Autonomous Selection of Meta-Agent Roles in Social Dilemma

Keiji Suzuki
Future University - Hakodate
Kameda-Nakano 116-2
Hakodate City, Hokkaido, 041-8655, Japan
suzukj@fun.ac.jp

Abstract

This paper shows the agent based simulation with changing roles in the social dilemma game, the tragedy of the common. The tragedy of the common is known to treat the problem that is how to manage the limited common resource. To control the usage of a common resource, the levy based control strategy is employed. The role of the meta-agent is introduced to charge the levy to other agents in the proposed simulation. The levy based control strategy can be expected to avoid the dilemma situations even if the agents obey the individual rationality. However, to introduce the meta-agents, it should be considered that who became the meta-agents and how to make the charging plan of levy against the activities of the agents. To solve the problems, it is proposed the activities of the agents that include the activities for playing the game and the activity as the role of the meta-agent. Namely, the proposed agents will become the meta-agent autonomously according to the situations. Concerning to adjusting the charging plan, the plan is adjusted by the GA based on the proposed evaluation methods. Throughout the experiments, the effects of the changing roles and the formation of the charging plan are examined.

Introduction

Agent based social behavior simulations are research field that treats complex game situations and examines artificial intelligence (Suleiman00) (Namatame02). Social dilemmas are one of the complex game situations and suite to examine the intelligence of agents (Conte97). In this paper, the Tragedy of the Common (Hardin68), which is one of the social dilemmas, is treated in the agent-based simulation. In this game, players use common limited resources to get the reward. If players behave based on the individual rationality, all players will face to tragedies loosing higher payoff. To avoid such tragedies, players have to make the relationship between other agents to prevent the selfish behaviors or change the problem structure, for example, changing the payoff functions. In the simulations of the Prisoner's Dilemma and the n-persons dilemma game, to avoid the selfish behavior, extended agent's abilities and several social norms were introduced (Liebrand96). The results of these approaches will be enough to avoid the dilemma situations if almost agents in the societies can possess the assumed abilities and norms. However, the simulated societies become more complex, that is, the properties of the agents become more heterogeneous, we should be prepare the another types of approach to avoid the dilemma situations. The changing problem structure can be thought as one of such other approaches. In this paper, to avoid the dilemma situations, the approach of the changing problem structure is employed. That is, the meta-agent is introduced to control the levy that changes the received rewards of the players (Yamashita01).

Because it is important who acts the role of meta-agents in real society, the agents in the social simulations should possess the decision mechanism that includes the selection of the role. Therefore, it is proposed that role of the meta-agent is treated as one of the activities of the players. Namely, the player will select the role of the meta-agent if the expected revenue exceeds than the expected rewards when the agent acts as the player. Another problem for introducing the meta-agent is how to set the plan of the levies charged to the activities of the players. To acquire the levy plan, the Genetic Algorithm is applied. Based on the evolutionary acquisition of the levy plan and the decision mechanism including the selection of the role, the effects of the interactions between the meta-agents and the agents are examined in the Tragedy of the Common game. In the next section, the problem structure of the tragedy of common is introduced. Then the proposed approach is described.

The tragedy of the common

The tragedy of the common (Hardin68) is famous game problem as one of the n-persons social dilemmas (Yao96). This game enables for us to analyze the behaviors of players sharing common limited resources. Owing to the common resources are limited, higher activity of agents to get the higher payoff will become to bring lower payoff. The one of the general form of the payoff function in the Tragedy of the Common (Hardin68) is as follows;

$$Payoff(a_i, TA) = a_i(|A| \times N - TA) - 2a_i \quad (1)$$

where, $Payoff(a_i, TA)$ is the payoff of agent i. a_i denotes the degree of activity of agent i. TA is the total activ-

Table 1: Example payoff table of the Tragedy of the Common

		Total activities of the agents except agent i									
		0	1	2	3	4	5	6	7	8	9
a^i	1	13	12	11	10	9	8	7	6	5	4
	2	24	22	20	18	16	14	12	10	8	6
	3	33	30	27	24	21	18	15	12	9	6

