The Eighteenth International Symposium on Artificial Life and Robotics 2013 (AROB 18th '13),
Daejeon Convention Center, Daejeon, Korea, January 30-February 1, 2013

The Evolution of Pre-play Communication in the Interactive Minority Game

Keita Nishimoto[1], Ivan Tanev[2], Katsunori Shimohara[2], Reiji Suzuki[1] and Takaya Arita[1]

[1]Graduate School of Information Science, Nagoya University, Japan
[2]Faculty of Science and Engineering, Doshisha University, Japan

[1]nishimoto@alife.cs.is.nagoya-u.ac.jp

Abstract: Minority Game (MG) is an *N*-person game which represents the collective behavior of agents in an idealized situation in which they have to compete for some finite resource. MG has been studied actively in various fields, but most studies have not focused on the communication among agents. To study the evolution of communication in the MG, we extended the standard MG to a new game named *Interactive Minority Game* (*IMG*) by incorporating the two aspects: a continuous strategy space and a pre-play communication stage. In order to understand basic behaviors of agents in the IMG, we prepare three agents each of which is equipped with a recurrent neural network (RNN) to adjust the tentative strategy value in the pre-play communication stage and evolved the connection weights of each RNN based on the payoff of IMG. As a result, we saw the emergence of various communications such as the adaptive adjustment behavior and oscillation of strategy values of each agent. Moreover, we found the strategy differentiation among agents where two agents adopt "high-risk high-payoff" strategy and the rest one adopts "low-risk low-payoff".

Keywords: Minority Game, Evolution of communication, Multi-agent simulation

1 INTRODUCTION

Minority Game (MG) is an *N*-person game which represents the collective behavior of agents in an idealized situation where they have to compete for some finite resource. Each of *N* agents chooses one out of two alternatives independently, and those who have chosen the minority choice among them win and are awarded a point [1]. Minority Game has been studied actively in various fields because of its simplicity and emergent characteristics (e.g., the emergence of cooperation among agents and the phase transition [2], [3]).

However, most studies have not focused on the communication among agents in the MG. In the real world, people in competitive situations do not make decisions based only on the past record, but also based on communication among them.

We focus on the communication among agents and propose a new model named *Interactive Minority Game (IMG)* [4], where agents make decisions based on dynamic interactions between them instead of the past record. There are several studies to incorporate communication among agents into the MG [5], [6]. Anghel et al. proposed a network-based Minority Game where agents are connected with a random network [5]. Agents employ a two-step decision making procedure and exchange the decision-making information with others through the links. First, each agent predicts what the minority choice will be based on its own strategy table. Then, it selects the agent which has made the most accurate predictions so far from among its neighboring agents including itself, and adopts its prediction as the final choice. As a result of simulations, they discovered that the scale-free imitation network emerged on the random network.

In their model, the information which agents exchange is only a binary decision and one-shot, and thus it cannot be regarded as a dynamic communication. We incorporate *pre-play communication stage* in which agents can modify their intentions continuously and dynamically observing others' intentions before their final decision making.

To incorporate the dynamic communication, we need to change the agents' intentions and payoff from binary to real values. In our model, each agent expresses its intention by strategy value a ($\in$ [-1, 1]) and receives the payoff depending on the value.

As a first step of our study [4], we dealt with the evolutionary dynamics of *social sensitivity* of agents and role switching. This paper focuses on the evolved pre-play communication among agents. We prepare three agents, each of which is equipped with a recurrent neural network (RNN) to adjust the tentative strategy value in the pre-play communication stage. We evolved the connection weights of each RNN with an evolution strategy (ES) based on the payoff of IMG played among these agents.

2 Interactive Minority Game (IMG)

In the MG with N (odd) agents proposed by Challet and Zhang [1], the payoff of an agent i choosing alternative A_i is calculated as follows:

$$payoff = -A_i\, sgn\left(\sum_{k=1}^{N} A_i\right), \qquad (1)$$
$$\left(A_i \in \{-1,1\}, payoff \in \{-1,1\}\right),$$

$$sgn(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & otherwise. \end{cases}$$

We propose the Interactive Minority Game by extending the MG on the following two points. First, we adopt a continuous strategy space instead of a binary one. The payoff of agent i with strategy value a_i is calculated as follows:

$$payoff = -a_i\, sgn\left(\sum_{k=1}^{N} sgn(a_i)\right), \qquad (2)$$
$$\left(a_i \in [-1,1], payoff = [-1,1]\right).$$

This equation represents the situation as follows: The possible signs of the strategy value (positive or negative) correspond to the alternatives in the standard MG. Agents win the game if the sign of their strategy is the minority sign in the group. Furthermore, the strategy's absolute value defines its "intensity". Higher intensity values lead to both higher risk and higher reward. The winning (losing) agents obtain a positive (negative) payoff equal to the absolute value of their strategies.

