

Emergence of Cooperative Coalition in NIPD Game with Localization of Interaction and Learning

Yeon-Gyu Seo

Department of Computer Science
Yonsei University
134 Shinchon-dong, Sudaemoon-ku
Seoul 120-749, Korea
kitestar@candy.yonsei.ac.kr

Sung-Bae Cho

Department of Computer Science
Yonsei University
134 Shinchon-dong, Sudaemoon-ku
Seoul 120-749, Korea
sbcho@csai.yonsei.ac.kr

Xin Yao

School of Computer Science
University of Birmingham
Edgbaston, Birmingham B15 2TT
United Kingdom
X.Yao@cs.bham.ac.uk

Abstract- The N-player iterated prisoner's dilemma (NIPD) game has been used widely to study the evolution of cooperation in social, economic and biological systems. Previous work on the NIPD game studied the impact of the number of players and the payoff function on the evolution of cooperation. This paper studies the localization issue in the NIPD game and investigates the impact of local interaction on genetically evolved strategies for the NIPD game. Our experimental results show that localization of interaction has a major impact on the evolution of cooperative coalitions, while localized learning makes the population oscillate. This paper also investigates the effect of the history length in the NIPD game. It is found experimentally that a longer history makes a population more stable, but it takes longer time to reach this stable state.

1 Introduction

The game-theoretic approach has been used to understand social and economic systems for a long time. One of the most well known games is the iterated prisoner's dilemma (IPD) game. In general, a player in IPD must choose one of the two decisions, defection (D) or cooperation (C). Table 1 shows the payoff matrix for the game. The game is repeated infinitely and none of the players know the end of game. No matter how many players cooperate, anyone of them will earn better payoff by defection. Therefore, defection may be a rational selection, and all players may select D and obtain payoff P. However, if all cooperate, they would get better score than all defection. This is the dilemma for players to face with in the IPD game.

Most work on evolving IPD game-playing strategies only considered two players, i.e., the 2IPD game. However, the 2IPD game cannot model many real-world complex problems in the social and economic systems. The NIPD game is a much more realistic model for many real-world problems. Table 2 shows an example of the payoff function in the NIPD game [YaDa94]. The basic principle in the 2IPD game still holds for the NIPD game: Defection is always better than cooperation for a single iteration, but mutual cooperation is better than defection in the long run. In the NIPD game, there are many issues to be considered, such as the pay-

off function [Ban94], noise [AxeDio88], population structure [NowMay92], localization [HoffWar96], the shadow of the future [Axe84, AxeDio88], the number of players [YaDa94], and so on. This paper concentrates on the localization issue. We are interested in how local interaction affects the evolution of game-playing strategies in the NIPD game where $N > 2$.

Table 1: Payoff matrix in 2IPD. $T > R > P > S$, $2R > T + P$

	Cooperate	Defect
Cooperate	R	S
Defect	T	P

Table 2: Payoff matrix in NIPD

No. of Cooperator	0	1	...	X	...	N - 1
Cooperate	C_0	C_1	...	C_x	...	C_{N-1}
Defect	D_0	D_1	...	D_x	...	D_{N-1}

The localization issue can be approached from two aspects: learning and interaction, where learning refers to the evolutionary learning process and interaction refers to the games played by individuals within a generation. These two factors have been reported to affect the emergence of cooperation [HoffWar96]. In this paper, we will investigate the emergence of cooperative coalitions and how localization affect them. Schelling [Sche78] defines the minimum coalition size as the least number of players that can begin getting interest by choosing specified action. A coalition represents a group of players who choose the non-favorable choice, e.g., cooperation in the IPD game [Sche78].

Generalization is a key issue in machine learning. Few work has been reported on the generalization ability of evolved NIPD strategies except for some limited results [DaYa95, DaYa97]. This paper will look at the issue how localization affects the generalization ability of evolved strategies.

The rest of this paper is organized as follows: Section

2 explains the encoding scheme used to represent NIPD game strategies, which is different from that used by Axelrod [Axe84]. Section 3 discusses several issues in localization. Section 4 presents our experimental studies. Finally, Section 5 concludes the paper with a few remarks.

2 Evolutionary Approach to NIPD Game

Genetic algorithms have been used to evolve strategies for the 2IPD game for many years [Axe87]. However, the work on evolving NIPD game strategies is relatively little [YaDa94, DaYa95, DaYa97]. One of the most important issues in evolving game strategies is their representation. Two different representations have been proposed [YaDa94], both of which are look-up tables that give an action for every possible contingency.

