Collective decision-making algorithm for the best-of-n problem in multiple options

Masao Kubo\textsuperscript{a}, Hiroshi Sato\textsuperscript{a}, Nhuhai Phung\textsuperscript{b} and Akihiro Yamaguchi\textsuperscript{c}

\textsuperscript{a}National Defense Academy of Japan, Yokosuka, Japan; \textsuperscript{b}Military Institute of Information Technology, Ha Noi, Viet Nam; \textsuperscript{c}Department of Information and Systems Engineering, Fukuoka Institute of Technology, Fukuoka, Japan

ABSTRACT

The best-of-n problem [Valentini G, Ferrante E, Dorigo M. The best-of-n problem in robot swarms: formalization, state of the art, and novel perspectives. Frontiers in Robotics and AI. 2017;4(9):1–18.] is a collective decision-making problem in which many robots (agents) select the best option among a set of \( n \) alternatives and is focused on the field of distributed autonomous robotic systems and swarm robotics. It is desirable to develop a collective decision-making algorithm that can work even when there are a lot of social–behavioural alternatives \((n \geq 2)\) to realize an intelligent system that can solve more complicated problems. However, previous studies mainly focused on binary collective decision-making scenarios \((n = 2)\). In this paper, we propose a collective decision-making algorithm for the best-of-n problem with a large number of options by using short-term experience memory with a trial and error approach at the group level. After proposing this decision-making process, we show typical behaviour. Next, we show the convergence of this algorithm when a quadratic function is used for the bias corresponding to the individual characteristic. Next, the bias distribution proposed shows that an equilibrium point where all candidates have the same number of supports is unstable, and consensus states are a stable fixed point. Therefore, dynamics is expected to converge towards a consensus. Simulation results and mathematical analysis show that the average time required to find the best option is nearly proportional to the number of options and does not depend on the number of robots.

1. Introduction

“Flocking” [1–4] is said to be a state in which many individuals are in physical proximity. In the following, a state where individuals are flocking and they can share information is called a group. Groups can expect to solve tasks that far exceed the capabilities of individuals, but this requires the group can reach consensus appropriately. The method of achieving this differs greatly depending on the presence or absence of a leader. The following deals with situations where there is no leader. This is commonly referred to as group decision-making or decentralized decision-making problems, and much research has been done.

The collective decision-making problems are generally organized into two categories: \textit{consensus achievement}, where agents aim at making a common decision, and \textit{task allocation}, where agents allocate themselves to different tasks, with the objective to maximize the performance of the collective [5]. In addition, consensus achievement is decomposed into two classes: \textit{continuous} and \textit{discrete}. An example of the first class, continuous, is the selection of a common direction of motion by a flocking of agents [2]. LCP algorithm [3] is known well. Here, we deal with the second class, discrete, which is generally called the best-of-n (BSTn) problem. In the BSTn problem, a group of robots has to select the best option among a set of \( n \) alternatives without a leader [6].

The preconditions for BSTn dealt with in this paper are as follows. First, this paper assumes that the options are symbolized beforehand and that each robot can explicitly select. The option selected by each robot can be transferred to other robots within the group, and the number of agents who choose an option can be counted by each robot. We suppose that information about the outcome of each option, such as its goodness, value, or effectiveness, is unknown in advance, and the robot only realizes the value of this option after executing that option.

We have summarized the related works in Table 1. The methods of the best-of-n problem can be roughly classified into value-sensitive methods or value-free methods. A value-sensitive method is a distributed decision-making method derived from social insects, in short, a method of dividing a group by the number of agents proportional to the goodness of options. For example, it is assumed that there are 100 agents and two options A and B, and the goodness is 60/100
Table 1. A summarization of the methods used in related studies.

| Related studies | n | Unanimous-ness | Methods and main results |
|-----------------|---|----------------|-------------------------|
| **Value-sensitive** | | | |
| Parker et al. [7] | 2 | Message flooding | The robot has six states and a group of the robots reaches a consensus in three steps. It collectively corrects individual errors by verifying an alternative with multiple robots. The first alternative that passes to the committing phase becomes the unanimous alternative by message flooding. Experiments using actual robots are conducted. |
| Montes et al. [8] | 2 | Majority-rule [9] | Proposing a method of a shortest-path problem by introducing Latency to the majority-rule presented by Kravipinsky and Redner [9]. The supervisor creates k teams of three randomly selected robots, and members of each team unify their options using a majority rule. The rule that the robot that chooses better options updates its own state more frequently converges to a state where everyone has better options. |
| Seeley et al. [10] | 2 | Honeybee Population model (HPM) | Solving deadlock problem using cross inhibition by stop signals in a nest-site selection scenario by honeybee swarms. |
| Pais et al. [11] | 2 | HPM | Improving the decision-making capabilities by introducing sensory noise to the cross inhibition of [10]. |
| Valenti et al. [6] | 2 | Weighted voter model | Proposing a weighted voter model for a site-selection scenario. The higher the quality of the option, the more likely it is to influence the votes of other agents. |
| Reina et al. [12] | 2 | HPM | For the problem of finding the shortest path (n=2), we propose a collective decision-making model based on the home selection behavior observed in honeybee colonies [10]. The correct decision is taken in at least 90% when n = 2, the resolution R > 0.15 for every group size except very small one. |
| Reina et al. [13] | ≥ 2 | HPM | Generalizing and extending the model of [12] to the best-of-n case. The correct decision is taken in at least 75% when n ≤ 7. |
| Hasegawa et al. [14] | ≥ 2 | The supervisor selects by majority decision | Applying a sensory discrimination mechanism using yes/no agents with differential thresholds to a model for making a collective decision among multiple options. The proportion of correct choices is more than 80%. |
| **Value-free** | | | |
| Iwanaga and Namatame [15] | 2 | Bias dist. | Proposing an individual decision-making method inspired by Schelling’s critical mass [16]. |
| Wessnitzer and Melhuish [17] | 2 | Message flooding | Proposing a method of a prey-hunting scenario using majority voting. Timers for switching targets like BRT are used. Experiments using actual robots are conducted. |
| Kubo et al. [18] | 2 | Phase transition | Proposing a method of a target-enclosing scenario to switch the direction of robots using phase transition [4]. |
| Phung et al. [19] | ≥ 2 | Bias dist. (Normal dist.) | Proposing a collective decision-making model, called BRT model, by introducing the give-up term. |
| Kubo et al. [20] | ≥ 2 | Bias dist. (Normal dist.) | Proposing a collective decision-making model by using UEM agents based on the BRT model [19]. |
| **The proposed** | ≥ 2 | Bias dist. (Quadratic function) | This proposes a collective decision-making model of BRT [19] with Short-term Experience Memory (SEM) that the time required to find a suitable option can be reduced than [19]. Also, the quadratic bias function is proposed to provide the analysis of BRT model. |

agents and 40/100 agents, respectively (x/y determine that x agents out of y agents say the option is good). Since the goodness of either option is unknown in the initial state, supporters of each option are about 50% of the population. After that, by comparing with other options, supporters of A will gradually increase, and in this case, it can be expected that 60 agents will issue A. However, it is difficult for all 100 agents to reach the state of indicating A, and an additional mechanism is needed to overwhelm the minority opinion. This valuesensitive method is not covered in this paper because the algorithm for this part is highly dependent on each application.

On the other hand, the value-free method is a method to expect to lead to behaviour suitable as a group, assuming that all members select the same behaviour [15, 17, 21]. For example, Iwanaga and Namatame [15] proposed an individual decision-making method inspired by T. Schelling’s critical mass [16], and they also clarified that the group quickly converges to the state where everyone selects one of the two options.

Also, we have proposed a value-free approach, BRT model (Bias and Rising Threshold model) in which by introducing the give up term to the individual decision-making process [15] even if there are more than two options [19, 20, 22–24]. At the time of convergence, the group knows the quality of the selected option at the moment. If the quality is high enough, the selected option will be the final result of the collective decision-making. Otherwise, individuals can quit selecting this option and start selecting another option to find better options. However, these previous studies dealt with the problem that the number of options n is relatively small, and its dynamics have not been clarified.

In this paper, we propose an extended version of the BRT model [19, 20, 22], and the characteristics are shown by mathematical discussions and computer simulations. For the analysis, we introduce the quadratic bias function for action selection of BRT model is proposed. In the BRT model, each agent selects one option out of a set of n-options and notify to the other agents. The agents change their choices based on the ratio of supporters for each option and their own bias θ. Here, we have the following two states, a state of the group where all agents select the same option as an agreed interval and the other state where they do not select
the same option as a competitive interval. Ordinary, an agreed interval and a competitive interval emerge iteratively until a good option is discovered.

