

How can agent know the global information without close coordination ?

Satoshi Kurihara
kurihara@ist.osaka-u.ac.jp
8-1, Mihogaoka, Ibaraki, Osaka, 567-0047 Japan
Osaka Univ., Japan

Shinya Sato
sato@ingrid.org
NTT Network Innovation Labs., Japan

Kensuke Fukuda
fukuda@nii.ac.jp
National Institute of Informatics, Japan

Toshiharu Sugawara
sugawara@t.ecl.net
NTT Communication Science Labs., Japan

ABSTRACT

In the multi-agent systems, even if each agent behaves selfish or cooperative, its behavior influences the other agents' behaviors and their behaviors influence its behavior cyclically. Therefore a certain information about the relative position of each agent among all the agents may be able to be extracted only from its behavior. If this information can be extracted, each agent has the possibility to improve its efficiency only by seeing its own behavior without knowing about the information of the other agents. In this paper, we verified this hypothesis by using the competitive multi-agent simulation environment named Minority Game, and confirmed that each agent's performance gain was actually possible. This analytical technique may be especially useful for massively multi-agent systems.

1. INTRODUCTION

Generally, in the multi-agent systems, each agent only has local view and cannot obtain the global information of the agents as a whole. In this case, for the each agent to be able to do is to infer the global view or its relative position by getting information from the other agents. But, there is no confirmation that the correct information can be obtained. At this point, in the multi-agent system, even if each agent behaves selfish or cooperative, its behavior influences the other agents' behaviors, in other words, its behavior changes the state of the environment formed by the agents. Then, the changed environment also influences each agent's behavior, and this cycle continues. Hence, it can be thought that the behavior of each agent may include some certain information about its relative position in the agents as a whole. If these information actually exist and each agent can extract one of these information, it can know its position by only checking its own behavior and may be able to improve its behavior efficiently. The purpose of this paper is to verify the possibility of this assumption by using the Minority Game (MG)[?]. In MG, a lot of agents play a simple game and even though they

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

AAMAS'06 May 8–12 2006, Hakodate, Hokkaido, Japan.
Copyright 2006 ACM 1-59593-303-4/06/0005 ...\$5.00.

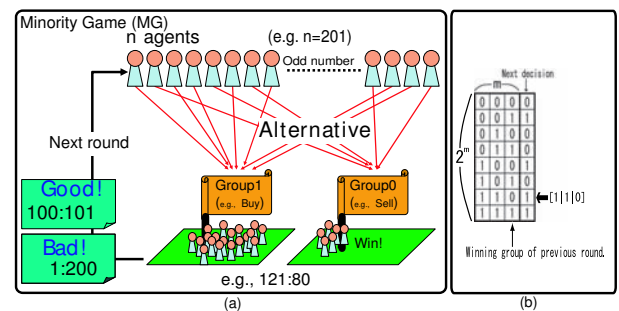


Figure 1: Rules and strategy table

have quite local view and behave quite selfish, certain cooperative behaviors are emerged. So, some information about each agent's performance must be included in the behavior of each agent in MG.

2. MINORITY GAME

Firstly, we review the rules of MG (see Fig. ??(a)). There are n (odd number) autonomous agents, each of which independently chooses between two alternatives (*group-0* or *group-1*) according to its own behavioral rules, that is, it does not see other agents' behaviors. After all of the agents choose one alternative or the other, and the agents that finish in the minority group are considered to be winners. This is one round of the game. Each winner agent is awarded one point, and the total numbers of winner agents is the profit in this round of the game. The each agent makes its selection by using one of multiple strategy tables that it holds. The entities in the table contain all combinations of m past history of the minority group choices and next decisions that corresponds to each of the combinations for the next round (see Fig. ??(b)). At the beginning of the game, each agent prepares s ($s \geq 2$) strategy tables, and the next decision entries of each strategy table is stored randomly by 0 or 1. After the game starts, the entries of each table cannot be modified. In the first round of the game a *m-past-winner-group-history* is randomly set, for example, to $[1|1|0]$, and each agent randomly chooses one of its holding s strategy tables and sees the next decision entry corresponding to $[1|1|0]$ (see Fig. ??(b)). And if the next decision entry is *group-1*, the agent selects *group-1*. Then if the agent becomes winner, one point is assigned to the agent and also to the selected strategy table. And if the agent loses, one point is deducted from the selected tables. After the score of the strategy

