

一般化メタ規範ゲームの解析

Analysis of General Meta-Norms Games

鳥海不二夫^{1*} 山本仁志² 岡田勇³
Fujio Toruimi¹, Hitoshi Yamamoto², and Isamu Okada³

¹ 東京大学

¹ The University of Tokyo

² 立正大学

² Risho University ³ 創価大学

³ Soka University

Abstract: In this paper, we build a Generic Meta-norms Game which integrates positive reaction such as reward and negative reaction such as punishment. We categorize the parts of the game using structural conditions and regions of attraction to cooperation. According the simulation results, some insights are derived such that one should use the second-level on rewarding or punishment to keep cooperation. Also, we found that games which includes meta-reward are more effective to realize cooperation dominant system easily.

1 はじめに

人間社会を単純なモデルで記述し、その本質的な性質を理解しようとしたときに、ゲーム理論の利用は有用である。特に、合理的な意思決定者の間に生じるジレンマの解決を表現するのに適したゲームとして、囚人のジレンマがある。囚人のジレンマゲームは2人非ゼロ和ゲームであり、プレイヤーは協調または裏切り行動を選択し、相手の行動に応じた得点を得ることが出来る。囚人のジレンマのナッシュ均衡は双方裏切りにあるが、パレート効率的ではない。囚人のジレンマでは、双方協調が望ましい状態であるが、どのように双方協調を実現するかが議論されることになる。

一方、囚人のジレンマを多人数に拡張したゲームとして、公共財ゲームがある。公共財ゲームは、社会的ジレンマを含む状況をモデル化したゲームであり、単純なモデルから多くの知見が得られるため、多くの研究で利用されている。

公共財への貢献を促進する仕組みとしては、Nowakらによる間接互惠性を用いる研究 [5, 4] や、Cohen ら、Santos らによる空間的な構造の特性を用いた研究 [2, 6]、Axelrod による協調しない者に対し罰則を与える規範ゲーム・メタ規範ゲームなどが提案されている [1]。このうち、Nowak や Cohen, Santos らのモデルは、行為者の意思決定は相手に対して協調するか否かだけである。

一方で、規範ゲームは、n 人囚人ジレンマの拡張として非協調者を罰するという行動原理を導入することで、集団における協調を促進しようとするものである。規範ゲームは社会にみられる多くの公共財問題に応用可能な基本モデルである。例えば二酸化炭素の排出抑制に関する国同士の利害関係問題や、ミクロなレベルでは住宅地におけるゴミ集積所の利用ルールなどがあげられる。規範ゲームが取り扱う本質的な目的は、いかに公共財へのフリーライドを抑制し協調を達成するかにある。そのため非協調者への懲罰が協調達成の仕組みとして導入されている。

一方、公共財ゲームにおいては協調者に対して評判を与えるという方法も議論されている。報酬は公共財ゲームにおいて、協調維持の代替案になりうるのだろうか。一般的には、報酬の効果は懲罰の効果よりも小さいといわれている。被験者実験を用いた分析 [7] では、協調者に対して報酬を与える場合と非協調者を罰する場合を比較して、最初のラウンドでは大多数の人は懲罰ではなく報酬を選択するが、最終的には多くの参加者は懲罰を選択すると報告されている。理論研究 [3] でも懲罰のほうが報酬よりも効果的だと指摘されている。それは、懲罰と報酬のコストの非対称性で説明できる。懲罰は協調が達成されたら必要なくなり、協調者は懲罰コストを払わなくて良くなる。報酬は、協調が維持される限り常に払い続ける必要がある。このような懲罰と報酬の関係性は、どのような条件下で成立するのであろうか。

メタ規範まで含めたゲームについては報酬の方が懲

*連絡先：東京大学大学院工学系研究科
東京都文京区本郷 7-3-1
tori@sys.t.u-tokyo.ac.jp

罰よりも効果的であるというシミュレーション研究がある [8]。一方で、報酬と懲罰が複合的に可能なゲームについてはまだ充分分析されていない。

そこで、本研究では協調行動に対する報酬、および裏切り行動に対する懲罰について、Axelrod のメタ規範ゲームを拡張した一般化メタ規範ゲームを用いて、報酬裏切りに対するメタ規範行動が協調の促進に与える影響をエージェントシミュレーションによって明らかにする。