The first representation is a generalization of the representation scheme used by Axelrod [Axe87] for the 2IPD game. In this scheme, each genotype is a lookup table that covers every possible history of the last few steps. A history in such a game is represented as a binary string of ln bits, where the first l bits represent the player's own previous l actions (most recent to the left, oldest to the right), and the other $n - 1$ groups of l bits represent the previous actions of the other players. For example, during a game of 3IPD with a remembered history 2 steps, i.e., $n = 3, l = 2$, one player might see this history:

$n = 3, l = 2$: Example history 11 00 11

The first l bits, 11, means this player has defected (a '1') for both of the previous $l = 2$ steps. The previous steps of the other players cooperated (a '0') on the previous l steps, and on the last the other players cooperated (0) on the most recent step, and defected (1) on the step before, as represented by 01.

For the NIPD remembering l previous steps, there are 2^{ln} possible histories. The lookup table genotype therefore contains an action (cooperate "0", or defect "1") for each of these possible histories. So we need at least 2^{ln} bits to represent a strategy. At the beginning of the game, there are no previous l steps of play from which to look up next action, so each genotype should also contain its own extra bits that define the presumed pre-game moves. The total genotype length is therefore $2^{ln} + ln$ bits.

The Axelrod-style representation scheme, however, suffer from two disadvantages. First, it does not scale well as the number of players increases. Second, it provides more information than is necessary by telling which of the other players cooperated or defected, when the only information needed is how many of the other players cooperated or defected. Such redundant information has been shown to have reduced the efficiency of evolution greatly [YaDa94].

Another representation scheme for the NIPD game strategies was first proposed by Yao and Darwen [YaDa94]. In this scheme, each individual is regarded as a set of rules stored in a lookup table that covers every possible history. An individ-

ual can remember the l most recent rounds. Such a history of l rounds is represented by:

- l bits for the player's own previous l moves, where a "1" indicates defection, a "0" cooperation; and
- another $l \log_2 n$ bits for the number of cooperators among the other $n - 1$ players, where n is the number of players in the game. This requires that n is a power of 2.

For the example, if we are looking at 8 players who can remember the 3 most recent rounds, then one of the players would see the history as:

History for 8 players, 3 steps: 001 111 110 101 (12 bits)

Here, the 001 indicates the player's own actions: the most recent action (on the left) was a "0", indicating cooperation, and the action 3 steps ago (on the right), was a "1", i.e., defection. The 111 gives the number of cooperators among the other 7 players in the most recent round, i.e., there were $111_2 = 7$ cooperators. The 101 gives the number of cooperators among the other 7 players 3 steps ago, i.e., there were $101_2 = 5$ cooperators. The most recent events are always on the left, previous events on the right.

In the above example, there are $2^{12} = 2048$ possible histories. So 2048 bits are needed to represent all possible strategies. In the general case of an N -player game with history length l , each history needs $l + l \times \log_2 n$ bits to represent and there are $2^{l+l \times \log_2 n}$ such histories. A strategy is represented by a binary string that gives an action for each of these possible histories. In the above example, the history 001 111 110 101 would cause the strategy to do whatever is listed in bit 1013, the decimal number of for the binary 00111110101.

Since there are no previous rounds at the beginning of a game, we have to specify them with another $l(1 + \log_2 n)$ bits. Hence each strategy is finally represented by a binary string of length $2^{l+l(\log_2 n)} + l(1 + \log_2 n)$.

3 Localization

The only interaction among players in the NIPD game is through playing games against each other in each generation. How well a player does depends on how strong/weak his opponents are. Localization restricts the opponents of a player to be selected only from his neighborhood (i.e., among his neighbors). This restriction will understandably influence how the player is to interact with others and what fitness he is going to get. Such restriction (localization) also influences evolutionary learning of game-playing strategies because fit strategies (fit in the context of localized interaction) will reproduce and propagate in the population. Very fit strategies without localization may not be fit at all with localization. We would guess that strategies learned with localization would be less robust than those learned without it since players have less exposure to a wide range of different opponents. However, the NIPD game with localization is a more realistic model of the real world. It would be interesting to know what difference between the two models.