Secondly, we add a pre-play communication stage before the agents confirm their strategy. During this stage, agents can continuously adjust their strategy. The tentative strategy of agent i at time step t (= 0, 1, .., T - 1) is represented as $a_i(t)$ ($a_i(0) = 0$). Each agent can adjust $a_i(t)$ gradually by $\varepsilon(t)$, after observing others' tentative strategies in the previous step. The final decision of agent i : a_i is defined as $a_i(T)$, and used for calculation of payoffs.

$$\begin{gathered} a_i = a_i(T), \\ a_i(t+1) = a_i(t) + \varepsilon(t), \\ (t = 0, 1, \cdots, T-1). \end{gathered} \qquad (3)$$

Note that if $a_i(t) + \varepsilon(t) > 1(< -1)$, then $a_i(t + 1) = 1$ (-1). In this study, we focus on the case of N = 3, the minimum number of agents for the MG. Fig. 1 shows an example game. The x-coordinate corresponds to the time step t and the y-coordinate represents $a_i(t)$ for each agent.

3 MODEL

3.1 Decision making mechanisms

Every agent is equipped with a recurrent neural network (RNN) to decide $\varepsilon(t)$ at each step. The reason why we choose to use RNNs is to enable agents to make decisions appropriately depending not only on the current inputs, but also on past inputs: RNNs can use their internal memory to process arbitrary sequences of inputs. Each RNN has three layers (5 input units, 6 hidden units, 4 output units), and the units use a sigmoid activation function ($f(x) = 1 / (1 + e^{-x})$). For simplification of the model, RNNs do not have bias units. Two output units in the output layer are recurrently connected to two input units in the input layer.

Every time step, the agent's RNN receives five input values: its own current strategy value, the distance from the strategy values of the other two agents to its own, and the values from the two output units from the previous step. Two units in the output layer generate the values *au* and *ad*, which determine $\varepsilon(t + 1) = \varepsilon(t) + (a_u - a_d) / 100$. The remaining two output units are connected one-to-one to two of input units.

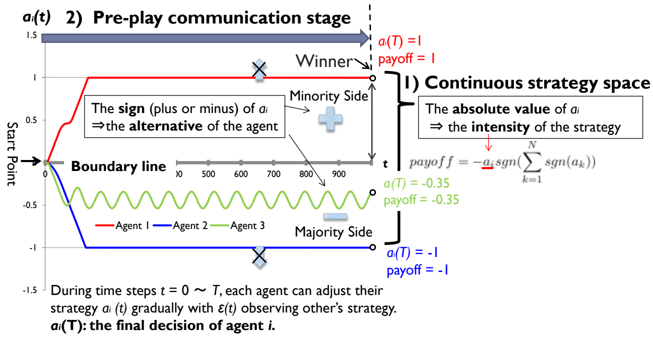

Fig. 1. A trial of IMG (N = 3, T = 1000).

3.2 Evolutionary algorithm

The full set of 54 connection weights in each RNN is encoded in the genotype of each agent and evolved using a simple type of Evolution Strategy (ES). The connection weights do not change during a trial. We assume three independent gene pools, each of which provides one agent in each trial, so the agents that interact in a game trial come from independently evolved gene pools. Each gene pool has n_p individuals.

1. Generating the initial population
   n_p genes are generated in each gene pool using uniform random number U as an initial population.
2. Evaluating individuals' fitness
   We assemble n_p groups of three individuals each of which randomly selected from each gene pool without duplications. Three individuals in each group play IMG and receive the payoff. Individuals picked out from pools are restored to their original pools after the trial. This procedure of group assembly and game trial is repeated R times. The fitness of each individual is defined as an accumulation of fitness over R trials.
3. Creating the next generation

The population in the next generation of each gene pool is composed as follows: First, we select n_e best individuals from n_p individuals (the elites), and preserve them to the next generation. Then, each of elites contributes two copies of themselves to the next generation, and small random values from a normal distribution N_R with a fixed standard deviation are added to each connection weight in the offspring. Finally, n_p-3n_e individuals with randomly generated genotypes are added to the population. These evolutionary operations for selection and reproduction are performed on each gene pool independently.

4 RESULTS

We evolved the population for 10000 generations. We used the following parameter settings: T = 1000, n_p = 40, n_e = 12, R = 40. Initial connection weights are drawn randomly from a uniform distribution U over [-1, 1], and mutation adds a random number from the normal distribution N_R (0, 0.2^2).

