協力の構築に必要とされる罰する者の役割に関する研究

〇大平哲史[*1*2]

[*1] 慶應義塾大学環境情報学部非常勤講師

[*2] 青山学院大学附置情報メディアセンター助教

[*1]tetsushi@sfc.keio.ac.jp, [*2]odaira@aim.aoyama.ac.jp

キーワード：協力の構築、罰する者の役割、空間囚人のジレンマゲーム

1　はじめに

非血縁関係下の協力の構築における、罰する者の必要性に疑問を呈する研究が多い。

例として、囚人のジレンマゲームでは、Dreber et al.(2008)がピア懲罰の導入によって、プール懲罰と同じく、プレイヤー全体の利得が減少することを示しており、さらに囚人のジレンマゲームの多人数版である公共財ゲームでは、Rand and Nowak (2011)が、罰は自然選択を通じて協力を促進せず、裏切り者による協力者に対する報復という非社会的な行為を助長するため、罰はプレイヤー自身を守るための自己中心的な手段に過ぎないと述べている。さらに、Nowak (2006)は、罰は協力の進化（協力者数の増加）に必要な機構ではなく、間接互恵性や群選択、ネットワーク互恵性（相互依存関係）といった、他の機構によって達成された協力の程度を促進するものであると示している。

こうした疑問を解消するため、著者は最近の研究(Ohdaira, 2016)で、各プレイヤーが相手と自分の利得の差が大きいほど、その利得の差に応じた確率で、一定の割合で自分の利得を犠牲にし、その犠牲と同等の罰を裏切り者の相手に与えるという、新しい罰の概念を提案した。犠牲と罰の値は、自分の利得に比例しているため、自分の利得が高いほど、裏切り者の相手に強力な罰を与えることができる。

確率的罰を考慮した研究は既に存在するが、Ohdaira (2016)のように、相手を罰する確率が、相手と自分の利得の差に応じて動的に変化するという点を考慮した研究は存在しない。具体的に、著者が Ohdaira (2016)で提案した罰と他の確率的罰との違いについて述べると、まず Chen et al. (2014)は、裏切り者を罰する責任を分配するための分かりやすい方法として、確率的罰を考慮しているが、各プレイヤー間でその確率に差はなく、動的に変化しない。決まった数のプレイヤーの分類とその分類に特定の確率を考慮した Chen et al. (2015)や、Perc and Szolnoki (2015)による確率を導入した不均一な罰に関する研究も、同じく固定された確率について扱っている。Szolnoki and Perc (2013)は、相手と自分の利得の差ではなく、グループ内での自分以外の条件的、非条件的な罰する者の数に応じた確率について考慮している。

Ohdaira (2016)は、提案した新しい罰の導入が、その機構によって自己中心的な行動を抑制し、協力が進化しやすいとされる、限られたプレイヤーが極めて多数の結びつきを持つ Scale-free と呼ばれるトポロジーのプレイヤー同士の相互依存関係だけではなく、さまざまなトポロジーのプレイヤー同士の相互依存関係において協力の進化を促進し、平均利得の向上にも寄与することを示している。

一方で、その結果は新しい罰の概念における中心的な制御パラメータである罰の係数、すなわち自分の利得の犠牲と、相手に与える罰の程度が一定という限られた場合のものであり、罰の係数を変化させたときも、新しい罰の導入によって協力の進化が促進され、プレイヤー全体の平均利得も増加するのか、またプレイヤー同士の相互依存関係の密度の違いにより、協力の進化とプレイヤー全体の平均利得の増加に最適な罰の係数に違いはあるのか、といった点が未解決である。

そのため、本研究はこうした未解決となっている問題について調べ、その結果を社会における相互依存関係へと一般化した場合、どのような示唆が得られるのかという点について述べていく。

2　モデル

本研究のモデルは、Nowak と May による、プレイヤー同士の相互依存関係を基盤とする空間囚人のジレンマゲーム(Nowak, 1992)に基づいており、各プレイヤーは裏切りか協力の 2 種類の戦略を持ち、相互依存関係のあるすべての相手プレイヤーと対戦を行って合計の利得を得る。すなわち、全プレイヤー数を N とし、対戦する 2 つのプレイヤーを i および j ($i \neq j$, $1 \leq i$, $j \leq N$)、それらが持つ戦略を $s(i)$ および $s(j)$、利得を $P(i)$ および $P(j)$ とすれば、$P(i)$ は（式 1）の利得行列 A を用いて（式 2）で表せる。ここで、$s(i)$ と $s(j)$ は単位ベクトルを用いて(1 0)または(0 1)で表せる。前者の場合はプレイヤー i の戦略が協力、すなわち協力者であることを、逆に後者の場合は裏切り、すなわち裏切り者であることを意味する。$O(i)$ は、プレイヤー i と相互依存関係のあるプレイヤーの集合を示す。

$$A = \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix} \quad (1 < b \leq 2) \tag{1}$$

$$P(i) = \sum_{j \in O(i)} s(i) A s(j)^T \quad (i \neq j,\ 1 \leq i, j \leq N) \tag{2}$$

