Efficient collective search by agents that remember failures

Masao Kubo
Department of Computer Science, National Defense Academy
Hashirimizu 1-10-20, Yokosuka, Kanagawa,239-8686, Japan

Nhuhai Phung
Department of Computer Science, National Defense Academy
Hashirimizu 1-10-20, Yokosuka, Kanagawa,239-8686, Japan

Hiroshi Sato
Department of Computer Science, National Defense Academy
Hashirimizu 1-10-20, Yokosuka, Kanagawa,239-8686, Japan

E-mail: masaok@nda.ac.jp, ed17006@nda.ac.jp, hsato@nda.ac.jp
http://www.nda.ac.jp/cs/stuff/masaok.html

Abstract

The BRT agent is an algorithm that can find appropriate collective behavior by changing the agreement contents in a trial and error manner. Computer experiments show that it is necessary to change the agreement contents the number of times that is proportional to the square of the number of choices. In this paper, we attempted to shorten this search time by introducing an agent that memorizes actions that were not able to achieve the expected effect of what they executed. As a result, we found that search time can be improved by just mixing a few this proposed taboo list agents.

Keywords: Swarm intelligence, Machine learning, Complex systems, Best-of-n problem

1. Introduction

Roughly speaking, if a set of required roles (for example, collecting foods and resting) and the relationship between these elements are embedded in individual behavior beforehand, for example they can keep energy of them constant [1]. On the other hand, it is anticipated that it will not be able to lead to some reasonable macro states of the swarm by the above division of labor methods such as when a new environment is encountered. They need methods how to emerge the swarm in a better macro state against such a situation [2]. To solve this issue we have to discuss individual's learning but also learning capability as swarm and no general models as to whether to discover, memorize, and learn new micro interactions leading have not been proposed. It has become a bottleneck in the construction of a more advanced self-organizing swarm system [4].

To this problem, a voting model (BRT model, Bias and Rising Threshold model) has been proposed that emerges a trial and error process at the macro level by momentary change of agreement of an agent group [6]. Every agent chooses one action from among the actions and notifies all of them. The agent changes its action in view of the number of supporters of each action. They keep voting until actions of the all members are same. This is a model that expands the 2-option consensus

building model against the agreement of [3] and others so that it can be used even in the case of multiple options and incorporates the passage of time. By appropriately designing the bias value θ which expresses the degree of preference of the behavior selected by itself, this algorithm has the feature of reaching consensus in almost constant time regardless of the number of agents in the group. A series of computer experiments shows that it is possible to find suitable consensus actions in a time proportional to approximately M^2 for the number of actions M.

This is inefficient in cases other than dynamic environments where the environment and tasks frequently change. Here, we introduce a derivation agent of BRT that memorizes a set of behavior that were once executed but not appropriate, and removes them from their own action selection. In the following, this function is called UEM (unsuitable experience memory), and an agent having this function is called an UEM agent. The memory used here is a short term, and it forgets probabilistically. As will be described later, if all agents have UEM they can expect to discover a good macroscopic behavior in proportion to M.

This paper is organized as follows. First, we will explain the BRT and its characteristics. Next we propose an UEM agent and estimate its characteristics. Finally, we verify the characteristics of the proposed method by computer experiment. As a result, we found that it is possible to reduce the time required for solution search even if all agents do not have UEM function.

2. Related works

2.1. *BRT model*

Now, we assumed that there are n agents $P_1, \ldots, P_n$. $A = \{a_1, \ldots, a_j, \ldots, a_M\}$ is set of agents's behavior candidates with $M \geq 2$. $A_i(t) \in A$ is the behavior of agent P_i at time t. The agent P_i has a bias θ_i ($0 < \theta_i < 1$). $n(a_j)$ is number of agents selecting behavior a_j.

As a result of everyone performing the same action, a macro state of the group occurs. And if this is suitable for the environment and tasks, this agreement behavior is assumed to be "suitable".

A_{goal} is the set of desirable behavior candidates ($A_{goal} \subset A$). In advance, the agent does not know A_{goal}. Only when all have agreed, it is understood that A_{goal} includes that agreed behavior.

At this time, the agent P_i decides the behavior $A_i(t+1)$ at the time t as follows. If

$$\frac{n(A_i(t))}{N} \geq \theta_i + \tau c\,(t\text{-}t_{i,last}) \qquad (1)$$

is satisfied, $A_i(t+1) = A_i(t)$. Otherwise, behavior other than $A_i(t)$ is stochastically chosen and becomes $A_i(t+1) \in A \backslash A_i$. Here, τ is a constant representing the increment of proponents number.