2 一般化メタ規範ゲーム

2.1 一般化メタ規範ゲームの枠組み

本研究では規範ゲームの一般化を行い、協調に対する報酬を表現可能とした一般化メタ規範ゲームの枠組みによって公共財ゲームを表現する。

公共財に対する協調・裏切りに対する報酬と懲罰を網羅的に表現するために、一般化メタ規範ゲームを図 1 のように表現する。一般化メタ規範ゲームの概要を以下に示す。

N 人のエージェントで構成される集団を考える。エージェント i は裏切るか協調するかの二つの行為を選択することができる。また他者の行為に対して、懲罰を与える、または報酬を与えることを選択できる。協調する確率を B_i 、懲罰を与える確率を V_i 、報酬を与える確率を L_i と表現する。

i が裏切ると、 i は T の利得を得ることができる。残りの $(N - 1)$ 人エージェントは H の損を被る。 j が i の裏切りを発見した場合、 j は自身の持つ懲罰確率 V_j によって i を罰する。 j が i を罰した場合、 i は P の利得を得る。 j は E の利得を得る。罰しなかった場合、 i, j の利得に変化はない。

ここまです規範ゲームである。メタ規範ゲームとは、エージェント j が i の裏切りを発見し、更に j が i を罰しなかったことをエージェント k が発見したときに k が j を罰するという構造を導入したものである。このとき、 k は j を V_k の確率で罰し、 j は P' の利得、 k は E' の利得を得る。また、 j が i を罰したことを k が発見し、 j に報酬を与えるメタ報酬ゲームを導入する。この際、 k は報酬を与える確率 L_k によって j に報酬を与え C' の利得を得る。 j は R' の利得を得る。

続いて i が協調した場合を考える。 i が協調すると、 i は F の利得を得る。残りの $(N - 1)$ 人エージェントは M の利得を得る。 j が i の協調を発見した場合、 j は自身の持つ報酬確率 L_j の確率によって i に報酬を与える。 j が i に報酬を与えた場合、 i は R の利得を得る。 j は C の利得を得る。

エージェント j が i に報酬を与えたことを k が発見したときに k が j に報酬を与えることができる。この

際、 k は報酬を与える確率 L_k によって j に報酬を与え C'' の利得を得る。 j は R'' の利得を得る。

またエージェント j が i に報酬を与えなかったことを k が発見したときに k が j を罰することができる。この際、 k は懲罰確率 V_k によって j に懲罰を与え E'' の利得を得る。 j は P'' の利得を得る。

なお、図 1 上ではエージェントが他者の行為を発見するかどうかの分岐は省略されている。

2.2 メタ報酬ゲームのエージェントベースモデル

2.2.1 公共財ゲームの概要

公共財ゲームにおいて、協調行動は他のユーザにとつての利益となり、協調行動を取ったエージェントはコストを払うことになる。したがって、各エージェントの行動選択は表 1 のような利得表を持つ。ただし、 N_c は協調した参加者の数、 F は協調のコスト、 M は他人が協調したことによって得られる報酬、 N_d は裏切った参加者の数、 T は裏切ることによって得られる報酬、 H は裏切られることによって受ける被害である。

メタ報酬ゲームは複数ステップにわたって N 人 4 人のジレンマゲームを繰り返し行うことになるが、その各ステップは以下の二つのフェーズからなる。

1. メタ報酬ゲーム
2. エージェント進化

各フェーズについて詳細について述べる。

2.2.2 メタ規範ゲームフェーズ

一般化メタ規範ゲームは以下のような流れで行われる。

まず、 N 人のエージェントが相互にリンクでつながっているネットワークを仮定する。メタ報酬ゲームは、このネットワーク上で行われる。本稿で行うシミュレーションでは、簡単のため完全グラフをネットワークとして用いている。各エージェント i はパラメータとして、協調率 B_i と返報率 L_i を持つ。協調率と返報率はそれぞれ長さ 3 のビット列によって表現される。すなわち、エージェントは長さ 6 のビット列によって表現される。このビット列をエージェントの遺伝子と呼ぶ。

ゲーム t において、エージェント i には発見率 $(0 \leq S_{it} < 1$ の実数) が与えられる。このとき、 $S_{it} < B_i$ ならばエージェントは協調し、そうでなければ裏切り行動を取る。協調行動を行った場合、エージェント i は協調コスト F を支払い、 i に隣接するすべてのエージェントは利得 M を取得する。一方、裏切り行動を取った場