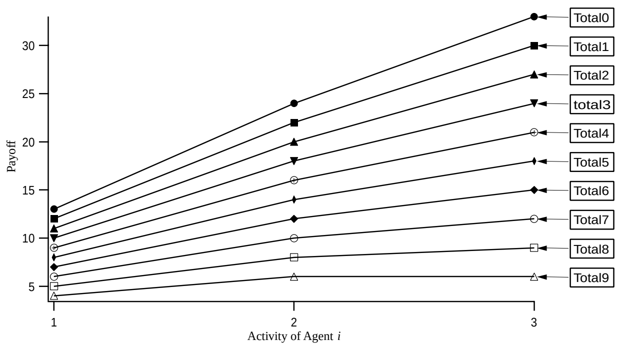

Figure 1: Graphical view of example payoff value in Table.1

ities, $\sum_j^N a_j$. A is the set of the activities. N is the number of the agents.

The example of the payoff function (Suzuki94) is shown as follows;

$$Payoff(a_i, TA) = a_i(16 - TA) - 2a_i \quad (2)$$

where, $TA = a_1 + a_2 + a_3 + a_4$. Here, 4 agents participate and 4 degrees of the activities, $a_i \in \{0, 1, 2, 3\}$, is supposed. The payoff values based on the function becomes like as Table 1.

Let's consider the game in which the player decides own activity based on the individual rationality. The game assumes the activity of agents consuming the limited common resources. Therefore, the payoff becomes will decrease when total activities increase. However, the agent i will increase the activity against any total activity of other agents, because the agent i can increase own payoff until the total activity of the agents reaching 11 in the example. Namely, the strategy of higher activity always dominates the strategy of lower activity. Thus all players will decide to increase their activities based on the individual rationality. Thus the decisions based on the individual rationality will cause the limited common resources being exhausted and the payoffs will become decrease accroding to the total activities. Finaly, the agents will be face to the tragedy. In the example, the tragedy arises when total activities reached 12. That is, the payoff can't incease even though the agent i increases the activity from 2 to 3.

The characteristic of the game is known that no technical solution exists. Therefore, to solve this game, players should change the individual rationalities to other types of rationality or problem structures should be changed to payoff function. One of the objectives of the agent-based simulations is examined what kinds of rationalities and extended problem structures can avoid social dilemmas like as the tragedies. In this paper, the architecture of the proposed agent based simulation is belonging to the extension of the problem structure. Namely, the meta-agent is introduced to prevent the agents based on the individual rationality causing the tragedy. The detail of the proposed approach is described in next section.

Approach

To solve the social dilemma, the levy based control strategy is introduced (Figure 2). In this strategy, the property of the levy plan is important. Apparently, it will be enable to fix the suitable levy plan by analyzing the game properties. Then the plan can be embedded in the simulation beforehand. However, one of the purposes of the agent-based simulations is the examination of the autonomous property of the agents in social environments. Therefore, the autonomous acquisition of the suitable levy plan is desirable. To acquire and control the levy plan according to the situations, the role of meta-agent is introduced. It is defined that the meta-agent can broadcast the levy plan and get the revenue by charging the levies to the activities of the agents. To realize the social simulation based on the meta-agent approach, we have to treat the problems, that is, who becomes the meta-agents and how to manage the levy plan. To treats the problems, it is proposed that the decision making includes not only the selection of the activity but also the selection of the role of meta-agent. Therefore, it is expected to appear the role of meta-agents autonomously according to the situation. To adjust the levy plan, the simple rules concerning the collection of the levies from the agents are proposed to prohibit the meta-agent behaving selfishly. The details of the architecture of the agents are described in the following subsections.

