

ジレンマゲームにおけるメタプレイヤーの自律的発現手法に関する研究

鈴木 恵二†

† 公立はこだて未来大学 〒041-8655 函館市亀田中野町 116-2

E-mail: †suzukj@fun.ac.jp

あらまし エージェントベースシミュレーションを用いて、エージェント間の様々な社会的インタラクションを考察していくためには多様な行動選択メカニズムの発展が必要となる。ここでは「共有地の悲劇」を取り上げ、プレイヤーの行動を変えさせるメタプレイヤーの役割を設定する。このメタプレイヤーをプレイヤー群から自律的に発現させるための意思決定機構と GA を用いた適応について報告する。このアプローチにより、プレイヤー群が悲劇的状况に陥らないようになることを示す。

キーワード エージェントベースシミュレーション、メタプレイヤー、共有地の悲劇

A study on autonomously emergent of meta-players in dilemma game

Keiji SUZUKI†

† Future University-Hakodate Kamedanakano 116-2, Hakodate, 041-8655 Japan

E-mail: †suzukj@fun.ac.jp

Abstract 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, 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 of the agents that includes the selection of roles. Concerning to adjusting the charging plans, evolutionary approach is employed. Throughout the experiments, the effects of the meta-agent selection and the formation of the charging plan are examined.

Key words Agent-based simulation, meta-player, dilemma

1. ま え が き

エージェントベースシミュレーションはより複雑なゲーム状況における研究やエージェントの知能化技術を吟味するために有効な手段となっている [1] [2]。社会的ジレンマ問題は複雑なゲーム状況の一つであり、エージェントの判断能力を吟味する上で適切な題材となっている [3]。ここでは社会的ジレンマの一つである共有地の悲劇 [4] を取り上げる。このゲームではプレイヤーは利得を得るために共有資源を利用する。このため各プレイヤーが利己的に振る舞うならば、共有資源は枯渇し、すべてのプレイヤーの利得は低いものになってしまう。このような悲劇的状况を避けるためには、各プレイヤーは他のプレイヤーと協調的な振舞いを取る関係を構築するか、問題構造を変更する、すなわち、利得関数の構造を変更する必要がある。

囚人のジレンマゲームや N 人囚人のジレンマゲームでは、こうした利己的な振舞いを避けるために、エージェントの能力を拡張し、様々な行動規範を導入する試みがなされている [8]。こ

のようなアプローチでは、すべてのエージェントが同じ能力を保持し規範が守られるならば、ジレンマ状況を回避するのに十分なアプローチとなっている。しかしながら、より複雑な状況ではエージェントの能力はヘテロなものとなり、上記の前提が守られるとは限らなくなる。よって、他のアプローチを考慮する必要がある。このようなアプローチの一つとして利得関数の構造を変更するアプローチがある。すなわち、メタエージェントを導入し、プレイヤーの行動に応じて徴収金を取ることにより、プレイヤーの利得構造を変更しようとするものである [5]。

しかし、このようなメタプレイヤーの導入には 2 つの問題がある。一つはだれがメタプレイヤーになるべきか、もう一つはメタプレイヤーの振舞い、すなわち課金のかけ方をどのように設定するのかという問題である。

一つ目の問題に対するアプローチとして、ここではプレイヤーに通常の行動選択に加え、役割変更の選択肢を加えることにより、状況に応じてプレイヤーからメタプレイヤーに、あるいはその逆に役割を自律的に変更するメカニズムの導入を試みる。

表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

2つ目のメタプレイヤーの振舞いをどのように決めるか、すなわち課金のかけ方をどのように決めるかについては、まずあえて、メタプレイヤーの目的を総徴収金額を最大化することとする。この目的のもと、遺伝的アルゴリズム (GA) を用いて、適応的に課金プランを構成する方法を提案するとともに、課金を徴収するルールの設定によりメタプレイヤーになって徴収金を得るためのメタプレイヤー間の競争によりメタプレイヤーの利己的な振舞いを抑制するとともに、適切な課金プランが形成されるようにする。

