Collective Intelligence with Selective Interactions and Reinforcement of Preferences

Akira Namatame,
Dept. of Computer Science, National Defense Academy, Yokosuka, JAPAN
Email: nama@nda.ac.jp

Abstract

There is no presumption that a collective action of interacting agents leads to collectively satisfactory results without any central authority. Agents normally react to others’ decisions, and the resulting volatile collective action is often far from being efficient. It is also common to exist a conflict between an individual and a collective. An agent wishes to maximize her own utility and a system designer wishes to implement a decentralized algorithm for maximizing the whole utility of a collective system. The system performance of interacting agents, however, crucially depends on the type of interactions among agents as well as how they adapt to others. We investigate how do interacting agents with heterogeneous preferences can generate an efficient collective action. We develop a model to examine the interaction between partner choice and an individual action. Agents choose their partners and also decide on a mode of behavior in interactions with partners. We show a collection of heterogeneous agents with diverse preference can realize the most efficient collective action by selecting their partners to interact. We also consider a collection of interacting agents with homogeneous preferences. We show a collection of homogeneous agents with the same preferences evolves into heterogeneous agents with diverse preferences by realizing the most efficient and equitable collective action.

Keywords: collective learning, reinforcement learning, efficiency, heterogeneity, selective interaction

1. Introduction

The performance of collective system consists of many interacting components, which we call agents, should be described on two different levels: the microscopic level, where the decisions of the individual agents occur and the macroscopic level where collective decisions can be observed. To understand the role of a link between these two levels remains one of the challenges of complex system theory. In examining collective, we shall draw heavily on the individual decisions. It might be argued that understanding how individual make decisions is sufficient to understand and improve collective action. These situations, in which an agent decision depends on the decisions of the others, are the ones that usually do not permit any simple summation or extrapolation to the aggregates. The greatest promise lies in analysis of situations where agents behave in ways contingent on one another, and these situations are central in theoretical analysis of linking micro to macro levels of collective decision.

In examining collective decisions, we shall draw heavily on the individual decisions. Indeed, an organization or society does not make decisions, individual do. It might be argued that understanding how individual make decisions is sufficient to understand and improve collective decision. In this paper, we take a different view. Although individual decision is nested within important to understand, it is not sufficient to describe how a collection of agents arrives at specific decisions. How do the heterogeneous and micro-world of individual behaviors generate and self-organize themselves to the global macroscopic orders of the whole? We aim at discovering fundamental local or micro mechanisms that are sufficient to generate the macroscopic structures and collective behaviors of interest. This type of self-organization is
often referred as collective orders emerge from the bottom up. In the simplest form our model agents are formulated with their own internal models and their rules of interaction, termed as local rules [13][14]. Among many factors that may influence the performance of the interacting system, we focus on collective learning.

The overall performance depends crucially on the type of interaction as well as the heterogeneity of agent preferences. It is also common that there exists a conflict between an individual and a group. An agent wishes to maximize her own utility and a system designer wish to implement a decentralized algorithm for maximizing the whole utility of a group [18][19]. While all agents understand the outcome is inefficient, acting independently is powerless to manage this collective about what to do and also how to decide. The question of whether interacting agents self-organize efficient macroscopic orders from bottom up depends on how they interact each other. We attempt to probe deeper understanding this issue by specifying how they interact each other. We consider two types of collective action, coordination and dispersion problems. With coordination problems an agent receives a payoff if she chooses the same action as the majority does. With dispersion problem an agent receives a payoff if she chooses the same action as the minority does. Agents myopically adapt their behavior based on their idiosyncratic rule to others’ behaviors. We analyze adaptive dynamics that relate the collective with the underlying individual adaptive rules. There are many parameters to be considered, among them, we examine the heterogeneity of utility and selective interaction among heterogeneous agents.