In addition, $t_{i,last}$ is the clock time at which the agent P_i last changed its behavior, and $(t\text{-}t_{i,last})$ is the time span, which the same action continues to be selected.

$$t_{i,last}(t+1) = \begin{cases} t+1 & A_i(t+1) \neq A_i(t) \\ t_{i,last} & \text{otherwise} \end{cases} \qquad (2)$$

c(t) is a function that is 1 when not achieving the desirable macro purpose as follows.

$$c = \begin{cases} 0 & \forall i, j, A_i(t) = A_j(t) \in A_{goal} \\ 1 & \text{otherwise} \end{cases} \qquad (3)$$

If the ratio of agents whom selects the same behavior as theirs is lower than $\theta_i + \tau c(t\text{-}t_{i,last})$, the agent selects a new behavior from other behavior candidates.
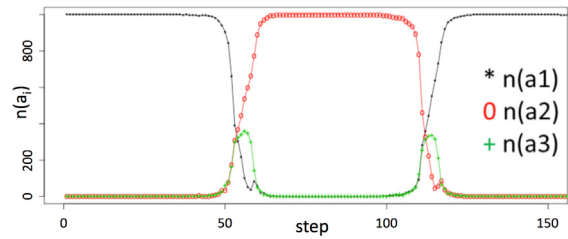

Figure 1. Transition Example (Staring from an agreed state)

2.2. Behavior example of BRT model and its search time characteristics

Here we illustrate the behavior of the BRT model. We will introduce its characteristics. Figure 1 is rendition of what was described in our past work [5]. Figure 1 shows a transition process from an agreed state. In the initial state, all of agents have agreed to a_0, but everyone agreed to another action (a_1) in about 65 steps. Again, 1000 agents agreed to a_0 action again in 125 steps. From the above it can be seen that agreed behavior can be switched from either state.

Here, we show that it is possible to agree with the desired behavior by computer experiment. The experimental results are shown in Figure 2. The + mark is the average time required for discovery, and the ○ mark is the value obtained by dividing it by M^2. Since the ○ mark is almost parallel to the x axis, it is mentioned that the average of the time required to discover the solution is proportional to M^2.
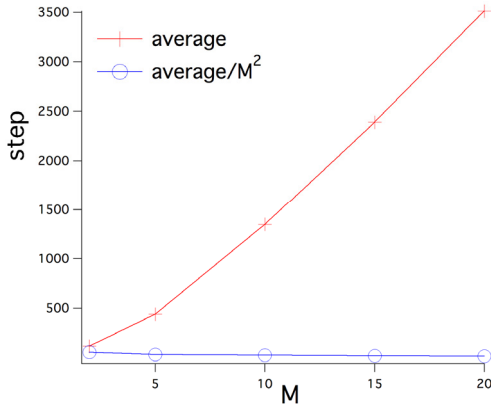

Figure 2. Characteristics of search time cost by BRT

3. The proposed agent (UEM agent)

The BRT agent remembers the last action that was agreed but did not suit, and select other behaviors. Namely this short-term memory is limited to the most recent behavior. Then they repeatedly agree on behaviors that were not suitable a while ago. If the environment has not changed since the last execution, this is not an efficient search. Therefore, we will introduce an agent that memorizes behavior that agreed but did not suit for a long time, and tried to make solution searching more efficient.

The agent has its list $L_i(t)$ that remembers actions that were agreed but not suitable, and adds it to this list with the probability p_{in} at the time when it is determined that it is inappropriate while agreeing.

$$L_i(t+1) = \begin{cases} L_i(t) \cup A_i(t) & \text{probability } p_{in} \\ L_i(t) & \text{otherwise} \end{cases} \quad (4)$$

where $A_i(t) = \{ \forall j, A_i(t) = A_j(t), A_i(t) \notin A_{goal} \}$.

When selecting a new action, select actions that are not included in $L_i(t)$ stochastically. Namely,

$$A_i(t+1) \in A \backslash L_i(t) \quad (5)$$

The agent stores it with Eq.5 and select a candidate with Eq.4. This BRT agent will be referred to as an UEM (unstable experience memory) agent below.

Table 1 Probability to support one unselected action.

Number of trial	BRT agent	UEM agent
1st	1/M	1/M
2nd	1/(M-1)	1/(M-1)
3rd	1/(M-1)	1/(M-2)
x-th	1/(M-1)	1/(M-(x-1))

3.1. *Approximation of characteristics of the proposed method*

Consider the ratio of UEM agents. Table 1 indicates the number of attempts to agree and the probability of selecting an arbitrary unselected action. When trying an agreement for the first time, the probability of selecting one action is (1/M) with BRT, UEM agent. Both agents know that one action was inappropriate at the second time, so both are 1/(M-1). However, in the third and subsequent times, the behavior selection probabilities of both agents will differ. The BRT agent remains 1/(M-1), but the UEM agent does not select actions that are already selected as inappropriate, so candidates are reduced step by step. The selection probability of x trial is 1/(M- (x - 1)). In order to shorten the search time it is necessary to reduce the number of times they agree with the action they tried already. This corresponds to the vote of UEM agent.