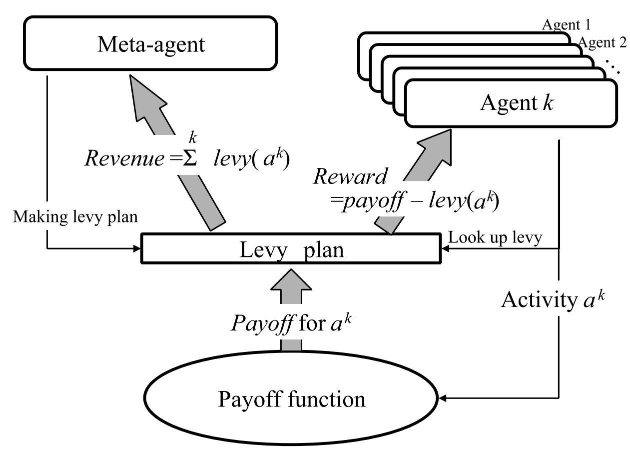

Figure 2: Schematic view of levy based control strategy.

Architecture of Agents

The group of agents is represented by N. $N = \{1, 2, \ldots, n\}$. $A = \{a^1, a^2, \ldots, a^k, a^{meta}\}$ denotes the activities of the agent. It is assumed a^{i+1} is higher activity than a^i in the tragedy of the Common game. The element of the activity, a^{meta}, means the agent acts as the role of the meta-agent. Namely, the agent can select the role of the meta-agent as one of the activities.

The decision making for selecting the role of the meta-agent is based on the levy plan. Each agent prepares the levy plan to evaluate the expected revenue for deciding the agent should become the meta-agent or not. If the agent selects the meta-agent role, the levy plan is broadcasted to other agents. The expected payoffs according to the activities as the player are evaluated with referring the broadcasted levy plans and the payoff function. According to the evaluations, the decision-making for selecting the activity carries out to maximize the expected incoming among the expected rewards and revenue. The details of the levy plan are described in the next subsection. The decision-making method are mentioned later.

Levy plan and charging rule

The levy plan consists of the values of levies, $\{Lv^1, Lv^2, \ldots, Lv^k\}$, corresponding to the activities, $\{a^1, a^2 \ldots, a^k\}$. Namely, the levy, Lv^i charges to the payoff of the activity, a^i.

The levies can change the original payoff values to the rewards of players. Therefore even if the players decide their activities based on the individual rationality, the suitable levy plan will prevent the activities exhausting the limited common resources. However, it is remained that the problem is how to set the suitable levy plan. The issue is connected to the planning policy of levies. In this approach, the individual rationality is employed for the planning policy. Namely, the objective of the meta-agent is to maximize the revenue. The individual rationality is simple policy. The policy doesn't ask the meta-agent to have a specific cooperative rationality. However, the policy of the meta-agent is afraid to increase the levies selfishly. To inhibit the selfish behavior of the meta-agent, the simple rule for charging the levy is introduced. That is, only if the payoff exceeds the levy value, $Payoff(a^i, TA) > Lv^i$, the meta-agent can receive the levy. Otherwise, the received value become 0. This simple rule can be expected to inhibit the selfish behavior of the meta-agent.

To search the suitable levy plan, the genetic algorithm is applied. The details of the evolution of the levy plan are mentioned later.

Decision making and revising process

To determine the activity of the agent, the revising process is prepared. In this game, the payoff is varied by the other agents activity, that is, TA. Therefore, it is difficult to determine the activity without the information of the other agent's activities. The revising process consists of the fixed length of steps. In each step, one of the agents randomly selected. The selected agents can decide the activity referring the other agent's activities and the broadcasted levy plans. To decide the activity, the agent calculates the expected rewards and the expected revenue. The expected rewards corresponding the activity, a^i can be obtained from the following equation.

