Give-and-Take in Minority Games

Akira Namatame and Hiroshi Sato
Dept. of Computer Science,
National Defense Academy
Yokosuka, 239-8686, JAPAN
E-mail: {nama, hsato}@nda.ac.jp

Abstract

There is no presumption that collective behavior of interacting agents leads to collectively satisfactory results. How well agents in adapting to their social environment is not the same as how satisfactory the social environment is that they collectively create. In this paper, we attempt to probe deeper understanding at this issue by specifying how agents interact by adapting their behavior. We consider the asymmetric coordination problems formulated as minority games. We address the following basic question: how do interacting agents realize an efficient coordination without any central authority by self-organizing macroscopic order from the bottom up? We introduce a new adaptive model based on the concept of give-and-take, in which agents yield to others if they gain and randomize their actions if they lose or do not gain. We show that both efficiency and equlity of colective behavior are siginificantly improved if agents adapt with the give-and-take strategy. We also investigate how agents co-evolve their give-amd-talke strategies from the bottom-up.

Keywords: *aymmetric coordination, give-and-take, social efficiency, coevolution, meta-rule*

1. Introduction

There are many situations where interacting agents can benefit from coordinating their actions. Social interactions pose many coordination problems to individuals. For example, individuals face problems of sharing and distributing limited resources in an efficient way. Consider a competitive network routing problem in which the paths from sources to destination have to be established by multiple agents. In the context these traffic networks, for instance, agents have to determine their route independently, and in telecommunication networks, they have to decide on what fraction of their traffic to send on each link of the network.

Coordination implies that increased effort by some agents leads the remaining agents to follow suit, which gives rise to multiplier effects. We classify this type of coordination as symmetric coordination [3]. Coordination is also necessary to ensure that their individual actions are carried out with little conflicts. We classify this type of coordination as asymmetric coordination [7]. Consider the following situation: A collection of agents have to travel using either the route A or route B. Each agent gains the payoff if he chooses the route which is also determined by what the majority does. This type of coordination is classified as symmetric coordination. On the other hand, each agent gains a payoff if he chooses the opposite route to what the majority does. This type of coordination is classified as asymmetric coordination.

Coordination problems are characterized with many equilibria, and they often face the problem of coordination failure resulting from their independent inductive processes [1][4]. An interesting problem is then under what circumstances will a collection of agents realizes some stable situations, and whether they satisfy the conditions of social efficiency. In recent years, this issue has been addressed by formulating the minority games (MG)[2][10]. However, the growing literature on the MG treats agents as automata, merely responding to changing environments without deliberating about individuals' decisions [13]. There is no presumption that the self-interested behavior of agents should usually lead to collectively satisfactory results [8][9]. How well each agent does in adapting to its social environment is not the same as how satisfactory a social environment they collectively create for themselves. An interesting problem is then under what circumstances will a society of rational agents realize social efficiency? Solutions to these problems invoke the intervention of an authority who finds the social optimum and imposes the optimal behavior to agents. While such an optimal solution may be easy to find, the implementation may be difficult to enforce in practical situations. Self-enforcing solutions, where agents achieve optimal allocation of resources while pursing their self-interests without any explicit agreement with others are of great practical importance.

Previous researches on the collective action or the problems of coordination include the standard assumption that agents use the same kind of adaptive rule. In this paper, we depart from this assumption by considering a model heterogeneous agent with respect to their meta-rules of interactions. We also use the term emergent to denote stable macroscopic patterns arising from the local rules of agents. The interaction of many individuals produces some kind of coherent, systematic behavior. Since it emerges from the bottom up, we describe it as an example of self-organization. The surprise consists precisely in the emergence of a macrostructure from the bottom up, which is from simple local rules that outwardly appear quite remote from the collective phenomena they generate. In short, it is not the emergent macroscopic object per se that is surprising, but the generative sufficiency of the simple local adaptive rules.

We are interested in the bottom-up approach for leading to more efficient coordination with the power of more effective learning at the individual levels [11]. Within the scope of our model, we treat models in which agents make deliberate decisions by applying rational learning procedures. We explore the mechanism in which interacting agents are stuck at an inefficient equilibrium.

While agents understand that the outcome is inefficient, each agent acting independently is power less to manage the collective activity about what to do and also how to decide. The design of efficient collective action is crucial in many fields. In collective activity, two types of activities may be necessary: Each agent behaves as a member an society, while at the same time, it behaves independently by adjusting its view and action. At the individual level, it learns to improve its action based on its own observation and experiences. At the same level, they put forward their learnt knowledge for consideration by others. An important aspect of coordination is the learning rule adapted by individuals.

2. Formalisms of the EL Farol Problem and Minority Games

The EL Farol bar problem and its variants provide a clean and simple example of asymmetric coordination problems [1][4]. Brian Arthur used a very simple yet interesting problem to illustrate effective uses of inductive reasoning of heterogeneous agents. There is a bar called El Farol in downtown Santa Fe. Many agents are interested in going to the bar each night. All agents have identical preferences. Each of agents will enjoy the night at El Farol very much if there are no more than the threshold number of agents in the bar. However, each of them will suffer miserably if there are more than the threshold number of agents. In Arthur's example, the total number of agents is $N=100$, and the threshold number is set to 60. The only information available to agents is the number of visitors to the bar in previous nights.

What makes this problem particularly interesting is that it is impossible for each agent to be perfectly rational, in the sense of correctly predicting the attendance on any given night. This is because if most agents predict that the attendance to be low (and therefore decide to attend), the attendance will actually be high, while if they predict the attendance will be high (and therefore decide not to attend) the attendance will be low. Arthur investigated the number of agents attending the bar over time by using a diverse population of simple rules adapted by agents. One interesting result obtained is that over time, the average attendance of the bar is about 60. Agents make their choices by predicting ahead of time whether the attendance on the current night will exceed the capability and then take the appropriate course of action. Arthur examined that the dynamic driving force behind this equilibrium.