本研究におけるすべてのプレイヤー数 N と利得行列 A におけるパラメータ b、初期状態における協力者数と裏切り者数の比率は、Ohdaira (2016)に従っている。すなわち、N=1000、b=1.5 であり、比率はほぼ一対一となっている。また、全 20 回のシミュレーションに関して、裏切り者と協力者の分布は、毎回ランダムに変化する。相互依存関係の空間構造についても、Ohdaira (2016)に従っており、N のプレイヤーからなる周期境界条件の 1 次元格子であり、 Watts and Strogatz (1998) に準じている。格子の各頂点がプレイヤーを表し、プレイヤー i は $k(i)$ の他プレイヤーと相互依存関係を持つ。$k(i)$ の平均値 $\langle k \rangle = (1/N) \sum_{1 \leq i \leq N} k(i)$ である。本研究では、Ohdaira (2016)で扱った、$\langle k \rangle$=4, 8 のレギュラー (Watts and Strogatz, 1998)、ランダム(Watts and Strogatz, 1998)、そして Barabási-Albert モデル (Barabási and Albert, 1999) と呼ばれる Scale-free の 3 種類の相互依存関係のトポロジーのうち、純粋に罰の係数の変化による影響を測定できる、レギュラーなトポロジーについて扱う。裏切り者と協力者の分布と異なり、全 20 回のシミュレーションに関して、相互依存関係のトポロジーは変化しない。なお、各トポロジーの詳細な構築方法は、Ohdaira (2016)の Methods に記されている。図 1 は $\langle k \rangle$=4 のレギュラーなトポロジーの場合を示しており、トポロジーを分かりやすくするために、N=12 としている。

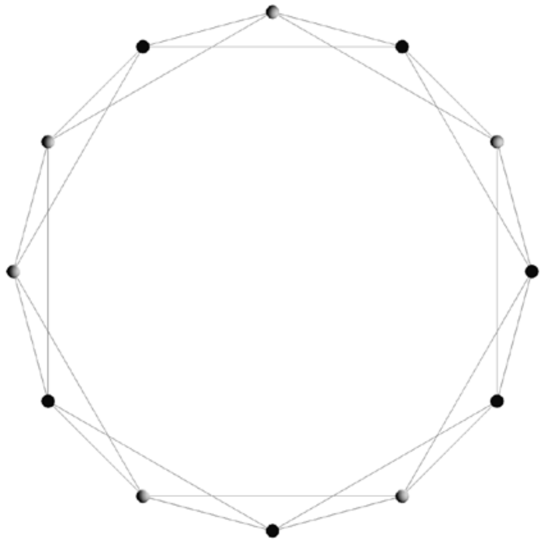

図 1：プレイヤー同士の相互依存関係がレギュラーなトポロジーである場合の初期状態。周期境界条件の 1 次元格子 ($\langle k \rangle$=4) である。トポロジーを分かりやすくするため N=12 としており、裏切り者 (グレー) 数と協力者 (黒) 数の比率はほぼ一対一である。

著者が、Ohdaira (2016)で提案した新しい罰の概念は、次のように定式化できる。プレイヤー i が、相互依存関係のある他プレイヤー $j \in O(i)$, $s(j)$=(0 1) (裏切り) に罰を与える確率は、プレイヤー i, j の利得の差に依存する。相互依存関係のある他プレイヤー $j \in O(i)$ の中で、$P(j) > P(i)$ かつ $s(j)$=(0 1)を満たすプレイヤー j の総数を $n(i)$、罰の係数を r ($0 \leq r \leq 1$)とし、次の 2 つの不等式、$P(i)(1-rn(i)) > 0$、$P(i) < P(j) \leq 2P(i)$ が成り立つとき、プレイヤー i はプレイヤー j に $rP(i)$ の犠牲と引き換えに、同等の罰を (式 3) で表される確率 $q_i(j)$ で与える。もし、$P(i)(1-rn(i)) > 0$ かつ $P(j) > 2P(i)$ であれば、$q_i(j)$=1 である。罰を与え、また罰を与えられた後のプレイヤー i, j の利得 $P'(i)$, $P'(j)$ は、次の (式 4) で表される。プレイヤー i が罰を与えること、そして罰を与えられたことによる利得の減少は独立して計算され、結果的に $P'(i) < 0$ となれば、$P'(i)$=0 となる。そのため、$P'(i)$, $P'(j)$ が負の値になることはない。また、r=0, 1 のときは、罰は機能しない。