Generally, if the number of options is large, the competitive interval is expected to continue for a long time. It is difficult to apply the proposed model to a group of robots that operate in a dynamically changing environment if it takes time to reach an agreement. Therefore, it is an extremely important task to clarify the relationship between the number of robots \( N \) and the number of options \( n \) to reach an agreement promptly. Therefore, in the following, we analyse the requirements and the approximate time to agree on suitable options.

The major difference between this paper and the previous models [19,22–24] is that, in addition to introducing a short-term memory mechanism, we propose a design method for bias \( \theta \), which can be expected to converge to an agreement. Even in the case of \( n \geq 2 \), when using this bias \( \theta \) design method, it can be guaranteed that the agreed state is a stable fixed point and the typical competitive states are unstable points.

Also, the time required to reach an agreement is almost independent of the number of agents \( N \). Furthermore, computer experiments confirmed that robots can reach the appropriate option when the number of options was 20 or less and that the average time required to reach the agreement increased almost linearly with an increase in the number of options \( n \).

The remainder of this paper is organized as follows. In Section 2, we explain the proposed method and its characteristics. In Section 3, we analyse the convergence and the optimality of the proposed model mathematically. In Section 4, we verify the characteristics and indicate the effectiveness of the proposed method using computer experiments.

Finally, we denote assumptions in this paper. The agents can communicate with nearby agents without any mistakes and noises. The communication speed is much faster than their other action selection processes. Therefore, the distribution of the ratio of supporters for each option used for their decision-making which is global information can be obtained by the local communication instantly and correctly. This assumption has been introduced for the reason to analyse the dynamics on the whole system without any noises as same as many past works [14,25–27]. The effects of a limited communication range of agents are discussed by [28].

## 2. Proposed method

### 2.1. Problem statement

The number of options is written as \( M \) instead of \( n \) in the following to avoid confusion. Consider the system of \( N \) agents and \( M \) options. Among \( M \) options, there is 1 superior option (suitable option) and \( M−1 \) inferior options (unsuitable option). The task is to discover the superior option by all of the agents.

In this paper, the most basic case is supposed, as follows. It is assumed here that no supervisor determines what to do. After each agent selects its action, it executes the selected action immediately. However, this paper deals with the most basic case and assumes an environment in which change occurs only when everyone’s options coincide. This situation corresponds, for example, to the case of transporting a heavy load that can only be moved by the combined efforts of all agents.

In addition, this paper supposes no supervisor who judges the suitableness of the options of a group of agents. Every agent evaluates the suitableness of their options on their own judgment function. On the other hand, in this paper, we also suppose that all agents have the same judgment function. If all agents attempt to perform the same option, this option will cause a change in the environment. The agents observe this change with their own sensors and obtain the same evaluation. However, in the experiment described below, the system will tell all agents if their actions match one of the pre-selected options, as described above. We do not deal with realistic and complex decisions, such as those that require a detailed judgment of the target situation, since these decisions vary from problem to problem, but there is research on the use of BRT in complex environments, such as [29].

On the other hand, whether everyone has taken the same measure is needed for the BRT algorithm described below. We assume that each agent observes others by their sensors and communicate with each other to find out.

### 2.2. The approach

BRT evaluates an option in an agreed state by all individuals like the method of Wessnitzer and Melhuish [17] then they judge collectively whether the option is good or not. When all their options match, we suppose that each agent evaluates the option based on its own criteria and decides that it is good if the criteria are exceeded. An option that everyone decides is good is called a suitable option. Therefore, if their criteria levels are low, multiple options may become suitable options, but here the option satisfying the criteria is one out of \( M \) to simplify the problem. Also, the judgment process of goodness which this agent performs is omitted because the evaluation process strongly depends on their applications.

### 2.3. Iwanaga and Namatame model [15]

The research by Iwanaga and Namatame [15,30] was inspired by critical mass created by Schelling [16]. As shown in Figure 1, they stated that individual decision-making depends not only on personal philosophy and
Figure 1. Role of social skin in individual decision-making [30].

personal preferences but also on the atmosphere of the whole group \( p(t) \). Each individual shall take the position of agreeing or disagreeing with an option and has a personal attribute bias value \( \theta_i \) (Figure 1 A), which is used in this decision-making process (Figure 1 C). Each individual observes the ratio of the members who agree with the option \( p(t) \) (Figure 1 B).

\[
\begin{align*}
    p(t) \geq \theta_i : & \text{Agreement} \\
    p(t) < \theta_i : & \text{Disagreement}
\end{align*}
\]  

As shown in Equation (1), an individual \( i \) agrees with an option if the percentage of agents who agree with this option \( p(t) \) (the agreement ratio) is higher than the bias value \( \theta_i \) at time \( t \). After all individuals have made their decisions, agreement ratio \( p(t+1) \) at time \( t+1 \) changes (Figure 1 D), so the individual makes its decision again (Figure 1 C).

Iwanaga and Namatame [15] proposed a method to analyse the dynamics of agreement ratio \( p(t) \) in this decision-making process. If the distribution of \( \theta \) is determined as shown in Figure 2, the cumulative distribution function \( F(\theta) \) forms an S-shaped curve as shown in Figure 3. If the agent makes a decision according to Equation (1), ratio \( p(t) \) at time \( t \) follows the following equation:

\[
F(\theta) = \sum_{\theta_i \leq \theta} n(\theta_i)
\]  

\[
p(t+1) = F(p(t))
\]  

\[
p^* = F(p^*)
\]

The solutions of Equation (4) \( p^* \) are called fixed points. There are two types of fixed points: stable point and unstable point. If bias value \( \theta_i \) is distributed in a bell shape, as shown in Figure 3, one unstable point \( (E_2) \) and two stable points \( (E_1 \text{ and } E_3) \) are generated. In other words, the behaviour of a group can be classified into the following three patterns.

(1) The case of \( p(t) = 0.5 \) corresponds to \( E_2 \) in Figure 3. The state of the group does not change because \( F(p(t)) = p(t) \). However, if \( p(t) \) changes for some reason, the state of the group will transit to one of the following two states.

(2) The case of \( p(t) > 0.5 \) \( (E_2 < p(t) \leq 1) \). In this case, because \( p(t+1) = F(p(t)) > p(t) \), the state of the group transits rapidly to the state where all agents agree with the option \( (p(t) = 1) \). In Figure 3, because of \( p(0) < p(1) < p(2) < p(3) \simeq 1 \), the agreed state is reached in approximately three units of time from the almost equilibrium state.

(3) The case of \( p(t) < 0.5 \) \( (0 \leq p(t) < E_2) \). Because \( p(t+1) = F(p(t)) < p(t) \), the state of group transits rapidly to the state where all agents disagree with the option \( (p(t) = 0) \).

If the cumulative distribution of bias values \( \theta_i \) is S shape, the number of iterations until the convergence does not depend on the number of individuals. Thus it is possible to expect that the convergence is achieved promptly even with a large number of individuals. However, in [15] there are no mathematical discussions of the convergence, and the scalability of the number of options \( M > 2 \) has not been considered.
2.4. BRT model with short-term experience memory (SEM)

Based on the above, in the following, we propose an extended model of our study in [19] in which agreement can be promptly reached despite a large number of options \((M > 2)\) and call it the BRT model.

Here, we assumed that there are \(N\) agents \(P_1, \ldots, P_i, \ldots, P_N\) \((1 \leq i \leq N)\). \(A = \{a_1, \ldots, a_j, \ldots, a_M\}\) is a set of options with \(M \geq 2\) \((1 \leq j \leq M)\). \(M\) is the number of options. \(A_i(t) \in A\) is the option of agent \(P_i\) at time \(t\). Agent \(P_i\) has a bias \(\theta_i\) \((0 < \theta_i < 1)\). Moreover, \(n_i(a_j)\) is the number of neighbours selecting option \(a_j\). When every agent chooses the same option, the group agrees to that option. This corresponds to, for example, in [17], the entire population heading to one of multiple feeding grounds.

If all agents select the same option, we suppose that the group agrees with that option. If the state of the group at this time is suitable for the environment, etc., this consensus option is called a suitable option. Here, we set the set of these suitable options as \(A_{goal}(A_{goal} \subset A)\). The agent does not know \(A_{goal}\) in advance but only when all agents have agreed, it is understood that \(A_{goal}\) includes that agreed upon option.