The Arthur's "El Farol" model has been extended in the form as Minority Games (MG), which show for the first time how equilibrium can be reached using inductive learning [2]. The MG is played by a collection of rational agents $G = \{A_i : 1 \le i \le N\}$. Without losing the generarity, we can assume N is an odd number. On each period of the stage game, each agent must choose privately and independently between two strategies $S = \{S_1, S_2\}$. We represent the action of agent A_i at the time period t by $a_i(t) = 1$ if he chooses S_1, and $a_i(t) = 0$ if he chooses S_2. Given the actions of all agents, the payoff of agent A_i is given by

[Payoff scheme 1]

(i) $u_i(t) = 1$ if $a_i(t) = 1$ and $p(t) = \sum_{1 \le i \le N} a_i(t) / N \le \theta$

(ii) $u_i(t) = 0$ if $a_i(t) = 0$ and $p(t) > \theta$ (2.1)

where θ is the capacity rate, and $\theta = 0.6$ with the El Farol problem, and $\theta = 0.5$ with the MG . Each agent first receives aggregate information $p(t)$ of all agents' actions, and then decides whether to choose S_1 or S_2. Each agent is rewarded with a unitary payoff whenever the side choosen happens to be chosen by the minority of the agents, while agents on the majority side receive nothing.

Since $A(t) \equiv \sum_{1 \le i \le N} a_i(t)$ represents the total number agents to choose S_1 (the total attendance) the time period t, the payoff scheme in (2.1) can be summarized as follows:

$$u_i(t) = -a_i(t)\,\mathrm{sgn}(A(t)) \tag{2.2}$$

The payoff function in (2.2) becomes to be the step function as shown in Fig.1.

We have another payoff scheme specifiesd as follows:

[Payoff scheme 2]

$$u_i(t) = -a_i(t) A(t) / N, \qquad A(t) = \sum_{1 \le i \le N} a_i(t) \tag{2.3}$$

The payoff function in (2.3), which is shown in Fig.2, is linear with respect to the proportion of the attendaces $p(t) = \sum_{1 \le i \le N} a_i(t) / N$. Each agent also gets aggregate information $p(t)$ which aggregate all agents' actions, and then he decides whether to chooses S_1 or S_2. Each agent is rewarded with a payoff which is linearly decreasing function of the proportion of the attendance, $p(t)$.With the payoff in scheme 1, whenever the side an agent chooses happens to be chosen by the minority of agents, they receive a unitary award, while agents on the majority side receive nothing. With the payoff in scheme 2, whenever the side an agent chooses happens to be chosen by the minority of the agents, they receive an award which is proportional to the level of the crowdedness.

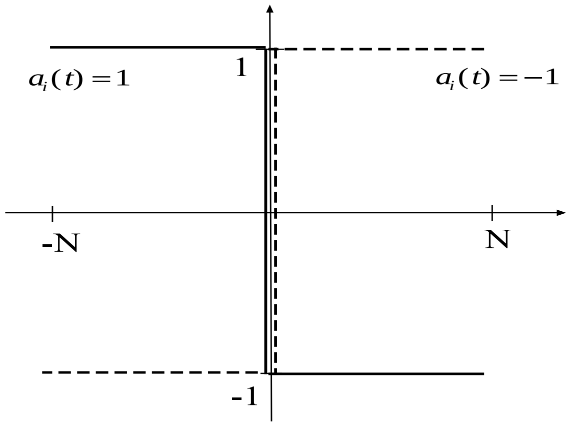


Fig1. The payoff scheme 1

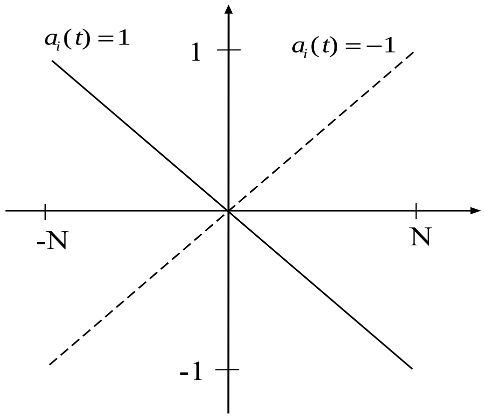


Fig 2. The payoff scheme 1

The MG game is characterized with many solutions. It is easy to see that this game has $\binom{N}{(N-1)/2}$ asymmetric Nash equilibria in pure strategies in the case where exactly *(N-1)/2* agents choose either one of the two sides. The game also presents a unique symmetric mixed strategy Nash equilibrium in which each agent selects the two sides with an equal probability. We analyze the structure of the MG to see what to expect. The social efficiency can be measured from the average payoff of one agent over a long-time period. Consider the extreme case where only one agent take one side, and all the others take the other side at each time period. The lucky agent gets a reward, nothing for the others, and the average payoff per agent is *1/N*. Equally extreme situation is that when *(N-1)/2* agents on one side, *(N+1)/2* agents on the other side where the average payoff is about 0.5. From the society point of view, the latter situation is preferable.

Several methods have been suggested to lead an efficient outcome when agents learn from each other [2][15]. All agents have access to public information of $p(\tau), \tau \le t$. The past history available at the time period t is represented by $\mu(t)$. How do agents choose actions under the common information $\mu(t)$? Agents may behave differently because of their personal beliefs on the outcome of the next time period $p(t+1)$, which only depends on what agents do at the next time period *t+1*, and the past history $\mu(t)$ has no direct impact on it.

3. Decomposition of Minority Games into 2x2 Asymmetric Games

The matching methodology also plays an important role in the outcome of the game. Agents interact with all other agents, which is known as the uniform matching, or they interact with a randomly chosen agent. Agents are not assumed to be knowledgeable enough to correctly anticipate all other agents' choices, however they can only access on the information about the aggregate behavior of the society with the random matching.