When the difference between the expected values of both at the x-th time is

$$\frac{u}{M-(x-1)} \quad (6)$$

where u (<n) is the number of UEM agents. This equation shows that if the population of UEM agents is large, it is less likely to choose actions that have already been tried. Also, as the denominator decreases as the number of attempts to reach an agreement progresses, it indicates that less agreement on already selected actions will be reduced. Therefore, even if the population of the UEM agent is small, the effect of shortening the search time can be expected.
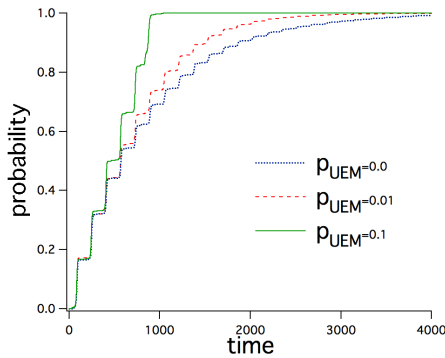
Figure 3. When changing the ratio of UEM agent.

4. The experimental results

4.1. *UEM agent ratio*

Figure 3 represents the search time as a cumulative probability density distribution, and the horizontal axis is the time required. The vertical axis represents the probability of finding a solution by this time. The solid line represents the case of $p_{UEM} = 0.1$. The search time has been found by 99% trial by 1000 steps. As estimated by the formula 6, even a small number of UEM agents can greatly improve search time.

The relationship between the ratio of UEM agent and the average search time is shown in Figure 4. The horizontal axis shows the ratio of the UEM agent, and the vertical axis shows the average search time. $p_{UEM} = 0.0$ requires an average of 834 steps, and $p_{UEM} = 1.0$ took an average of 462 steps. With $p_{UEM} = 0.001$, it is 810 steps, and search time can be improved just by mixing minute UEM agents. As we increased the proportion of UEM agents, dramatic improvements were seen up to $p_{UEM} = 0.1$ (meaning that $p_{UEM} = 0.1$ required an average of 486 steps).

From these two experiments, we found that the proposed UEM agent can improve the search time of BRT agent.

5. Conclusion

In this paper, in order to improve the efficiency of the trial and error search behavior that the BRT agent group emerges, we proposed a derived type of BRT agent from the candidates by recording the failed experience called UEM agent. If all flocks are composed of UEM agents, we can expect that the search time increases linearly with M. However, even in the case of mixing with the BRT agent, we estimate that even a small number of UEM agents are effective, and verified the trend in computer experiments. As a result, it turned out to show the behavior as expected.
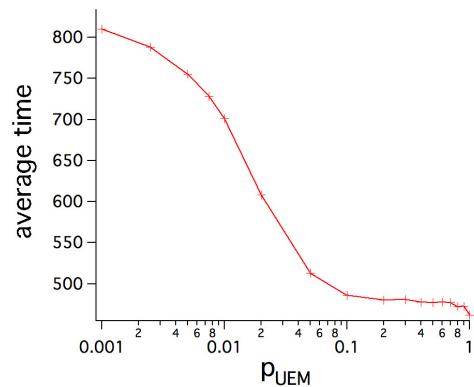

Figure 4. Relation between ratio of UEM agent and search time.

References

1. E. Bonabeau, G. Theraulaz, J. L. Deneubourg, Fixed Response Thresholds and the Regulation of Division of Labor in Insect Societies, *Bulletin of Mathematical Biology*, Vol.60(1998)753- 807.
2. G. Valentini, E. Ferrante, M. Dorigo, The Best-of-n Problem in Robot Swarms: Formalization, State of the Art, and Novel Perspectives, *Frontiers in Robotics and AI*, Vol.4, 9(March 2017)1-18.
3. A. Namatame (In Japanese), *Strategic Decision Making*, Asakura Publisher (2001)119-149.
4. M. Kubo, N. H. Phung, H. Sato, A. Namatame, Direction switch behavior to enclose a pack of targets based on phase transition, *Proceedings of swarm 2015(2015)*155-161.
5. N.H. Phung, M. Kubo, H. Sato, S. Iwanaga, Agreement algorithm with trial and error method at macro level, *Proceedings of Swarm 2017(2017)*82-85.
6. N.H. Phung (In Japanese), *A study on extension of enclosing algorithm for multiple mobile agents*, Master's thesis at national defense academy of Japan(2017)