Figure 4 shows the decision-making flow of agent \(P_i\) based on BRT model with short-term experience memory. The agent \(P_i\) decides its option \(A_i(t + 1)\) at time \(t\) as follows:

\[
A_i(t + 1) = \begin{cases} 
A_i(t) & \text{if } n_i(A_i(t))/N_i \geq \theta_i + \tau \cdot c_i(t) \cdot (t - t_{last}(t)), \\
\text{Prob}(a_j) & \text{otherwise},
\end{cases}
\]

Equation (5)

\[
1 - b_j = \sum_{j \in \{1, \ldots, M\}} (1 - b_j),
\]

\[
n_i(A_i(t))/N_i \geq \theta_i + \tau \cdot c_i(t) \cdot (t - t_{last}(t))
\]

Equation (6)

\[
A_i(t + 1) = \begin{cases} 
A_i(t) & \text{if } n_i(A_i(t))/N_i \geq \theta_i + \tau \cdot c_i(t) \cdot (t - t_{last}(t)), \\
\text{Prob}(a_j) & \text{otherwise},
\end{cases}
\]

Equation (7)

\[
\tau = \text{constant representing the prediction rate of the increase in the number of supporters.}
\]

\[
t_{last}(t) = \begin{cases} 
t & A_i(t) \neq A_i(t - 1), \\
t_{last}(t - 1) & \text{otherwise}
\end{cases}
\]

Equation (8)

\[
c_i(t) \text{ is a function that is equal to 0 when consensus option is a suitable option and is defined as follows: }
\]

\[
c_i(t) = \begin{cases} 
0 & \text{agreement}(t) = \text{true } \land \text{ } A_i(t) = a_x \in A_{goal} \\
1 & \text{otherwise}
\end{cases}
\]

Equation (9)

\[
\text{agreement}(t) = \text{agreement}(t) = \begin{cases} 
\text{true } & \forall i, \exists j, A_i(t) = a_j \\
\text{false } & \text{otherwise}
\end{cases}
\]

Equation (10)

\[
\text{If the ratio of agents who select the same option is lower than } \theta_i + \tau \cdot c_i(t) \cdot (t - t_{last}(t)), \text{ the agent selects a new option } a_i (a_i \neq A_i(t)) \text{ with probability } \text{Prob}(a_i). \text{ The second term on the right side of Equation (6) is a term with a value increasing over time, called BRT term. We will explain in detail the BRT term in the next section.}
\]

Agent \(P_i\) has a list \(B_i = \{b_j : 1 \leq j \leq M\}\) \((0 \leq b_j \leq 1)\) to memorize options unfit. At the initial time, \(B_i\) is set as follows:

\[
b_j = 0, \quad j \in \{1, \ldots, M\}
\]

Equation (11)

At time \(t\), if all agents agree with option \(a_u (a_u \in A, \text{ i.e. } u \in \{1, \ldots, M\})\), i.e. \(\text{agreement}(t) = \text{true}\), but this is not suitable, \(B_i\) is updated as follows:

\[
b_j = \begin{cases} 
b_j \cdot \beta & \forall j \neq u \\
1 & \text{otherwise}
\end{cases}
\]

Equation (12)

The most recently unsuitable option stored in is set \(b_u = 1\). The memories of other options are forgotten.
little by little. $\beta$ $(0 \leq \beta < 1)$ is a constant representing good of the memory. In this paper, we adopt $\beta = 0.2$.

In previous research [19], if the condition equation (6) is not satisfied, then an option other than $A_i(t)$ is stochastically chosen and becomes $A_i(t + 1) \in A \setminus A_i(t)$. It may repeatedly select the option that was not suitable a while ago. In this study, we introduce $B$ to memorize options that were collectively agreed upon but not suitable to improve the efficiency of the suitable option discovery. Because $\beta$ is small, this short-term memory is limited to be able to remember the very recent option.

### 2.5. Proposed distribution of bias $\theta_i$

In previous research [19,22], bias value $\theta_i$ has been generated using the bell-shaped curve based on a Gaussian distribution. In this paper, we adopt a distribution by a quadratic function which makes the mathematical discussions more easily.

We set $\theta_i$ as follows:

$$
n(\theta_i) = \begin{cases} 
0 & \theta_i \leq 0 \\
k_1 N \theta_i^2 & 0 < \theta_i \leq \frac{1}{M} \\
k_1 N \left( \theta_i - \frac{2}{M} \right)^2 & \frac{1}{M} < \theta_i < \frac{2}{M} \\
0 & \frac{2}{M} \leq \theta_i
\end{cases}$$

where $N$ is the number of agents, $n(\theta_i)$ represents the number of agents taking $\theta_i$, $\mu = 1/M$ is the expectation of the distribution, and $k_1 = \frac{3M^3}{2}$ is a normalization term. Thus we derive the following equation:

$$
\frac{\int_0^1 n(\theta_i) \, d\theta_i}{N} = 1
$$

Figure 5 illustrates three examples of the distributions of $\theta$ determined by Equation (13). As can be seen, the shape of the bias distribution is a peak determined by the number of options $M$ only.

### 2.6. Behaviour and characteristics of the BRT model up to the agreement

In the following, the behaviour of BRT will explain separately for (1) reaching an agreement and (2) finding a suitable agreement option.

Here, we explain the function of the BRT term of Equations (5) and (6) to realize the best option discovery ability. Agents repeat the same decision-making mechanism, but apparently, group behaviour can be divided into three phases: the competitive phase, the evaluation phase, and the revoting phase.

First, Figure 6 shows the typical behaviour for the agreement of a group of the proposed agents. The behaviour consists of three phases: competitive phase, evaluation phase, and revoting phase. The vertical axis represents the ratio of agents selecting each option and the horizontal axis represents the elapsed time step. At the initial step, each agent randomly chooses its option.

Next, agent $P_i$ selects its behaviour according to Equation (5). It decides whether to continue or quit the current option $A_i(t)$ with its personal attribute ($\theta_i$) and surrounding atmosphere ($n_i(A_i(t))/N$) and informs its decision to the whole. As time passes, a certain option is selected by all and becomes an agreement option (option $a_1$ in this figure).

We call this the end of competitive phase as Figure 6 shows and $T_v$ is the time required for the competitive phase. Next, since it is necessary to evaluate whether this agreement option is suitable for the group, this agreement option is maintained for a while even after a lapse of time. This is called evaluation phase, and $T_a$ is the time during the evaluation phase.

If agents accept the current option as a suitable option for them, the best-of-$n$ procedure is finished successfully. On the other hand, if the current agreement option is unsuitable for them, the agents switch to other options and a Revoting phase will start. By repeating this process, we can expect to realize the ability to discover the suitable option among many options.

![Figure 5. Examples of bias $\theta$ distribution.](image)

![Figure 6. Three phases of agreeing with a collective option.](image)
Next, we explain how BRT term can make an agreed state. If Equation (6) is satisfied, agent $P_i$ continues to select the current option $A_i(t)$. By the time elapsed, the value of BRT term on the right side of Equation (6) increases by $\tau$. When the value of the right side of Equation (6) exceeds the left value $(n_i(A_i(t))/N)$, the agent $P_i$ stops selecting its current option and selects another option as $A_i(t+1)$. Repeating this, $(n_i(A_i(t))/N)$ becomes 1 and the whole group will select only one option as its agreement option.

Figure 7 shows a simple example of this collective decision-making process. $A = \{a_1, a_2\}$ is a set of agent options with $M = 2$. There are $N$ agents $P_1, \ldots, P_N$ with their bias values $\theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \ldots, \theta_i, \ldots, \theta_N\}$. In the figure, only $P_1, P_2, P_3, P_4$ are represented for simplicity. The height of the blue vertical bar represents the bias value. At initial time, we assume that the agents selecting option $a_1$ are $P_1, P_2, \ldots$, and the agents selecting option $a_2$ are $P_3, P_4, \ldots, n(a_i)/N$ is the ratio of agents selecting the option $a_i$. As shown in Figure 7(B), at $t = k$, $(\tau \cdot k)$ is added to the bias of each agent $\theta_i$ (Equation 6). At this time, since the value of $(\theta_4 + \tau \cdot k)$ of agent $P_4$ exceeds the value of $n(a_2)/N$, at next time step $t = k + 1$ Figure 7(C), agent $P_4$ stops selecting option $a_2$ and starts selecting option $a_1$. Thus $n(a_1)/N$ increases while $n(a_2)/N$ decreases accordingly. In this way, as the time elapsed the group will collectively select option $a_1$.