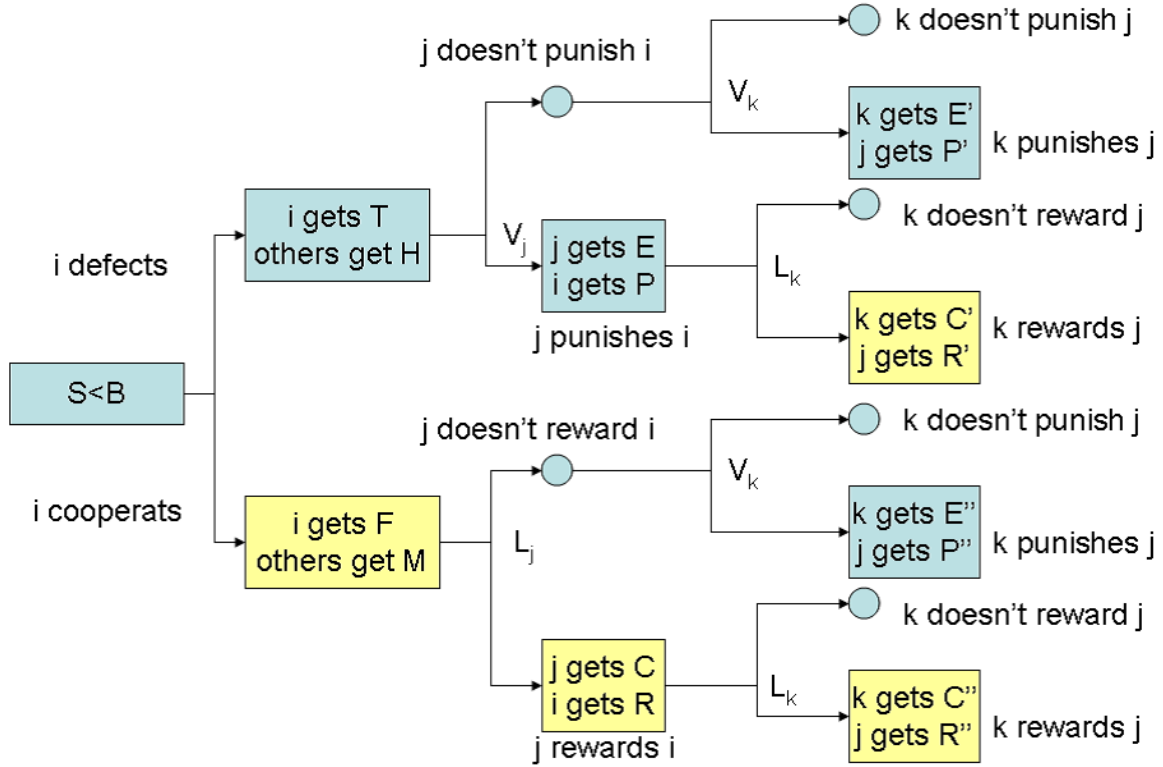


図 1: 一般化メタ規範ゲーム

表 1: 公共財ゲームにおける利得表

	協調 (C)	裏切り (D)
協調 (C)	$-F + (N_c - 1)M \setminus -F + (N_c - 1)M$	$-F - N_d H \setminus N_c M + T$
裏切り (D)	$T + N_c M \setminus -F - N_d H$	$T - N_d H \setminus T - N_d H$

合, エージェント i は利得 T を得て, 他の隣接するすべてのエージェントは裏切られたコスト H を支払う.

次に, 協調または裏切りを行ったエージェント i と隣接するエージェント j が, 確率 S_{it} で協調または裏切りに気づくかどうかを判断し, 気づいた場合は図 1 に従って報酬または懲罰を与える.

さらに, エージェント j と隣接するエージェント k が確率 S_{jt} でエージェント j の行動に気づき, 図 1 に従ってメタ報酬またはメタ懲罰を与える.

なお, 本論文の中では簡単のため,

$$R = R' = R'' = -P = -P' = -P'' \quad (1)$$

$$C = C' = C'' = E = E' = E'' \quad (2)$$

$$V_i = L_i \quad (3)$$

とする.

上記の操作を 1 ステップに 4 回繰り返し行い, 各エージェントの利得の合計を求め, 当該利得を各エージェ

ントの適応度とする.

2.2.3 エージェント進化フェーズ

各ステップの最後に遺伝的アルゴリズムによってエージェントの戦略を進化させる.