$$q_i(j) = \frac{P(j) - P(i)}{P(i)},\ P(i) > 0 \tag{3}$$

$$\begin{aligned} P(i)' &= P(i) - rP(i) \\ P(j)' &= P(j) - rP(i) \end{aligned} \tag{4}$$

すべてのプレイヤーの利得を計算した後、以下の (式 5) に示す通り、プレイヤー i は相互依存関係のある他プレイヤー $j_{\max} \in i \cup O(i)$ の戦略を、次の世代における対戦の戦略とする。プレイヤー $j_{\max}$ が複数存在する場合、プレイヤー i は、それらが持つ戦略からランダムに次の世代における対戦の戦略を選択する。すべてのプレイヤーの戦略は同時に更新され、各回のシミュレーションは 300 世代に到達するまで続く。

$$\begin{aligned} s(i)' &= s(j_{\max})\ \ j_{\max} \in i \cup O(i) \\ P(j_{\max})' &= \max(P' \in i \cup O(i)) \end{aligned} \tag{5}$$

3 結果

以下の結果を示す各図は、$\langle k \rangle$=4, 8 の順番で、罰の係数 r=0.2, 0.4, 0.6, 0.8 の各場合について、協力者数とプレイヤー全体の平均利得が、世代とともにどのように変化するかを示す。なお、$\langle k \rangle$=4, 8 のいずれについても、プレイヤー同士の相互依存関係がレギュラーなトポロジーである場合、罰を導入しないと、$\langle k \rangle$=4 の場合は協力者数とプレイヤー全体の平均利得の増加速度が著しく遅くなり、$\langle k \rangle$=8

の場合は協力者数とプレイヤー全体の平均利得が全く増加しないことが分かっている(Ohdaira, 2016)。

まず、$\langle k \rangle$=4 の場合の結果について、図 2 は r の値別に世代ごとの協力者数を、図 3 は r の値別に世代ごとのプレイヤー全体の平均利得をそれぞれ示している。すべての r の値に関して、最終的な協力者数とプレイヤー全体の平均利得の増加が見られ、新しい罰の導入によって、協力の進化が促進されることが分かる。

図 2 より、r=0.2 のときが最も協力者数の増加速度が遅く、r=0.4 になると協力者数の増加速度はかなり速くなるが、一方で r=0.6 では r=0.2, 0.4 の中間の増加速度となり、r=0.8 で最大の増加速度となることが分かる。図 3 に示す、プレイヤー全体の平均利得の増加速度も協力者数の増加速度と傾向が類似しており、r=0.2 のときが最も遅く、r=0.4 になるとかなり速くなる一方、r=0.6 では r=0.2, 0.4 の中間となり、r=0.8 で最大となる。そして、r=0.2 のときは過剰な罰の増加により、一時的にプレイヤー全体の平均利得の減少が生じている。