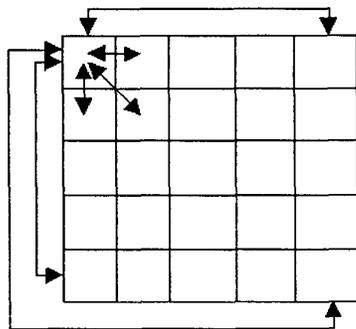
Related work on the NIPD game with localization includes Nowak and May's work on spatial evolution [NowMay92]. In their model, a population of players are distributed on squares of a torus who are only capable of the always defection (AD) and always cooperation (AC). Each player interacts with his eight neighbors only and imitates the strategy of any better performing one. Cooperative behavior can be sustained in clusters of players that insulate cooperators from hostile ADs under certain payoffs.

Huberman and Glance [HubGlan93] criticized synchronous evolution. They showed that the result of asynchronous evolution is different from that of the synchronous one. Warning and Hoffmann [HoffWar96] also experimented with localized interaction and learning between players on torus employing Moore machine to play games and report that localized learning and interaction improve the cooperativity of population.

In this paper, the model as shown in Figure 1 is used. The next section will present some experimental results on how the neighborhood size and the history length affect the emergence of cooperative coalition and the generalization ability of evolved strategies.



(a) One-dimension case, where neighborhood size of 2



(b) Two-dimension case, where neighborhood size of 1

Figure 1: Localization of interaction and learning

There are several methods of selecting $N - 1$ opponents among a player's neighbors in the NIPD game. Random selection will be used in our study. In other words, if one player

for a game is determined, the other $N - 1$ players are selected at random among his neighbours. It is worth pointing out that such random selection leads to sparse evaluation of a player since he does not see all different groups of opponents. What he learns will be biased towards what he has seen. This is very similar to the situations in our social and biological environments.

4 Evolutionary Process

Evolutionary process in localization is like as Figure 2. For evaluation, one strategy is randomly selected, then $N - 1$ other strategies are selected at random among his neighbors. Also, for localization of learning, a weak strategy can have the chance to change himself into better one that is picked out among his adjacent neighbors. Here, it is important to select the candidate to replace a weak strategy by. If we have a weak strategy changed only by his best neighbor, the population might show fast convergence, after all, the evolution would stop at near local optima rather than global optima. Therefore we have a weak strategy changed by his neighbor chosen among his adjacent top strategies and we make top strategies in the local area survive in the next generation for evolution of population.

5 Experimental Results

For the experiments, we used the population size 100, crossover rate 0.6, mutation rate 0.001. Two-point crossover with elite preserving strategy was adopted. We consider two factors in our experiments, learning and interaction, which might have different effects on the evolution of cooperative coalition. The payoff function is $2x - 2$ for cooperation and $2x + 1$ for defection [YaDa94], where x is the number of cooperators in Table 2. Also, we have the average number of games per player 20 at each generation. The history size has been varied from one to three to investigate the effect of history size.

High level of localization of interaction can improve cooperativity, but it can do that only with localized learning [HoffWar96]. Cooperative coalition size of the general IPD game are shown in Figures 3, 4 and 5. Figures ?? and 9 show the variation of coalition size in case of localizing interaction. We can see that a cooperative coalition is large when the neighborhood size is small. As the neighborhood size increases, it becomes more difficult to evolve cooperative coalition, i.e., there were less number of cooperators in the population. Figure ?? also shows how a longer history length encourages the emergence of cooperators.

