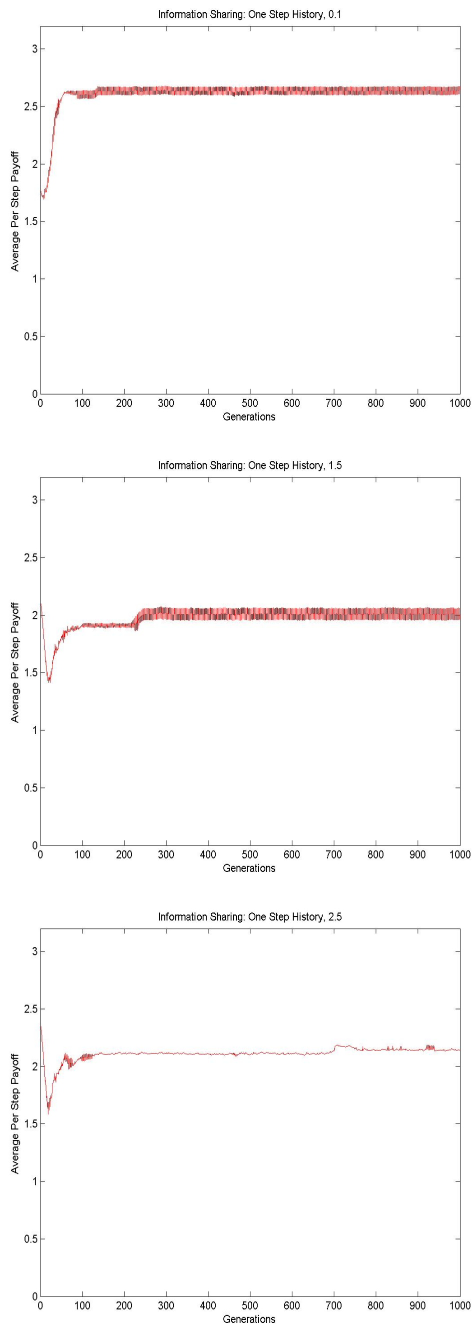

Fig. 5. The average payoff in 30 runs at One history step considering Information Sharing and different temptation levels: 0.1, 1.5 & 2.5

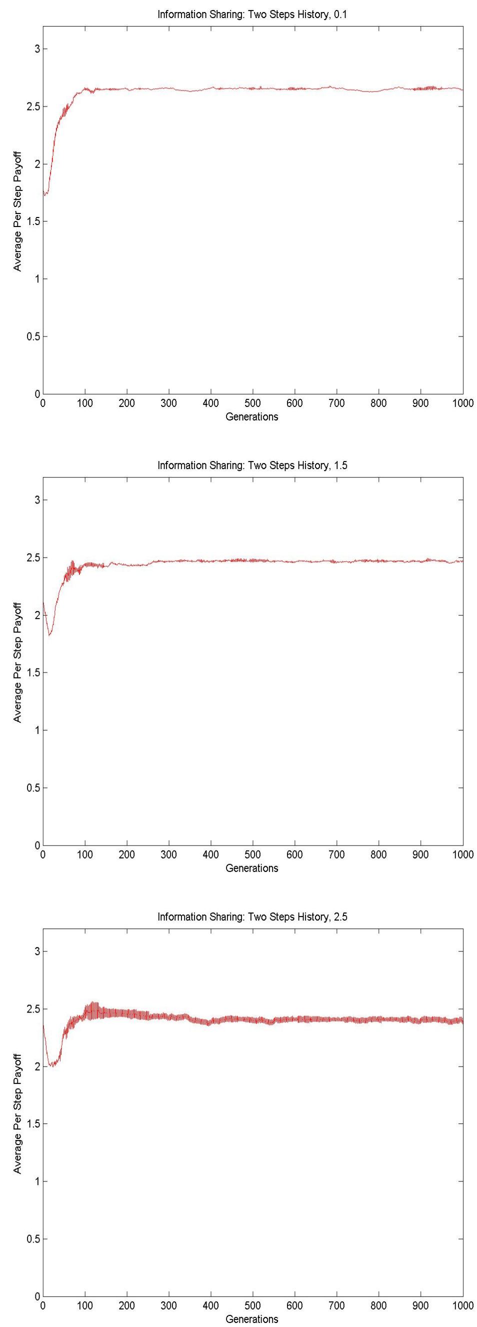

Fig. 6. The average payoff in 30 runs at Two history steps considering Information Sharing and different temptation levels: 0.1, 1.5 & 2.5

vious scenario has no effect on the best strategy actions for low temptation levels. The strategy gives priority to its own game history, so if the shared information indicates that the population was a cooperative one and the game history holds any defection, the strategy will be to defect and if the shared information was that the population is defective and the game history was mutual cooperation, the strategy will be to cooperate.