First, we focus on how pre-play communication among agents develops during the early stages of the evolution process. Fig. 2 represents the average fitness of each gene pool and the average fitness of all individuals from the 0th generation to the 99th generation. We see a rapid increase in fitness in all gene pools. The average fitness reached approximately -5 at the 99th generation.

Fig. 3 shows an example communication at the 50th generation. We see that the strategy value of one agent reached the upper limit and that of another agent reached the lower limit, while the remaining agent's value remained near the boundary line between two signs (a_i (t) = 0). We focus on the agent whose strategy value remained near the boundary line. In the situation shown in Fig. 3, the strategies of the other two agents are on the upper and the lower area respectively, and they did not change their strategy values. Thus, the focal agent could not avoid ending up on the majority side, and so its payoff falls below 0 regardless of which side it picks. The optimal behavior thus is to choose a strategy value as close to the boundary line as it can, and receive payoff of near 0. It was often observed in the simulations that the final strategies of the three agents settled on these three positions in the strategy space: the upper limit, the lower limit (high risk and high return strategy) and around the boundary line (low risk and low return strategy). This can be regarded as strategy differentiation among agents.

The average payoff of each agent over R trials becomes near 0, if 1) the final strategy of the agent nearby the boundary line stays very close to it, and 2) its sign splits fifty-fifty between positive and negative. In this scenario, the situation in which the final strategies fall in the same area of the strategy space (yielding payoff far below 0) is avoided. In Fig. 2 we see an increase in average fitness from the initial generation to the 20th generation, likely due to the emergence of the strategy differentiation among agents.

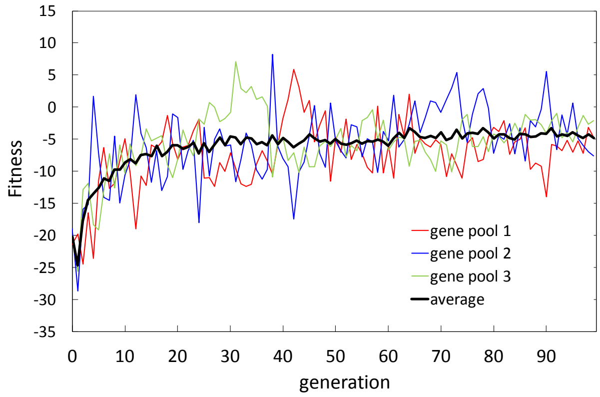

Fig. 2. Evolution of fitness.

Fig. 4 shows a typical communication from the 70th generation in which Agent 1 and Agent 3 can be seen to respond each other's behavior. In Fig. 3, all agents initially lower their strategy values. Agent 3 (at around step 40), and Agent 1 (at around step 80) can be seen to switch their directions and start increasing their strategy values, in order to avoid the situation in which the strategy values of all agents remain negative, in other words, all lose. Once the strategy value of Agent 1 surpasses 0 (at around step 150), Agent 3 switches its direction again correspondingly. The most likely explanation for these behaviors of Agent 1 and Agent 3 is that they changed the increase or decrease in their strategy values in response to the strategy values of the others.

Fig. 5 shows the communication among agents at the 80th generation. We can observe that agents interacted with each other more actively than was the case in the earlier generations. This sort of oscillation was often observed in the simulation.

Next, we classified the types of the pre-play communication among agents into the following three types: (a) Oscillation type (All agents cross the boundary line (a_i (t) = 0) more than one time in the communication stage.) (b) Fork type (All agents do not satisfy the condition of (a), and each strategy value is as follows: one's strategy value is more than 0.8, another's strategy value is more than -0.1 and less than 0.1, and the other's strategy value is less than -0.8), (c) Others. Fig. 6 shows the occurrence rate of each communication type from the 0th generation to the 99th generation. From this figure, we see that the Fork type occurs first, followed by the Oscillation type. We conducted ten trials and confirmed this tendency in nine trials among them. This increase in communicative complexity should be worth noting as in the field of animal signaling it has been believed that complex communication cannot be evolved in conflicting situations [7].

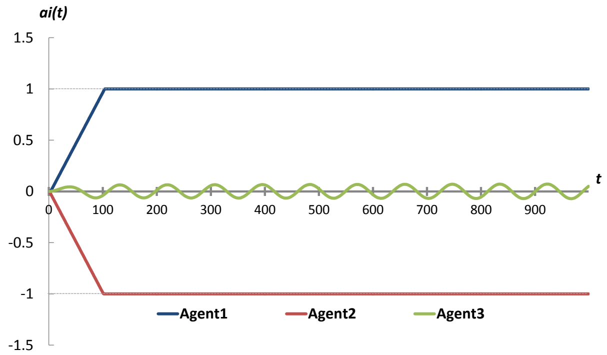

Fig. 3. Communication among agents at the 50th generation.