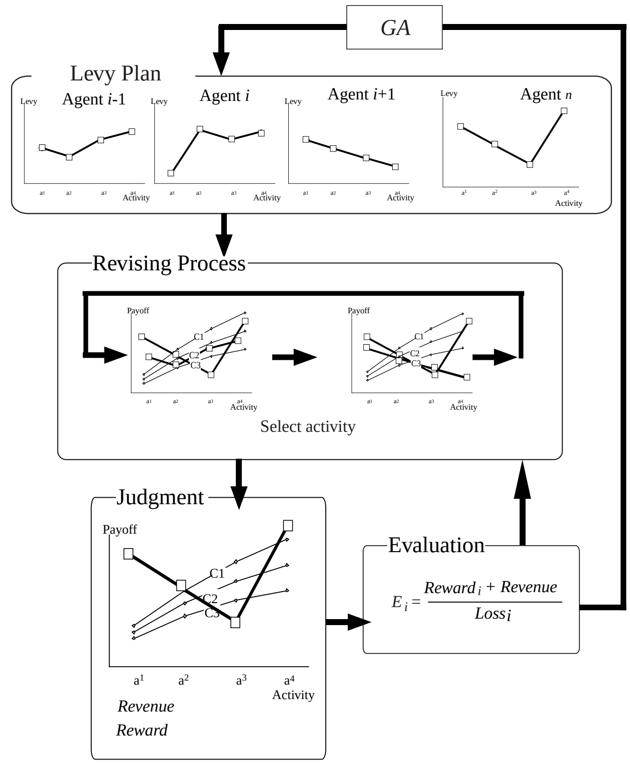

Figure 3: Proposed agent-based simulation process.

$$Reward(a^i) = Payoff(a^i, TA) - Lv^i_{min} \tag{3}$$

Where, Lv^i_{min} is minimum levy value among the broadcasted levy values. Namely, in the proposed method, more than one agent can become the meta-agent. Therefore, some levy plans are simultaneously broadcasted. For example, the levy plans, Lv_k and Lv_h, are broadcasted from the agent k and h respectively. Among the several levy values corresponding the activity a^i, it is supposed only the minimum levy value can be charged to the activity and only the proposed agent of the levy value receives the incoming levy. For example, if $Lv^i_k \in Lv_k$ is smaller than $Lv^i_h \in Lv_h$, the agent k can receive the incoming levy Lv^i_k. If no meta-agent exist, Lv^i_{min} is treated as 0. The competition between the meta-agents is also expected to prohibit the selfish behavior of the meta-agents.

Concerning the expected revenue, the value can be obtained from the following equation.

$$Revenue = \sum_i Lv'^i |A^i|, \tag{4}$$

$$where \quad Lv'^i = \begin{cases} Lv^i & \text{if } Lv^i < Lv^i_{min} \\ 0 & \text{otherwise} \end{cases}$$

In the above equation, $|A^i|$ means the number of agents taking the activity a^i.

Because the competition between the meta-agents, the expected revenue can be added if the prepared levy value, Lv^i, is smaller than Lv^i_{min}.

Throughout the calculations for the expected rewards and the expected revenue, the activity is selected to maximize the incoming value. If the expected revenue exceeds the expected rewards, the agent selects the role of the meta-agent. Otherwise, the agent acts as the player selecting the activity corresponding the maximum expected reward.

In the revising process, the agent decides the activity in turn. The length of the revising process is assumed to be enough for that one agent can revise the own decision in several times. When the revising process is over, the final decisions of the agents are applied to the game. Then the rewards and the revenue are decided. To avoid the tragedy situations, it is required that the suitable levy plans are proposed in the revising process. To acquire the suitable levy plan, the genetic algorithm is applied.

Evolution of levy plan

To make the suitable levy plan, the genetic algorithm is applied. Each agent has the population of chromosomes. The chromosome consists of the blocks. Each block is decoded to the value of levy in the levy plan. In the following simulation, the value of the levy is encoded by binary representation in the block. The levy plan decoded from the chromosome is applied to the game. After the game, the agent with the applied levy plan gets the reward and revenue. The evaluation of the levy plan is based on the reward, revenue, and the following loss value. The loss ,$Loss_i$, of agent i is determend as follows:

$$Loss_i = \begin{cases} Loss_i^p & \text{if } a_i \neq a^{meta} \\ Loss_i^m & \text{if } a_i = a^{meta} \end{cases} \quad (5)$$

$$Loss_i^p = \sum_j (Lv_i^j - Lv_{min}^j)|A^j| \quad (6)$$

$$Loss_i^m = \sum_j Y_j |A^j| \quad (7)$$

$$where \quad Y_j = \begin{cases} (Lv_i^j - Payoff(a^j, TA))|A^j| \\ \quad \text{if } Lv_i^j > Payoff(a^j, TA) \\ 0 \quad \text{otherwise} \end{cases}$$

In the above equations, $Loss_i^p$ is the loss value for the agent i as the player. Therefore, the levy plan of the agent i isn't applied to the game. The situation can be considered that the levy plan of the agent i is inferior to the applied levy plans of others. Thus, the deference between the levy plan of the agent i and the applied levy plans of others is considered as the loss value like as the equation 6. $Loss_i^m$ is the loss value for the agent i playing the role of the meta-agent. In this case, the loss is occurred if $Lv_i^j > Payoff(a^j, TA)$.

