

## 空間配置型 N 人版囚人のジレンマゲーム

田村 誠\*  
Makoto TAMURA

### 1. はじめに

繰り返し囚人のジレンマ(Iterated Prisoner's Dilemma: IPD)は、抽象化されたシンプルな状況でありながら利己的な個人による協調の進化に対する説明としてよく知られている。とりわけ Axelrod(1984)が行なった、二種類の戦略が繰り返し対戦トーナメントは有名である。このトーナメントでは TFT(Tit for Tat)戦略が二度にわたって優勝した為、TFT が IPD の最強戦略と解釈されてきた。しかし、その後 Nowak and Sigmund(1993)らによって TFT 戦略には幾つかの弱点も指摘されている(例えばノイズの有無など)。Axelrod も記述しているが、本来、ある戦略が勝つか否かはあくまでも相手戦略との相互関係に依存するのである。さらに、人間社会、あるいは生態系などに対して囚人のジレンマゲームをより妥当なモデルとする為には、参加するプレイヤーを増やすことが考えられる。

こうした問題意識から本稿は以下のような構成で議論を進める。まず、マルチエージェントシミュレーションの特性を生かし、戦略模倣ゲームとしてモデルを定式化して空間上に複数戦略が同時に存在する IPD を実行する。従来、数理生態学の分野を中心に研究されている空間配置型 N 人版 IPD モデルでは、TFT 戦略を基本にして、前回の相手の戦略に対して確率的に戦略が反応するように設定されていた(Nowak and May(1992,1993)、Grim(1995)、Nakamaru et al(1998)など)。これに対して本稿は、各エージェントの戦略を明示的に規定し、より複雑な戦略を同時に存在させている。

次に、先行研究において安定的に強いと指摘されている戦略の有効性を再検証する。この際に、ABS(Agent Based Simulator)を利用することで視覚的にもこのような戦略の動態が観察できることは本稿の一つの特徴に挙げられる。果たして TFT 戦略が勝ち残るのか、それとも他の戦略が台頭するのであろうか。そして、もしその他の戦略が勝ち残るのであれば、如何なるメカニズムが作用しているのか議論していく。

---

\*東京大学大学院総合文化研究科広域科学専攻 Department of General Systems Studies,  
Graduate School of Arts and Sciences, Univ. of Tokyo.  
E-mail: tamura@sanshiro.c.u-tokyo.ac.jp

## 2. モデルの概要

ここでは、本稿で用いるシミュレーションモデルの概要を述べる。

### 2.1. 囚人のジレンマ

本モデルにおける各エージェントの利得構造は表 1 のような初期設定としている。数学的には、この 4 つの値が  $T > R > P > S$  かつ  $2R > T + S$  を満たす場合を囚人のジレンマと呼ぶ。この囚人のジレンマゲームは非零和非協力ゲームであり、最適戦略が存在しないことが知られている<sup>1</sup>。戦略の有効性は相手が取る戦略に依存する。

表 1 モデルにおける利得構造

	協調(C)	裏切り(D)
協調(C)	3(R),3(R)	0(S),4(T)
裏切り(D)	4(T),0(S)	1(P),1(P)

R: Reward, S: Sucker, T: Temptation, P: Punishment

### 2.2. モデルルール

本モデルはセルオートマトンの環境にエージェントを配置し、N 人版囚人のジレンマ・シミュレーションを行った。以下のルールでモデルは実行される。

1. 各エージェントは空間上にランダムに配置される。
2. 外生的に設定される対戦回数にしたがって、図 1 に示すような周囲の 8 近傍(ムーア近傍)と繰り返し対戦を行なう。
3. 対戦回数を経過すると、8 近傍 + 自分の中で最も点数の高い戦略を摸倣する。同点の戦略がある場合は、確率的にランダムに摸倣する(2、3 の動作を 1 世代と見なすこともできる)。
4. 新たなゲームが開始され、再び周囲と繰り返し対戦を行なう(2、3 を繰り返す)