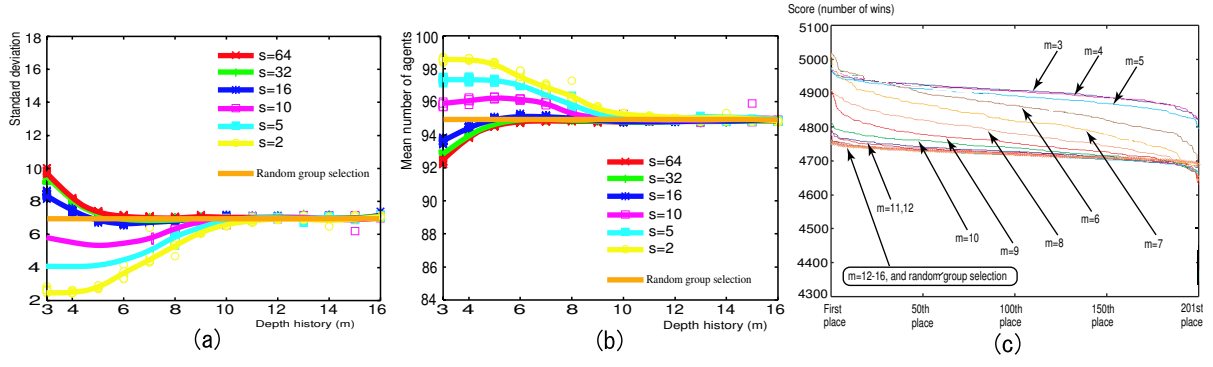


Figure 2: (a) Standard deviations of the number of winning agents. (b) Mean numbers of winning agents

tables of all the agents are updated, then the m -past-winning-group-history is updated from $[1|1|0]$ to $[1|0|1]$ because group-1 becomes winner-group in this round of the game. In the second and subsequent rounds, each agent selects the strategy table that has the highest score. The round is repeated a predetermined number of times, and the final result of the game is the total number of points acquired by winner-agents.

2.1 Coordinated behaviors

Then, the following overall order is emerged through such the simple rules. The game was played for the number of rounds described below; One trial for each parameter pair, (s, m) , $\{s|2, 5, 10, 16, 32, 64\}$, $\{m|3 \text{ to } 16\}$, is 10,000 rounds of the game by 201 agents; ten trials were conducted for each pair. Fig. ??(a) shows the standard deviation of the number of winning agents, and Fig. ??(b) shows the mean numbers of winning agents. The horizontal lines in Fig. ??(a) and (b) represent the standard deviation and mean value when all of the agents choice group-0 or group-1 randomly for each round. These graphs show that, for the lower values of s , especially for $s = 2$, the standard deviation became lowest and the mean number of winning agents became highest when m was from about three to five. Fig. 2(c) shows the rankings of the 201 agents by average score after 10,000 rounds ($s = 2$). In the case where each agent randomly selected group-0 or group-1, all of the agents were able to get approximately 4750 points. In contrast to this, the mean score was high for small m ($m = 3$ to 5), and although some few differences between agents can be seen in the scores, all or almost all of the agents were able to achieve stable high scores. These simulation results mean that some kind of cooperative behavior among the agents was driving the winning-group ratio closer to 100:101 in small s , especially in $s = 2$, and m . Especially, the most interesting characteristic is that, although we would expect behavior based on longer histories to be more efficient, m larger than 10 produce results that are the same as those of random behavior. In this paper, we mainly ran the game by several number of agents like $n = 101$, 201, and 301, and since our interesting is in the cooperative behavior of agents, we ran the game by $s=2$ (each agent has two strategy tables).