Using the reward, the revenue and the loss, the evaluation of the chromosome, which generates the levy plan, is defined as follows:

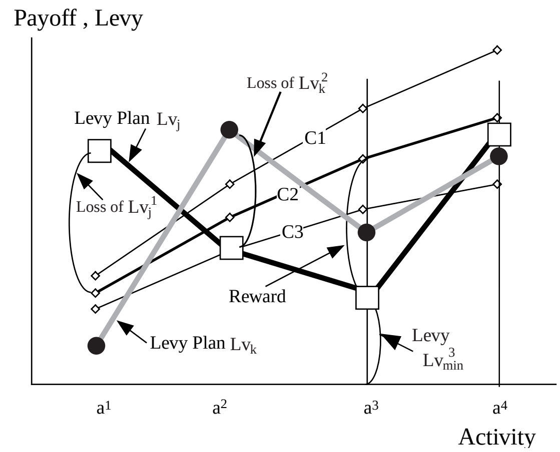

Figure 4: Relation between the payoffs and levy plan concerning reward, levy and loss.

$$E_i = \frac{Reward(a_i) + Revenue}{Loss_i} \quad (8)$$

In the simulations, the game is iterated in several times against to the same levy plan, because the averaged evaluation value would like to be considered.

After the evaluation values are fixed for all of the chromosomes, the genetic operations are applied. Namely, the crossover, the mutation and reproduce are applied in turn.

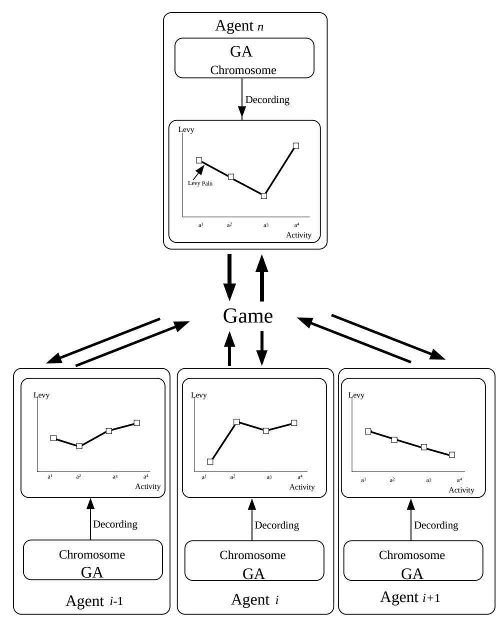

Figure 5: Each agent applies genetic algorithms to acquire suitable levy plans.

Simulation

To confirm the effectiveness of the proposed methods for avoiding the tragedy situation in the social dilemmas, the simulation is executed. The payoff function is set as follows;

$$Payoff_i = a^i(|A| \times N' - \sum_{j}^{N'} a^j) - 2a^i \tag{9}$$

where N' is the number of agents excepting the number of meta-agents.

The number of agents is 7. The activities for each agent are 7 and a^{meta}. Each agent has 30 chromosomes. According to the decoded plan, the game, The tragedy of the common, is iterated 10 times. In each game, the time of the revising process is 21. Namely, 3 times for revising chances are given for each agent. The averaged evaluation values in the iterated games are given as fitness of the chromosomes. The crossover and mutation are applied the chromosomes. The crossover rate is 1.0 and the mutation rate is 0.05. Under these parameters, the evolution of the levy plans in the agents are proceeded until 200 generations. The calculation time is about 9 hours for one simulation with Pentium 4, 1.5 GHz. The example of the decision process is shown in Fig.6.

In this figure, a^{meta} means that the agent selects the role of meta-agent and others are quantity of the activity as the players. From the decision process, two agents selected the role of meta-agent autonomously.