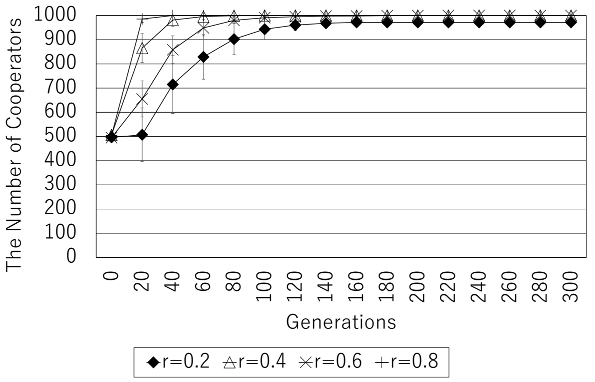

図 2 : $\langle k \rangle$=4 の場合に関する罰の係数 r の値別の世代ごとの協力者数

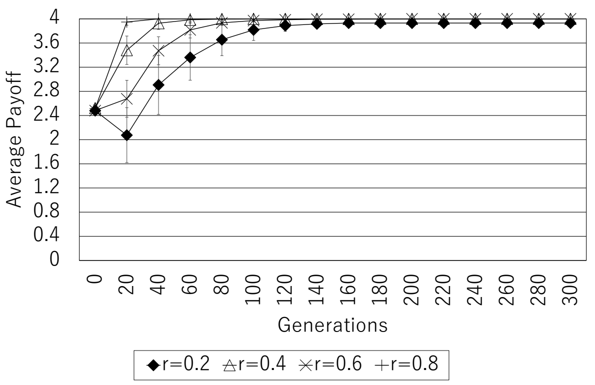

図 3 : $\langle k \rangle$=4 の場合に関する罰の係数 r の値別の世代ごとのプレイヤー全体の平均利得

次に、$\langle k \rangle$=8 の場合の結果について、図 4 は r の値別に世代ごとの協力者数を、図 5 は r の値別に世代ごとのプレイヤー全体の平均利得をそれぞれ示している。$\langle k \rangle$=4 の場合と同じく、すべての r の値に関して、最終的な協力者数とプレイヤー全体の平均利得の増加が見られ、新しい罰の導入によって協力の進化が促進されることが分かる。

しかし、その結果を詳しく見ると、$\langle k \rangle$=4 の場合とは逆に、図 4 に示す通り、r=0.2 のときが最も協力者数の増加速度が速く、r=0.4 になるとやや減少し、r=0.6, 0.8 に関しては、r=0.8 の方がわずかに速いものの、その差は大きくなく、しかも一時的に大きく協力者数が減少している。一方で、プレイヤー全体の平均利得と協力者数、それぞれの増加速度の傾向が類似する点については、$\langle k \rangle$=4 の場合と同じである。すなわち図 5 に示す通り、r=0.2 のときが最も速く、r=0.4 になるとやや減少し、r=0.6, 0.8 に関しては、r=0.8 の方がわずかに速いものの、その差は大きくなく、一時的な協力者数の大きな減少に伴い、プレイヤー全体の平均利得も一時的に大きく減少している。

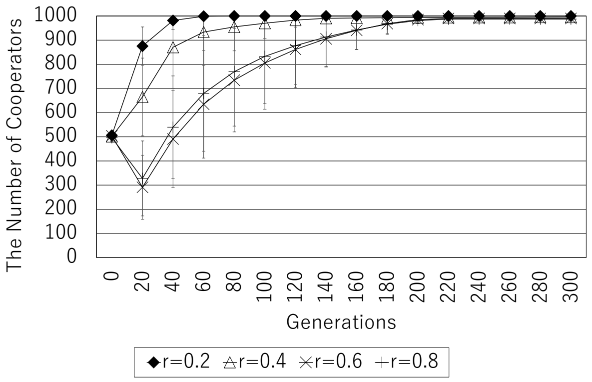

図 4 : $\langle k \rangle$=8 の場合に関する罰の係数 r の値別の世代ごとの協力者数

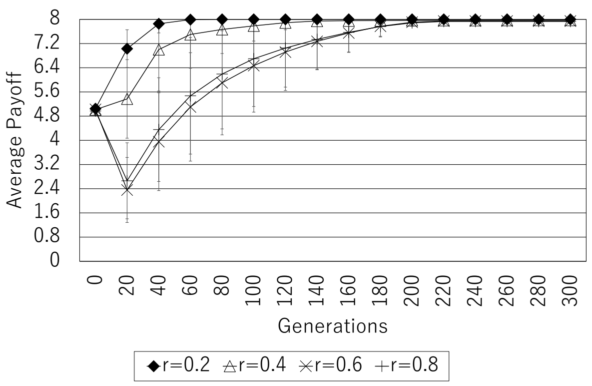

図 5 : $\langle k \rangle$=8 の場合に関する罰の係数 r の値別の世代ごとのプレイヤー全体の平均利得

以上の結果から、$\langle k \rangle$=4 の場合は基本的に r の値が増加すると協力者数の増加速度、プレイヤー全体の平均利得の増加速度共に速くなるが、r=0.4, 0.6 の間で逆転があり、必ずしも r の値の増加に伴って協力者数とプレイヤー全体の平均利得の増加速度が速くなるわけではないことが分かる。そして

$\langle k \rangle=8$ の場合は、$\langle k \rangle=4$ の場合とは対照的に、r の値が増加するほど、協力者数とプレイヤー全体の平均利得の増加速度が遅くなることが分かる。次節では、こうした現象が生じた理由について考察する。