Agents are rewarded a unitary payoff whenever the side chosen happens to be chosen by the minority of the population. The El Farol prolem and Minority Games have a common feature that an agent's utility depend on the number of total participants. We now show the MG can be represented as 2x2 games in which an agent play with the aggregate of the society of the population *N* with payoff matrix in Table 1. Let suppose each agent plays with all other agents individually with the payoff matrix in Table 1. The payoffs of agent A_i from the play with all other agents with S_1 and S_2 are given:

$$\overline{U}_i(S_1) = -n + N - n - 1 = -A(t) - 1$$

$$\overline{U}_i(S_2) = n - (N - n - 1) = A(t) + 1 \qquad (3.1)$$

where *n* represents the number of agents to choose S_1. Deviding the above payoffs by *N*, we obatin the average payoff of each interaction with one agent as:

$$U_i(S_1) = \overline{U}_i(S_1)/N \cong -A(t)/N$$

$$U_i(S_2) = \overline{U}_i(S_2)/N \cong A(t)/N \qquad (3.2)$$

We denotethe proportion of agents to choose S_1 at the time period *t* by $p(t) = \sum_{1\le i\le N} a_i(t)/N$. The payoffs of agent A_i from the play with one randomly chosen agent (random matching) with S_1 and S_2 are given:

$$U_i(S_1) = 1 - \sum_{1\le i\le N} a_i(t)/N = 1 - p(t) \qquad U_i(S_2) = \sum_{1\le i\le N} a_i(t)/N = p(t) \qquad (3.3)$$

The above payoff scheme is described in Fig.3.

Table 1 The payoff matrix of the minority games

Own's strategy \ The other's strategy	S_1 (go)	S_2 (stay)
S_1 (go)	0, 0	1, 1
S_2 (stay)	1, 1	0, 0

Table 2 The payoff matrix of the general minority games

Own's strategy \ The other's strategy	S_1 (go)	S_2 (stay)
S_1 (go)	0, 0	θ, 1−θ
S_2 (stay)	1−θ, θ	0, 0

The El Farol model is about the equilibrium, the MG is about fluctuations, these two models can be treated with the following generic formulation: Let suppose each agent play the two-person game using the payoff matrix in Table 2 with all other agents or the aggregate of the society. The

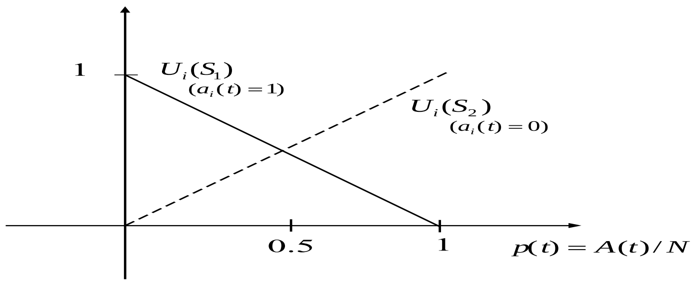


Fig3:The payoff scheme of the payoff matrx in Table1

payoff when agent A_i chooses S_1 and the aggregate chooses S_2 is given by θ ($0<\theta<1$), and the payoff when agent A_i chooses S_2 and the aggregate chooses S_1 is given by $1-\theta$. The El Farol problem can be modeled with $\theta=0.6$, which is the ratio of the bar, and the MG is formulated with $\theta=0.5$.

Social efficiecy of the MG also depends on the payoff scheme. First of all, we obatin with the payoff scheme 1. Let suppose there exists some central authority, and it leads that a little bit larger number of agents than $N\theta$ choose S_1 if $\theta\geq 0.5$, and a little bit fewer agent than $N\theta$ choose S_1 if $\theta<0.5$. In this case the average payoff per agent is obtained as $Max(\theta,\ 1-\theta)$. Similary, if the central authority leads that a little bit fewer number of agents than $N\theta$ choose S_1 if $\theta\geq 0.5$, and a little bit larger than $N\theta$ choose S_1 if $\theta<0.5$. In this case the average payoff per an agent is obtained as $Min(\theta,\ 1-\theta)$. Then we have the following average payoffs as the best case and the worst case:

[The payoff of the best case]

$$\underset{0\leq\theta\leq 1}{Max}(\theta,\ 1-\theta) \tag{3.4}$$

[The payoff of the worstcase]

$$\underset{0\leq\theta\leq 1}{Min}(\theta,\ 1-\theta) \tag{3.5}$$

Thefore the average payoffs of the best case and the worst case become the same at $\theta=0.5$.

We now consider with the payoff scheme 2 in (2.3). The expected payoff of an agent who chooses S_1 is given $1-\theta$, and that of an agent who chooses S_2 is θ, where θ denotes the proportion of agents who choose S_1 .Therefore the average payoff of an agent is given by $2\theta(1-\theta)$, which takes the maimum value 0.5 at $\theta=0.5$. The average payoff under the payoff scheme 2 is also shown in Fig.4 as the function of the capacity rate θ.

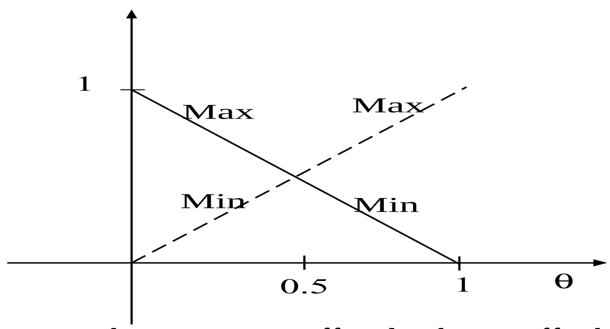


Fig4: The average payoff under the payoff scheme1

4. Models of Individual Learning

Game theory is typically based upon the assumption of a rational choice. In our view, the reason for the dominance of the rational-choice approach is not that scholars think it to be realistic. Nor is game theory used solely because it offers good advice to a decision maker, because its unrealistic assumptions undermine much of its value as a basis for advice. The real advantage of the rational-choice assumption is that it often allows deduction. The main alternative to the assumption of rational choice is some form of adaptive behavior. The adaptation may be at the individual level through learning, or it may be at the population level through differential survival and reproduction of the more successful individuals. Either way, the consequences of adaptive processes are often very hard to deduce when there are many interacting agents following rules that have nonlinear effects.

We specify how agents adapt their behavior in response to others' behavior in strategic environments. Among the adaptive mechanisms that have been discussed in the learning literature are the following [5][6][12][14]. An important issue in strategic environment is the learning strategy adapted by each individual.

(1) Reinforcement learning

Agents tend to adopt actions that yielded a higher payoff in the past, and to avoid actions that yielded a low payoff. Payoff describe choice behavior, but it is one's own past payoffs that matter, not the payoffs of the others. The basic premise is that the probability of taking an action in the present increases with the payoff that resulted from taking that action in the past. [6]

(2) Best response learning

Agents adopt actions that optimize their expected payoff given what they expect others to do. In this learning model, agents choose best replies to the empirical frequencies distribution of the previous actions of the others.