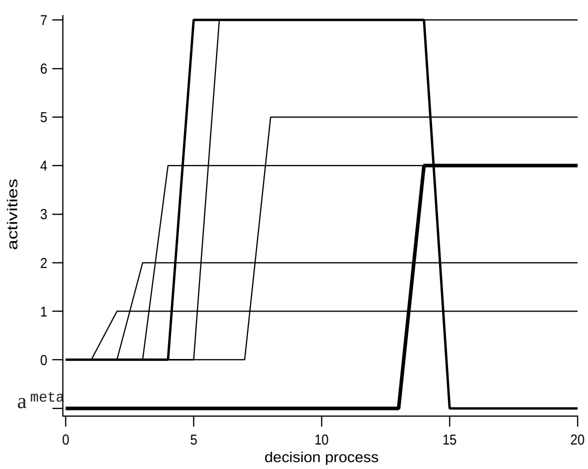

Figure 6: Example decision process in the Tragedy of the Common game

To comfirm the effect of the meta-agents, the limitation of the number of the meta-agent is evaluation. In this experiment, the case of only one agent can become meta-agent and the case of two are compared. Concerning two cases, the averaged reward and averaged lost are compared. The averaged reward is shown in Fig.7 . The averaged lost is shown in Fig.8. From both figures, the case of 1 meta-agent become higher reward and lower lost. Therefore, the number of meta-agents can be condidered as suitable one in this problem.

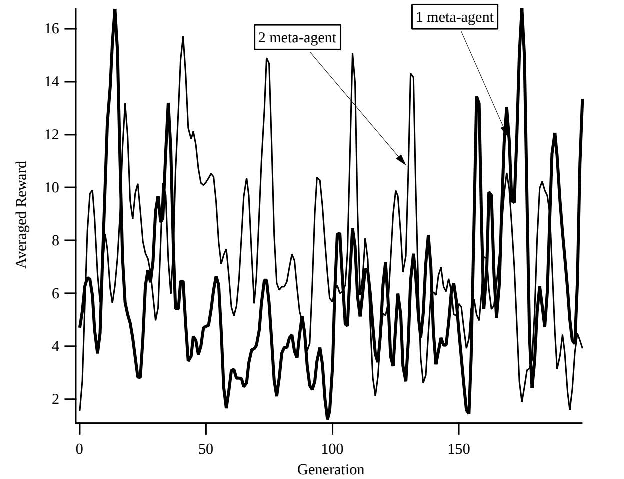

Figure 7: Progress of averaged reward concerning the number of meta-agents.

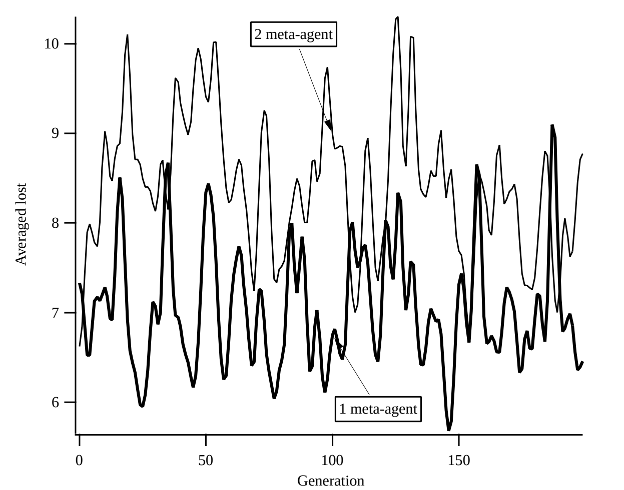

Figure 8: Progress of averaged lost concerning the number of meta-agents.

Concerning to the activities of the agents, the effect of the number of the meta-agent-players are compared. The averaged selection rate of the activities is made from the generation 180 to 200. Fig.9 shows the case of no-meta-agent-players. Namely, the agents determine the activity from only own expected payoff prediction. In this case, apparently tragedy situation occurred. Fig.10 shows the case of the limited to one meta-agent and Fig.11 shows the limited to two meta-agents. In both case, terrible tragedy situations are avoided comparing to the no-meta-agent case. Especially, the limited to one meta-agent is more distributed the activities than the case of two meta-agents. Therefore, the limited one meta-agent-player is effective to avoid the tragedy situation than multiple meta-agent case.