We examine the interaction between partner choice and individual action. Agents choose their partners and also decide on a mode of behavior in interactions with these partners. We show a collection of heterogeneous agents with diverse preferences can realize the most efficient collective action by selecting their partners to interact. We also investigate how do interacting homogeneous agents with the same preference can generate an efficient collective action We show a collection of homogeneous agents evolves into heterogeneous agents with diverse preferences by reinforcing their preferences so that they can realize the most efficient and equitable collective action.

2. Two Types of Coordination Problems

There are many situations where interacting agents can benefit from coordinating their actions. Examples where coordination is important include trade alliance, the choice of compatible technologies or conventions. The design of efficient collective action from bottom up becomes to a crucial issue in many disciplines [6][10][12]. Agents face problems of sharing and distributing limited resources in an efficient way, and there are many situations where interacting agents can benefit if they take the same action. In these situations, they may lead the remaining agents to follow suit, which is defined as multiplier or bandwagon effects. These strategic situations are modeled as coordination games. Coordination usually implies the increased effort by some agents if he selects the same strategy as the majority does [2][3][7]. There is another type of an coordination problem in which an agent receives gain if he selects the same strategy as the minority does [17]. The EL Farol bar problem and its variants may provide a clean and simple example of this type [1][8]. They have to coordinate their decisions with others in order to improve their utility. The coordination problem implies that increased effort by some agents leads the remaining agents to follow suit, which gives multiplier effects. On the other hand, in the dispersion problem, agents receive payoff if they select the opposite strategy as the majority does.
We investigate these two types of coordination formulating as the coordination game and the dispersion game. A much studied class of games is the set of coordination games in which both agents gain payoffs only when they can coordinate by choosing the same action. A complementary class that has received little attention is the set of games in which both agents gain payoffs only when they can coordinate by choosing the distinct action. The latter cases have sometimes been called the battle-of-sex games or anti-coordination games. Most discussion of these games has focused on two-person games, where the coordination games and the dispersion games become the same by renaming the strategy of one of the two agents [9]. However, they differ with many agents. The dispersion problem with many agents which has been also called the minority games, generates more complex behavior of interest. The traditional game theory, however, is silent on how agents know which equilibrium should be realized when a coordination game has multiple equally plausible equilibria [16]. The game theory does not answer in explaining how agents should behave in order to overcome an inefficient equilibrium situation. Coordination problems are characterized with many equilibria, and then coordination failures occur resulting from their independent inductive processes [1]. A solution to this kind of a problem invokes the intervention of an authority that finds the social optimum and imposes the optimal behavior to agents. While such an optimal solution may easy to find, the implementation may be difficult to enforce in practical situations. We also face situations where the unique or few optimal solutions exist are rare, and we usually face infinite number of solutions.

(1) The coordination problem
As a specific situation, we consider the following coordination problems with binary choices. There is a collection of agents \( G = \{A_i; 1 \leq i \leq N\} \) with two strategies, \( S_1 \) and \( S_2 \). Given the actions of all agents, the payoff of each agent is given by

\[
U(S_i) = ap(t) - c, \quad U(S_j) = bp(t)
\]

where \( p(t) \) represent the proportion of agents to chooses \( S_1 \) at time \( t \). Each agent first gets aggregate information \( p(t) \) about all other agents’ actions, and then he decides whether to choose \( S_1 \) or \( S_2 \). Each agent is rewarded with more payoff whenever the side she chooses happen to be chosen by the majority of the agents, while agents on the minority side gets less payoff. All agents have access to public information on the record past histories. We represent the past history available at the time period \( t \) by \( p(t) \). How do agents choose their actions under this common information? Agents may behave differently because of their personal beliefs on the outcome of the next time period, which only depend on what agents do at the next time period \( t+1 \), and the past history has no direct impact on it.