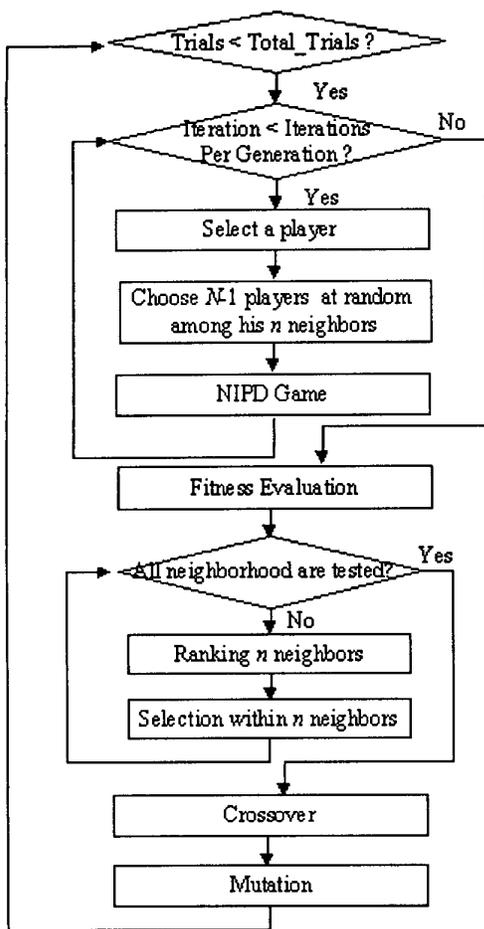


Figure 2: Evolutionary process in localization

Figure 10 shows the result of learning under localization. It is interesting to note that the population always moved away from its initial state after learning. It has been shown that the number of players (i.e., N) in the NIPD game has a major impact on the evolution of cooperation [YaDa94]. This paper shows that when neighborhood size is small, such as 2 (in 4IPD), 4 (in 8IPD) and 8 (in 16IPD), a high level of cooperative coalition can be evolved irrespective of the number of players. As the neighborhood size becomes large, such as 30, the cooperative coalition becomes more difficult to evolve. It depends on the number of players and the history length heavily, as shown in Figures 6, 7, 8 and 9.

It is interesting to note in Figure ?? that a short history length encourages the quick emergence of cooperation, but such cooperation breaks down against after while. It is not very stable. This is probably caused by the random selection

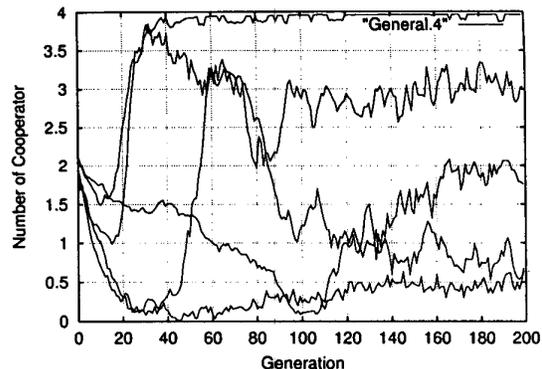


Figure 3: In case of general 4IPD game and history length 1

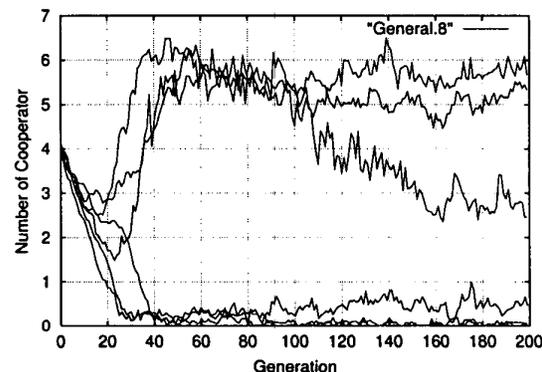


Figure 4: In case of general 8IPD game and history length 1

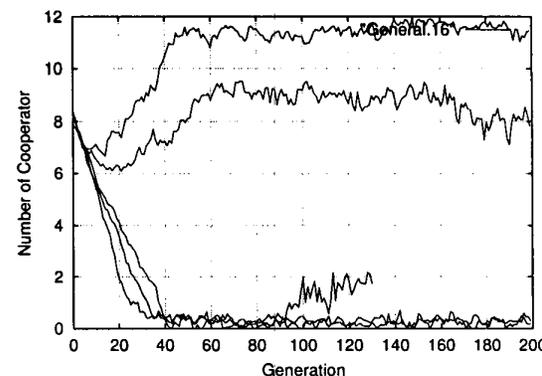
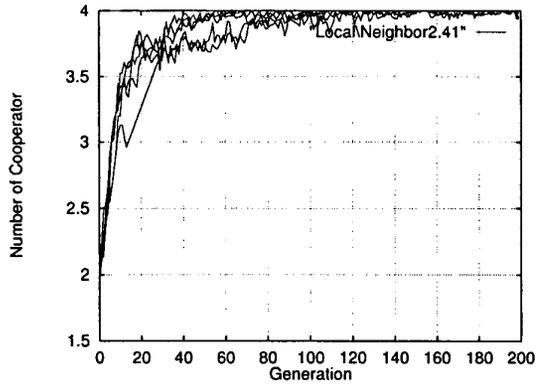
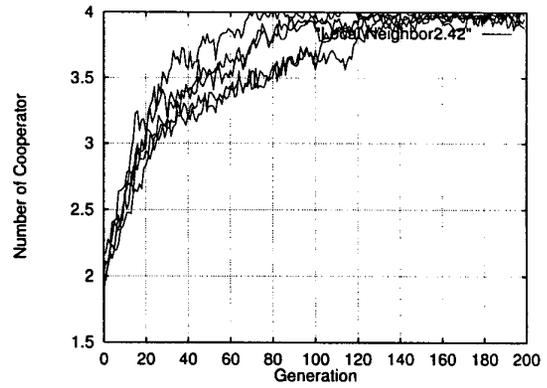