From previous experiments, it is known that the behaviour changes considerably according to $\tau$ and the number of options $M$. In the case where $M$ is large, the group needs more time to compete for an agreement. In the case of small $\tau$, its voting is done gradually, agents will take much time to reach an agreed state. If $\tau$ is large, the agents may give up selecting the option that they selected at the initial time then it is difficult to reach an agreement.

Namely, in the case of an agent with $\theta = 0$, if no agreement is reached within

$$\frac{1}{\tau}$$

steps from Equation (5), the support will be changed to another option. Therefore, it is expected that it will be difficult to reach an agreement if $\tau$ is not small enough.

2.7. Typical behaviour to discover the suitable option

This section describes the revoting phase and until the behaviour of discovering the suitable option. If the selected agreement option is not suitable $A_i(t) \neq A_{goal}$, the value of $c_1(t)$ will remain at 1 (Equation 9) and all the agents will keep selecting to evaluate this option for a while. As time passes, the right side of Equation (6) continues to increase steadily. Since each agent has its own $\theta_i$, the timing of reaching 1 also varies. An agent whose value on the right side of Equation (6) exceeds 1 will finish selecting the current option and select another option. This will trigger to revote to converge another option. By repeating this process, a group can switch the agreement option until the suitable option is discovered.

Here, by the computer simulations we show the typical behaviour until discovering a suitable option. A suitable option is discovered by a repeat of the agreement process in the last section.

We set the number of agents $N = 1000$, the number of options $M = 3$ (i.e. $\mu = 1/3$), $\tau = 0.003$. Two typical initial conditions are considered: (i) each agent selects its option randomly and (ii) all agents select an unsuitable option at the initial time. We assume that $A_{goal} = \{a_3\}$.

Figure 7. An example of time progress in the decision-making process.
Figure 8. Starting randomly. Transition examples to discover a suitable option. The horizontal axis is time and the vertical axis is $n(A_i), i = 1, 2, 3, N = 1000, \tau = 0.003$.

Figure 8 shows an example behaviour in the first case. The horizontal axis is time and the vertical axis is $n(A_i), i = 1, 2, 3$. All the agents selected and agreed with option $a_2$ approximately at time step 10. Since $a_2$ is not the suitable option, the second term of Equation (6) always increases. As such, everyone changed its opinion and agreed with another option, $a_1$, at time step round 160, option $a_2$ at time step round 310 again, and option $a_1$ at time step round 470. They finally discovered the suitable option $a_3$ at time step round 620.

Figure 9 shows an example in the second case. The group agreed on option $a_1$ in the initial state, and then it switched its option at time step round 160, 310, 470 and discovered the suitable option at time step round 620.

From the above, in both of these cases, we found that the group can make the collective decision and discover the suitable option using the proposed method even when the number of options is more than 2.

2.8. The advantage of the SEM mechanism

Figure 10 shows the cumulative probability distribution of the time it takes to find the suitable option when $\beta$ in Equation (12) is varied from 0.0 to 0.8. The number of options $M=5$, the number of agents $N=30$. At least 10,000 trials were conducted for each $\beta$. The line of $\beta = 0.0$ (dotted line) corresponds to the conventional method [19], and the lines of $\beta > 0.0$ (solid line) correspond to the BRT with SEM. In both cases, the CDF became a staircase depending on the number of times the agreement was made. On the other hand, as a result of this experiment, we found that the time required for discovery was shorter for the model with SEM as the number of agreements increased.

In the conventional method ($\beta =0.0$), there are more than 5% of trials that require more than 3500 steps before everyone agrees on the suitable option. This time is unavoidable if the environment is highly dynamic, where the suitable option at the moment is different.
from the one agreed upon in the previous round. However, if the environment is not such a highly dynamic one, it is reasonable not to choose an option that was judged to be unsuitable several times before. We have found that by adjusting $\beta$, the SEM mechanism can significantly reduce the time for discovery.

3. Mathematical analysis

In this chapter, we analyse the convergence of the proposed model and show the stability of each agreed state of any $M$ and the instability of the equilibrium states in which all options are equally supported. Furthermore, using the case of $M = 3$, we show that there are other fixed points when $M > 2$ while this point is also unstable. From this, it expects that the model will converge to a consensus state over time while this model will not be caught at least the state where all options are equally supported. Next, we estimate the time to find a suitable option and show that this time does not depend on the number of agents.

3.1. The stability of the agreed states and the equilibrium state in time evolution dynamics of agent ratio

3.1.1. The agreed state and the equilibrium states in time evolution dynamics of agent ratio

Here, we discuss the time evolution dynamics of the ratio of agents who select each option and the stability of the state in which agents agree to a specific option. In this section, $\tau = 0$ is set for simplicity. Let $x_j(t)$ be the ratio of agents selecting option $a_j$ in step $t$.

$$x_j(t) = \frac{n(a_j)}{N}, \quad (j = 1, \ldots, M - 1) \quad (16)$$

Since the sum of the agents’ ratios is 1, the state of the discrete dynamical system that describes the time evolution of $x_j(t)$ is represented by the $M - 1$ dimension vector $X(t)$.

$$X(t) = (x_1(t), x_2(t), \ldots, x_{M-1}(t)) \quad (18)$$

Let $G_j$ be the mapping that determines $x_j(t + 1)$ from $X(t)$ as follows:

$$x_j(t + 1) = G_j(X(t)), \quad (j = 1, \ldots, M - 1) \quad (19)$$

The time evolution mapping $G_j$ is defined as the following equation:

$$G_j(X(t)) = x_j(t)F(x_j(t)) + \frac{1}{M - 1} \left( \sum_{1 \leq k < M, k \neq j} (x_k(t) \left( 1 - F(x_k(t)) \right) \right), \quad (j = 1, \ldots, M - 1) \quad (20)$$

Here, $F$ is the cumulative distribution function of the parameter $\theta$ of each agent. In Equation (20), the first term on the right side corresponds to the ratio of agents who also select option $a_j$ in the next step among the agents who select the option $a_j$. The second term corresponds to the ratio of agents who change to the option $a_j$ in the next step among agents who have selected the other than $a_j$. The sum of these two terms is the ratio of agents who select the option $a_j$ in the next step.

Regarding the first term, $F(x_j(t))$ is the ratio of agents whose $\theta$ parameter is smaller than $x_j(t)$ and it includes the ratio of agents who select the option other than
Therefore, we can obtain the ratio of agents who select the option $a_j$ and whose parameter $\theta$ is smaller than $x_j(t)$ by multiplying $F(x_j(t))$ by $x_j(t)$. This is the percentage of agents who do not change the option $a_j$ in the next step. For the second term, by multiplying $1 - F(x_k(t))$ by $x_k(t)$, we can also obtain the ratio of agents who changes the selecting option from $a_k$ in the next step, since $1 - F(x_k(t))$ is the ratio of agents whose parameter $\theta$ is greater than $x_k(t)$. Dividing by $M - 1$ gives the ratio of agents who change from the option $a_k$ to $a_j$, since there are $M - 1$ possible options to change. By summing this ratio for the options other than $a_j$, the ratio of agents who change to option $a_j$ from the other options in the next step can be obtained.

The state $X$ is given by the following vector:

$$X(t + 1) = G(X(t))$$  \hspace{1cm} (22)

Let $X^*_j$ be the agreed state in which all agents agree to the option $a_j (j = 1, 2, \ldots, M)$. The agreed state $X^*_j (j = 1, 2, \ldots, M - 1)$ is defined as a vector with the $j$th dimension being 1 and the other dimensions being 0. For example, the agreed state $X^*_1$ is given by the following vector:

$$X^*_1 = (1, 0, \ldots, 0)$$ \hspace{1cm} (23)

The state $X^*_M$ agreed to $a_M$ has all dimensions of 0 and is given by the following vector:

$$X^*_M = (0, 0, \ldots, 0)$$ \hspace{1cm} (24)

Next, let $X^*_E$ be the equilibrium state in which all options are selected with the same ratio $M^{-1}$. The equilibrium state $X^*_E$ is given by the following vector:

$$X^*_E = (M^{-1}, M^{-1}, \ldots, M^{-1})$$ \hspace{1cm} (25)

It is trivial that the agreed states $X^*_1, X^*_2, \ldots, X^*_M$ are equivalent in stability due to the symmetry of the dynamics. Therefore, in the next section, we analyse the stability of the agreed states $X^*_1$ and $X^*_M$, and the equilibrium state $X^*_E$.