この役割変更を含む行動選択メカニズムと課税プランの遺伝的獲得を導入したエージェントの能力を共有地の悲劇問題に適用することにより吟味することとする。

以下ではまず、共有地の悲劇について記述し、その後提案するアプローチの詳細を示す。

2. 共有地の悲劇

共有地の悲劇 [4] は N 人社会的ジレンマの一つとして知られている [6]。このゲームは限られた資源を共有するプレイヤーの振舞いを分析するものである。共有資源が限られているので、高い利得を得ようとするエージェントが増えるほど、利益が減っていく特徴がある。共有地の悲劇における利得関数は以下のようになる [4]。

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

ここで $Payoff(a_i, TA)$ はエージェント i の利得。 a_i はエージェントの活動度。 TA は全体の活動度、 $\sum_j^N a_j$ 。 A は活動度の集合。 N はプレイヤーの数である。

具体例として、 $N=4$ 、活動度が4つの場合、 $a_i \in \{0, 1, 2, 3\}$ 、からなる場合を記す [7]。また表1にこの場合の利得表を示す。

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

このゲームではプレイヤーは有限の共有資源を消費する活動を行う。従ってすべてのプレイヤーの活動度が増えるに従って利得は減少する。しかしながらより高い活動が常に低い活動度に対して支配的であるためにプレイヤーはより高い活動度を選択しようとする。よって利己的に振舞っている限りはプレイヤーはより高い活動度を選択しようとし、その結果、共有資源は枯渇することになり、悲劇的な結末を迎えることになる。この例では、活動度の総和が12に達したとき、悲劇が起き、プレイヤー i が活動度を2から3へ上げたとしても利得が上らなくなる。

このゲームの特徴は、技術的解が存在しないことである。そのため、このゲームを解くためにはプレイヤーの個人合理性を変更するか、利得関数の構造を変更するしかない。エージェント

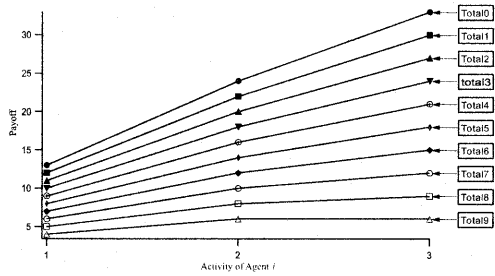


図1 Graphical view of example payoff value in Table.1

ベースシミュレーションの目的は、どのような合理性の基準が適切か、あるいは問題構造の変更が意義あるものとなるのか検証することにある。ここでは問題構造を拡張する方法、すなわち、メタエージェントの導入により利己的に振舞うエージェントにより悲劇が生じないようにすることを試みる。次節で提案するアプローチを示す。

3. アプローチ

ここでは社会的ジレンマに対して、課金による制御戦略を導入する (図2)。この方法では、活動度に対する課金プランが重要になる。事前にゲームの設定から適切な課金プランを分析し設定することは可能である。しかしここでは、シミュレーション中のエージェントが状況に応じて適応的に適切な課金プランが獲得できるかどうかについて試みる。課金プランを獲得し適用する役割としてメタプレイヤーを導入する。メタプレイヤーは課金プランをプレイヤーに告知し、各プレイヤーの活動度に応じて課金を課す役割を担うものとする。このようなメタプレイヤーの導入にあたっては、誰がメタプレイヤーとなるか、またどのように課金プランを設定するかが問題となる。誰がメタプレイヤーとなるかという問題に対するために、プレイヤーは自分の活動度を定めるだけでなく、メタプレイヤーとなるかプレイヤーとなるかの役割の選択も行うものとする。この役割選択により自律的にメタプレイヤーが発現することが期待される。課金プランの獲得に当たっては、メタプレイヤーの利己的な振舞いを抑制するためにプレイヤーへの課金ルールを設定する。エージェントの構造に関する詳細を次に示す。

3.1 エージェントアーキテクチャ

エージェント集団を N とする。 $N = \{1, 2, \dots, n\}$ 。 $A = \{a^1, a^2, \dots, a^k, a^{meta}\}$ をエージェントの取る活動度とする。活動度の大小関係を $a^{i+1} > a^i$ と仮定する。また a^{meta} はエージェントがメタプレイヤーとして行動することを表すものとする。すなわち、メタプレイヤーとしての振舞いはエージェントの一行動として位置づけられる。

メタプレイヤーとなることを選び取るかどうかは課金プランに拠るものとする。全てのエージェントは各々課金プランを有し、プレイヤーの活動度の状況から期待徴収額を計算し、メタプレイヤーになるべきかどうかを判断するものとする。もし、あるエージェントがメタプレイヤーになることを選んだ場合には、そ