エージェントは, 隣接するエージェントおよび自身自身から適応度に応じた割合で親となるエージェントを二体選択する. ここで, エージェント i が選択される確率 Π_i は以下の式によって決定される.

$$\Pi_i = \frac{(v_i - v_{min})^2}{\sum_j (v_j - v_{min})^2} \quad (4)$$

なお, v_i はエージェント i の適応度, v_{min} は全エージェント中もっとも小さい適応度を持つエージェントの適応度である.

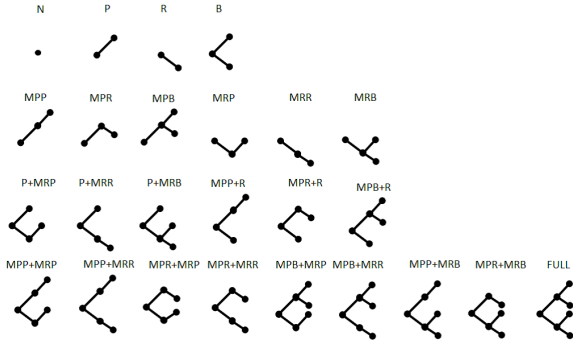


図 2: 全ゲームパターン

得られた一对の親エージェントをあらわす遺伝子を交叉させ、新たな遺伝子を得る。エージェントは親エージェント同士の交叉から得られた新たな遺伝子を自分自身の遺伝子として進化する。なお、交叉には一様交叉を採用した。

また、0.01 の確率で各遺伝子座の値がランダムに突然変異する。突然変異は当該遺伝子座の値が反転されることで表現する。

3 一般化メタ規範ゲームのシミュレーション

3.1 メタ規範ゲームの種類

公共財ゲームにおいて、報酬と懲罰の効果を確認するため、メタ規範ゲームにおける報酬、懲罰、メタ報酬、メタ懲罰の有無が、協調の進化にどのような影響を与えるかを確認する。

メタ規範ゲームに存在する規範行動は、

- 協調に対する報酬 (R)
- 裏切りに対する懲罰 (P)
- 裏切りに対する懲罰に対する報酬 (PR)
- 裏切りに対する懲罰しないことに対する懲罰 (PP)
- 協調に対する報酬に対する報酬 (RR)
- 協調に対して報酬しないことに対する懲罰 (RP)

が存在する。すべての規範行動の組み合わせを考えると、25 パターン存在する¹。全パターンについて図 2 に示す。なお、メタ規範があるゲームは MPP, MPR のようにあらわし、メタ懲罰、メタ報酬の双方がある場合は MPB, MRB のように表す。

¹協調に対する報酬 (R) が存在しなければ、協調に対する報酬に対する報酬 (RR)、協調に対して報酬しないことに対する懲罰 (RP) については考慮する必要がないため、2⁶ パターンではない。

表 2: 設定パラメータ

パラメータ	値
協調コスト F	-3
協調による利得 M	1
裏切りによる利得 T	3
裏切りられたコスト H	-1
返報コスト $C = E$	$c=2$
報酬利得および懲罰被害 $P = R$	$0.1 < r < 5.0$

3.2 シミュレーション設定

本シミュレーションにおける設定パラメータを表 2 に示す。協調したエージェントは、コスト $F = -3$ をかけ、他のエージェントに利得 $M = 1$ を与える。一方裏切ったエージェントは、利得 $T = 3$ を得て、他のエージェントに $H = -1$ の被害を与える。

規範ゲームについては、報酬、懲罰共にコスト $c = 2$ を消費し、報酬であれば対象に r の報酬を与え、懲罰であれば対象に r の被害を与える。

このとき、 r の値を 0.1 刻みで変化させることで、報酬懲罰の量と協調率との関係を確認する。

なお、各シナリオについて 100 回行い、その平均を各シミュレーションの結果として用いた。

3.3 規範ゲームにおける協調の進化

3.3.1 メタ規範を持たないゲーム

まず、メタ規範を持たない 4 種類のゲーム、すなわち N, P, R, B ゲームについて協調がどのように変化するかを確認する。

図 3 に、メタ規範の無い N, P, R, B ゲームにおける報酬懲罰量 r による協調率の変化を示す。これより、メタ規範の無いゲームにおいてはすべてのゲームにおいて協調が進化しないことが明らかとなった。