### 3.1.2. Stability analysis of the agreed states and the equilibrium state

First, we show that $X^*_1$, $X^*_M$, and $X^*_E$ are fixed points of the dynamics of Equation (22).

$$G_1(X^*_1) = 1 \times F(1) + \frac{1}{M - 1}$$

$$G_1(X^*_1) = \left( \sum_{1 \leq k \leq M, k \neq j} (0 \times (1 - F(0))) \right) = F(1) = 1$$ \hspace{1cm} (26)

$$G_j(X^*_j) = \frac{1}{M - 1} (x_j(t)(1 - F(x_j(t)))) = \frac{1}{M - 1} (1 \times (1 - F(1))) = 0 \hspace{1cm} (j = 2, \ldots, M - 1)$$ \hspace{1cm} (27)

$$G_j(X^*_M) = \frac{1}{M - 1} (1 \times (1 - F(1))) = 0 \hspace{1cm} (j = 1, \ldots, M - 1)$$ \hspace{1cm} (28)

$$G_j(X^*_E) = M^{-1} F(M^{-1}) + \frac{1}{M - 1} \left( \sum_{1 \leq k \leq M, k \neq j} (M^{-1}(1 - F(M^{-1}))) \right) = M^{-1} F(M^{-1}) + M^{-1} (1 - F(M^{-1})) = M^{-1} \hspace{1cm} (j = 1, \ldots, M - 1)$$ \hspace{1cm} (29)

Here, the following equations are used.

$$F(0) = 0, \hspace{0.5cm} F(1) = 1$$ \hspace{1cm} (30)

From the above equations, the agreed state $X^*_j, (j = 1, \ldots, M)$ and the equilibrium state $X^*_E$ are fixed points of the mapping $G$.

Next, we show the Jacobian matrix $T(X)$ of $G$.

$$T(X) = \frac{\partial G(X)}{\partial X} = \left( \begin{array}{ccc} \frac{\partial G_1(X)}{\partial x_1} & \cdots & \frac{\partial G_1(X)}{\partial x_{M-1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial G_{M-1}(X)}{\partial x_1} & \cdots & \frac{\partial G_{M-1}(X)}{\partial x_{M-1}} \end{array} \right)$$ \hspace{1cm} (31)

$$\frac{\partial G_j(X)}{\partial x_j} = \frac{\partial}{\partial x_j} \left( x_j(t) F(x_j(t)) + \frac{1}{M - 1} \right) \left( \sum_{1 \leq k \leq M, k \neq j} (x_k(t)(1 - F(x_k(t)))) \right)$$ \hspace{1cm} (32)

Since Equation (22) is the dynamics of discrete-time, the fixed point $X^*_1$ is stable if the absolute value of each eigenvalue of $T(X^*_1)$ is less than 1. Similarly, if the absolute value of each eigenvalue of $T(X^*_M)$ is less than 1, then the fixed point $X^*_M$ is also stable. On the other hand, the fixed point $X^*_E$ is unstable, if there are one or more eigenvalues of $T(X^*_E)$ which absolute value is greater than 1. As will be described later, the eigenvalues are real numbers in these three cases.
Equation (32) is calculated as
\[
\frac{\partial G_j(X)}{\partial x_l} = \frac{\partial G_j(X)}{\partial x_l} = F(x_l) + x_l F'(x_l) + \frac{1}{M-1} (F(x_M) - 1 + x_M F'(x_M)) \tag{33}
\]
for \(l = j\), and
\[
\frac{\partial G_j(X)}{\partial x_l} = \frac{\partial}{\partial x_l} \left( \frac{1}{M-1} ((x_l(1 - F(x_l))) + (x_M(1 - F(x_M))) \right) = \frac{1}{M-1} (F(x_l) - x_l F'(x_l) + F(x_M) + x_M F'(x_M)) \tag{34}
\]
for \(l \neq j\).

Substituting \(X_1^*\) for \(X\) in Equations (33) and (34), we obtain \(T(X_1^*)\) as the following matrix:

\[
T(X_1^*) = \begin{pmatrix}
F'(1) + \frac{M-2}{M-1} & 0 & \cdots & 0 \\
\frac{1}{M-1} (-F'(1) - 1) & -1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{M-1} (-F'(1) - 1) & 0 & \cdots & 0 \\
\end{pmatrix}
\]

Since \(T(X_1^*)\) is a lower triangular matrix, its eigenvalues are diagonal elements.

\[
\lambda_1 = F'(1) + \frac{M-2}{M-1} \quad (M \geq 2)
\]

\[
\lambda_j = -\frac{1}{M-1} \quad (M \geq 3, j = 2, \ldots, M-1)
\]

If the absolute values of these eigenvalues are less than 1, the fixed point \(X_1^*\) is stable. Since \(0 \leq F'(x)\) by definition, we obtain the stability condition of the agreed state \(X_1^*, \ldots, X_M^*\) as the following inequality:

\[
0 \leq F'(1) \leq \frac{1}{M-1} \tag{38}
\]

\(T(X_M^*)\) is also calculated by substituting \(X_M^*\) for \(X\) in Equations (33) and (34).

The eigenvalues of \(T(X_M^*)\) are equal to the eigenvalues of \(T(X_1^*)\) (Equations 36 and 37). Therefore, the agreed states \(X_1^*, X_2^*, \ldots, X_M^*\) are stable under the same condition shown in Equation (38).

Next, we calculate \(T(X_E^*)\) by substituting \(X_E^*\) for \(X\) in Equations (33) and (34).

\[
\frac{\partial G_j(X_E^*)}{\partial x_l} = \frac{1}{M-1} (MF(M^{-1}) - 1 + F'(M^{-1})) \quad (j = 1, 2, \ldots, M-1) \tag{40}
\]

\[
\frac{\partial G_j(X_E^*)}{\partial x_l} = \frac{1}{M-1} (-F'(M^{-1}) - M^{-1} F'(M^{-1})) + F(M^{-1}) + M^{-1} F'(M^{-1}) = 0 \quad (l \neq j, l, j = 1, 2, \ldots, M-1) \tag{41}
\]

Then we obtain \(T(X_E^*)\) as the following matrix:

\[
T(X_E^*) = \begin{pmatrix}
\frac{1}{M-1} (MF(M^{-1}) - 1 + F'(M^{-1})) & 0 & \cdots & 0 \\
0 & \ddots & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & \frac{1}{M-1} (MF(M^{-1}) - 1 + F'(M^{-1})) \\
\end{pmatrix}
\]

Since \(T(X_E^*)\) is a diagonal matrix, all eigenvalues are equal and given by Equation (40). From \(F(M^{-1}) \geq 0\) and \(F(M^{-1}) \geq 0\), the equilibrium state \(X_E^*\) becomes unstable when the following inequality is satisfied:

\[
\frac{1}{M-1} (MF(M^{-1}) - 1 + F'(M^{-1})) > 1 \tag{43}
\]

Simplifying this inequality, we obtain the instability condition of \(X_E^*\) as follows:

\[
F'(M^{-1}) > M(1 - F(M^{-1})) \tag{44}
\]
3.1.3. Characteristics of the agreed states and the equilibrium state for any $M$

The proposed bias distribution Equation (13) is as follows:

$$f(\theta) = \begin{cases} 0 & (\theta \leq 0) \\ \frac{3M^3}{2} \theta^2 & (0 < \theta \leq \frac{1}{M}) \\ \frac{3M^3}{2} \left( \theta - \frac{2}{M} \right)^2 & \left( \frac{1}{M} < \theta \leq \frac{2}{M} \right) \\ 0 & (\frac{2}{M} < \theta) \end{cases} \quad (45)$$

For the given bias distribution $f(\theta)$, $F(x)$ and $F'(x)$ are defined as follows:

$$F(x) = \int_0^x f(\theta) \, d\theta \quad (46)$$
$$F'(x) = f(x) \quad (47)$$

The values of $F(x)$ and $F'(x)$ can be calculated as follows:

$$F(M^{-1}) = \int_0^{M^{-1}} \frac{3M^3}{2} \theta^2 \, d\theta = \left[ \frac{M^3}{2} \theta^3 \right]_0^{M^{-1}} = \frac{1}{2} \quad (48)$$

$$F'(M^{-1}) = f(M^{-1}) = \frac{3M}{2} \quad (49)$$
$$F'(1) = f(1) = 0 \quad (50)$$

From Equation (50), the stability condition Equation (38) of the agreed states is satisfied as follows:

$$0 \leq F'(1) = 0 < \frac{1}{M-1} \quad (51)$$

From Equations (48) and (49), the instability condition (Equation 44) of the equilibrium state is also satisfied as follows:

$$F'(M^{-1}) = \frac{3M}{2} > M(1 - F(M^{-1})) = \frac{M}{2} \quad (52)$$

Therefore, the agreed states are stable and equilibrium state is unstable for the proposed bias distribution $f(\theta)$ shown in Equation (45).