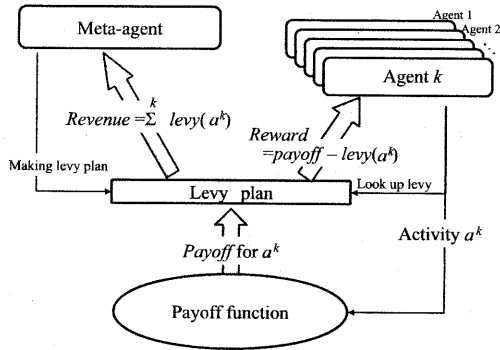


図2 Schematic view of levy based control strategy.

のエージェントの課金プランが全てのエージェントに告知される。一方、プレイヤーを選び取ったエージェントは、課された課金プランに従って自分の利得が最大となるように活動度を選択するものとする。各エージェントは、メタプレイヤーとなってプレイヤーから徴収金を集めるか、もしくはプレイヤーとして活動度を選択するかのいずれかを選び取ることになるが、この選択は期待されるどちらかの報酬が最大となるかで決めるものとする。すなわち、エージェントはその役割にかかわらず報酬、すなわち徴収金か利得のどちらか、を最大化するように利己的に振る舞うものとする。次に課金プランの詳細を示す。

3.2 課金プランと徴収ルール

課金プランは、活動度 $\{a^1, a^2, \dots, a^k\}$ に応じた課金値 $\{Lv^1, Lv^2, \dots, Lv^k\}$ から構成される。すなわち、活動度 a^i に対して課金値 Lv^i が課金される。

この課金が課されることによりもとの利得関数からの利得が変更されることになる。したがってたとえプレイヤーが利己的に活動度を選択するものとしても、適切な課金プランが設定されていれば、共有資源の枯渇を避けることが可能となる。しかしながら適切な課金プランをどのように形成するかが問題となる。この問題は課金プランを設定するためのポリシーに依存する。ここではプランの形成におけるポリシーとして個人合理性を仮定する。すなわち、メタプレイヤーの目的は課金から得られる徴収の和を最大化するものとする。個人合理性のみを仮定することによって、メタプレイヤーにそれ以外の特別な特性を持たせないこととする。しかしながら個人合理性のみを設定することによって、メタプレイヤーが利己的に振舞い、課金額を増加し続ける恐れがある。このような行動を抑制するために単純な徴収ルールを設定する。徴収ルールとして、徴収が可能なのは利得が課金を上回っている場合、 $Payoff(a^i, TA) > Lv^i$ の場合に限るものとする。もし、課金額が利得を上回った場合にはメタプレイヤーおよびプレイヤーの利益は共に0とする。この単純ではあるが、徴収ルールによりメタプレイヤーの利己的な振舞いが抑制されると期待される。課金プランは、各々のエージェン

ト毎に遺伝的アルゴリズム (GA) により形成されるものとする。課金プランの GA による形成については後述する。

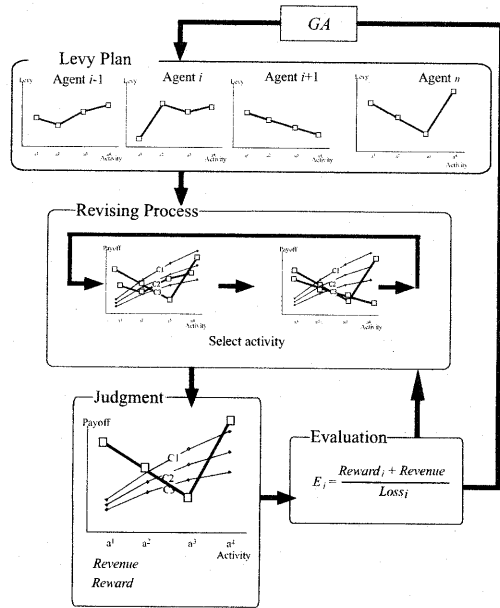


図3 Proposed agent-based simulation process.

