

繰り返し囚人のジレンマゲームにおける非報復戦略の可能性

藤原 紫王里* 中桐 斉之**

兵庫県立大学環境人間学部環境人間学科

1、はじめに

人類及び人類以外の社会においては、様々な協力の形式が存在する。囚人のジレンマ (PD) ゲームにおいても、個人(プレイヤー)たちが自者と他者を有利にしようとする際、互いに協力し合うことがある。ロバート・アクセルロッドが行ったノイズなし繰り返し囚人のジレンマゲームにおける戦略の研究では、それぞれの戦略が他の戦略を相手にして得た得点を総合するコンピュータを使った試合を行い、第一回目の試合では、Axelrod と Hamilton[1]が考案したしっぺ返し戦略 (TFT) が勝利することが分かった。しかしノイズあり繰り返し囚人のジレンマゲームにおいては、Novak と Sigmund[2]が考案したパプロフ戦略 (PAV) が最強であると証明されている。

しかしながらこの二つの戦略はどちらも「報復戦略」である。現実世界では、すべて協力する協調戦略 (AC) を選択する生物は多数存在する。しかしながら、AC はすべて裏切る裏切り戦略 (AD)、および PAV より劣っていると考えられている。また、裏切られたら裏切り返すという行動は、人類の道徳基準より程遠い。一方、協調戦略の相手にしてもらいたい行動をとる「黄金律」は、人類の理想の戦略と言える[3]。

本研究では、一次元格子に四つの戦略 (AC, AD, TFT, PAV) を配置し、繰り返し囚人のジレンマゲームを行った。各戦略は、同じ戦略を持つ仲間たちで形成される独自のコロニーを保有しているとする。各コロニーのパッチ数を変えてシミュレーションを行った先行研究においては、コロニー内のセルはすべて同じ戦略を持ち、 $m \geq 9$ のとき最終平衡で AC がすべてのプレイヤーを占領する。しかし、同じ集団内であってもすべての仲間が同じ戦略を選択するとは限らない。本研究では一定の確率で仲間のコロニーが他の戦略を保有するノイズを加えることとした。

2、モデル

二人のプレイヤーが繰り返し囚人のジレンマゲームを行う。1STEP の間にプレイヤーは協調 (CO) と裏切り (DE) の二つの手からどちらか一つを選択する。ここで一定の確率で選択肢を間違えるゲームノイズ (x) を加えた。互いのプレイヤーの手により獲得する得点が決定する。本研究では、表 1 の利得行列を適用した。

表 1. 利得行列

	CO	DE
CO	3	0
DE	5	1

格子セルを、AC、AD、TFT、PAV のうちどれか一つの戦略を持ったコロニー (パッチ) を表すと考える。各コロニーでは、同一の戦略を持った $m+1$ の仲間が存在する。

初期条件として一次元格子に 4 つの戦略を等確率で配置する。次に、ランダムに一つのプレイヤー (α)、隣接したセルのどちらか一方のプレイヤー (β) をランダムに選択する。それぞれのプレイヤーは、隣接したプレイヤーと同コロニー内セルとゲームをする。同コロニー内セルはプレイヤーと同じ戦略を持つが、コロニー内ノイズ (y) により一定の確率で他の戦略に変わる。



図 1. ゲームを行うセルの選択

例をあげると、図 1 において初めに α が選択され、次に両隣のセルから $1/2$ の確率で右隣の β が選択される。 α は図 1 の斜線のプレイヤーとゲームをし、 β は灰色のプレイヤーとゲームをする。選択されたプレイヤーは両隣とゲームを行った後、同コロニー内のプレイヤーと m 回ゲームを行う。プレイヤー同士の勝敗は、 $2+m$ 回行ったゲームで得た合計点で競う。一つのゲームにつき、 N 回の繰り返し囚人のジレンマゲームを行う。得点が低かったプレイヤーおよびコロニーは、得点の高かったプレイヤーの戦略に塗り替えられる。両者の得点が同じだった場合、 $1/2$ の確率でどちらかのプレイヤーの勝利となり、もう一方のプレイヤーの戦略を塗り替える。

The possibility of non-retaliation strategy
in repeated prisoner's dilemma game

*Shiori Fujiwara

School of Human Science and Environment, University of Hyogo

**Nariyuki Nakagiri

School of Human Science and Environment, University of Hyogo

3、シミュレーション結果と考察

コロニー内パッチ数 (m) とコロニー内ノイズ (y) を変えてシミュレーションを行った。ゲームノイズ (x) は 0.1 で固定する。今回のシミュレーションでは、負けた戦略が勝った戦略に塗り替えられるため、図の 2、4 のように最終的に一つの戦略が他の戦略を支配する。横軸は空間、縦軸は時間を表す。