3.1.4. Characteristics of the agreed states and the equilibrium state for $M = 2$

For the case of $M = 2$, the time evolution dynamics of the agent ratio is given by the following one-dimensional mapping:

$$G_1(x) = xF(x) + (1-x)(1 - F(1-x)) \quad (53)$$

Here, $G_1(x)$ has the following symmetry for any given $F(x)$.

$$G_1(x) = 1 - G_1(1-x) \quad (54)$$

If $F(x)$ has similar symmetry,

$$F(x) = 1 - F(1-x), \quad (55)$$
then

$$G_1(x) = F(x). \quad (56)$$

Therefore, for the case of $M = 2$, if the cumulative distribution function has the symmetry shown in Equation (55), the time evolution dynamics is given by the cumulative distribution function.

The stability condition of the agreed state (Equation 38) is

$$0 \leq F'(1) < 1. \quad (57)$$

The instability condition of the equilibrium state (Equation 44) is

$$F'(2^{-1}) > 2 - 2F(2^{-1}). \quad (58)$$

If $F(x)$ has the symmetry of Equation (55), the instability condition (Equation 58) becomes

$$F'(2^{-1}) > 1. \quad (59)$$

since $F(2^{-1}) = 2^{-1}$.

The proposed bias distribution (Equation 13) for $M = 2$ is as follows:

$$f(\theta) = \begin{cases} 0 & (\theta \leq 0) \\ 12\theta^2 & (0 < \theta \leq \frac{1}{2}) \\ 12(\theta - 1)^2 & \left( \frac{1}{2} < \theta \leq 1 \right) \\ 0 & (1 < \theta) \end{cases} \quad (60)$$

In this case, the stability condition of the agreed state (Equation 57) and the instability condition of the equilibrium state (Equation 59) are satisfied, since

$$F'(1) = f(1) = 0 \quad (61)$$
and

$$F'(2^{-1}) = f(2^{-1}) = 1 \left( \frac{1}{2} \right)^2 = 3. \quad (62)$$

3.1.5. Characteristics of the agreed states and the equilibrium state for $M = 3$

For the case of $M = 3$, the time evolution dynamics of the agent ratio is given by the next two-dimensional mapping.

$$\begin{cases} G_1(x_1, x_2) = x_1F(x_1) + \frac{1}{2}(x_2(1 - F(x_2))) \\
(1 - x_1 - x_2)(1 - F(1-x_1 - x_2)) \end{cases} \quad (63)$$
$$\begin{cases} G_2(x_1, x_2) = x_2F(x_2) + \frac{1}{2}(x_1(1 - F(x_1))) \\
(1 - x_1 - x_2)(1 - F(1-x_1 - x_2)) \end{cases} \quad (63)$$

We can find a fixed point other than the agreed states and the equilibrium state by solving the following
The intersection points of this intersection line and the curved surface is green plane is $x = 2$ when $x > 0.5$ other than the fixed point (the equilibrium point) of $x_1 = x_2 = 1/3$.

Figure 12 shows the graph of $G_1(x, x) - x$. Here, the plot range is $0 \leq x \leq 0.5$, since $x_3 = 1 - x_1 - x_2$ is negative when $x > 0.5$.

For the case of $x_1 = x_2 = x \in [1/3, 1/2]$ and $x_3 = 1 - 2x \in [0, 1/3]$, expanding Equation (64) by using the proposed bias distribution Equation (45), the following quartic equation is obtained.

$$(3x - 1)(135x^3 - 225x^2 + 129x - 25) = 0$$

Equation (65) has two real number solutions and two complex number solutions, and the real number solution other than $x = 1/3$ is as follows:

$$x^* = \frac{1}{9} \left( 5 + \frac{-25 + 3\sqrt{105}}{5^{2/3}} \right)^{1/3} = 0.4784402 \ldots$$

The Jacobian matrix equation (31) for $M = 3$ can be obtained as follows:

$$T(x_1, x_2) = \begin{pmatrix} g_{1,1}(x_1, x_2) & g_{1,2}(x_1, x_2) \\ g_{2,1}(x_1, x_2) & g_{2,2}(x_1, x_2) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\begin{align*}
g_{1,1}(x_1, x_2) &= \frac{\partial G_1(x_1, x_2)}{\partial x_1} \\
&= F(x_1) + x_1 F'(x_1) + \frac{\beta(x_1, x_2)}{2} \\
&= 1 - F(x_2) - x_1 F'(x_2) + \frac{\beta(x_1, x_2)}{2} \\
g_{1,2}(x_1, x_2) &= \frac{\partial G_1(x_1, x_2)}{\partial x_2} \\
&= 1 - F(x_1) - x_2 F'(x_1) + \frac{\beta(x_1, x_2)}{2} \\
g_{2,1}(x_1, x_2) &= \frac{\partial G_2(x_1, x_2)}{\partial x_1} \\
&= 1 - F(x_1) - x_1 F'(x_1) + \frac{\beta(x_1, x_2)}{2} \\
g_{2,2}(x_1, x_2) &= \frac{\partial G_2(x_1, x_2)}{\partial x_2} \\
&= 1 - F(x_2) - x_2 F'(x_2) + \frac{\beta(x_1, x_2)}{2}
\end{align*}$$

$$\beta(x_1, x_2) = \frac{d(x_M F(x_M))}{dx} = F(1 - x_1 - x_2) - 1 + (1 - x_1 - x_2) F'(1 - x_1 - x_2)$$

When $x_1 = x_2 = x$, the eigenvalue of the Jacobian matrix of Equation (67) are as follows.

$$\lambda_1 = -1 + 3\alpha(x)$$

$$\lambda_2 = -1 + \alpha(x) + 2\beta(x, x)$$

$$\alpha(x) = \frac{d(x_F(x))}{dx} = F(x) + x F'(x)$$

Therefore when $M = 3$, the conditions that the absolute values of the eigenvalues $\lambda_1$ and $\lambda_2$ are greater than 1 is obtained as follows.

The condition of $|\lambda_1| > 1$ is

$$\alpha(x) < 1/3$$

or

$$\alpha(x) > 1.$$
If either of the eigenvalues $\lambda_1$ and $\lambda_2$ has an absolute value greater than 1, the fixed point becomes unstable. For the fixed point $x^*_u$, $|\lambda_1| > 1$, since $\alpha(x^*_u) = 1.596477$. Because of $-3 - 2\beta(x^*_u, x^*_u) = -1.008658 \ldots, 1 - 2\beta(x^*_u, x^*_u) = 2.991341 \ldots, |\lambda_2| > 1$ is not satisfied. Therefore, the fixed point $x_1 = x_2 = x^*_u$ is unstable.

In this section, we have clarified the conditions for the stability of the agreed states and the instability of the equilibrium state for any $M$ and any bias distribution. We also used the case of $M = 3$ to show that a new fixed point exists when $M > 2$, and showed that this fixed point is unstable in this case of $M = 3$. From these results, it can be expected that the ratio of agents converges to the agreed state, which are stable, over time. While the rigorous analysis is difficult when $M \geq 4$, at least the following two propositions have been clarified. (1) An equilibrium state in which all options have equal support is out of balance due to instability. (2) Agents converge to an agreement in the neighbour of the agreed state.

### 3.2. Approximation of time required to find the suitable option

Here, we consider the average time required for discovering the suitable option $\hat{T}$. As mentioned earlier, sometimes a competitive state and an evaluation state have emerged iteratively until the suitable option is discovered. The notation used below is summarized in Table 2.