(2) The dispersion problem
There is another type of interaction in which we have to utilize different methodology. We consider a competitive routing problem of networks, in which the paths from sources to destination have to be established by independent agents. For example, in the context of traffic networks, agents have to determine their route independently. In telecommunication networks,
they have to decide on what fraction of their traffic to send on each link of the network. The utility of each agent is determined what the majority does, and each agent gains utility only if she chooses the opposite route of the majority does. The utility function of an agent if she chooses $S_1$ or $S_2$ is given as follows: Given the actions of all agents, the payoff of each agent is given by

$$ U(S_1) = a(1 - p(t)), \quad U(S_2) = bp(t) $$  \hfill (2.2)

Let us analyze the structure of the dispersion problem to see what to expect. The efficiency can be measured from the average payoff of one agent over the long-time period. Consider the extreme case where only one agent take one action $S_1$, and all the other agents take the other strategy $S_2$. The lucky agent gets a reward $a$, nothing for the others. We consider another extreme situation when a proportion of agents to choose $S_1$ is $a/(a+b)$ and the rest of agents choose $S_2$. In the latter case, the most efficiency is achieved with the average payoff $ab/(a+b)$. The MG game is characterized with many solutions. It is easy to see that this game has asymmetric Nash equilibria in pure strategies in the case where a proportion of agents to choose $S_1$ is given by $a/(a+b)$. The game also presents a unique symmetric mixed strategy Nash equilibrium in which each agent chooses $S_1$ by $a/(a+b)$ and $S_2$ by $a/(a+b)$. With this mixed strategy, each agent can expect the payoff $ab/(a+b)$ on each time period. The variance is a measure of the degree of equity. The higher the variance, the higher magnitude of the inequality and the aggregate welfare is at the lower level even if the efficiency level is high.

An interesting problem is then under what circumstances will a collection of agents realizes some particular stable situations, and whether they satisfy the conditions of efficiency? Coordination problems are characterized with many equilibria, and then coordination failures occur resulting from their independent decision makings. A solution to problems of this kind invokes the intervention of an authority that finds the social optimum and imposes the optimal behavior to agents. While such an optimal solution may easy to find, the implementation may be difficult to enforce in practical situations. We also face many situations where the unique or few optimal solutions exist are rare, and we usually face infinite number of solutions.

3. **Heterogeneity in Preferences and Adaptation with Best-response**

In human societies, an essential element is that individuals differ from each other. This diversity comes into play in many instances of collective behavior. In this section, we depart from the assumption of homogeneity with respect to the payoff. In particular, we consider continuum games with infinite number of heterogeneous agents with respect to their preferences.

The coordination problem with N-persons formulated with the payoff function in (2.1) or (2.2) can be decomposed into the interaction problem between an individual and the aggregate with the payoff matrix given in Table 1 or Table 2. The heterogeneity among agents can be described by their payoff parameter $\theta$ in Table 1 or in Table 2. The payoff matrix can be replaced by a one-dimensional vector of the payoff parameter one for each agent, which allows
enormous simplification in the ensuring analysis. In our model, agents have idiosyncratic payoff parameter $\theta$. The diversity of a collection of heterogeneous agents is characterized by the distribution function of the payoff parameter. We consider a collection of heterogeneous agents $G = \{A_i: 1 \leq i \leq N\}$, and denote the number of agents with the same parameter value $\theta$ in $G$ by $n(\theta)$. We define the density of $\theta$ by $f(\theta)$, which is obtained by divided $n(\theta)$ by the total number of agent $N$, i.e.,

$$f(\theta) = n(\theta) / N.$$ (3.1)

As specific examples, we consider three distribution functions shown in Fig.1. In Fig.1(a), all agents have the same payoff parameter value of $\theta = 0.5$, and in Fig.1(b), a half of agents have the payoff parameter value of $\theta = 0$, and the rest of them have of $\theta = 1$. In Fig.1(c), we show a collection of heterogeneous agents, and the density function is given as the normal distribution.