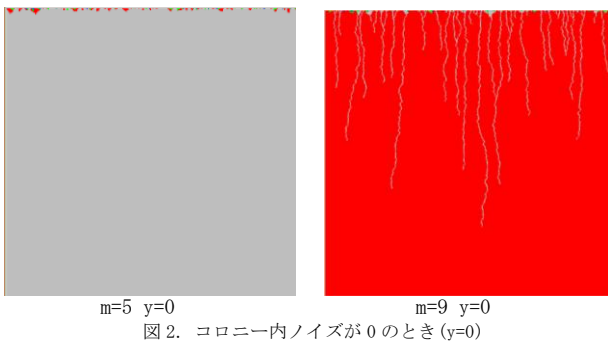


図 2. コロニー内ノイズが 0 のとき ($y=0$)

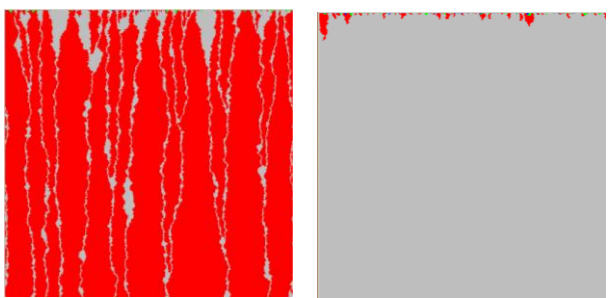


図 3. コロニー内パッチ数が 6 ($m=6$)

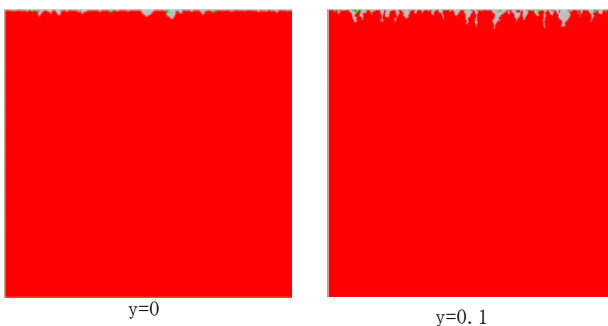


図 4. コロニー内パッチ数が 10 ($m=10$)

$y=0$ のときの典型的なダイナミクスは図 3 に示す。 $m \geq 9$ のとき、最終平衡では AC が占領する。

図 3 の (a) はコロニー内ノイズが 0 のときの結果である。AC の中に PAV がわずかに生き残る現象が見られるが、この現象について考察する。表 2 の平均利得票によると、図 6 の (c) A の平均得点は $2 \times 3.65 + 6 \times 2.602 = 22.912$ から与えられる。一方、(c) B の平均利得は $1.65 + 2.89 + 6 \times 2.89 = 21.88$ を得る。つまり PAV は AC に勝利する。次に (d) A の場合、平均利得は $3.65 + 2.602 + 6 \times$

$2.602 = 21.864$ を得る。これは AC の平均利得を下回っているため、AC が勝利する。従って、PAV の単独プレイヤーは AC の領域の中をランダムウォークする。格子が有限の場合、一つの PAV セルだけが最終平衡で残る。^[3]

しかし図 3 の (b) のとき、AC が打ち負かされ、PAV が占領する。 $y=0.1$ のときは、1/10 の確率でコロニー内に他の戦略を持った仲間が存在するため AC に限り必ず平均得点が低くなり、この現象が発生する。

図 4 で示すダイナミクスでは、 $m \geq 10$ 、 $y=0.1$ のとき、最終平衡では AC が占領する。

表 2. 各戦略の平均利得票

	opponents			
	AC	PAV	TFT	AD
A	$3-x-x^2$	$1.5+1.5x$	$3-4x+4x^2+2x^3$	$4x+x^2$
C	(2.89)	(1.65)	(2.642)	(0.41)
P	$4-3.5x$	$3-5x+11x^2-8x^3$	2.25	$0.5+3.5x$
A	(3.65)	(2.602)	(2.25)	(1.202)
V				
T	$3+x-6x^2+2x^3$	2.25	2.25	$1+2x+2x^3$
F	(3.402)	(2.25)	(2.25)	(1.202)
T				
A	$5-6x+x^2$	$3-1.5x$	$1+7x-10x^2+2x^3$	$1+3x-x^2$
D	(4.41)	(2.85)	(1.602)	(1.29)

(c)



(d)



図 6. シミュレーション結果の考察

動物の形成するコロニーには定説通りにいかなる個体も一定数存在する。本研究では、コロニー内に他の戦略をとる個体がいっても、仲間の数が十分多ければ問題ないことを説明することができる。

【参考文献】

- [1] R. Axelrod : The Evolution of Cooperation, Basic Books, New York, 1984
- [2] M. A. Nowak and K. Sigmund : Evolution of indirect reciprocity by image scoring, Nature Vol. 393, pp. 573-577, 1998
- [3] 上原隆司, 鈴木佳祐, 吉村仁, 秦中啓一, 内藤広海 : 格子上囚人のジレンマゲームにおける多様な勝者, 第 76 回情報処理学会全国大会講演論文集, 1, 287-288, 2014