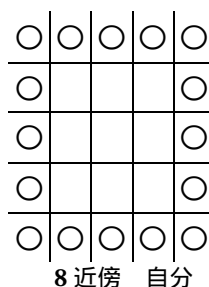


図 1 各エージェントの対戦相手

各エージェントは誰とでも対戦するのではなく、自分の周囲にいる相手と繰り返しゲームを行なう。ここでは対戦回数を 10 回とし、その後得点の高い戦略をコピーし新しいゲーム

<sup>1</sup> より厳密に言えば、ナッシュ均衡は「双方裏切り」であるのに対して、パレート最適は「双方協調」となり、両者にズレが生じている。

ムが始まるように設定した。つまり対戦回数を経過すると新たな世代によるゲームが開始する。このように、各エージェントはそれぞれ異なる戦略を持った個体群として、ゲームを開始し、選択圧によって戦略の分布が変化する。戦略分布は時間とともに変化し、個体自体は学習しないが社会は学習することになる。

### 2.3. 戦略の種類

本稿で使用するモデルは、表 2 のような戦略(エージェント)を想定している。本稿と同様に空間配置型の IPD を作成した Grim(1995)は、単に前回の相手の戦略に対して確率的に協調と裏切りを決めるような戦略を想定している<sup>2</sup>。これは確率変数で割り振っているものの、TFT が基本となっており本質的な戦略の種類は一つである。Pavlov のようにそれ以前の記憶が残っている戦略や、相手の手番に左右されない繰り返し戦略など戦略の行動様式が複数あるという点は本稿の特色と言えよう。

表 2 各戦略の概要

戦略名	内容
All-C	常に協調。
All-D	常に裏切り。
Joss	TFT 同様相手に裏切られると即座に裏切り返すが、相手が協調した後では必ず協調するのではなく、9割協調して1割裏切る。
TFT	Tit for Tat。最初は協調し、あとはオウム返し。相手が協調したら自分も協調し、裏切れば裏切り返す。
T2FT	堪忍袋。基本的に TFT だが、二回続けて裏切られないと裏切り返さない。
HardTFT	TFT をやや厳しくする。初めの 2 回は協調。前 2 回のうち一度でも裏切られたら、裏切り返す。
Random	協調と裏切りを確率的にランダムに選択する。
2CforD	最初は協調し、相手が 2 度協調してきたら裏切る。
Tester	最初は裏切り、相手の出方を調べる。相手が裏切ったら次回自分は協調し、その後は TFT。相手が協調したら、次回次々回は協調、その後一回おきに裏切る。
Friedman	最初は協調。相手が一度でも裏切ると、その後ずっと裏切り続ける。
SoftFriedman	最初は協調。相手が二度続けて裏切ると、その後ずっと裏切り続ける。
per-CD	CD CDCD... と繰り返す。
per-DC	DC DCDC... と繰り返す。
per-CDD	CDDCDDCDD... と繰り返す。
per-CCD	CCDCCDCCD... と繰り返す。
Pavlov	後述

<sup>2</sup> Grim(1995) や Nakamaru et al(1998)では、 $\langle p, q \rangle$  ( $p$  は相手が協調した時に協調、 $q$  は裏切ったときに協調する確率) という表現を用いて戦略を規定している。

## Pavlov 戦略

TFT 戦略は、ノイズによって協調と裏切りを交互に繰り返してしまう、All-C に対して採  
取できない(常に協調)、といった欠点がある。これらを補う戦略として、Nowak and  
Sigmund(1993)が提唱した Pavlov 戦略の概要を表 3 に示す(有田(2000)より)。例えば、相  
手が協調することを期待して自分が協調を出したときに裏切られた、あるいは(相手が協調  
を出さずと思って)裏切りを出したときに相手も裏切ってきたときは期待通りいかなかったケ  
ースである。このような失敗の選択をしてしまったときに前回の選択を変更する。逆に前  
回自分が裏切りで相手が協調した場合はうまくいったので、手を変える必要はない。Pavlov  
はノイズに強い戦略といわれる。