したがって、公共財ゲームにおいて懲罰報酬を用いるには、メタ規範が必要と考えられる。

3.4 懲罰によるゲーム

次に、第一段階では懲罰のみを持つゲーム、すなわち P, MPP, MPR, MPB について協調の進化を確認する。

図 4 に、それぞれのゲームの結果を示す。これより、メタ規範の無い P ゲームは協調が進化しないが、懲罰およびメタ懲罰のみを持つ MPP ゲームについても協調はそれほど進化しない。

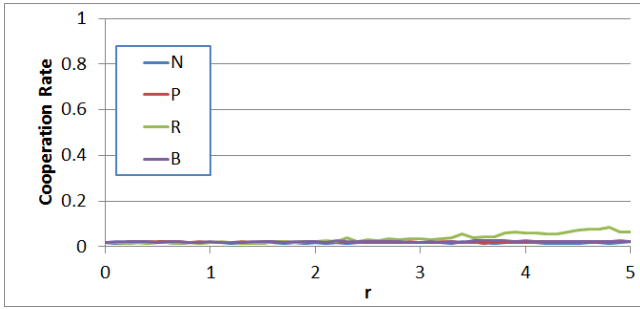


図 3: メタ規範無し

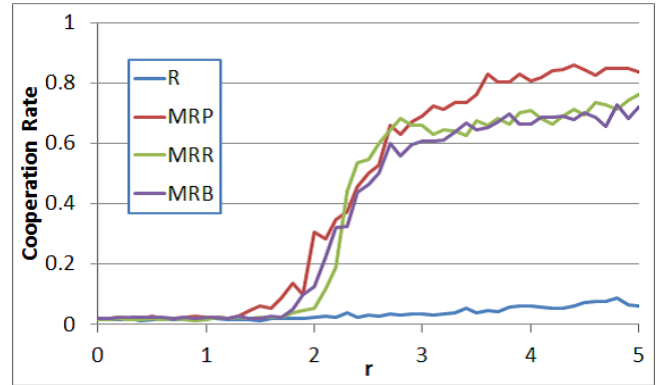


図 5: 報酬のみのゲーム

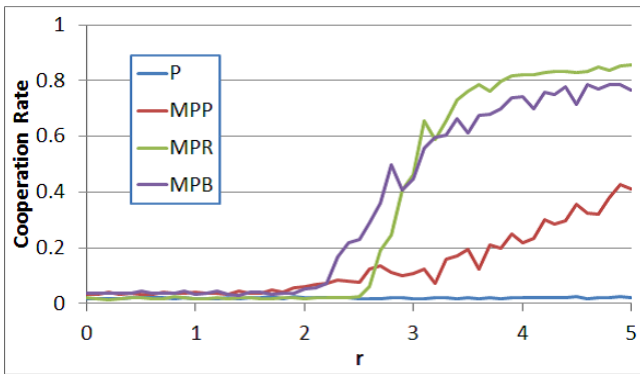


図 4: 懲罰のみのゲーム

一方懲罰を与えたことに対して報酬を与えることが可能な MPR, MPB ゲームについては r が十分大きければ協調が支配的になることが明らかとなった。

3.5 報酬によるゲーム

次に、第一段階で報酬のみを持つゲーム、すなわち R, MRP, MRR, MRB ゲームについて協調の進化を確認する。

図 5 に、それぞれのゲームの結果を示す。これより、報酬構造を持つゲームでは、メタ規範の無い R ゲームを除き、いずれも r が十分大きければ協調が支配的となることが分かる。

3.6 報酬懲罰を含むゲーム

最後に、報酬懲罰を双方含むゲームについてみる。まず、このうちメタゲームを含まない報酬または懲罰ゲームを含む $P + MRP, P + MRR, P + MRB, MPP + R, MPR + R, MPB + R$ ゲームについて見る。この結果は図 6 に示す。これより、メタ報酬を含むゲームについては協調が進化するが、メタを含まない報酬のみを含むゲーム、すなわち $MPP + R, MPR + R, MPB + R$ ゲー

ムについては、協調が支配的にはならない。 MPR, MPB ゲームについては協調が支配的になることが分かっていることから、逆にメタを含まない報酬機能を導入することで、逆に協調の進化を阻害することが明らかとなった。