4　考察

罰の係数 r の値が増加すると、不等式 $1-rn(i)>0$ が成立するためには、$n(i)$の値が減少しなければならないため、自分の周囲に高利得の裏切り者が多数存在する、いわゆる搾取されているプレイヤーi は、そうした裏切り者を罰することができなくなる。すなわち、r の値の増加に伴い、罰を与えることができるプレイヤーi の数は減少する。特に $r \geq 0.5$ となると、$n(i)=1$ となるため、各プレイヤーの罰がターゲットとなる裏切り者の利得を効果的に減少させないと、協力者数が増加しにくくなる。

$\langle k \rangle=4$ の場合、プレイヤー同士の相互依存関係が疎であるため、概ね r の増加に伴い協力者数とプレイヤー全体の平均利得の増加速度が速くなるが、$r=0.6$ において落ち込みが見られるのは、$n(i)=1$ である状況では、より強力な罰が必要となるためである。一方で$\langle k \rangle=8$ の場合、r の増加に伴い、協力者数とプレイヤー全体の平均利得の増加速度が遅くなる。$\langle k \rangle=8$ というプレイヤー同士の相互依存関係が密な状態では、$n(\mathrm{i})=1$ となる $r \geq 0.5$ の領域において、協力者数とプレイヤー全体の平均利得の増加速度は大きく減少する。

以上のことから、$\langle k \rangle$の値が小さいときは、協力者数とプレイヤー全体の平均利得の増加には、すべてのプレイヤーが裏切り者を罰する必要はなく、限られた数のプレイヤーが、自らの利得を大きく犠牲にして、搾取されている協力者を守るように、強力な罰を裏切り者に与えればよいことが分かる。一方で$\langle k \rangle$の値が大きいときは、協力者数とプレイヤー全体の平均利得の増加には、$\langle k \rangle$の値が小さいときよりは、裏切り者を罰する犠牲は少なくて済む一方、より多くのプレイヤーが裏切り者を罰する責任を共有する必要があることが分かる。

5　おわりに

本研究は、Ohdaira (2016)が提案した新しい罰の概念における中心的な制御パラメータである罰の係数、すなわち自分の利得の犠牲と、相手に与える罰の程度を変化させたときも、新しい罰の導入によって協力の進化が促進され、プレイヤー全体の平均利得も増加すること、またプレイヤー同士の相互依存関係の密度の違いにより、協力の進化に最適な罰の係数に違いがあることを示した。

そして、その結果を社会における相互依存関係へと一般化した場合、人々の相互依存関係が疎な場合、限られた数の人達が自らの利得を大きく犠牲にして、搾取されている人達を守るように、裏切り者に対して強力な罰を与えればよいこと、逆に密な場合は、疎な場合と比較してより多くの人達が、自らの利得を相対的に小さく犠牲にして、裏切り者を罰する責任を共有する必要があることが示唆される。これは、既存の協力の進化に関する知見に重大な貢献をもたらすものである。

今後の研究では、各プレイヤーが自分の好むプレイヤーとのみ相互依存関係を維持し、好まないプレイヤーとの相互依存関係を破棄する機構について検討していく予定である。

参考文献

Barabási, A. L. and Albert, R. (1999) Emergence of scaling in random networks. *Science*, 286(5439), 509-512.

Chen, X., Szolnoki, A. and Perc, M. (2014) Probabilistic sharing solves the problem of costly punishment. *New J. Phys.*, 16, 083016.

Chen, X., Szolnoki, A. and Perc, M. (2015) Competition and cooperation among different punishing strategies in the spatial public goods game. *Phys. Rev. E*, 92(1), 012819.

Dreber, A., Rand, D. G., Fudenberg, D. and Nowak, M. A. (2008) Winners don’t punish. *Nature*, 452(7185), 348-351.

Nowak, M. A. and May, R. M. (1992) Evolutionary games and spatial chaos. *Nature*, 359(6398), 826-829.

Nowak, M. A. (2006) Five rules for the evolution of cooperation. *Science*, 314(5805), 1560-1563.

Ohdaira, T. (2016) Evolution of cooperation by the introduction of the probabilistic peer-punishment based on the difference of payoff. *Sci. Rep. (Nature Publishing Group)*, 6, 25413.

Perc, M. and Szolnoki, A. (2015) A double-edged sword: Benefits and pitfalls of heterogeneous punishment in evolutionary inspection games. *Sci. Rep. (Nature Publishing Group)*, 5, 11027.

Rand, D. G. and Nowak, M. A. (2011) The evolution of antisocial punishment in optional public goods games. *Nature Communications*, 2, 434.

Szolnoki, A. and Perc, M. (2013) Effectiveness of conditional punishment for the evolution of public cooperation. *J. Theor. Biol.*, 325, 34-41.

Watts, D. J. and Strogatz, S. H. (1998) Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684), 440-442.