(3) Evolutionary learning

Agents who use high-off payoff strategies are at a productive advantage compared to agents who use low-payoff strategies, hence the latter decrease in frequency in the population over time (natural selection). In the standard model of this situation agents are viewed as being genetically coded with a strategy and selection pressure favors agents that are fitter, i.e., whose strategy yields a higher payoff against the population.

(4) Social learning

Agents learn from each other with social learning. For instance, agents may copy the behavior of others, especially behavior that is popular to yield high payoffs (imitation). In contrast to natural selection, the payoffs describe how agents make choices, and agents' payoff must be observable by others for the model to make sense. The crossover strategy is also another type of social learning.

These learning models can be represented on the spectrum in Figure 3. The reinforcement learning and social learning based on give-and-take take limiting cases representing at the right-most and left-most points of the spectrum.

Fig.5 : The spectrum of learning models

5. Collective Behavior with Best-Response Learning

With the assumption of rationality, agents are assumed to choose an optimal strategy based on a sample of information about what other agents have done in the past. Agents are able to calculate best replies and learn the strategy distribution of play in a society. They gradually learn the strategy distribution in the society. With best-response learning, each agent calculates the best strategy based on information about the current distributional patterns of strategies [5]. At each period of time, each agent decides which strategy to choose given the knowledge of the aggregate behavior of the population. Each agent thinks strategically, knowing that everyone else is also making a rational choice given its own information.

An important assumption is how they receive knowledge of the current strategy distribution. For simplicity we assume that there is no strategic interaction across time; an agent’s decision depends only on current information and not on any previous actions. The dynamics for collective decision of agents are described as follows: Let *p(t)* be the proportion of agents who have chosen S_1 at time *t*. Let $U_i(S_k)$ be the expected payoff to A_i when S_k, *k=1,2,* is chosen. The best-response is then given as follows:

If $U_i(S_1) > U_i(S_2)$ then choose S_1

If $U_i(S_1) < U_i(S_2)$, then choose S_2 (5.1)

The expected payoffs of agent A_i are obtained as

$$U_i(S_1) = \theta(1 - p(t)), \quad U_i(S_2) = (1-\theta)p(t) \tag{5.2}$$

The best-response adaptive rule of agent A_i is then obtained as follows:

(i) If $p(t) < \theta$, then choose S_1

(ii) If $p(t) > \theta$, then choose S_2 (5.3)

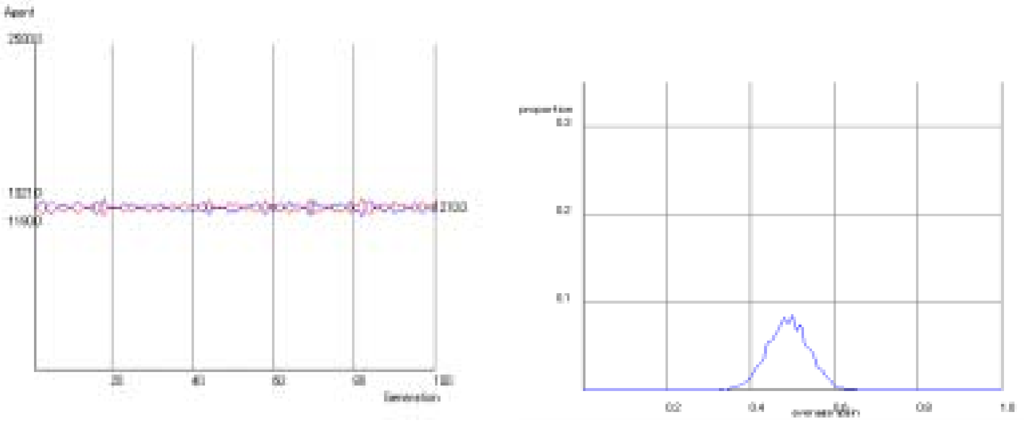
The aggregate information *p(t)*, the current status of the collective decision, provides a significant effect on agents' rational decisions.

The result of the learning with the global best-response strategy is simple. Starting from any initial condition *p(0)*, it cycles between the two extreme situations where all agents choose S_1 or S_2. Under this cyclic behavior, no agent gains resulting in a huge waste.This result has a considerable intuitive appeal since it displays situations where rational individual action, in pursuit of well-defined preferences, lead to undesirable outcomes.

6. Collective Behavior with Mixed Strategies

In this section, we provide simulation results when each agent adapts the same strategy *RND(x)*, which represents the mixed starategy $\boldsymbol{x}=(x, 1-x)$ of chossing S_1 with the probability x and S_2 with *1-x*. In the section 2, we showed that the MG can be analyzed by the 2x2 games. The rational behaviour of an agent in the MG becomes to be the same as the one when each agent interacts to all other agents with the payoff matrix in Table 1. The payoff matrix in Table 1 has the unique symmetric mixed strategy Nash equilibrium in which each agent selects the two sides with the equal probability. If all agents adapt the mixed Nash equilibrium strategy, *RND(0.5)*, each agent can expect the payoff 0.5 of each time period, and the society payoff follows a binomial distribution with the mean equal to *N/2* and the variance *N/4*. The variance is also an measure of the degree of social efficiency. The higher the variance, the higher magnitude of the fluctuations around *N/2* and the corresponding aggregate welfare loss.

We consider a population of agents with *N=2,500* with the capacity rate θ=0.5. In Figure 6, we showed the simulation result when all agents adapt the same mixed Nash equillibrium strategy *RND(0.5)*. The Fig 4(a) shows the number of agents having chosen S_1 and S_2 over time, and it is shown that the average number of agents who choose S_1(Go) converges to the capacity of the bar, indicating that collective behavior satisfies the constraint. In Fig.4 (b), we showed the proportion of agents with the same average payoff. The avergae payoff per agent ranges from 0.7 to 0.3, and the diffrence of payoff for lucky agents and that of unlucky agetns becomes large. This indicate that the social inequality spreads throughout the society.