それ以外のゲームについては、図 7 に示す。これより、何らかの形でメタ規範を懲罰報酬の双方で含むゲームについては、協調が支配的になることが明らかとなった。

4 考察

4.1 協調が支配的となる条件

全 25 パターンのシミュレーションの結果、協調が支配的となる条件は、

- メタ規範行動が存在する

かつ、

- 報酬行動、またはメタ報酬行動が含まれる
- メタ行動の無い報酬行動は含まれない

のいずれかを満たすことが必要であると明らかとなった。

ここで、協調が支配的とはならない、メタ行動の無い報酬行動は含まれる場合、すなわち $MPP + R, MPR + R, MPB + R$ について考える。 MPR, MPB では協調が促進されているのに対し、同じゲームでも報酬を与える行動が追加されることによって裏切りが支配的となる。なぜこのようなことが起きるのだろうか？

ここで、返報行動率 $V_i = L_i$ について見る。図 8 に、平均返報率が r によってどのように変化したかを示す。これより、 r が変化しても返報率がほぼ 0 のままであり、協調行動に対する報酬や、裏切り行動に対する懲罰が行われないことが分かる。したがって、協調行動を行うインセンティブがなくなり、結果として協調が進化しない。

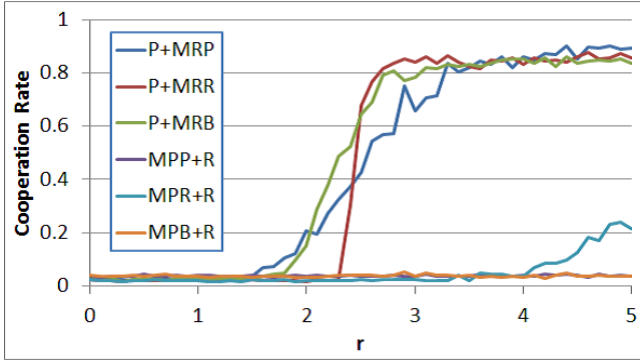


図 6: 懲罰・報酬ゲーム (メタ無しを含む)

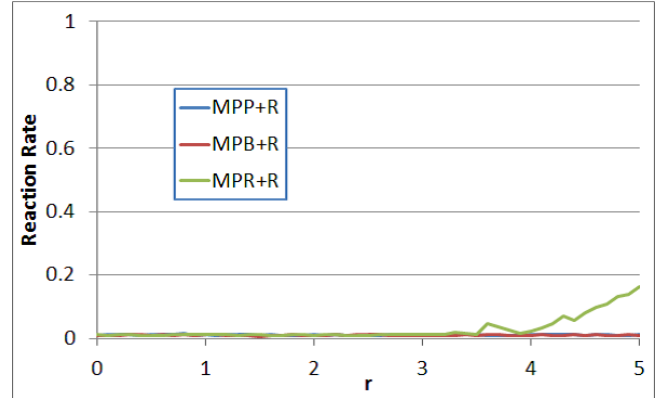


図 8: 報酬行動ありのゲームにおける返報率

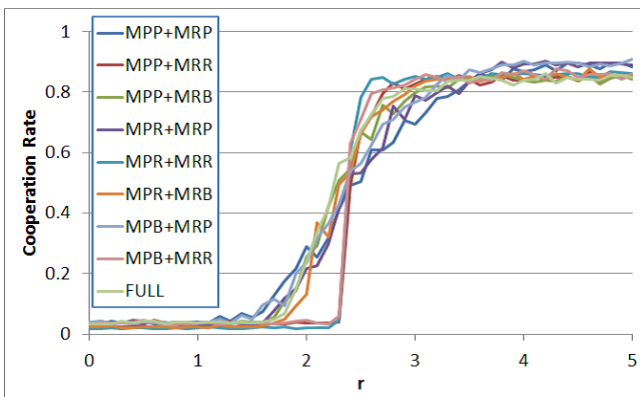


図 7: 懲罰・報酬ゲーム

では、なぜ返報行動は行われぬのだろうか。これには、報酬行動がポイントとなる。他のエージェントが協調行動を取った場合、返報性の高いエージェントは当該エージェントにコスト c を払って報酬 r を与えることになる。ただし、メタ報酬制度はないためコスト c を払った見返りはない。そのため、返報行動を取るメリットがなくなり、結果として返報行動を持たないようなエージェントが有利となる。すなわち、

1. 協調行動をとるエージェントが増加
2. 返報行動を行うエージェントの利得が減少
3. 返報行動を行わないエージェントが増加
4. 協調行動をとるエージェントが減少