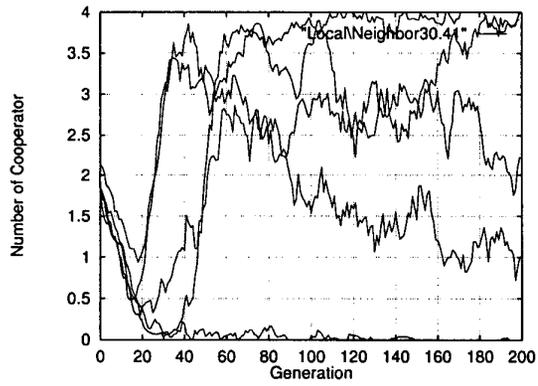
Figure 5: In case of general 16IPD game and history length 1



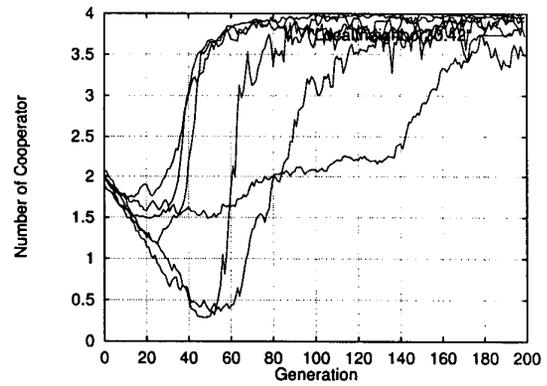
(a) neighborhood size 2



(a) neighborhood size 2



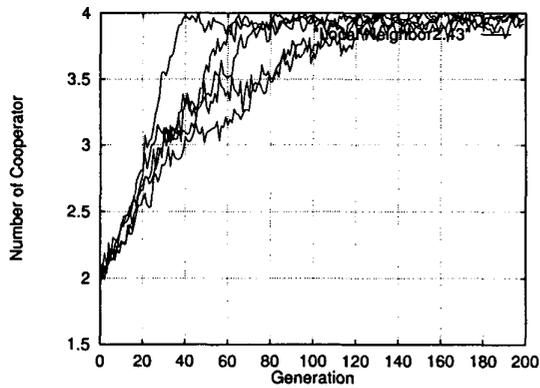
(b) neighborhood size 30



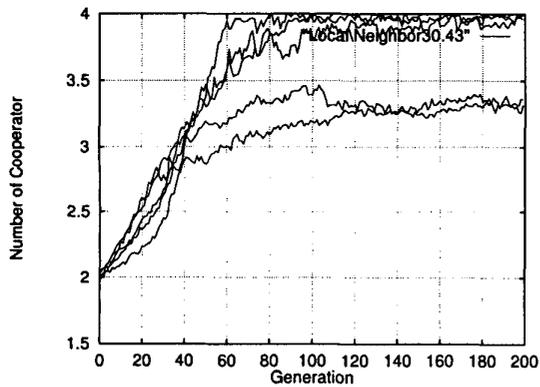
(b) neighborhood size 30

Figure 6: Variation of cooperative coalition size in localization of interaction. in case of 4 players and history length 1

Figure 7: Variation of cooperative coalition size in localization of interaction. in case of 4 players and history length 2

