

囚人のジレンマゲームのシミュレーション解析 — 協調戦略の可能性 —

原田凌太† 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

人類及びその他様々な生物の社会においては、様々な協力の形式が存在する。囚人のジレンマゲームにおいても、プレイヤーたちがゲームを有利に進めようとする際に互いに協力し合うことがある。ノイズなし囚人のジレンマゲームにおける戦略の研究[1]においては、Axelrod と Hamilton の考案したしっぺ返し戦略(以下 TFT と呼ぶ)が勝利することが分かった。しかし、ノイズあり囚人のジレンマゲームにおいては、Novak と Sigmund[2]の考案したパブロフ戦略(以下 PAV と呼ぶ)が最強であるとされている。

しかしながら、この両戦略はどちらも裏切りを前提とした報復戦略の形式である。現実世界においては常に協力する協調戦略(以下 AC と呼ぶ)を選択する生物が多数存在するが、AC は常に裏切る裏切り戦略(以下 AD と呼ぶ)、および PAV より劣っているとされている。しかし、常に裏切るという行動や裏切られたら裏切り返すという行動は、人間の道徳基準には程遠い。一方で、常に相手に対して協力の選択をとる協調戦略は、人類において理想の戦略であると言える[3]。

本研究では、二次元格子上に4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。また、ゲームを行う際に、必ずしもプレイヤーの意図通りに進むとは限らないため、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。

本研究では以下の表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1つを、AC, AD, TFT, PAV のうちいずれか1つの戦略を持ったプレイヤーを表すと考える。

初期条件として二次元格子上に4つの戦略をランダムに等確率で配置する。次にランダムに1人のプレイヤー(α)を選択し、その上下左右いずれかのプレイヤー(β)をランダムに選択する。各プレイヤーは、それぞれの上下左右の隣接したプレイヤーと囚人のジレンマゲームを行う。

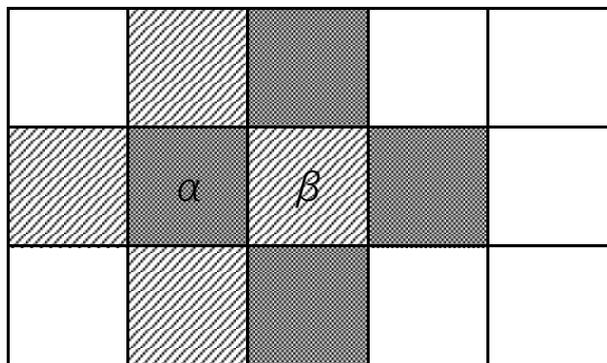


図 1. ゲームを行うセルの選択

例えば、図1において初めに α が選択されたとすると、次にその上下左右の隣接するセルから1/4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対戦を行い、 β は網掛けのプレイヤーと対戦を行う。それぞれのプレイヤーが上下左右4マスと対戦をして得られた得点の合計を、各プレイヤーの持ち点とし、その点数を比べることで α 、 β の勝敗を決定した。一セットにつき、この対戦を10000回行い、それを100セット行った。得点の低かったプレイヤーは、得点の高かったプレイヤーの戦略に塗り替えられる。両プレイヤーの得点が同じであ

Simulation analysis of the prisoner's Dilemma-Possibility of cooperation strategy-

†Ryota Harada

School of Human Science and Environment, University of Hyogo

‡Nariyuki Nakagiri

School of Human Science and Environment, University of Hyogo

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。

3. シミュレーション結果と考察

ノイズ率を0~1に変化させ、コンピュータによる実験を行った。結果を図2に示す。それぞれのノイズ率における各戦略の残存数は以下の図2の通りである。

	0	0.1	0.2	0.3	0.4
AC	1133.5	1087.3	1101.1	649.8	621.8
AD	3657	3690.2	3609	5221.3	5533.7
TFT	3192.8	3134.1	3179.3	2270.8	2075.6
PAV	2016.7	2088.4	2110.6	1858.1	1768.9

	0.5	0.6	0.7	0.8	0.9	1
AC	2448.3	5831.7	5871.1	5909.1	5916.7	5683.1
AD	2410.1	513.8	516.4	505.6	505.7	515.6
TFT	2482.2	1614.2	1578	1569	1598	1740.2
PAV	2659.4	2040.3	2034.5	2016.3	1979.6	2061.1

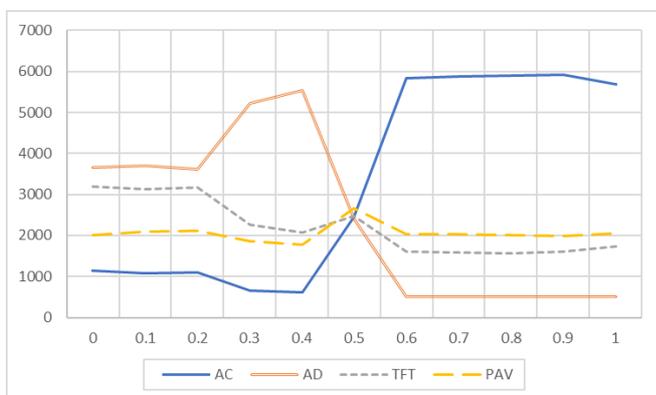


図2. 各戦略の数

相手プレイヤーの信頼度が高く、ノイズ率が低い(0.1~0.4)の時、ACは非常に弱く、数は全体の5~10%ほどである。しかし、ノイズ率が0.5の時を境に協調戦略の数は大きく増加し、0.6以降は安定して最大の数となる。ADはACとは逆の増減の仕方をしており、ノイズ率が0.5の時を境に大きく減少する。先行研究において、有効とされていたTFT, PAVはノイズ率0.5の時全戦略数がほぼ同数になるときを除いて、数が約15~20%と安定した数となっている。本研究のように、ノイズ率が一定に設定された状況においては最強であるとは言い難いが、実際の社会となると、ノイズ率が一定であるわけではない。

そういった状況においては、安定した残存数となるTFT, PAVは確かに有効な戦略である。しかし、明確に相手に対する信用度が低い、つまりノイズ率が高いと判断できる際には、ACは最も有効な戦略であるということが、本研究から示された。

発表では、相互作用の範囲を広げた時の結果についてもあわせて報告する。

【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