But as the temptation level increases to 1.5, the information sharing effect becomes more significant where if the shared information showed that the population is cooperative and there is a mutual defection in the game history, the strategy action

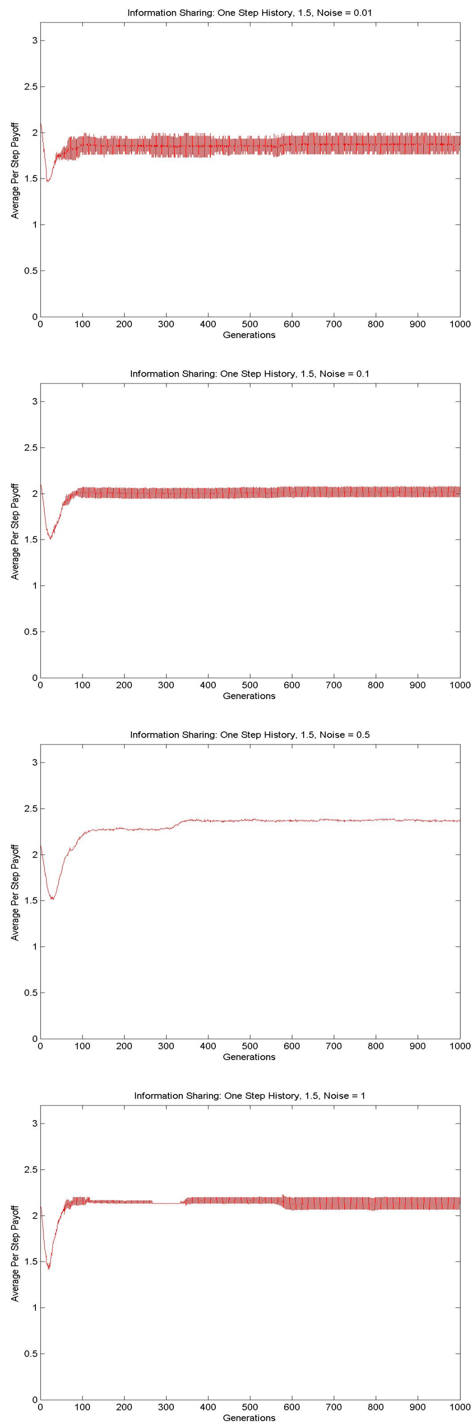

Fig. 7. The average payoff in 30 runs at One history step considering different noise levels in Information Sharing and temptation level of 1.5

TABLE II

A SAMPLE OF THE BEST STRATEGY FOR DIFFERENT TEMPTATION LEVELS WITH INFORMATION SHARING AND HISTORY OF 1.

Temptation level	Strategy
0.1	c d d d c c d d
1.5	c d c c c d d d
2.5	c d c d c d c c

will be to cooperate giving a second chance for cooperation, and also if the general impression was that the population is defective and there is any defection in the game history, the strategy will be to defect even if the player himself was the one who defected in the last round and the opponent was cooperative. This indicates why, with information sharing, the cooperation level becomes lower.

At the highest temptation level of 2.5, the Information sharing effect becomes lower as the direct payoff of the players will affect their actions in a more significant way. So if the shared information was that the population is cooperative, and the game history was a mutual defection then the strategy action is to defect.

TABLE III

A SAMPLE OF THE BEST STRATEGY FOR DIFFERENT TEMPTATION LEVELS WITH DIFFERENT NOISE LEVELS IN SHARED INFORMATION, HISTORY OF 1 AND TEMPTATION 1.5.

Noise level	Strategy
0.01	c d c c c d d d
0.1	c d c d c d d c
0.5	c d c d c d d d

By introducing noise to the shared information, the information sharing began to lose its significance, so for temptation level 1.5 and noise level of 0.1 (Table III), if the shared information indicates that the population is cooperative and the game history holds a mutual defection then the strategy action will be to defect, neglecting what was announced about the population nature and not providing an opportunity for cooperation to evolve. Another noteworthy item is for mutual defection of the player combined with prior information that the population is primarily defective, yet the strategy action selected is to cooperate. This may be a random mutation. As the noise level increases, the impact of the information sharing - which caused a drop in the cooperation level - is reduced, and the cooperation level starts to increase again as shown in Figure 7. However, we also tried a noise level of 1 (reversing the information completely), the level of cooperation is reduced (Figure 7). This scenario is symmetric to the 0 noise level scenario, as such the information sharing impact returns.