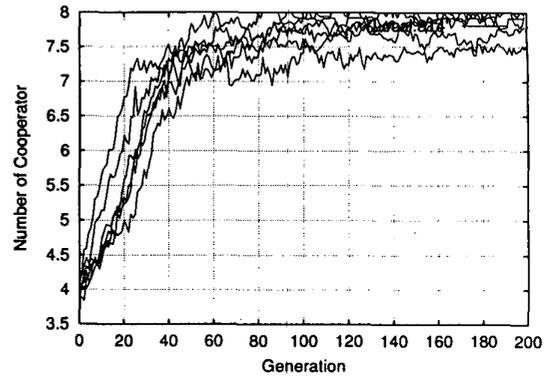


(a) neighborhood size 2

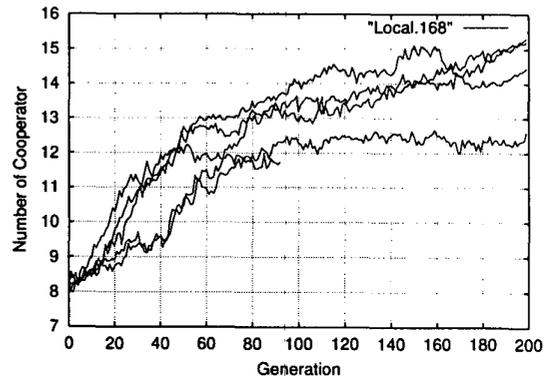


(b) neighborhood size 30

Figure 8: Variation of cooperative coalition size in localization of interaction. in case of 4 players and history length 3



(a) 8 players and neighborhood size 4



(b) 16 players and neighborhood size 8

Figure 9: The number of player and localization of interaction.

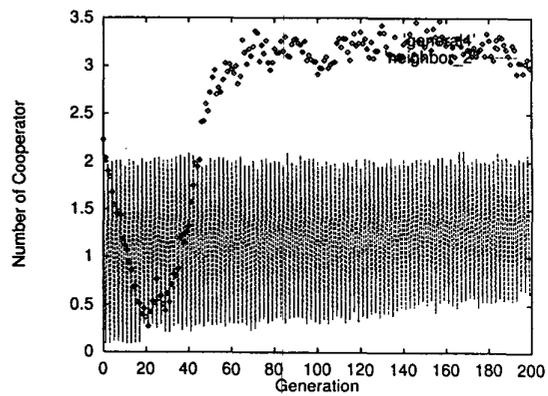


Figure 10: Localization of learning. Dotted line is for 2 neighbors and large dot is for the general case, when the number of players is 4.

method we used for choosing opponents. On the other hand, a longer history length slows down the emergence of evolution. However, once cooperation emerges it stays longer and is more stable.

In order to investigate the generalization ability of evolved strategies, we would like to play the evolved strategies against a number of other strong strategies. We use two different methods to choose opponent strategies: (1) a few best strategies in the initial population, and (2) evolved strategies themselves. In the first method, we generate the population at random and then make an individual play the NIPD game against other strategies. After the end of game, we sort and choose top 15 strategies in the population as the test strategies.

Table 3 shows the opponent strategies which are selected from the initial population, when the number of players is 8 and history length 1. Tables 4 and Table 5 show the results of the evolved strategies playing against best random strategies and themselves, respectively. As shown in these tables, the generalization ability of evolved strategies with localization is not very good because they were evolved in a small cosy coalition rather than in a diverse and hostile environment. They have learned to cooperate well with their similar kind, but did poorly against defectors. Figure 11 shows the convergence of population with localization.

Table 3: Opponent strategies selected from the initial population

Opponent Strategy
1110 1100 1111 1001 1001
1010 1001 0111 1101 0110
1100 1111 1001 1101 0100
0010 1111 1001 1101 1000
1101 1010 1111 1111 1010
0000 1111 1111 1101 1110
0001 1011 1101 1100 1100
0111 0111 0111 1111 0010
0110 1110 0101 1011 1010
0011 1011 1111 1111 1011
1101 1101 1101 1111 0100
1010 0011 1111 1100 1011
1101 0100 0011 1100 0100
1110 1100 1101 1101 1000
0011 1110 1011 1110 0111

6 Concluding Remarks

This paper studies the cooperative coalition and relationship between generalization and localization in the NIPD game. It is shown that localization encourages cooperative coalition. The more localized the interaction is, the easier it is

Table 4: The Result is when evolved strategies play against the random strategies selected from the initial population.

Evolved Strategy	Mean	S.D.	Mean of Opp.	S.D.
00011100101001000001	4.364	0.268	5.336	0.183
00001001000010001011	4.137	0.194	5.575	0.160
01001000011111010011	4.366	0.201	5.446	0.197
00011000010100111100	4.268	0.241	5.486	0.211
00100100111000110000	4.064	0.180	5.452	0.146
00000000100101000000	4.219	0.200	6.315	0.157
00111011011110101000	4.795	0.191	5.025	0.124
00100011010010100100	4.769	0.158	5.497	0.136
01101110100011011110	4.856	0.201	4.929	0.209
00000100010011000110	4.286	0.229	5.965	0.196

Table 5: The result is when evolved strategies play against themselves.