![Distribution Functions](image)

(a) The identical density function with the one-peak  
(b) The density with the two-peaks  
(c) The normal density function
In examining collective decisions, we shall draw heavily on the individual adaptive
decisions. Within the scope of our model, we treat models in which agents make deliberate
decisions by applying rational procedures, which also guide their reasoning. In order to describe
the adaptation process at the individual level, we may have two fundamental models, global
interaction and local interaction. It is important to consider with whom an agent interacts and
how each agent decides his action depending on others’ actions. Agents may adapt based on the
aggregate information representing the current status of the whole system (global information).
In this case, each agent chooses an optimal decision based on aggregate information about how
all other agents behaved in the past. In many situations, agents are not assumed to be
knowledgeable as to correctly guess or anticipate other agents’ actions, or they are less
sophisticated and that they do not know how to calculate best replies [16]. With local adaptation
each agent is modeled to adapt to local information [2][5]. As a specific model, we consider the
lattice structure as shown in Fig.2. We arrange a collection of heterogeneous agents in the area of
50×50 (2500 agents in total) with no gap, and four corners and an edge of an area connect it with
an opposite side. The consequence of their actions also gives an effect on agents with whom not
directly linked. The hypothesis of local adaptation also reflects limited ability of agents’ parts to
receive, decide, and act based upon information they receive in the course of interaction. The
main point is that an agent's decision depends on what it knows about others.

![Fig.2: Location and the way of interaction among agents](image)

[Global adaptation with coordination problem]
Agents adopt actions that optimize their expected payoff given what they expect others to do. In
this model, agents choose the best replies to the empirical frequencies distribution of the previous
actions of the others. We obtain the adaptive rule of each agent as her best response. Let’s denote
the proportion of agents having chosen $S_1$ at time $t$ in a population by $p(t)$. An agent with
the payoff parameter $\theta$ in Table 2 calculates her expected utilities as follows:

$$
U(S_1) = p(t)(1-\theta), \quad U(S_2) = (1-p(t))\theta
$$

(3.1)

By comparing the expected utilities under $S_1$ and $S_2$, the optimal adaptive rule of an agent is
obtained as the function of the aggregate information on collective $p(t)$ and her idiosyncratic
payoff parameter $\theta$ as follows:

(i) If $p(t) \geq \theta$, choose $S_1$
(ii) If $p(t) < \theta$, choose $S_2$

(3.2)
[Local adaptation in the coordination problem]
The local adaptive rule of an agent \( A_i \) is obtained as follows. Let’s denote the proportion of the neighbors of an agent having chosen \( S_1 \) at time \( t \) by \( p_t(t) \). The optimal adaptive rule with local adaptation is obtained as follows:

(i) If \( p_t(t) \geq \theta \), choose \( S_1 \)
(ii) If \( p_t(t) < \theta \), choose \( S_2 \)

\[ (3.3) \]

[Global adaptation with the dispersion problem]
We obtain the adaptive rules of agents with asymmetric interaction. Let’s denote the proportion of agents having chosen \( S_1 \) at time \( t \) in a population by \( p(t) \). An agent with the payoff parameter \( \theta \) in Table 3 calculates her expected utilities as follows:

\[ U(S_1) = (1 - p(t))\theta, \quad U(S_2) = p(t)(1-\theta) \]

By comparing the expected utilities under \( S_1 \) and \( S_2 \), the optimal adaptive rule of an agent is obtained as the function of the aggregate information on collective \( p(t) \) and her idiosyncratic threshold \( \theta \) as follows:

(i) If \( p(t) \leq \theta \), choose \( S_1 \)
(ii) If \( p(t) > \theta \), choose \( S_2 \)

\[ (3.4) \]

[Local adaptation with the dispersion problem]
The adaptive rule with local adaptation is obtained as follows: Let’s denote the proportion of the neighbors of agent \( A \) who have chosen \( S_1 \) at time \( t \) by \( p_t(t) \). The optimal adaptive rule with local adaptation is obtained as follows:

(i) If \( p_t(t) \leq \theta \), choose \( S_1 \)
(ii) If \( p_t(t) > \theta \), choose \( S_2 \)

\[ (3.5) \]

4. Selective Interaction of Agents with Heterogeneous Preferences
The crucial concept for describing heterogeneity of agents is their preference characterized by their payoff parameter \( \theta \). Heterogeneity of preferences makes it possible to have a different type interaction, selective interaction. This is possible because that each agent has a different payoff parameter. We classify heterogeneous agents into the following two types:

Type 1: Agent with the payoff parameter \( \theta \leq 0.5 \). (An agent prefers the strategy \( S_1 \) to \( S_2 \).)
Type 2: Agent with the payoff parameter \( \theta > 0.5 \). (An agent prefers the strategy \( S_2 \) to \( S_1 \).)