表 3 Pavlov 戦略と TFT 戦略

前回の自分の手	相手の前回の手	Pavlov	TFT
C	C	C	C
C	D	D	D
D	C	D	C
D	D	C	D

これら 16 戦略を用いて、シミュレーションを試行した結果を次節で記述する。

### 3. 試行結果

本節では 2 節で述べたモデルを用いて、戦略数の推移について検証する。空間の大きさを  $12 \times 12$  としているので、各戦略は初期状態で  $9(=144/16)$  個ずつ配置されている。図 2 にモデルの実行例を示す。各戦略は色で区別され、また数値は各エージェントの得点を表す<sup>3</sup>。最初はランダムに戦略が配置されているが、世代交代が進むにつれてムーア近傍でのゲームというローカルな情報によって、いくつかの集団が形成される。その際に、得点の低い戦略は淘汰され全滅する。

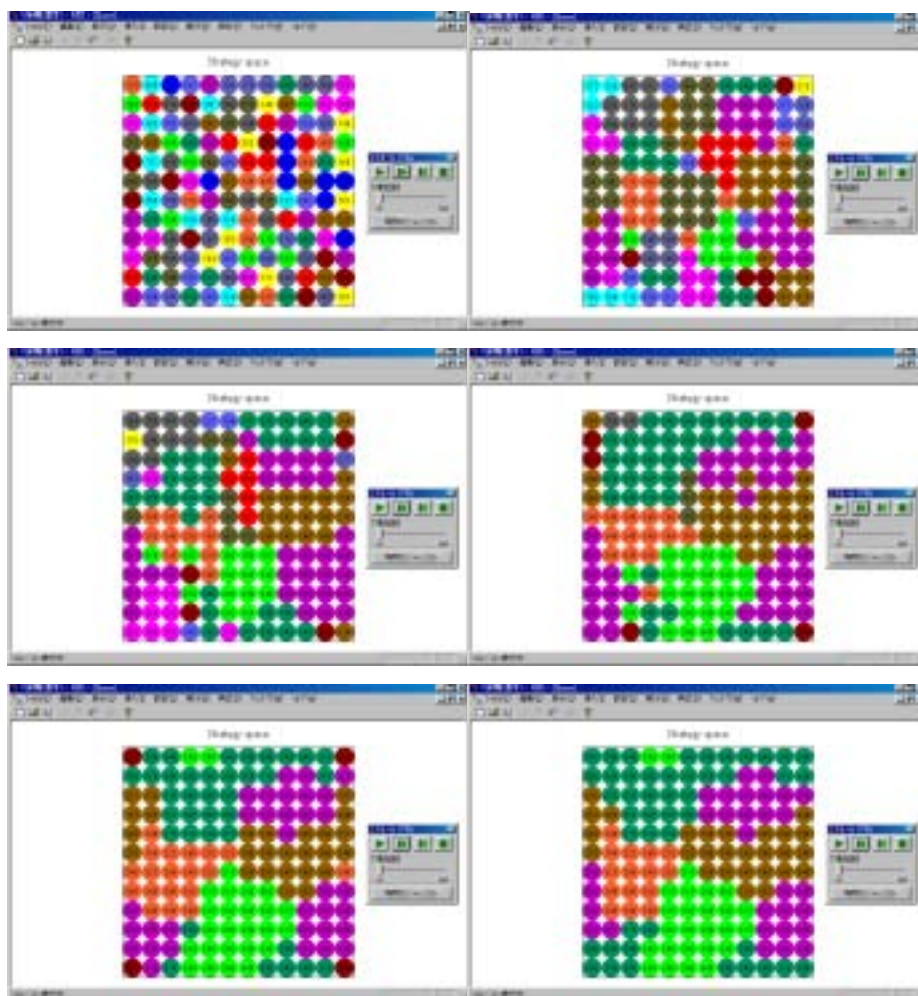


図 2 モデルの実行例  
(左上 右上 左中の順に 10,20,30,40,50,100 ステップ後を示す)

<sup>3</sup> 主な戦略の配色は以下のとおりである。青：All-C、赤：Joss、黄緑：TFT、黄：Random、水色：2 CforD、ピンク：Tester、緑：Friedman、紫：softFriedman、オレンジ：Pavlov。

図 2 の例では、はじめは All-D が最も多くこの他 Tester や per-CDD などの繰り返し戦略も存在するが、次第に駆逐され最終的には Friedman、softFriedman、HardTF、TTFT、Pavlov が共存する状態になった。このとき 10 ステップ後は全体の戦略の 64%、20 ステップ後は 56%に一度低下し、76%、89%に再び上昇し、50 ステップ後は 93%が協調を選択するようになっている。