(a) (b)

Fig.6: The simuulation result under Nash equilibrium startegies: (a)The number of agents to choose S_1 and S_2, (b) the proportion of agents with the same average payoff under the mixed Nash strategy

7. Collective Behavior with Give-and-Take Strategies

Several learning rules have been found to lead an efficient outcome when agents learn from each other [2][15]. In this section, we propose the give-and-take strategy which departs from the conventional assumption such that agents update their behaviors in order to improve their measure functions such as payoffs. It is commonly assumed that agents tend to adopt actions that yield a higher payoff in the past, and to avoid actions that yield a low payoff. With the give and take learning, on the contrary, agents are assumed that they yield to others if they receive the payoff by taking the opposite starategy at the next time period, and they choose randomly if they do not gain the payoff. We formalize the payoff scheme with give-and-take strategy as foilows:

(i) If $a_i(t) = 1$ (Choose S_1) and $p(t) < \theta$ (Minority), then agent A_i gains the one unit

(ii) If $a_i(t) = 0$ (Choose S_2) and $p(t) > \theta$ (Minority), then agent A_i gains the one unit

(iii) If $a_i(t) = 1$ (Choose S_1) and $p(t) > \theta$ (Majority), then agent A_i gains nothing

(iv) If $a_i(t) = 0$ (Choose S_2) and $p(t) < \theta$ (Majority), then agent A_i gains nothing (7.1)

We denote the status of the collective choice by all agents with the following state variable $\omega(t)$ as follows:

$$\omega(t) = 1 \ \ if \ \ A(t) \geq N\theta \qquad \omega_2(t) = 0 \quad if \quad A(t) < N\theta \tag{7.2}$$

Each agent receives common information on $\omega(t)$ which aggregate all agents' actions of the last time period, and then decides whether to choose S_1 or S_2 at the time period *t+1* by considering whether he is rewarded at the previous time *t*: The action $a_i(t+1)$ of agent A_i at the next time period *t+ 1* is determined by the following rules:

(i) $(\omega(t) = 0) \wedge (a_i(t) = 1) \Rightarrow a_i(t+1) = 0$

(ii) $(\omega(t) = 1) \wedge (a_i(t) = 0) \Rightarrow a_i(t+1) = 1$

(iii) $(\omega(t) = 1) \wedge (a_i(t) = 1) \Rightarrow a_i(t+1) = RND(x)$

(iv) $(\omega(t) = 0) \wedge (a_i(t) = 0) \Rightarrow a_i(t+1) = RND(y)$ (7.3)

where *RND(x)* represents the mixed starategy $\boldsymbol{x}=(x, 1-x)$ of chossing S_1 with the probability *x* and S_2 with 1-*x*.

[give-and-take strategy without delibration]

If agents with give-and take learning adapt the random strategies *RND(x)=RND(y)* in (5.3) as the mixed starategies $\boldsymbol{x}=\boldsymbol{y}=(0.5, o.5)$ of the payoff matrix Table 1, we define as give-and-take strategy without careful delibrattion. We consider a population of agents with *N=2,500* with the capacity rate θ=0.5. In Figure 7, we showed the simulation result when all agents adapt pure give-and take learning. It is shown that the number of agents who choose S_1(Go) converges to about 2*N/3*, larger than the capacity. In

[give-and-take strategy with Delibration]

We now consider how agents choose the radom strategy *RND(x)* when they are in the majority by choosing S_1, the radom strategy *RND(y)* when they are in the majority by choosing S_2, so that they eventually to converge to the capacity. The expected number of agents to choose S_1 at the next time *t+1* if they use the rules in (5.3) is given as

$$A(t+1) = xA(t) + N - A(t) \tag{7.4}$$

Therefore, if they choose *RND(x)* so that

$$A(t+1) = N\theta \tag{7.5}$$

, then we can obtain *x* as

$$x = \{N\theta - (N - A(t))\} / A(t) \tag{7.6}$$

Here we assume the condition $N - A(t) \le N\theta$ is satisfied. In the case when the attendance is more that the capacity (over crowded), the number of agents to stay at home is smaller than the capacity. This assumption is easily satisfied with $\theta = 0.5$.

Similarily we obtain the mixed strategy *RND(y)* when they are in the majority by choosing S_2 . The expected number of agents to choose S_1 at the next time t+1 if they use the rules in (5.3) is given as

$$A(t+1) = y(N - A(t)) \tag{7.7}$$

Then, if we choose *RND(y)* so that the following condition is satisfied.

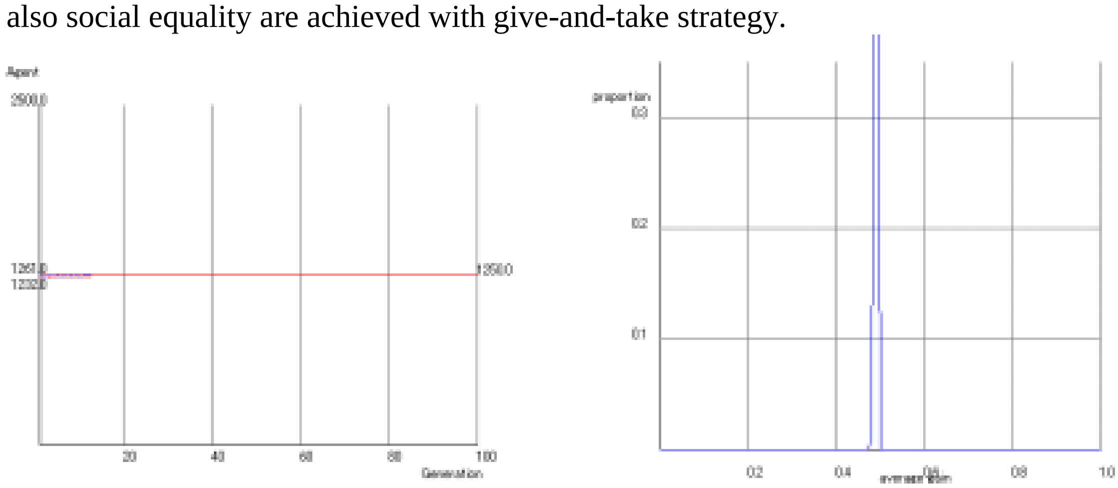
$$A(t+1) = N\theta \tag{7.8}$$

Then we obtain as y as

$$y = N\theta / (N - A(t)) \tag{7.9}$$

Here we assume the condition of $N - A(t) \ge N\theta$ is satified. In the case when the attendance is below than the capacity, the number of agents to stay at home is greater than the capacity. This assumption is also easily satisfied with $\theta = 0.5$.