3.3 意思決定と再考プロセス

エージェントの行動を決定するために再考プロセスを用意する。このゲームにおいては、利得は他のエージェントの活動度の総和、すなわち TA によって変化する。したがって行動を決定するためには、他のエージェントの状況を知る必要がある。再考プロセスは一定の繰り返しのステップからなる。各ステップでは、ランダムにエージェントが選択され、行動を選択する。選択されたエージェントはその時点での他のプレイヤーの活動度と課金プランから自分の行動を決定する。この行動選択においてエージェントは、メタプレイヤーになった場合の期待総徴収額と、プレイヤーとして適切な活動度を選択した場合の期待利得をそれぞれ計算する。このときの活動度 a^i への期待利得は以下の式で表される。

$$Reward(a^i) = Payoff(a^i, TA) - Lv_{min}^i \quad (3)$$

ここでまず、メタプレイヤーは複数存在する場合があることを注意しておく。 Lv_{min}^i はメタプレイヤー達から出された課金プランのうちで、活動度 a^i に対して最小となっている課金額を示す。すなわち、課金プランが複数存在した場合には各活動度 a^i において、最少額となっている課金を適用し、徴収額はその最小課金額を提示したメタプレイヤーに支払われる。例えば、 $Lv_k^i \in Lv_k$ が $Lv_h^i \in Lv_k$ よりも少額であった場合には、メタプレイヤーとなっているエージェント k が課金額 Lv_k^i を徴収することができる。もし、だれもメタプレイヤーになっていない場合には、 Lv_{min}^i を0として扱う。この徴収に関するルールはメ

タブレイヤの利己的な振舞いを抑制するものとして期待される。一方、各エージェントが有する課金プランに従ってメタプレイヤになった場合の期待総徴収額は以下の式で表される。

$$\text{Revenue} = \sum_i Lv^{i^*} |A^i|, \quad (4)$$

$$\text{where } Lv^{i^*} = \begin{cases} Lv^i & \text{if } Lv^i < Lv_{min}^i \\ 0 & \text{otherwise} \end{cases}$$

ここで、 $|A^i|$ は活動度 a^i を取っているプレイヤの数である。徴収金は他のメタプレイヤの示した課金額よりも少額であった場合に得ることができる。

以上の期待利得と、期待総徴収額の計算を通して、エージェントは報酬を最大化するように行動を決定する。もし、期待総徴収額がプレイヤとして振る舞った場合の期待利得よりも大きい場合にはエージェントはメタプレイヤとしての行動を選択する。逆の場合には、プレイヤとして利得が最大となる活動度を選択する。

再考プロセスではこの行動選択をランダムにエージェントを選びながら一定期間繰り返すものとする。そして再考プロセスの期間が終了した時点で、各エージェントが最後に選択した行動を最終決断として固定し、実際の報酬の決定を行う。

この再考プロセスを通して、適切な課金プランを有するエージェントがメタエージェントとして現れれば、悲劇的状况を避けることができると期待される。次にこの課金プランを獲得させるための GA の適用方法を示す。

3.4 課金プランの進化的獲得

適切な課金プランを獲得するために GA を適用する。各エージェントはそれぞれ遺伝子集団を有するものとする。遺伝子はブロックからなり、各ブロックは課金プラン内の課金額を表すものとする。後述するシミュレーションでは、課金額を二進表現により表すものとした。遺伝子表現からデコードされた課金プランはゲームに適用されて利得や徴収額としての報酬を獲得することになる。課金プランの評価としてこの報酬に加え、損失値を考慮するものとした。エージェント i の損失値は次の式で与えられる。

$$\text{Loss}_i = \sum_j (Lv_j^i - Lv_{min}^j) |A^j| \quad (5)$$

上記の式において、 Loss_i^j はエージェント i がプレイヤとして行動した場合の損失値であり、保持する課金プランが他のエージェントが有する課金プランよりも適切さにかけていたためにメタプレイヤに成れなかったことに対する損失と考える。よって実際にメタプレイヤにより掛けられた課金プランと保持する課金プランとの差を損失値として計算する。

以上の利得、総徴収額、損失値を使って遺伝子を評価する。評価関数は以下の通りである。

$$E_i = \frac{\text{Reward}(a_i) + \text{Revenue}}{\text{Loss}_i} \quad (6)$$

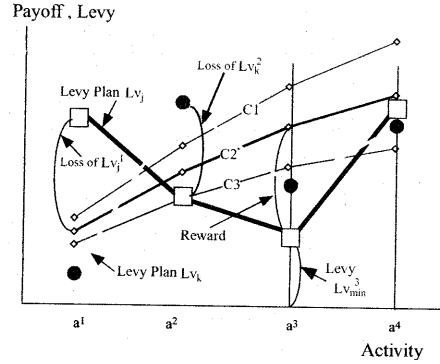


図4 Relation between the payoffs and levy plan concerning reward, levy and loss.