上記の戦略数の変化を定量的に分析する。図 3 は 16 戦略が存在するシミュレーションを 30 回実行したときの各世代における戦略数の平均値を示したものである。

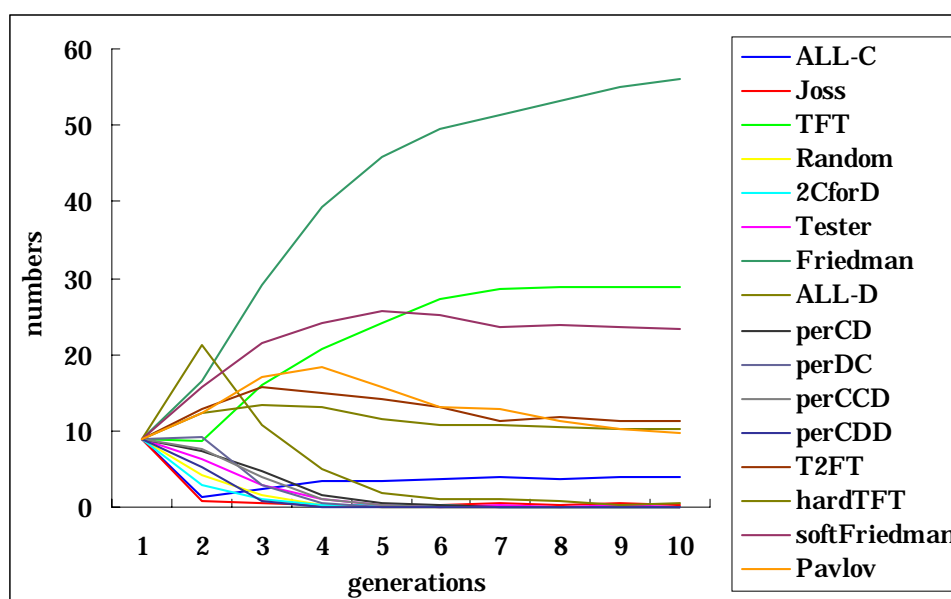


図 3 戦略数の推移(16 戦略)

10 世代後に平均的に一番多く生き残るのは Friedman であり、次いで TFT、softFriedman、T2FT、hardTFT、Pavlov、All-C の順となった。この他の戦略はほぼ全滅に近い状態になった。30 回の試行で Friedman は一度も全滅することはない、TFT、softFriedman は 4 回、Pavlov、T2FT、hardTFT は 13 回全滅している。それ以外の戦略は All-C が生き残ることが何回かあった他はほぼ毎回全滅してしまった。個々の戦略の動態を観察すると、Friedman はすぐに増加し始め、TFT はそれにやや遅れて増え始める。TFT は All-D が衰退し協調的戦略が増えることで自らの勢力を伸ばしているようである。All-D は一度増加するがすぐに他の戦略に滅ぼされ(D C という戦略になると得点が入らない)、最終的に勝ち残る場合は少ない。この他、Pavlov は全滅しにくい戦略であるが、大勢を占めることはなかった。また、全体としては協調行動に収斂している。

このように先行研究の指摘に対して、Friedman が最も大きな勢力となった。Friedman が勝ち残るのは偶然的なものであるか、必然的であるのかをさらに検証する必要がある。

そこで、基本設定として用いた利得構造を変化させたときの動態を観察した。ここでは表 1 の T を 4 から 5 に変更し、より裏切りのインセンティブが高まるように設定した。その他の条件は図 3 と全く同じである(試行回数も同じく 30 回としている)。図 4 から分かるように、さらに Friedman が独り勝ちする傾向が強くなった。なお、表 1 の P を 1 から 2 に変更し、ジレンマ状況をやや弱めた実験も行ったが、Friedman が大勢を占めるという結果は変わらなかったため省略する。

