1T-1

Sequential actions in iterated prisoner's dilemma game

Zhejing Luo†, Reiji Suzuki† and Takaya Arita†
Graduate School of Information Science, Nagoya University†

1. Introduction

The iterated prisoner's dilemma (IPD) game is used extensively in studies on the evolution of cooperative behaviors [1]. In many published studies, the action of the players lasts one round. However, it is also natural to assume a situation in which a player can declare which action to play in future rounds to its opponent. The purpose of this paper is to clarify how such a behavior can affect cooperation. We extend the deterministic strategy of IPD game [2] by introducing a two-round cooperative sequential action to provide an perspective on how introducing sequential action affects the strategy evolution in the IPD game. The preliminary results with the spatial evolutionary model of IPD strategies with sequential actions show that the sequential action can significantly facilitate the evolution of cooperation.

2. Model

Each player is located in a cell in a two-dimensional square lattice with periodic bounding conditions (torus). The size of the torus is 20x20, therefore the total number of players is 400.

The typical IPD payoff matrix (Table 1) is used in this research. We define the sequential action as a series of nonstop identical actions (Cooperate or Defect) performed by a player for a limited number of rounds. In this paper, we assume a two-round cooperative sequential action, which is denoted as C2. We also assume that the opponent knows which action (the first or the second) in C2 is played by the player. Each cooperative action in C2 is seen as the same as the one-turn cooperation (C1) in payoff calculation.

Table 1. Payoff Matrix

Player	Opponent	
	C	D
C	(3, 3)	(0, 5)
D	(5, 0)	(1, 1)

Each player has an unique strategy denoted by a numerical string, which determines how it would respond depending on the possible history of actions in the last round between itself and the opponent, as shown in Table 2. In the table, "0" represents a one-turn cooperation (C1), "1" represents a one-turn defect (D), "2" represents the first cooperative move in a two-turn sequential cooperation (C2-1), and "4" represents the second (C2-2). Each digit in the strategy, composed of 0, 1 or 2, represents the action played in the initial round or the action played after the corresponding history of actions. When C2-1 (2) is used, the player chooses a cooperation (C2-1) in the current round, and also plays a cooperation in the next round (C2-2). The strategy string's length is 5 for baseline IPD games, and 13 after C2 is introduced as an available action (Table 2).

The player's strategy in a population of *N* agents is randomly generated at the beginning of the first iteration of interaction and evolution processes. The fitness of the player is defined as its average payoff obtained in an iteration. We assumed the two different cases of interactions: local and non-local interactions. In the local case, each player plays a certain rounds of IPD games with its standard 8 Moore neighbors. In the non-local case, each agent plays with 8 randomly selected opponents. After the all IPD games, all players check the fitness of itself and its opponents, imitate the strategy of the player with the highest fitness, and start a new iteration. This process repeats until all iterations are finished.

Table 2. An example of the 13-digit strategy "0110100021012". The shadowed 8 situations would only appear after C2 is introduced into the IPD games.

		1	2	3	4	5	6	7	8	9	10	11	12	13
Last round	Self	/	0	1	0	1	0	1	0	1	4	4	4	4
	Opp.	/	0	0	1	1	2	2	4	4	0	1	2	4
This round		0	1	1	0	1	0	0	0	2	1	0	1	2

A random assignment of action is embodied in the three processes of the model: random generation of actions in the initial strategies, random noise when determining each action during the IPD game, and random mutation of each action (digit) during the strategy imitation. When a random action (digit) is chosen randomly in these processes, we adopted the following probability: 50% for C and 50% for D in the baseline (standard) IPD game without C2; and 57.14% for D, 28.57% for C1 and 14.29% for C2-1 in the extended IPD games with C2, in order to make the probability for an occurrence of cooperative actions the same between these settings.

3. Experiments

We used the following settings: agent number N = 400 (20x20 grid), rounds per iteration: 100, total iterations: 1000, mutation probability: 0.01, noise probability p_n: 0.05. With these parameters, we examined specifications of the model: B: non-local baseline IPD, C2: non-local baseline+C2 IPD, BL: local baseline IPD and C2L: local baseline+C2 IPD.

After introducing C2 action into the IPD games, we observe a rise in the frequency of cooperation (Fig. 1).

The average frequency of cooperation over 1000 iterations rises from 0.153% to 0.802% in non-local IPD, and from 3.34% to 60.09% in local IPD. In non-local experiments, the percentage of C2 action increases as the frequency of cooperation rises; but the opposite phenomenon appeared in local experiments: a rise in the C2 action percentage often accompanies a drop in the frequency of cooperation. In local experiments, the percentage of C2 fluctuates widely between 2.44% and 72.70%, while the percentage of C1 and C2 is stable around 83.15% with a standard deviation of 4.25%.