このシミュレーション過程では同じ課金プランを使って、数回ゲームが行われるものとし、その時の平均評価値を遺伝子の評価値として適用するものとする。すべての遺伝子の評価値が定まった後に、遺伝的操作である、交差、突然変異、そして世代交代を適用する。

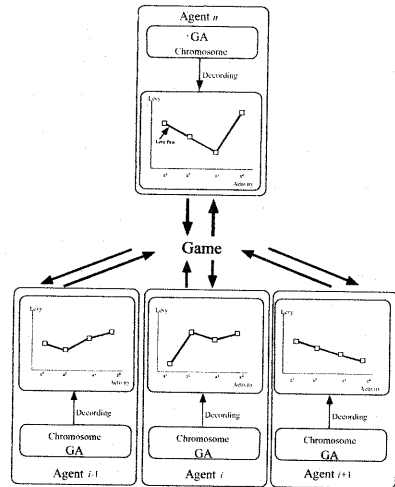


図5 Each agent applies genetic algorithms to acquire suitable levy plans.

4. シミュレーション

提案手法の有効性を検証するためにシミュレーションを行った。利得関数は以下のように設定した。

$$\text{Payoff } f_i = a^i (|A| \times N' - \sum_j a^j) - 2a^i \quad (7)$$

ここで N' はメタプレイヤの数を除いたエージェントの数、すなわち、プレイヤの数である。

ここでエージェントの数は6と12、30の場合について検証

を行った。エージェントの行動選択は、活動度が6と a^{meta} となる。各エージェントは遺伝子集団数として30を有する。課金プランをデコードしたうえで、共有地の悲劇のゲームを10回繰り返すものとした。各ゲームにおいて再考プロセスの長さは各エージェントが4回意思決定を行うことができるステップ数とした。繰り返しゲームによって与えられる平均評価値は課金プランの元となる遺伝子のフィットネス値として使われる。遺伝操作として交差率1.0、突然変異率0.05とした。上記のパラメータを用いて50世代まで世代交代を行った。

12エージェントによる再考プロセスの様子(図.6)と獲得された課金プランを図.7に示す。

この図において縦軸上の $meta$ はエージェントがメタプレイヤーの役割を選択したことを示し、他の値は活動度を表す。この再考プロセスにおいて、5つのエージェントがメタプレイヤーの役割を自律的に選び取っている様子が示されている。ステップの初期においては、高い活動度を選び取るエージェントが現れるが、メタプレイヤーの出現によりそのような活動が抑制されていることがわかる。次にステップ20において一度メタプレイヤーの役割を担うエージェントがいなくなり、それまでメタプレイヤーとなっていたエージェントが、活動度5を選び取りフリーライダー化している点に興味深い。その後は、ステップ30までの間に、メタプレイヤーが次々に現れ、3つのエージェントが互いにメタエージェントとして競合していることが見て取れる。最終的に3つのメタプレイヤーのうち、一つがプレイヤーに戻り、2つのメタプレイヤーが残った結果となっている。

獲得された課金プランは最も高い活動度を抑制するように、高い課金が課せられているのがわかる。一方、低い活動度にも比較的高い課金が課せられ、活動度4において最も低い値となっている特徴が得られた。この課金プランは、悲劇的状况となる高い活動度を避けさせる一方、活動度があまりにも低くなり、メタプレイヤーが十分な徴収が得られなくなることのないように計画されていると考えることができる。すなわち、活動度4に低い課金を設定することで、すべてのプレイヤーにその活動度を選択させ、低い課金ながらも総和として高い徴収金額を得ようとしている戦略が獲得されたと考えられる。

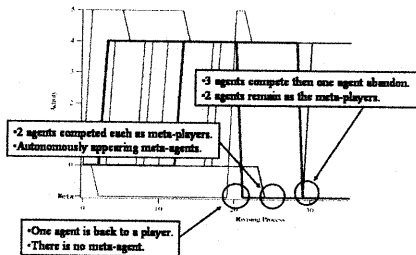


図 6 Example decision process in the Tragedy of the Common game.