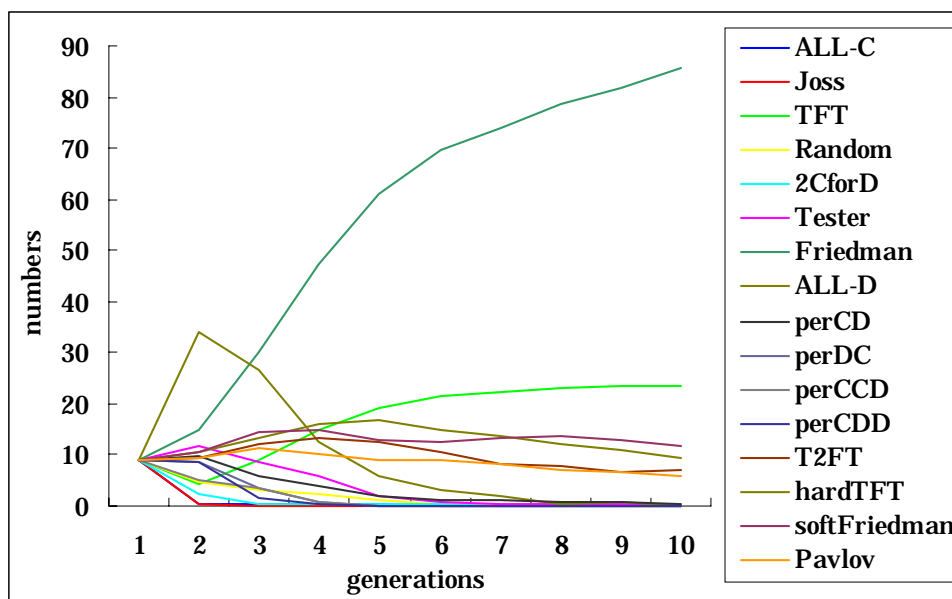


図 4 戦略数の推移(利得変更後。16 戦略)

次に空間に配置する戦略を Friedman ともう 1 種類だけにして、他の 15 戦略とそれぞれ戦わせることにした(利得は表 1 と同じ)。All-C や TFT のように協調戦略同士の間には、両者の利得に差が出ないので均衡状態となる。一方、All-D などの搾取的戦略に対して Friedman は強く、その戦略を全滅させることがしばしば見られた<sup>4</sup>。すなわち全ての戦略に対して Friedman が負けることはなかった。

以上のように、本モデルにおいて Friedman 戦略が安定的に勝ち残ることが明らかになった。次節でその背景にあるメカニズムについて考察する。

<sup>4</sup> この他に、TFT と Pavlov を対戦させる実験も行っている。TFT が Pavlov を全滅させる状況が観察された。Pavlov よりも TFT の方が集団安定的であると考えられる。

#### 4. 考察

前述の Grim(1995)では、TFT より寛容な GTFT という戦略が生き残っており、寛容な戦略の安定性が指摘されている。しかし、Grim で想定されている戦略は前回の相手の手番に確率的に反応する TFT、GTFT および All-D であり、Friedman や Pavlov といった戦略は設定されていない。本稿のように自らの戦略が明確に規定されている戦略が多数存在する場合には、Friedman が勝ち残りやすい状況が生まれるというのは興味深い結果である<sup>5</sup>。

一方、TFT は Friedman ほどには勝ち残らなかった。ただし、ここで注意しなければいけないのは、Axelrod(1984)のトーナメントでも TFT 戦略は最強の戦略ではなく、平均得点が最高だったということである。つまり、TFT は勝つ戦略というよりも、負けない戦略という方が適当であろう。そのため本稿において、TFT は最強戦略となっていないものの、多くのケースで生き残る戦略となっている。

Friedman 戦略は、Axelrod(1984)の第 1 回トーナメントで 15 戦略中 7 位、第 2 回トーナメントでは 63 戦略中 52 位だった。これに対して TFT は二大会とも優勝し、Axelrod(1984) は良い成績の教訓として nice(上品さ、有徳)、forgiveness(寛容)、retaliatory(報復的)を提示した<sup>6</sup>。上品であるが寛容ではないにもかかわらず、本モデルで Friedman が強いのは報復的性質が功を奏したことが予想される。