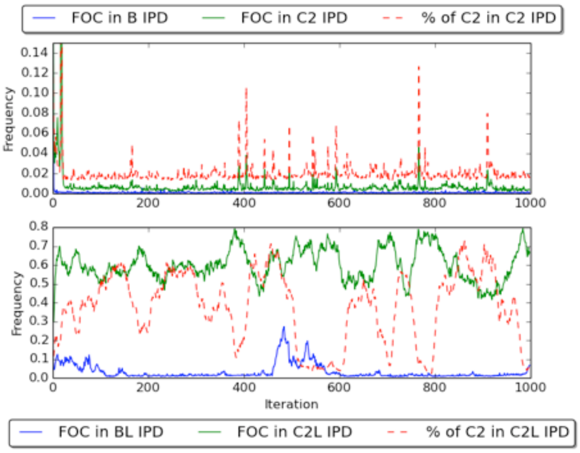

Fig.1 The frequency of cooperation (FOC) and the percentage of C2 sequential action in non-local (Top) and local games (Bottom) over 1000 iterations

Introducing C2 sequential action into the IPD simulation model as an action option in addition to the original one-turn C and D actions changed the dominant strategies emerged in the simulation, which is expected to the main reason that introduction of C2 facilitated cooperation. Table 3 shows the top dominant strategies over 1000 iterations in each case. In non-local baseline IPD games, ALL-D (11111) was always the dominant strategy, because noises or mutation breaks down mutual cooperation easily. After introducing C2, the dominant strategies responded in the same way (defect) when both players chose a one-turn action, while their response towards C2 actions were more cooperative, which might have increased the proportion of cooperation very slightly.

In local baseline IPD games, defect-oriented strategies such as 11**1 tended to dominate the population. After introducing C2, the agents evolved strategies that made them respond cooperatively to the opponent's C2 actions. From Table 3 bottom, we found the all top dominant strategies share the following properties: agent cooperates in the initial round, plays a sequential cooperation C2 after mutual C1, and plays cooperative actions after both agents played C2-2. This allowed the agents to establish robust cooperative relationships against noises quickly because mutual cooperation with C2 actions is expected to be more robust against noises than that with C1 actions. It also should be noticed that this effect was obvious in local games, when agents always play with the same opponents, whom likely possess strategies similar to the opponent's because of the spatial propagation.

However, mutual cooperation with C2 actions has a cost in that tended to be exploited by defectors. Thus, we observed a kind of cyclic evolution process of the strategies: → cooperators with C2 actions → defect-oriented strategies → cooperators with C1 actions, as indicated in the fluctuation of the proportion of C2 actions among all cooperative actions in Fig. 1 bottom.

Table 3. Top 5 dominant strategy listed by frequency over 1000 iterations of non-local (top) and local (bottom) IPD games before and after introducing C2. (# represents a cooperative action, including C1 & C2)

	B	freq.	C2	freq.
1	11111	100%	1111101#101#1	19.4%
2			1111111#1#101	9.9%
3			11111011101#1	9.8%
4			1111101#1#1#1	8.5%
5			111111101#000	5.6%
Total		100%		53.2%

	BL	freq.	C2L	freq.
1	11001	46.8%	0210000021#1#	5.1%
2	11101	27.4%	02100100#1##0	5.0%
3	01001	10.1%	0210100021012	4.7%
4	01111	8.6%	02010000110#0	2.3%
5	11111	3.1%	0211#100#0020	2.1%
Total		96.0%		19.2%

Conclusion

We discussed whether and how a sequential action of cooperation facilitates cooperative behaviors in spatial IPD games. We found that it can boost cooperation by making cooperative relationship robust against noises. Future work includes more detailed analysis, effects of bonus and penalty on sequential actions, etc.

References

[1]. P. J. Darwen and X. Yao. On Evolving Robust Strategies for Iterated Prisoner's Dilemma. *Progress in Evolutionary Computation*, Vol. 956, pp. 276–292, 1995.

[2]. H. Ishibuchi and N. Namikawa. Evolution of Iterated Prisoner's Dilemma Game Strategies in Structured Demes under Random Pairing in Game-Playing. *IEEE Trans. on Evolutionary Computation*, Vol. 9, No. 6, pp. 552-561, 2005.