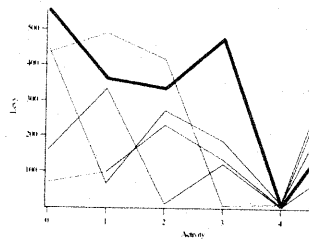


図 7 The applied tax plans in the decision process. The bold line of tax plan is applied finally.

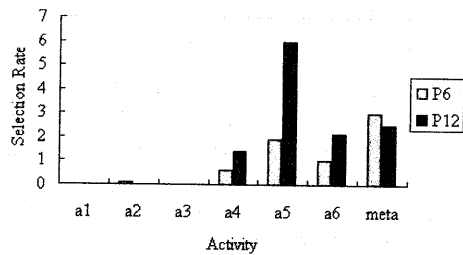


図 8 Averaged selection rate of activities among 6 and 12 players.

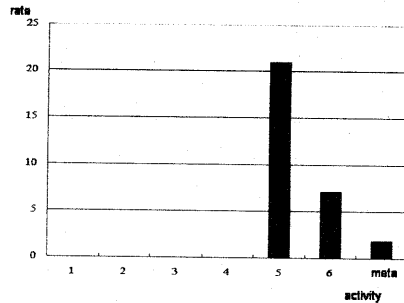


図 9 Averaged selection rate of activities among 30 players.

40から50世代における平均活動度の値を図.8および図.9に示す。この結果より、エージェントの数によらず悲劇的状况が避けられていることが確認される。

また図.10にエージェント数30におけるプレイヤーとメタプレイヤーの平均報酬値を比べた結果を示す。この結果よりプレイヤーとメタプレイヤーとの間に獲得された報酬の差はほとんどなく、メタプレイヤーの総徴収額を最大化する目的に基づいた利己的振

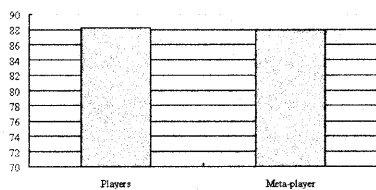


図 10 Averaged rewards of players and metaplayers

る舞いが、導入した徴収ルールとメタプレイヤーとして振舞うための競合により抑制されていることがわかり、有効性が確認された。

5. ま と め

本稿では共有地の悲劇を取り上げ、社会的ジレンマゲームに役割変更を含めたエージェントベースシミュレーションを提案した。共有資源の使用を制御するためにメタプレイヤーを導入した。メタプレイヤーの役割は、利己的に振る舞おうとするプレイヤーの活動度に応じて課金することである。このようなメタプレイヤーの導入においては、誰がメタプレイヤーとなるべきか、またどのようにメタプレイヤーのポリシーを設定するかが問題となる。これらの問題に対して、各エージェントに課金プランを保持させるとともに、エージェントの行動選択にメタプレイヤーの役割を選択できるものとして付け加えた。適切な課金プランを得るために GA を採用し、進化的操作により課金プランを形成させるとともに、単純な徴収ルールとメタプレイヤーに成ろうとするエージェント同士の競合を損失値という形で導入した。メタプレイヤーの目的は総徴収額を最大化しようとするなかで、シミュレーション実験より適切な課金プランが形成され、自律的なメタプレイヤーの発現と悲劇的状況の回避がなされることが確認された。

文 献

- [1] Suleiman, R., Troitzsch, K. G., Gilbert, N. eds. 2000. *Tools and Techniques for social Science Simulation*, Springer.
- [2] Namatame, A., Terano, T., Kurumatani, K., eds. 2002. *Agent-Based Approaches in Economic and Social Complex Systems*, IOS Press.
- [3] Conte, R., Hegselmann, R., Terna, P. eds. 1997. *Simulating Social Phenomena*, Springer.
- [4] Hardin, G. 1968. The Tragedy of the Commons, *Science* 162:1243.
- [5] 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
- [6] Yao, X. 1996. Evolutionary stability in the N-person prisoner's dilemma. *BioSystems*, 37, 189-197.
- [7] Suzuki, M. 1994. *New Game Theory*, Soukei-Shobou, (in Japanese)
- [8] Liebrand, W., Messick, D. eds. 1996. *Frontiers in Social*