Friedman が勝ち残りやすくなった原因をさらに検討してみよう。Joshi(1987)は囚人のジレンマゲームを N 人版に拡張し、複数の戦略の存在比率、集団の大きさ、ゲームの利得行列などをパラメータとして各戦略の適応度を計算し、その比較により戦略の侵入の可能性などを論じている。そして、常に裏切り戦略が占めているときには、ある割合以上にならないと TFT 戦略が共存できないことや集団サイズが大きくなるとこの割合が増加し、侵入が困難になることを示した。つまり、囚人のジレンマゲームをただそのまま N 人に一般化すると二人版に比べて協調関係が成立しにくくなることを明らかにした。

二人版 IPD では裏切り者は明白だが、多人数版 IPD では裏切り者の特定が難しくなる。また、裏切り返すという制裁の効果が裏切り者だけに向かわなくなり、無関係な協力者の利益まで減らしてしまう可能性がある。空間上に存在することで報復相手が必ずしも報復し返してこない場合がある。Friedman は積極的に自ら裏切ることはしないが、一度協調関係が崩れると敏感に反応し、報復を行なう。しかし、その仕返しを必ずしも自らが受けなくても済むために勝ち逃げできる。それ故、このような匿名性の高い社会モデルにおいては、Friedman のような報復的な戦略が勝ち残ったと言える。

これまで Friedman が勝ち残った幾つかの原因を考察してきた。IPD は比較的長い間研究されており、本稿で想定した戦略以外にもまだ数々の戦略が考案されている。その中か

---

<sup>5</sup> Matsushima and Ikegami(1998)は、All-C,Friedman,TFT,Pavlov,Punish-once を用いてノイズのない 3 人ゲームを行った。このゲームでは Friedman の防衛的性質の安定性を論じているが、本稿のような空間配置型モデルで Friedman が勝ち残るとい報告は少ない。

<sup>6</sup> さらに 4 つ目の教訓として、clarity(明快さ)を挙げている。



ら Friedman を越える戦略を見付け、安定的な戦略の特性を探ることは今後の発展の一つに挙げられる。それとともに、今回の試行の背後に潜むメカニズムをさらに解明することは重要である。

また、N 人数版囚人のジレンマは公共財供給問題(フリーライダー問題)に適用することができる。仮に公共財から得られる利益が大きい場合には、社会の協力者たちが費用を出し合って「裏切り者を摘発し、裏切り者だけに有効な制裁をする専門家および組織<sup>7</sup>」を用意する誘引が発生する。したがって、監視エージェントの導入などして匿名性を排除するといったモデル化も考えられる。

以上のように、N 人版 IPD には様々な方面からの拡張や発展が考えられる。これらは今後の研究に譲るが、本稿が一連の分析への一助となれば幸いである。

## 参考文献

- 有田 隆也, 『人工生命』, 科学技術出版, 2000 年.
- Axelrod,R., *The Evolution of Cooperation*, Basic Books, 1984(松田裕之訳 『つきあい方の科学』, ミネルヴァ書房, 1998 年.).
- Grim,P., "The Greater Generosity of the Spatialized Prisoner's Dilemma," *Journal of Theoretical Biology*, Vol.173, pp.353-359, 1995.
- Joshi,N.N., "Evolution of Cooperation by Reciprocation within Structured Demes," *Journal of genetics*, Vol.66, No.1, pp.69-84, 1987.
- Matsushima,M. and Ikegami,T., "Evolution of Strategies in the three-person Iterated Prisoner's Dilemma Game," *Journal of Theoretical Biology*, Vol.195, pp.53-67, 1998.
- Nakamura,M., Nogami,H. and Iwasa, Y., "Score-dependent Fertility Model for the Evolution of Cooperation in a Lattice," *Journal of Theoretical Biology*, Vol.194, pp.101-124, 1998.
- Nowak,M.A. and May,R., "Evolutionary games and spatial chaos," *Nature*, Vol.359, pp.826-829, 1992.
- Nowak,M.A. and May,R., "The spatial dilemmas of evolution," *Int. J. Bifurcation and Chaos*, Vol.3, pp.35-78, 1993.
- Nowak,M.A. and Sigmund, K., "A Strategy of Win-Stay Lose-Shift that Outperforms TIT-FOR-TAT in the Prisoner's Dilemma Game", *Nature*, Vol.393, No.6432, pp.56-58, 1993.

---

<sup>7</sup> 納税問題であれば、いわゆるマルサがこれに対応する。