The prior experiment is illustrated by Figure 13. This is the result of measuring $T_D$, that is, the sum of the time of the competitive phase and the evaluation phase, by performing 10,000 experiments each by changing the combination of $\tau$, $M$, and $N$.

$$T_D := T_v + T_a$$  \hspace{1cm} (77)

In this figure, the whiskers of each bar indicate the corresponding standard deviation and the variance of $T_D$ is very small in any pairs of $M$ and $\tau$. Therefore, in the following, we assume that the time required for an agreement is a constant $\hat{T}_D$.

Assume that there is only one suitable option, $a_s$, among $M$ options. When voting for the first time, the probability of selecting $a_s$ is $1/M$. From the second agreement, the agents know that one option was not $a_s$. We assume the next agreed option is determined almost at random. Under these assumptions, Table 3 indicates the estimation time required to discover the suitable option. Thus, the time required for discovery $\hat{T}$ is the following:

$$\hat{T} = T_D' \cdot \left( \frac{1}{M} + \sum_{i=2}^{\infty} \left( 1 - \frac{1}{M} \right)^{i-2} \frac{i}{M-1} \right)$$  \hspace{1cm} (78)

where

$$\bar{E} = \frac{1}{M} + M - 1$$  \hspace{1cm} (79)

This equation is the average number of agreements required for the discovery. Thus the average number of agreements required for the discovery can be estimated to be almost linear for large $M$.

The time required for agreeing to $\hat{T}_D$ is the time taken from the initial time to the time when the competitive phase ends. $T_D$ is different each time it is tried depending on the initial conditions. The characteristics are analysed below. At a certain $q$th attempt, agent $P_i$ exits the agreement when both sides of the Equation (6) are equal, so the agreed time $T_{Di}^q$ is the following expression. $\theta_i^q$ is the bias of agent $P_i$ during this trial.

$$\theta_i^q + \tau \cdot T_{Di}^q = 1$$  \hspace{1cm} (80)

An agent that has the largest bias at $q^{th}$ run $\theta_i^q = \max(\theta_i^q)$ will be the first agent who quit the agreement so the shortest time span $T_D^q = \min(T_{Di}^q)$ is obtained as follows:

$$T_D^q \geq \frac{1 - \theta_i^q}{\tau} \geq \frac{1 - \theta_i^\max}{\tau}$$  \hspace{1cm} (81)

The bias of the proposed method, $\theta_i^q$, is stochastically generated between 0 and $\frac{2}{M}$. Therefore, if the number of agents $N$ is large enough, the following equation holds:

$$\theta_i^\max \approx \frac{2}{M}$$  \hspace{1cm} (82)

So, we proposed $\hat{T}_D$ for approximation of $T_D$ as follows:

$$\hat{T}_D := \frac{1 - \frac{2}{M}}{\tau}, \text{ where } M > 2$$  \hspace{1cm} (83)

Using this and Equation (78), (83), we can derive the approximate value of the average time required for one agreement $\hat{T}$ as follows:

$$\hat{T} = \frac{1 - \frac{2}{M}}{\tau} \cdot \left( \frac{1}{M} + M - 1 \right)$$  \hspace{1cm} (84)

$$= \frac{M + \frac{3}{M} - \frac{2}{M^2} - 3}{\tau}$$  \hspace{1cm} (85)
From this, we can expect that the average time required for discovery $\bar{T}$ increases linearly as the number of options $M$ increases. Also, this estimation shows that $\bar{T}$ does not depend on the number of population $N$. In the following, the accuracy of this approximation will verify by computer experiments.

4. Experimental results

In this section, when there are many options, we will conduct a verification experiment of a BRT model with an SEM mechanism using a new bias generation method. In particular, the estimation performance of Equation (83) and (85) will be clarified by computer experiments.

4.1. Scalability in group size $N$

Here, we clarify the scalability characteristics of the group size $N$ in an agreed time. To confirm that the time required for an agreement $\bar{T}_D$ does not depend on the group size $N$, we conduct verification experiments.
### Table 3. Probability and time required to discover the suitable option.

| Number of agreements | Number of candidates | Probability of discovery | Time required for discovery |
|----------------------|----------------------|--------------------------|----------------------------|
| 1st                  | $M$                  | $\frac{1}{M}$            | $T_D$                      |
| 2nd                  | $M−1$                | $\left(1 - \frac{1}{M}\right) \cdot \frac{1}{M−1}$ | $2T_D$                     |
| 3rd                  | $M−1$                | $\left(1 - \frac{1}{M}\right) \cdot \left(1 - \frac{1}{M−1}\right)$ | $3T_D$                     |
| $p^0$                | $M−1$                | $\left(1 - \frac{1}{M}\right) \cdot \left(1 - \frac{1}{M−1}\right)^2$ | $i \cdot T_D$             |

![Figure 14](image1.png)

**Figure 14.** Average time required for making an agreement $T_D$. The red dotted line represents $T_D = 1/τ$. $M = 2, 5, 10, 12, 14, 16, τ = 0.003, 10,000$ trials.

Across different pairs of group sizes $N$ and the number of options $M$.

First, we set

$$N = k_2 \cdot M$$  \hspace{1cm} (86)

and run the simulations with $k_2 = 1, 2, 4, 6, 8, 10, 20, 40, 60, 80, 100$ and measure the time required for an agreement $T_D$. $τ = 0.003$ and assume that the agents select their options randomly at the initial time. The result of 10,000 trials is shown in Figure 14. The horizontal axis represents $k_2$, the vertical axis represents the average time required for an agreement $T_D$. Red dotted line shows $T_D = 1/τ$.

From the figure, except for the very small value of $k_2$ ($k_2 = 1, 2$), the average time required for an agreement $T_D$ is almost constant for the increase of $N$.

When $k_2$ is very small that cannot be certified as a group, the time required for an agreement is larger than when $k_2$ is large. If $k_2$ is small, i.e. $N$ is small, as shown in Figure 15, the generated bias is not correctly shaped in the form of the proposal distribution. So this is the reason why the group could not make agreements promptly.

From the above, if $τ$ is small enough for the number of options $M$, the proposed method can enable groups to make agreements promptly and the time required for an agreement does not depend on the group size.

#### 4.2. Best option discovering

Next, we show that a group can discover the suitable option using the proposed method by computer experiments. The experimental setups are as follows: the number of agents $k_2 = 10$, $τ = 0.003$. Assuming that each agent randomly selects its option at the initial time, and we verified whether it was possible to select the suitable option by changing the number of options $M$. One option $a_x$, which belongs to $A$, was selected in advance and set to the suitable option ($|A_{goal}| = 1$). With each value of $M$, the number of cases where all of the agents discovered the suitable option $a_x$ was counted for all 10,000 trials, and the cumulative distribution of success ratios of the discovery was determined. The experimental results are shown in Figure 16 and Figure 17.

We define success ratio $E_{success}$ is the proportion of cases when they found the suitable option in 10,000 trials:

$$E_{success} = \frac{n_{success}}{n_{trials}}$$  \hspace{1cm} (87)

where $n_{success}$ is the number of cases when agents find the suitable option and $n_{trials}$ is the number of trials.

Figure 16 shows the cumulative distribution of success ratios of the discovery across different $M$. The
horizontal axis represents the time required for the discovery, the vertical axis represents the cumulative distribution of success ratios. As can be seen, in all of the cases of different $M$, the cumulative distribution of success ratios of the discovery were very high values ($\geq 99\%$). Therefore, the group could discover the suitable option $\alpha_k$ even if $M$ is large.

Figure 17 shows the average time required for the discovery and the estimation $\hat{T} = \frac{M + \frac{3}{2} \cdot \frac{2}{M^2}}{\tau}$ in Equation (85). The horizontal axis represents the number of options $M$, the vertical axis represents the average time. The black line is the result of regression analysis on the relationship between the average time required for the discovery $\bar{T}$ and $M$ using Equation (85). The figure shows the estimation fits well. The coefficient of determination $R^2$ is 0.994 $\approx 1$ means the residual is very small and denotes that the experimental values are almost the same as the estimated.

From the above, we can conclude that if $\tau$ is small enough, the proposed method can enable the groups to discover the suitable option and the average time required for the discovery is nearly proportional to the size of $M$ and the time does not depend on the number of population $N$.