We also consider a population of agents with *N=2,500* with the capacity rate θ=0.5. In Figure 8 shows the simulation result when all agents adapt the give-and take learning rules in (5.3). Fig 8(a) shows the number of agents having chosen S_1 and S_2 over time, and it is shown that the average number of agents who choose S_1(Go) converges to the capacity, indicating that collective behavior satisfies the constraint. Fig.8 (b) shows the proportion of agents with the same average payoff. The majority of agents receive the avergae payoff 0.5. This result indicates that not only social efficiency, but also social equality are achieved with give-and-take strategy.



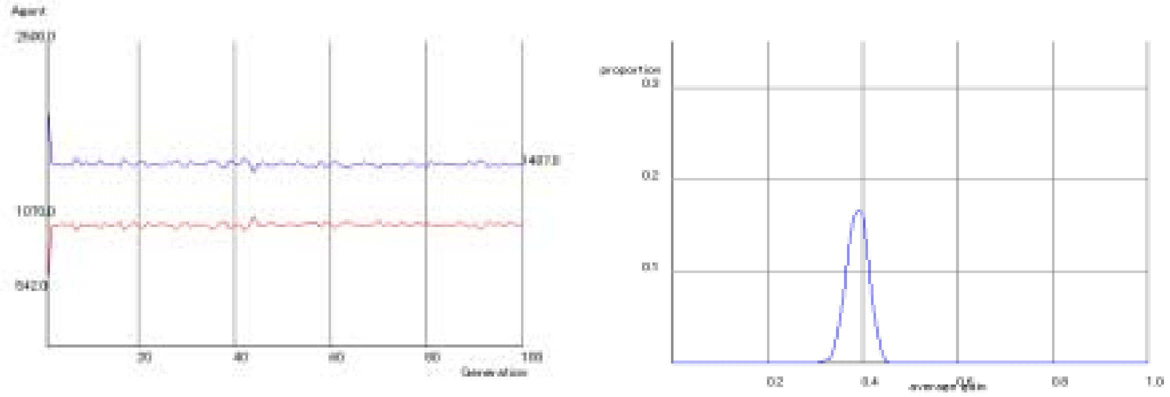
(a) (b)

Fig.7: Simulation result of give-and-take strategy with delibration (θ=0.5) (a)The dynamic change of numbers of agents

of having choosen S_1 and S_2 (b) The proportion of agents with the same payoff

As shown in Section 3, the the average payoff per agent is given by $2\theta(1-\theta)$, which takes the maimum vallue at the capacity rate $\theta=0.5$. We evaluate the performace of give-and take learning with with the capacity rate θ=0.6, the capacity rate of the El Farol problem. In Figure 8, we showed the simulation result. The Fig 8(a) shows the number of agents of having chosen S_1 and S_2 over time. It is shown that the average number of agents who choose S_1(Go) converges to the capacity given by $N\theta$.. In Fig.8(b), we showed the proportion of agents with the same average payoff. The majority of agents received the payoff less than 0.5, which is the average payoff at the efficient collective behavior. The deviation of the averarge payoff become to be large compared with Fig. 8(b).

We evaluate the performace of give-and take learning with with the capacity rate θ=0.4, and the simulation result is shown in Figure 10. As shown in Fig 8(a) the number of agents to choosS_1 (Go) eventualy converged to the capacity $N\theta$. In Fig.9(b), we showed the proportion of agents with the same average payoff and the majority of agents received the payoff less than 0.5, which is the average payoff at the efficient collective behavior. The deviation of the averarge payoff become to be large compared with Fig. 7(b). This result indicate that the problems of inefficiency and uneuity become to be crucial if the capacity rate θ deviates from 0.5, and increasing the asymmetry of the minority and majority sides. This implies that we may need the central authority in order to achieve both social efficiency and equity in asymmetric situations.



(a) (b)

Fig.8: Simulation result of give-and-take strategywith delibration (θ=0.6)(a)The dynamic change of numbers of agents of having chosen S_1 and S_2 (b) The proportion of agents with the same payoff

8. Evolutionary Learning with Local Matching

In this section, we investigate evolutionary learning where agents learn from the most successful neighbours, and they co-evolve their strategies over time. Each agent adapts the most successful strategy as guides for their own decision (individual learning). Hence their success depends in large part on how well they learn from their neighbours. If the neighbour is doing well, its strategy can be imitated by all others (collective learning). In an evolutionary approach, there is no need to assume a rational calculation to identify the best strategy. Instead, the analysis of what is chosen at any specific time is based upon an implementation of the idea that effective strategies are more likely to be retained than ineffective strategies [15]. Moreover, the evolutionary approach allows the introduction of new strategies as occasional random mutations of old strategies. The evolutionary principle itself can be thought of as the consequence of any one of three different mechanisms. It could be that the more effective individuals are more likely to survive and reproduce. A second interpretation is that agents learn by trial and error, keeping effective strategies and altering ones that turn out poorly. A third interpretation is that agents observe each other, and those with poor performance tend to imitate the strategies of those they see doing better.

In this section we consider the local matching as shown in Fig. 10, where each agent is modeled to be matched with his 8 neighbours. Each agent is modeled to be matched several times with the same neighbour, and the rule of the strategy selection is coded as the list as shown in Figure 11.