3. AGENT BEHAVIOR AND POWER-LAW

For each agent, its holding strategy tables are very important to decide its behavior, so, in [?], we investigated how the strategy tables are used by each agent to analyze each agent detail behavior. Then, up to now, we have discovered the following curious fact, While the high winning ratio agents used both strategy tables, there was no typical fixed period for the continuous use of one strategy

table. That is, a self-similarity was visible in these results. Fig. ?? shows, on log-log scales, per-agent histograms of the periods over which either of the strategy tables was continuously used. Results for $m=3, 7$, and 14 are given for (a) $n = 101$, and (b) $n = 301$. Power-law can be seen in both cases of $m=3$ (graphs were nearly straight lines) and this also reflects the presence of a self-similarity. To the contrary, when $m=14$ power-law is not seen. Interestingly, the histograms for agents with $m=7$ show a mix of the two types of results; some agents are similar to the graphs for $m=3$, while others are similar to those for $m=14$. It can be thought that the straight lines are for *cooperative agents*, while the curves are for *non-cooperative agents*. So, we calculated the all slopes of the histograms in Fig. 3 (see Fig. ??). As you can see, the consistent relation can be seen that is the slope of high winning-ratio agent is more horizontal (nearer to 0) than that in low winning-ratio agent. That is, it can be thought that each agent can know its rough position among the group by seeing whether the way of its own strategy table selection obeys the power-law or not.

4. IMPROVEMENT OF WINNING-RATIO USING POWER-LAW FEATURE

The easiest way to change the agent's behavior is to renew its holding strategy tables. So, the following experiments in which each agent renews its strategy tables by three rules were done. As for the experiment, we ran the game by 101 agents ($n = 101, s=2$). After the game is repeated 10,000 times (10,000 times is one turn), each agent's efficiency is evaluated by the following three kinds of criteria, and each agent which is identified as inefficient renews its strategy tables, then next 10,000 times is started cyclically. In the experiment we repeated the renewing process 10 times that is we ran the game total 100,000 times. We prepared these three evaluation-rules:

(evaluation-1) Since the agent with its power-law slope being nearly 0 has high winning-ratio, we randomly select u pieces of agents whose power-law slope become lower than the previous turn, and selected u pieces of agents renew their strategy tables. In this evaluation, since the calculation of each agent's power-law is computable only from its behavior, each agent need not know the total of agents in the game. However since this evaluation uses only the change of slope even the high winning-ratio agent may renew its strategy tables.

(evaluation-2) We randomly select u pieces of agents whose power-law slope become lower than the certain value (0.75), and selected agents renew their strategy tables. In this evaluation, we use the following knowledge that is, there are consistent relation between the

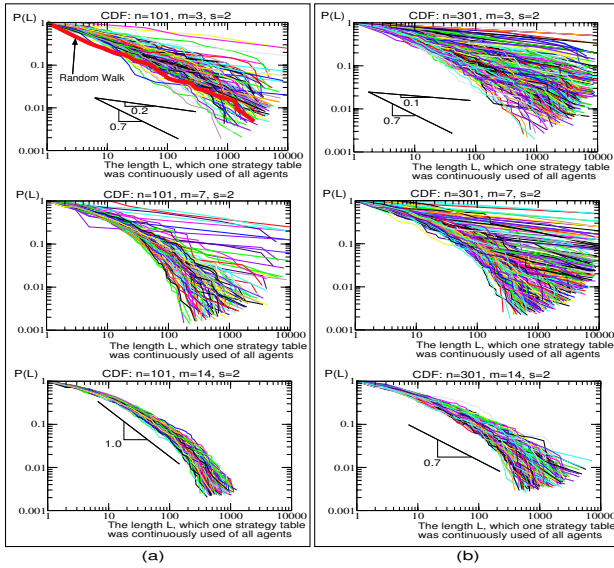


Figure 3: No. of steps L over which either strategy table was in continuous use for (a) $n=101$ and (b) $n=301$ with (upper) For $m=3$ (all 101 agents and 301 agents), (middle) For $m=7$ (all 101 agents and 301 agents) (lower) For $m=14$.

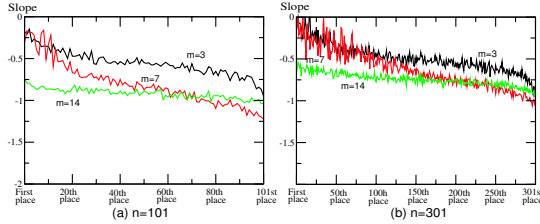


Figure 4: Slopes of the histograms in Fig. 5.

power-law slope and the winning-ratio in each agent and this consistent relation does not depend on the total of participating agents in the game. In this evaluation, each agent can know its approximate position without knowing the total of participating agents, and agents only having low winning-ratio can renew their strategy tables.