Evolved Strategy	Mean	S.D.	Mean of Opp.	S.D.
00011100101001000001	14.000	0.000	14.000	0.000
00001001000010001011	14.000	0.000	14.000	0.000
01001000011111010011	14.000	0.000	14.000	0.000
00011000010100111100	14.000	0.000	14.000	0.000
00100100111000110000	14.000	0.000	14.000	0.000
00000000100101000000	14.000	0.000	14.000	0.000
00111011011110101000	14.000	0.000	14.000	0.000
00100011010010100100	14.000	0.000	14.000	0.000
01101110100011011110	14.000	0.000	14.000	0.000
00000100010011000110	14.000	0.000	14.000	0.000

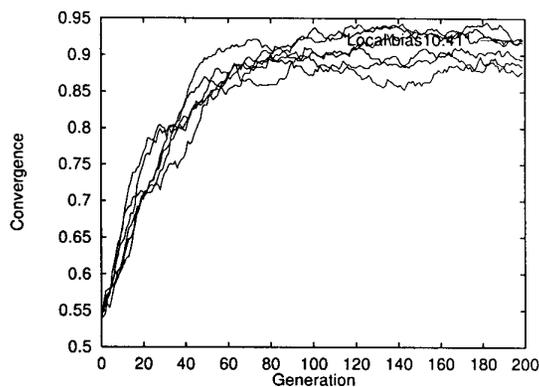


Figure 11: Convergence of Localization.

to evolve cooperation. However, the history length plays an important role here. When the history length is short, such cooperation could be rather short and unstable. A longer history length makes cooperation more stable but the cooperation takes longer time to emerge.

It is also shown in this paper that the number of players (i.e., N) in the NIPD game is not an important parameter when localization is introduced, provided the neighbourhood size is sufficiently small. Localization has a negative impact on the generalization ability of evolved strategies because they were not exposed to a diverse and hostile environment. They perform well against players of their kind, i.e., cooperators, but poorly against defectors.

In the future, we will investigate how generalization ability of evolved strategies may be improved even when localization is used.

Acknowledgments

This work has been supported partially by a grant (975-0900-004-2) from the Korea Science and Engineering Foundation and a grant from the Australian Research Council.

Bibliography

- [Axe84] Axelrod, R., *The Evolution of Cooperation*. Basic Books, New York, 1984.
- [Axe87] Axelrod, R., The evolution of strategies in the iterated prisoner's dilemmas, *Genetic Algorithms and Simulated Annealing*, Chapter 3, pp. 32–41, 1987.
- [AxeDio88] Axelrod, R. and Dion, D., The further evolution of cooperation. *Science*, Vol. 242, 1385–1389, 1988.
- [Ban94] Banks, S., Exploring the foundation of artificial societies: Experiments in evolving solutions to iterated N-player prisoner's dilemma. *ALife IV*, 337–342, 1994.
- [DaYa95] Darwen, P. and Yao, X.: On evolving robust strategies for iterated prisoner's dilemma. *Progress in Evolutionary Computation, Lecture Notes in Artificial Intelligence, Vol. 956*, Springer-Verlag, Berlin, pp.276-292, 1995.
- [DaYa97] Darwen, P. and Yao, X.: Speciation as automatic categorical modularization. *IEEE Transactions on Evolutionary Computation*, Vol. 1, No. 2, pp.101-108, 1997.
- [HoffWar96] Hoffmann, R. and Warning, N., The localization of interaction and learning in the repeated prisoner's dilemma. *Santa Fe Institute Working Paper no.96-08-064*, 1996.
- [HubGlan93] Huberman, B. A. and Glance, N.S., Evolutionary games and computer simulations, *Proc. of National Academy of Sciences (USA)*, Vol. 90, pp. 7715–7718, 1993.
- [NowMay92] Nowak, M. A. and May, R. M., Evolutionary games and spatial chaos. *Nature*, Vol. 359, 1992.
- [Sche78] Schelling, T. C.: *Micromotives and Macrobehaviour*. New York, 1978.
- [YaDa94] Yao, X. and Darwen, P.: The experimental study of N-player iterated prisoner's dilemma. *Informatika*, Vol. 18, 435–450, 1994.