5. Conclusion

In this paper, we propose a BRT model with short-term memory, which is an agent decision-making algorithm for the best-of-n problem. BRT can generate group-level trial and error dynamics. By repeating the competitive and evaluation phases, the group can find the best option from many options. In this study, we proposed a bias generation method using a quadratic function, clarify the stability of the agreed states and the unstable condition of the equilibrium states. When $M = 2.3$, we show that the agreed states are stable and that the equilibrium states are unstable. In addition, although a new equilibrium state occurs when $M > 3$, it is expected that the proposed method will converge to the agreed state because the equilibrium state is unstable. We also proposed a method for estimating the time required to find an appropriate option. We validated this estimation by computer experiments and showed that the time needed to find a suitable option increases almost linearly with the increase in the number of options. Then, it found that the time does not depend on the number of robots.

Application to real multi-agent systems such as the cooperative behavior of multiple robots and application to the analysis of the collective behaviour of living beings are future tasks.

Acknowledgements

The authors would like to thank the reviewers of this submission for useful comments. M.K appreciates his beloved wife, Ikuko.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The authors appreciate the support of JSPS 21K04582.

Notes on contributors

Masao Kubo is an associate professor of the department of Computer Science of National Defense Academy of Japan. He received his Doctor of Engineering degree from Hokkaido University. He studies on Multi-agent system and Swarm intelligence.

Hiroshi Sato is an Associate Professor in the Department of Computer Science at the National Defense Academy in Japan. He received a BE degree in Physics from Keio University in Japan and an ME and DE in Computer Science from Tokyo Institute of Technology in Japan. His research interests include agent-based simulation, evolutionary computation, deep learning, and machine learning. He is a member of the Japanese Society for Artificial Intelligence (JSAI), Society of Instrument and Control Engineers (SICE), The Institute of Electronics, Information and Communication Engineers (IEICE), etc. He was the editorof IEICE and SICE.

Nhuhai Phung received his PhD degree from National Defense Academy of Japan in 2020. He is currently a researcher with the Institute of Information Technology, AMST, Hanoi, Vietnam. His research interests include swarm robotics systems, collective decision-making, computer vision, and deep learning for medical image analysis.

Akihiro Yamaguchi received his Doctor of Science degree from Hokkaido University, Japan, in 1997. He is currently a professor of the department of information and systems engineering, Fukuoka Institute of Technology. His research interests include neural networks, non-linear dynamics, and applications of chaotic systems.

References

[1] Floreano D, Mattiussi C. Bio-inspired artificial intelligence theories, methods, and technologies. Cambridge (MA): MIT Press; 2008.
[2] Reynolds CW. Flocks, herds and schools: a distributed behavioral model. ACM SIGGRAPH Computer Graphics. 1987;21(4):25–34.

[3] Olfati-Saber R, Murray RM. Consensus problems in networks of agents with switching topology and time-delays. IEEE Trans Automat Contr. 2004;49(4):1520–1533.

[4] Vicsek T, Czirok A, Ben-Jacob E, et al. Novel type of phase transition in a system of self-driven particles. Phys Rev Lett. 1995;75:1226–1229.

[5] Valenti G, Ferrante E, Dorigo M. Self-organized collective decision making: the weighted voter model. Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems; 2014. p. 45–52.

[6] Parker CAC, Zhang H. Cooperative decision-making in decentralized multiple-robot systems: the best-of-n problem. IEEE/ASME Trans Mechatron. 2009;14:240–251.

[7] Montes de Oca MA, Ferrante E, Scheidler A, et al. Majority-rule opinion dynamics with differential latency: a mechanism for self-organized collective decision-making. Swarm Intelligence. 2011;5(3–4):305–327.

[8] Krapivsky PL, Redner S. Dynamics of majority rule in two-state interacting spin systems. Phys Rev Lett. 2003;90:238701.

[9] Seeley TD, Visscher PK, Schlegel T, et al. Stop signals provide cross inhibition in collective decision-making by honeybee swarms. Science. 2012;335(6064):108–111.

[10] Pais D, Hogan PM, Schlegel T, et al. A mechanism for value-sensitive decision-making. PLoS ONE. 2013;8:e73216.

[11] Reina A, Valenti G, Fernandez-Oto C, et al. A design pattern for decentralised decision-making. PLoS ONE. 2015;10:e0140950.

[12] Reina A, Marshall JAR, Trianni V, et al. Model of the best-of-N nest-site selection process in honeybees. Phys Rev E. 2017;95:052411.

[13] Hasegawa E, Mizumoto N, Kobayashi K, et al. Nature of collective decision-making by simple yes/no decision units. Sci Rep. 2017;7:1.

[14] Iwanaga S, Namatame A. The complexity of collective decision. Nonlinear Dynam Psychol Life Sci. 2002;6(2):137–158.

[15] Schelling TS. Micromotives and macrobehavior. New York: Norton & Company; 1978.

[16] Wessnitzer J, Melhuish C. Collective decision-making and behaviour transitions in distributed ad hoc wireless networks of mobile robots: target-hunting. In: Banzhaf W, Ziegler J, Christaller T, et al., editors. Advances in artificial life. ECAI. 2003. p. 893–902. (Lecture Notes in Computer Science; 2801).

[17] Kubo M, Phung NH, Sato H, et al. Direction switch behavior to enclose a pack of targets based on phase transition. Proceedings of Swarm. 2015;2015:295–298.

[18] Phung NH, Kubo M, Sato H, et al. Agreement algorithm with a trial and error method at the macrolevel. Proceedings of Swarm. 2017;2017:82–85.

[19] Kubo M, Phung NH, Sato H. Efficient collective search by agents that remember failures. Proceedings of ICAROB 2018; 2018.

[20] Scheidler A, Brutschy A, Ferrante E, et al. The k-unanimity rule for self-organized decision making in swarms of robots. IEEE Trans Cybern. 2016;46(4):1175–1188.

[21] Phung NH. A study on extension of enclosing algorithm for multiple mobile agents [Master’s thesis]. National Defense Academy; 2015. (in Japanese).

[22] Phung NH, Kubo M, Sato H. El Farol bar problem by agreement algorithm based on trial and error behavior at the macro level. Proceedings of the 22nd Asia Pacific Symposium on Intelligent and Evolutionary Systems; 2018.

[23] Phung N, Kubo M, Sato H, et al. Agreement algorithm using the trial and error method at the macrolevel. Artificial Life Robotics. 2018;23:564–570.

[24] Reina A, Valenti G, Fernandez-Oto C, et al. A design pattern for decentralised decision-making. PLoS ONE. 2013;8:e73216.

[25] Hasegawa E, Mizumoto N, Kobayashi K, et al. Nature of collective decision-making by simple yes/no decision units. Sci Rep. 2017;7:1.

[26] Schelling TS. Micromotives and macrobehavior. New York: Norton & Company; 1978.

[27] Wessnitzer J, Melhuish C. Collective decision-making and behaviour transitions in distributed ad hoc wireless networks of mobile robots: target-hunting. In: Banzhaf W, Ziegler J, Christaller T, et al., editors. Advances in artificial life. ECAI. 2003. p. 893–902. (Lecture Notes in Computer Science; 2801).

[28] Kubo M, Phung NH, Sato H, et al. Direction switch behavior to enclose a pack of targets based on phase transition. Proceedings of Swarm. 2015;2015:295–298.

[29] Phung NH, Kubo M, Sato H, et al. Agreement algorithm with a trial and error method at the macrolevel. Proceedings of Swarm. 2017;2017:82–85.

[30] Kubo M, Phung NH, Sato H. Efficient collective search by agents that remember failures. Proceedings of ICAROB 2018; 2018.

[31] Scheidler A, Brutschy A, Ferrante E, et al. The k-unanimity rule for self-organized decision making in swarms of robots. IEEE Trans Cybern. 2016;46(4):1175–1188.

[32] Phung NH. A study on extension of enclosing algorithm for multiple mobile agents [Master’s thesis]. National Defense Academy; 2015. (in Japanese).

[33] Phung NH, Kubo M, Sato H. El Farol bar problem by agreement algorithm based on trial and error behavior at the macro level. Proceedings of the 22nd Asia Pacific Symposium on Intelligent and Evolutionary Systems; 2018.

[34] Phung N, Kubo M, Sato H, et al. Agreement algorithm using the trial and error method at the macrolevel. Artificial Life Robotics. 2018;23:564–570.

[35] Reina A, Valenti G, Fernandez-Oto C, et al. A design pattern for decentralised decision-making. PLoS ONE. 2013;8:e73216.

[36] Hasegawa E, Mizumoto N, Kobayashi K, et al. Nature of collective decision-making by simple yes/no decision units. Sci Rep. 2017;7:1.

[37] Schelling TS. Micromotives and macrobehavior. New York: Norton & Company; 1978.