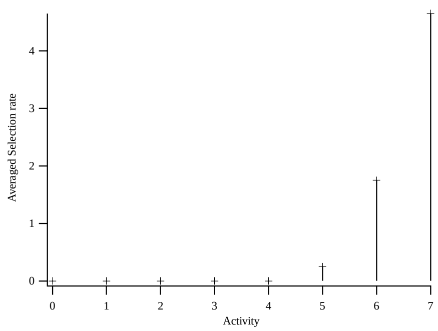

Figure 9: Averaged selection rate of activities with no meta-agent-player.

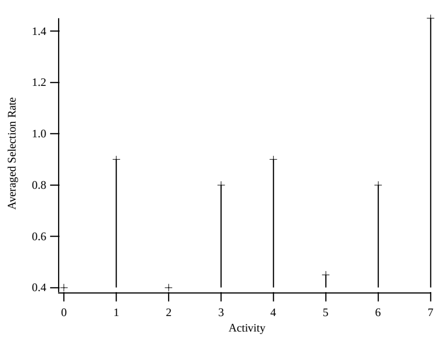

Figure 10: Averaged selection rate of activities with limited one meta-agent-player.

Conclusion

This paper shows the agent based simulation with changing roles in the social dilemma game, the tragedy of the common. To control the usage of a common resource, meta-agents are introduced in the game. The objective of the meta-agents is to get the revenue from the agents according to their activities. To introduce the meta-agents, it should be considered that who became the meta-agents and how to make the charging plan of levy against the activities of the agents. Therefore, it is proposed the decision making

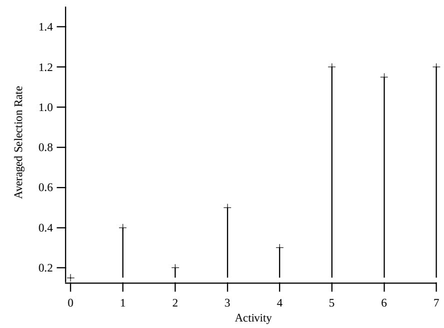

Figure 11: Averaged selection rate of activities with limited two meta-agent-player.

of the agents that includes the selection of roles. Namely, the agents have the levy plan and the payoff prediction. The agent determines the activity including selection of the meta-agent role by comparing the expected payoff and the expected revenue. To make the suitable levy plan and the payoff prediction, the coevolution of both internal plans are employed. From the experiment, the proposed autonomous selection of meta-agent well performed to avoid the tragedy situation. In this simulation, the GA is applied to acquire the levy plans, however, the other adaptation techniques seems applicable. For example, the reinforcement learning will be useful for getting the suitable levy plan.

Concerning the number of meta-agents, limited one meta-agent-player is more effective than the case of competitive two meta-agents. The effect of the competition between the multiple meta-agents will be required to analyze in future work.

References

Suleiman, R., Troitzsch, K. G., Gilbert, N. eds. 2000. *Tools and Techniques for social Sience Simulation*, Springer.

Namatame, A., Terano, T., Kurumatani, K., eds.2002. *Agent-Based Approaches in Economic and Aocial Complex Systems*, IOS Press.

Conte, R., Hegselmann, R., Terna, P. eds. 1997. *Simulating Social Phenomena*, Springer.

Hardin, G.1968. The Tragedy of the Commons, *Science* 162:1243.

Yamashita, T., Suzuki, K., Ohuchi, A. 2001. Distributed Social Dilemma with Competitive Meta-players, *Int. Trans. in Operational Research* Vol.8, No.1:75–88

Yao, X. 1996. Evolutionary stability in the N-person prisoner's dilemma,*BioSystems*, **37**, 189–197.

Suzuki, M.1994, *New Geme Theory*, Soukei-Shobou, (in Japanese)

Liebrand, W., Messick, D. eds. 1996. Frontiers in Social Dilemmas Research, *Springer*.