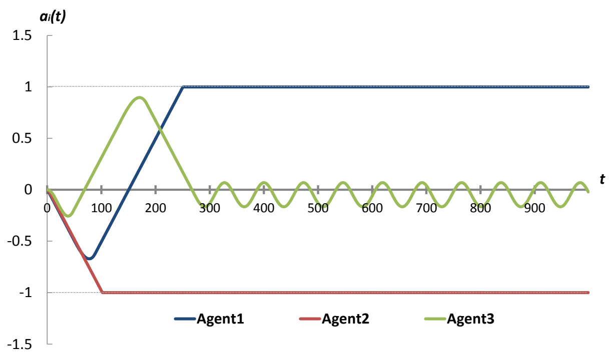

Fig. 4. Communication among agents at the 70th generation.

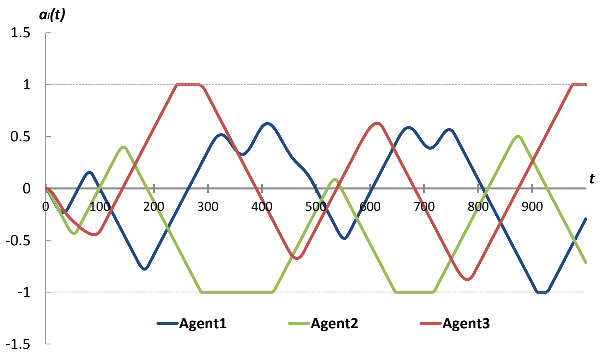

Fig. 5. Communication among agents at the 80th generation.

4 CONCLUSIONS

We extended the standard MG to a new game named *Interactive Minority Game (IMG)* by incorporating the two aspects: a continuous strategy space and a pre-play communication stage. As a result of evolutionary simulations in which each agent's RNN evolves based on the payoff of 3-person IMG, we discovered that the strategy differentiation among agents where two agents adopt "high-risk high-payoff" strategy and the rest one adopts "low-risk low-payoff". We also saw the emergence of various communications such as the adaptive adjustment behavior and oscillation of strategy values of agents.

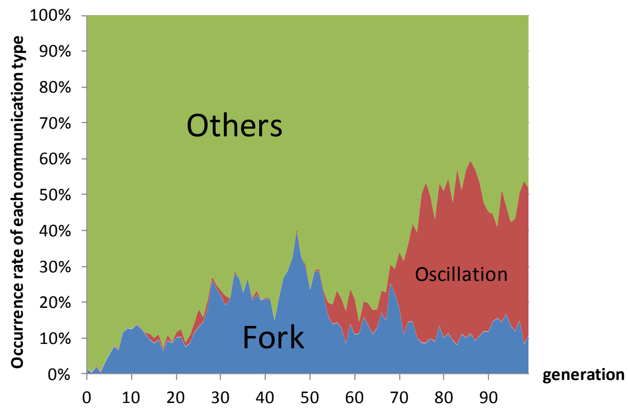

Fig. 6. Occurrence rate of each communication type in each generation.

REFERENCES

[1] Challet D and Zhang YC (1997), Emergence of cooperation and organization in an evolutionary game, *Physica A* 246: 407-418.
[2] Savit R, Manuca R and Riolo R (1999), Adaptive Competition, Market Efficiency, and Phase Transitions, *Phys. Rev. Lett* 82(10): 2203–2206.
[3] Challet D and Marsil M (1999), Phase transition and symmetry breaking in the minority game, *Phys. Rev. E* 60: R6271-R6274.
[4] Nishimoto K et al., Coevolutionary dynamics between roles and social sensitivity in an extended minority game, *Proceedings of ALIFE13*, Michigan, USA, July 19-22, 2012: 226-233.
[5] Anghel M, Toroczkai Z, Bassler KE and Korniss G (2004), Competition in social networks: emergence of a scale-free leadership structure and collective efficiency, *Phys. Rev. Lett.* 92: 058701.
[6] Remondino M and Cappellini A (2005), Minority game with communication of statements and memory analysis: a multi agent based model, *International Journal of Simulation: Systems, Science & Technology* 6(5): 42-53.
[7] Maynard-Smith J and Harper D (2003), *Animal signals*, Oxford University Press, Oxford, U.K.