という流れで、協調行動をとるエージェントがいなくなる。

逆に、懲罰行動のみがあるような $P + MRP, P + MRR, P + MRB$ では、協調行動が増えれば懲罰行動は行わないですむため、返報行動を行うエージェントの利得も減少しない。

以上より、公共財ゲームにおいては、中途半端な報酬行動の導入は協調の促進を阻害することが示された。

4.2 報酬懲罰量と協調の進化

報酬懲罰量 r の変化と各ゲームにおける協調率を確認する。このとき、現実社会でのシステムを考えると、出来るだけ少ない報酬懲罰量 r で協調が支配的になることが望ましい。そこで、協調が十分に支配的になったと判断できる協調率 \bar{B} が、

$$\bar{B} = \frac{1}{N} \sum B_i > 0.8 \quad (5)$$

となるための最小の r を確認した。

その結果、 $r < 3$ の条件下で協調率 $\bar{B} > 0.8$ となったゲームは、 $MPR + MRR, MPB + MRR, P + MRR, MPP + MRR, MRP + MRR, FULL$ のみであった。これより、報酬したことへの報酬、すなわち MRR が含まれるゲームであれば、少ない懲罰報酬量で協調が支配的となることが分かった。

以上より、協調の促進にはメタ報酬が重要な役割を果たすことが明らかとなった。したがって、公共財ゲームを基本とするような社会システムを考える際には、メタ報酬ゲームを基本に、メタ懲罰を補助的に導入することが望ましいと考えられる。

5 結論

本論文では、Axelrod のメタ規範ゲームを拡張した一般化メタ規範ゲームで考えられる全 25 種類のゲームについて、シミュレーションによって協調が促進されるゲームがどのような条件を持ったものかを確認した。協調が支配的となる条件は、

- メタ規範行動が存在する

かつ、

- 報酬行動、またはメタ報酬行動が含まれる

- メタ行動の無い報酬行動は含まれない

のいずれかを満たす必要があることが明らかとなった。また、少ない報酬懲罰量で協調が支配的になるためには、メタ報酬ゲームの存在が重要であることが示された。

今後の課題としては、本モデルの数学的な解析があげられる。メタ規範ゲームがどのような動学によって協調の進化を促進するのかを数理的に解き、そのメカニズムを明らかにする必要がある。

また、今回はモデル上での話しであったが、メタ規範ゲームにはWEB上のコミュニティサイト、ソーシャルメディアをモデル化できる可能性がある。どのようなソーシャルメディアがメタ規範ゲームのどのようなゲームに該当し、それぞれのソーシャルメディアにどのようなシステムを導入することでより協調が促進されるかを確認することも今後の課題である。

謝辞

本研究は科研費(24300064)の助成、およびJibeMobile株式会社との共同研究による助成を受けて行われたものである。

参考文献

- [1] R.M. Axelrod. An Evolutionary Approach to Norms. *American Political Science Review*, Vol. 80, No. 4, pp. 1095–1111, 1986.
- [2] M D Cohen, R L Riolo, and R Axelrod. the Role of Social Structure in the Maintenance of Cooperative Regimes. *Rationality And Society*, Vol. 13, No. 1, pp. 5–32, 2001.
- [3] Christian Hilbe and Karl Sigmund. Incentives and opportunism: from the carrot to the stick. *Proc. R. Soc. B*, Vol. 277, pp. 2427–2433, 2010.
- [4] M a Nowak and K Sigmund. The dynamics of indirect reciprocity. *Journal of theoretical biology*, Vol. 194, No. 4, pp. 561–574, 1998.
- [5] Martin Nowak and Karl Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, Vol. 393, No. June, pp. 573–577, 1998.
- [6] Francisco C Santos, Flávio L Pinheiro, Tom Lenaerts, and Jorge M Pacheco. Role of diversity in the evolution of cooperation. *Journal of theoretical biology*, Vol. 299, pp. 88–96, September 2011.
- [7] Matthias Sutter, Stefan Haigner, and Martin G. Kocher. Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations. *Review of Economic Studies*, Vol. 77, No. 4, pp. 1540–1566, 2010.
- [8] 鳥海不二夫, 山本仁志. ソーシャルメディアにおける協調の進化. *情報処理学会論文誌*, Vol. 53, No. 11, pp. 2507–2515, nov 2012.