(evaluation-3) We select u pieces of agents having low score, and selected agents renew their strategy tables. In this evaluation, it is necessary for each agent to know the following global information, each agent's score and the number of total of participating agent. This evaluation can be thought as the best way to select the agent to renew its strategy table. We prepared this evaluation for the comparison to above evaluation-1 and -2.

Fig. ?? shows the score of 101 agents of nine kinds of evaluation settings. We ran the three kinds of games ($u=5$, $u=10$, and $u=20$) for each evaluation. And as mentioned above, in each situation, strategy table renewing process was repeated 10 times and score of agents after the each renewing process was plotted to the graph of each evaluation setting. As a result, in all settings it can be confirmed that the winning-ratio has improved in all almost agents. First of all, $u=5$ was the best in all three kinds of evaluation. The reason can be thought as follows; if we select too many agents, many renewed strategy tables may destroy the current good strategy table's combination of high winning-ratio agents and thus the effect of renewing may be lost. As for the evaluation-1 and -2,

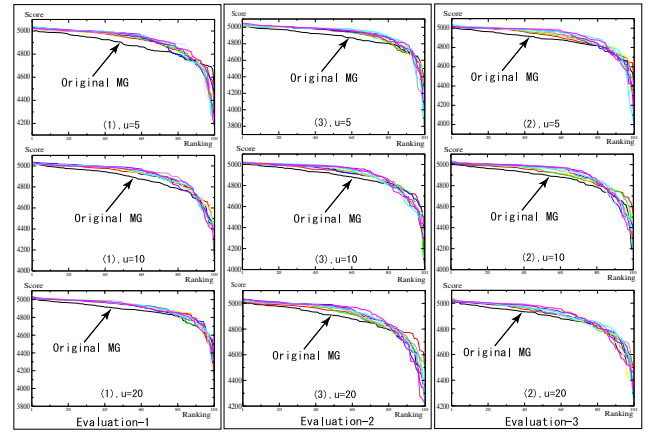


Figure 5: Score of 9 kinds of evaluation settings

evaluation-2 was better than evaluation-1 as we expected because in evaluation-2, only the agents having low winning-ratio are able to renew their strategy tables. Consequentially, evaluation-2 was the most effective though it had been expected that evaluation-3 was at first most effective. While the evaluation-3 uses the several global information, it pays attention only to the score transition of agent. On the other hand, the evaluation-1 and -2 pay attention to the agent behavior. So, in the following two situations, "the score is good by chance even if the evaluation of behavior is bad" and "the score is bad by chance even if the evaluation of behavior is good", agent should not renew its strategy tables in the former situation, and should renew them in the latter situation, thus this experiment result shows paying attention to agent behavior is effective as the evaluation methodology.

5. CONCLUSION

In this paper, we verified whether the efficiency of each agent behavior was able to be improved only by seeing its behavior in MG. In MG, important information for the agent to know the efficiency of its behavior was power-law. Power-law is most typical factor of scale-free network. For example, in the Internet it is known that power-law can be seen in the flow of packets when the number of packets of flowing is the maximum in the traffic [?]. This means if the flow of traffic can be controlled to follow of power-law when the flow of traffic becomes an inefficiency situation due to some troubles, the flow of traffic may be corrected to an efficient situation. In the future work, we also plan to apply this coordination methodology to an actual system like the ubiquitous information communication systems and to verify whether an efficient performance improvement is possible.

6. REFERENCES

- [1] D. Challet, M. Marsili and R. Zecchina, Phase Transition in a Toy market (1999); on the major Minority Game web site: <http://www.unifr.ch/econophysics/minority/>
- [2] S. Kurihara, K. Fukuda, T. Hirotsu, O. Akashi, S. Sato and T. Sugawara, How collective intelligence emerge in complex environment?, *Proceedings of Bio-ADIT 2004*, pp. 498 - 511, (2004)
- [3] K. Fukuda, M. Takayasu, and H. Takayasu, Analysis of Minimal Model of Internet Traffic, *Traffic and Granular Flow'01* (M. Fukui, Y. Sugiyama, M. Schreckenberg, and D. E. Wolf Eds.), Springer, (2002).