We also classify interaction types into the following three types:
(1) Random interaction: Each agent has a chance to interact with neighbors of any type.
(2) Homogeneous interaction: Each agent interacts with neighbors of the same type.
(3) Heterogeneous interaction: Each agent interacts with neighbors of the opposite type.
Type1: An agent who prefer $S_1$ to $S_2$  ●Type2: An agent who prefer $S_2$ to $S_1$

(a) Random interaction    (b) Homogeneous interaction    (c) Heterogeneous interaction

Fig.2: Configurations of interaction patterns

In this section, we evaluate the collective action of interacting heterogeneous agents by its efficiency by obtaining the average payoff of the population. For a population of identical agents (Case 1), all agents gain the same payoff and equity is high, but the efficiency is moderate. In global adaptation with diverse population, efficiency is moderate. In local adaptation with random neighbors, efficiency is low, and diversity tightens gap in efficiency with the homogeneous neighbors, efficiency becomes high.

< The coordination problem >

(1) Random interaction
Case 1: $U = 0.25$,  Case 2: $U=0.5$

(2) Homogeneous interaction
Case 1: $U = 0.5$,  Case 2: $U=1$

(3) Heterogeneous interaction
Case 1: $U = 0$,  Case 2: $U=0$

< The dispersion problem >

(1) Random interaction
Case 1: $U = 0.25$,  Case 2: $U=0.5$

(2) Homogeneous interaction
Case 1: $U = 0$,  Case 2: $U=0$

(3) Heterogeneous interaction
Case 1: $U = 0.5$,  Case 2: $U=1$

We describe the selective interaction process of agents with heterogeneous preferences: An agent of Type 1 with the payoff parameter $\theta \leq 0.5$ chooses $S_1$, since such an agent prefers the strategy $S_1$ to $S_2$. On the other hand, an agent with the payoff parameter $\theta > 0.5$, chooses $S_2$, since such an agent prefers the strategy $S_2$ to $S_1$. In Fig.3, we represent the selection process of an agent. Each agent interact with neighbors by choosing her preferred strategy. If she receives the average payoff per one neighbor more than 0.5, she remains the same location, otherwise she moves to an another location in order to interact with different neighbors.

We show the simulation result in Fig.4, when each agent faces the coordination problem. At beginning we set two types agents randomly in the lattice of 50x50. After a few hundred of the repetition, a collection of heterogeneous agents could self-organize both their locations and their choice of actions so that they can realize homogeneous interaction shown in Fig.3(b), and they could archive the most efficiency.
We show the simulation result in Fig.5, when each agent faces the dispersion problem. At beginning we set two types agents randomly in the lattice. After a few hundred of the repetition, a collection of heterogeneous agents could self-organize both their locations and their choice of actions by realizing interactions shown in Fig.3(c), and they could archive the most efficiency.
5. Collective Learning of Agents with Homogeneous Preference

Game theory is typically based upon the assumption of a rational choice [9]. 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 the adaptive behavior with reinforcement learning. They collectively reinforce their preferences over possible actions. Agents will try any of two strategies, and repeat those that led to high payoffs in the past. Propensity of trying an strategy is increased according to the associated payoff. Agents tend to adopt actions that yielded a higher payoff in the past, and to avoid actions that yielded a low payoff. Payoff describes 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 [2].