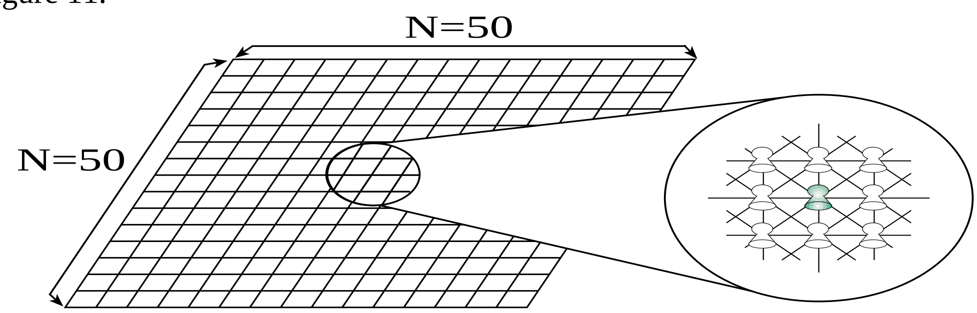


Fig 9: A Society of agents with loccal interactions: Each agent interact with 8 neighbours

A part of the list is replaced with that of the most successful neighbour. An agent's decision rule is represented by the N binary string. At each generation gen, $gen \in [1,..,lastgen]$, agents repeatedly play the game for T iterations. An agent A_i, $i \in [1...N]$, uses a binary string i to make a decision about his action at each iteration t, $t \in [1...T]$. A binary string consists of 22 positions (genes). Each position p_j, $j \in [1, 22]$, is represented as follows. The first and second position, p_1 and p_2, encodes the action that the agent takes at iteration $t=1$ and $t=2$. A position p_j, $j \in [3,6]$, encodes the history of mutual hands (cooperate or defect) that agent i took at iteration $t-1$ and $t-2$ with his neighbor (opponent). A position p_j, $j \in [7,22]$, encodes the action that agent i takes at iteration $t>2$, corresponding to the position p_j, $j \in [3,6]$.

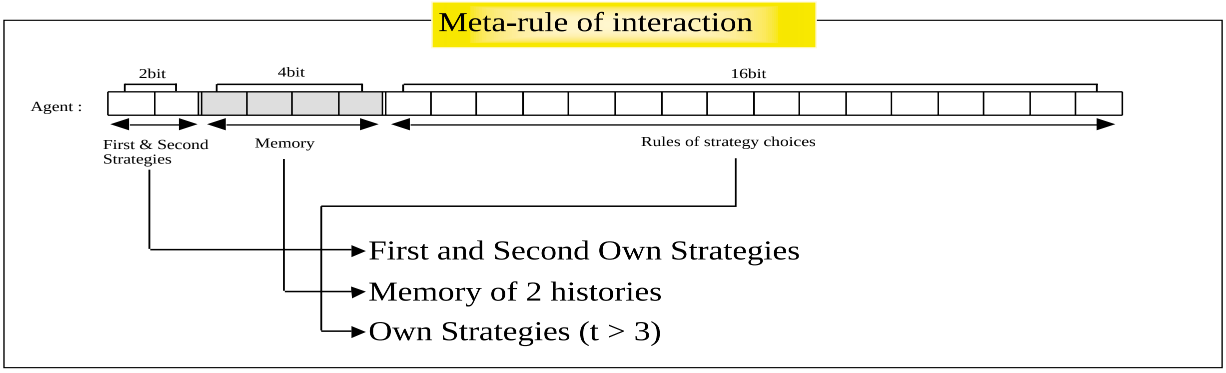
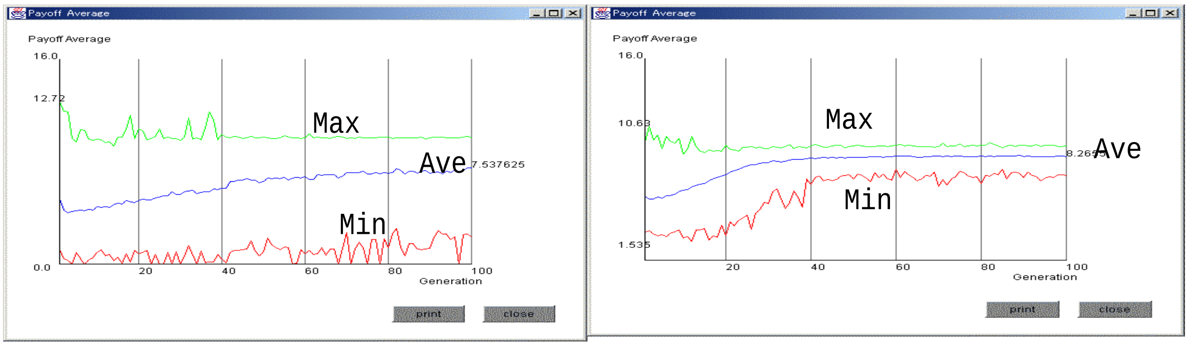


Fig 10: The representation of the meta-rule a the list of strategies

Each agent interacts with the agents on all eight adjacent squares and imitates the strategy of any better performing one. In each generation, each agent attains a success score measured by its average performance with its eight neighbours. Then if an agent has one or more neighbours who are more successful, the agent converts to the strategy of the most successful of them or crosses with the strategy of the most successful neighbour. Neighborus also serve another function as well. If the neighbor is doing well, the behaviour of the neighbour can be shared, and successful strategies can spread throughout a population from neighbour to neighbour [11].

We also consider the error at the choice of the startegy. Agents choose their staetegy which is specified by the meta-rule. However, we assume there exists small probabilitty of choosing the wrong strategy. We showed the si,ulation results without any erro and with the error rate 5% in Fig. 11. Consequently, we can conclude that evolution learning leads to a more efficient situation in the strategic environments. Significant differences were observed when agents have small chances of making mistakes. As shown Fig 11 (a), the highest payoff and the lowest payoff become to be close, which imply that each agent acquires the almost the same.



(a) Withou any mistake (b) Error rate: 5%

Fig 11: The average payoff with evolutionary learning

At beginging, wach agent has a diffrent meta-rule which is specified by the 16 bit information. In Fig 12, we showed the meta-rules acquired by 400 agents, which are aggregated into 15 types. The numbers of the blanket represent the numbers of agents who acquiredthe same types of the meta-rules. Those 15 meta-rules have also the commonality as shown in Fig.12. If agents choose $S_1(0)$ and their oppnent chooses $S_2(1)$ at the previous time period, then they choose $S_2(1)$. If agents choose $S_2(1)$ and their oppnent chooses $S_1(0)$ at the previous time period, then they choose $S_1(0)$. These rules represent if they gain then they change their staetegy, and this is the principle of give-and-take strategy as discussed at the section 7. From this result we can conclude that the evolutionary learning of meta-rules with some mistakes help agents to acquire give-and-take rules in the long run, which lead a society of agents to be both efficient and equitable.