Overall, information sharing has its highest impact on the dynamics of the IPD game when the temptation level is medium. With low temptation level, the dynamics do not change. With high temptation level, it seems that the short term eagerness of players to maximize their immediate utilities is more important than information sharing. When noise in information sharing is introduced, it seems that there is a slight gain in the level of cooperation.

V. Conclusion and future work

This paper introduced the concept of Information Sharing to the Iterated Prisoner's Dilemma Game. Information Shared is the globally available information about the most common action - cooperate or defect - pursued by the population in the previous generation. Experiments were conducted to determine the impact of information sharing on the evolution of strategies and in particular the evolution of cooperation as a strategy.

Perhaps surprisingly, the addition of information sharing led to a decrease in cooperation over the non-information sharing case. Whether this is a global phenomenon or dependent on parameters of the experiment is currently unclear. The information sharing impact may differ by varying the population size and/or the length of the games (number of rounds). Moreover, changing the type of the shared information and the basis on which this information is shared could also affect the dynamics. Initial exploratory experiments have indicated a coupling between information sharing and population size in terms of overall player strategy.

The addition of noise to Information Sharing - effectively uncertainty or error in the communication of what was the previous generation's most common strategy - was also examined. It was discovered that noise ameliorated the impact of information sharing, such that at high enough noise levels cooperation returned to its previous, pre Information Sharing, high levels.

Considerable further work is possible in this area. A deeper examination of the impact of information sharing and the strategies that emerge for its exploitation, a richer or perhaps more timely update of global information about the community of player's actions, as well as examining how information sharing interacts with spatialised or graph structured versions of the IPD, all appear fruitful avenues of pursuit. In particular immediate emphasis is being given to the interplay of population size and rounds played with information sharing; as well as alternate mechanisms and criteria for information sharing.

Acknowledgements

This work is supported by an Australian Research Council (ARC) discovery grant number DP0667123.

References

[1] R. M. Axelrod, "Effective choice in the prisoner's dilemma," *Journal of Conflict Resolution*, vol. 24, no. 1, pp. 3–25, March 1980a.

[2] ——, "More effective choice in the prisoner's dilemma," *Journal of Conflict Resolution*, vol. 24, no. 3, pp. 379–403, September 1980b.

[3] ——, *The Evolution of Cooperation*. New York: Basic Books, 1984.

[4] ——, *The Evolution of Strategies in the Iterated Prisoner's Dilemma. In Genetic Algorithms and Simulated Annealing*. Morgan Kaufmann, 1987.

[5] ——, *The Complexity of Cooperation*. New Jersey: Priniceton University Press, 1997.

[6] S. Y. Chong and X. Yao, "Behavioral diversity, choices, and noise in the iterated prisoner's dilemma," *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 6, pp. 540–551, December 2005.

[7] P. Darwen and X. Yao, "On evolving robust strategies for iterated prisoner's dilemma," in *In Progress in Evolutionary Computation, Lecture Notes in Artificial Intelligence*, vol. 956. Verlag, Heidelberg, Germany: Springer, 1995b, pp. 276–292.

[8] P. J. Darwen and X. Yao, "Why more choices cause less cooperation in iterated prisoner's dilemma," in *Proceedings of the 2001 Congress on Evolutionary Computation*. Piscataway, NJ, USA: IEEE Press, May 2001, pp. 987–994.

[9] C. Goh, H. Quek, E. Teoh, and K. Tan, "Evolution and incremental learning in the iterative prisoner's dilemma," in *The 2005 IEEE Congress on*, vol. 3. IEEE Press, September 2005, pp. 2629–2636.

[10] K. Lindgren, "Evolutionary phenomena in simple dynamics," *Artificial Life II*, vol. 10, pp. 295–312, 1991.

[11] J. C. Oh, "Emergence of cooperative internet server sharing among internet search agents caught in the n-person prisoner's dilemma game," *Knowledge and Information Systems*, vol. 7, pp. 23–55, 2005.

[12] D. H. Sonnenwald, "Challenges in sharing information effectively: Examples from command and control," *Information Research*, vol. 11, no. 4, July 2006.

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

[14] 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.

[15] X. Yao and P. J. Darwen, "How important is your reputation in a multi-agent environment," in *Proc. of the 1999 IEEE Conference on Systems, Man, and Cybernetics*. Piscataway, NJ, USA: IEEE Press, October 1999, pp. 575–580.