In this section, we consider a collections of homogeneous agents with the payoff parameter $\theta = 0.5$ as shown in the density function by Fig.1(a). Therefore all agents have the same payoff matrix. We show a collections of homogeneous agents evolve into a collection of heterogeneous agents with the density function Fig.1(b) in order to realizing the most efficiency. We now describe the reinforcement process. Each agent repeats $T(=10)$ interactions with neighbors. If agents gain the average payoff per one neighbor for one interaction more than 0.5 by choosing $S_1$, they increase $1-\theta$ by $\Delta \theta = 0.01$, and decrease $\theta$ by $\Delta \theta = 0.01$ in her payoff matrix in Table 3. Similarly, agents who gain the average payoff more than 0.5 by choosing $S_2$, they increase $\theta$ by $\Delta \theta = 0.01$, and decreases $1-\theta$ by $\Delta \theta = 0.01$

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Reinforcement of learning the payoff-matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Coordination problem</td>
<td>(b) Dispersion problem</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Choice of agent $A$</th>
<th>Choice of other Agents</th>
<th>$S_1$</th>
<th>$S_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>$1-\theta + \Delta \theta$</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$S_2$</td>
<td>0</td>
<td>$\theta - \Delta \theta$</td>
<td></td>
</tr>
</tbody>
</table>

$(K=10, \Delta \theta = 0.01)$

In Fig.6, we represent the processes of collective reinforcement learning. Each agent interacts with neighbors by choosing her preferred strategy. If she receives the average payoff per one neighbor more than 0.5, she reinforce her payoff parameter in Table 3(a) if they face the coordination problem, and in Table 3(b) if they face the dispersion problem.
We show the simulation results in Fig.7, when each agent face the coordination problem. In Fig.7(a), the initial proportion of agents to choose S1 at beginning is set to p(0)=0.6. A collection of homogeneous agents with the same payoff parameter $\theta = 0.5$, gradually self-reinforce their payoff parameters, and after the repetition of T=600, all agents have the same preference of $\theta = 1$. In Fig.7(c), the initial proportion p(0) is set to p(0)=0.4. Homogeneous agents of the same payoff parameter $\theta = 0.5$ self-reinforce their payoff parameters, and after the T=600 repetition, all agents have the same parameter $\theta = 0$. In Fig.7(b), the initial proportion p(0) is set to p(0)=0.5, and after the T=600 learning process, a half of agents have the same payoff parameter $\theta = 1$, and the rest of agents have $\theta = 0$. In all three cases, a collection of homogeneous agents collectively reinforce their preferences and evolve into heterogeneous agents, so that the most efficient and equitable collective action can be self-organized. The result of collective learning depends on the initial situation.

Fig.7: The distributions of the payoff parameter $\theta$ after collective reinforcement learning (The coordination problems)
We show the simulation results in Fig.8, when each agent faces the dispersion problem. In Fig.8 (a), the initial proportion of agents to choose S1 at beginning is set to $p(0)=0.5$. Homogeneous agents of the same payoff parameter $\theta = 0.5$ self-reinforce their payoffs, and after the T=600 learning process, a half of agents have the same parameter $\theta = 1$, and the rest of agents have $\theta = 0$. This result implies that a collection of homogeneous agents collectively reinforce their preference levels so that the most efficient and equitable collective action can be self-organized. The performance of collective learning under the dispersion problem, however, heavily depends on the initial situation. In Fig.8(b), the initial proportion of agents to choose S1 at beginning is set to either $p(0)=0.4$ or $p(0)=0.6$. They are remain as homogeneous agents with the same payoff parameter $\theta = 0.5$ by failing to collectively reinforce their payoff parameter.

![Fig.8: The distributions of the payoff parameter $\theta$ after collective reinforcement learning (The dispersion problem)](image)

6. Conclusion
In this paper we addressed the issue of collective decisions by agents in which they have to realize both efficient and equitable utilization of limited resources. Agents normally react to aggregate of others’ decisions, and the resulting volatile collective decision is often far from being efficient. By means of experiments, we showed that the overall performance depends crucially on the types of interacting decisions as well as the heterogeneity of agents in term of their preferences. We considered two different types of interaction which are formulated as the coordination problem and the dispersion problem. We showed that the most crucial factor that considerably improves the performance of the system of interacting agents are the endogeneous selection of the partners and reinforcement of preferences at individual levels.

References