TYPE1 =	0,1,0,1,1,1,0,0,0,1,0,1,1,1,0,0	(10)
TYPE2 =	0,1,0,1,1,1,0,0,0,1,0,1,0,1,0,0	(31)
TYPE3 =	1,1,0,0,1,1,0,0,0,1,0,1,0,1,0,0	(72)
TYPE4 =	0,1,0,0,1,1,0,0,0,1,0,1,0,1,0,0	(61)
TYPE5 =	0,1,0,0,1,1,0,0,0,1,0,1,1,1,0,0	(10)
TYPE6 =	1,1,0,0,1,1,0,0,0,1,0,1,1,1,0,0	(13)
TYPE7 =	0,1,0,0,1,0,0,0,0,1,0,1,0,1,0,0	(40)
TYPE8 =	1,1,0,1,1,1,0,0,0,1,0,1,0,1,0,0	(37)
TYPE9 =	1,1,0,0,1,0,0,0,0,1,0,1,0,1,0,0	(43)
TYPE10 =	0,1,0,1,1,0,0,0,0,1,0,1,0,1,0,0	(31)
TYPE11 =	1,1,0,1,1,0,0,0,0,1,0,1,0,1,0,0	(35)
TYPE12 =	0,1,0,0,1,0,0,0,0,1,0,1,1,1,0,0	(2)
TYPE13 =	1,1,0,1,1,1,0,0,0,1,0,1,1,1,0,0	(12)
TYPE14 =	0,1,0,1,1,0,0,0,0,1,0,1,1,1,0,0	(1)
TYPE15 =	1,1,0,0,1,1,0,0,0,1,1,1,0,1,0,0	(2)

Fig 12: The types of the meta-rules acquired by 400 agents: The numbers of the blanket represent the numbers of agents who acquiredthe same types of the meta-rules.

	memory				
	t−2		t−1		
Position	own	opp	own	opp	action
#2	0	0	0	1	1
#3	0	0	1	0	0
#5	0	1	0	0	1
#7	0	1	1	0	0
#8	0	1	1	1	0
#9	1	0	0	0	0
#10	1	0	0	1	1
#11	1	0	1	0	0
#12	1	0	1	1	1
#14	1	1	0	1	1
#15	1	0	1	0	0
#16	1	1	1	1	0
0	0	1	0	1	0/1

Fig 13: The commonalities of the meta-rules

9. Conclusion

The interaction of heterogeneous agents produces some kind of coherent, systematic behavior. We investigated the macroscopic patterns arising from strategic interactions of heterogeneous agents who behave based on local rules. In this paper we addressed questions such as: 1) how a society of selfish agents self-organizes, without a central authority, their collective behavior to satisfy the constraints? 2) How does learning at individual levels generate more efficient collective behavior? 3) How does co-evolution in a society put in its indivisible hand to promote self-organization of emerging collective behaviors? In previous works in the area of collective behavior, the standard assumption has been that agents use the same kind of adaptive rule. In this paper, we departed from this assumption by considering a model heterogeneous agent with respect to their meta-rule of making decisions at each time period. Agents use ad hoc meta-rules to make their decision based on past performance. We also considered several types of learning rules for agents to update their meta-rule to make decisions. We considered specific strategic environments in which a large number of agents have to choose one of two sides independently and those on the minority side win, which is known as minority game. A rational approach is helpless in our minority game by generating a large-scale social inefficiency. We introduced a new learning model at the individual level, give-and-take strategy in the situation where every agent should make his decision based on the past history of the collective behavior. It is shown that emergent collective behavior is more efficient than that generated from the mixed Nash equilibrium strategies. We also proposed collaborative learning based on Darwinism. It is shown that in strategic environments where every agent has to keep improving their meta-rule in order to survive, if agents learn from each other, then the social efficiency is realized without central of authority.

10. References

[1]Arthur W.B. “Inductive reasoning and Bounded Rationality”, *American Economic Review*, Vol.84, pp406-411, 1994.

[2]Challet,D& Zhang.C.," Emergence of Cooperation and Organization in an Evolutionary Game, Physica A246, 1997.

[3]Cooper, R: *Coordination Games,* Cambridge Univ. Press, 1999.

[4]Fogel B., Chellapia K., “Inductive Reasoning and Bounded Rationality Reconsidered”, *IEEE Trans. of Evolutionary Computation*, Vol.3, pp142-146, 1999.

[5]Fudenberg, D., and Levine, D., *The Theory of Learning in Games*, The MIT Press, 1998

[6]Hart,S.,and Mas-Colell, *A*, Simple Adaptive Procedure Leading to Correlated Equilibrium, *Econometrica*, 2000.

[7]Iwanaga S., Namatame. A: “Asymmetric Coordination of Heterogeneous Agents”, IEICE Trans. on Information and Systems, Vol. E84-D,, pp.937-944, 2001.

[8]Iwanaga S., Namatame. A: “The Complexity of Collective Decision, Journal of Nonlinear Dynamics and Control”, Vol..6, No.2, pp137-158, 2002.

[9]Kaniovski, Y., Kryazhimskii, A., & Young, H. Adaptive Dynamics in Games Played by Heterogeneous Populations. *Games and Economics Behavior*, *31*, 50-96,2000.

[10]Marsili Matteo.,"Toy Models of Markets with Heterogeneous Interacting Agents", Lecture Notes in Economics and Mathematical Systems, Vol.503, pp161-181, Springer, 2001

[11]Namatame, A,and Sato, H., "Collaborative Learning in Strategi Environments", Proc. of AAAI Spring Symposium, 2002

[12]Samuelson,,L., *Evolutionary Games and Equilibrium Selection*, The MIT Press, 1998.

[13]Sipper M., *Evolution of Parallel Cellular Machines*: The Cellular Programming Approach, Springer, 1996.

[14]Young, H.P., *Individual Strategy and Social Structure*, Princedom Univ. Press, 1998.

[15]Weibull J., *Evolutionary Game Theory,* The MIT Press, 1996.

[16]Wolpert, D, Wheeler. K., and Tumer, K, "Collective Intelligence for Control of Distributed Dynamical Systems,"Europhysics Letters , Vol. 49, No. 6, 2000.

[17]Wolpert, D, and Tumer, K," Optimal Payoff Functions for Members of Collectives," In Advances in Complex Systems, 2001.