5K-3

格子上囚人のジレンマゲームにおける多様な勝者

上原 隆司[2], 鈴木 佳祐[1], 吉村 仁[2], 泰中 啓一[1], 内藤 広海[1]

静岡大学工学部システム工学科[1]

静岡大学創造科学技術大学院[2]

1. 導入

人類および人類以外の社会において、様々な協力の形式が存在する。囚人のジレンマ(PD)ゲームは、個人(プレーヤー)たちが自身と他者を有利にしようとする状況に協力が生じる場合があることを明白に例証する。これまでのところ、利己主義な戦略に対して使用する有効な戦略が多く報告されている。Axelrod と Hamilton が考案したしっぺ返し(TFT)戦略は、反復型囚人のジレンマゲームのコンテストにおいて優勝した。また、Nowak と Sigmud はパブロフ(PAV)戦略を紹介した。ゲームがノイズ(エラー)を含む場合、PAV は TFT に勝利することができる。これらの戦略は、特に格子上で協力の発展について説明することが可能だ。しかしながら、TFT と PAV の両戦略は裏切られたら裏切り返すという「報復主義」に基づく戦略であり、人類における道徳基準よりほど遠い。

恐らく、他人にしてもらいたい行為をする「黄金律」は道徳基準として最も広く容認されるだろう。PD ゲームの場合では、すべて協力する(AC)戦略が黄金律に最も近い戦略である。しかしながら、AC は他の戦略、特にすべて裏切る(AD)戦略および PAV より劣っていると考えられている。AC が勝者であると報告した著者はほとんどいない。今回、我々が行った反復型囚人のジレンマゲームのコンピュータシミュレーションでは、最初に 4 つの戦略(TFT, AD, PAV, AC)が格子上にランダムに分配される。最終平衡では一つの戦略がセルを全て占領することができる。パラメータ値に依存するため、一次元格子上でもすべての戦略は勝利者になり得る。これらの勝利者と法則は数学的な理論に基づいて説明することができる。

2. モデルと手法

反復型囚人のジレンマ(PD)ゲームは、一組のプレーヤーによって行われる。1STEP で、プレーヤーはそれぞれ、協力するかあるいは裏切るかの 2 つのオプションから 1 つを選択する。互いのプレーヤーが協力すれば、両者は支払い R を得る。片方が協力し、もう片方が裏切った場合、前者（後者）は支払い S（T）を得る。互いのプレーヤーが裏切れば、両者は支払い P を得る。今回のシミュレーションでは、以下のスタンダードモデルを適用した。

$$(T,\ R,\ P,\ S) = (5,\ 3,\ 1,\ 0) \quad (1)$$

格子セルを AC、TFT、PAV、AD のうち 1 つの戦略を持ったコロニー（パッチ）を表すと考える。各パッチでは、同一の戦略を持った m+1 のプレーヤーが存在する。

初期条件として、格子に 4 戦略をランダムに配置するが、それぞれは等確率(1/4)で配置される。ランダムに 1 つの格子セル(α)を選択し、次に、隣接した格子セル(β)を 1 つ選択する。それぞれのプレーヤー(セル)は、隣接した格子と同パッチ内セルとゲームをする。例えば、図 1(a)の一次元格子の場合、セルαのプレーヤーは、セルβのプレーヤーおよび青いセルのプレーヤーとゲームを行い、次にα内の同パッチのプレーヤーとゲームを行う。つまり、プレーヤーはそれぞれ m+2d 回ゲームを行う。尚、d は空間の次元を表す。

Multiform winners in Prisoner's Dilemma game on lattice

[1]Keisuke Suzuki, Kei-ichi Tainaka, Hiromi Naitoh
Department of Mathematical and Systems Engineering, Shizuoka University
[2]Takashi Uehara, Jin Yoshimura
Graduate School of Science and Technology, Shizuoka University

(a) (b)

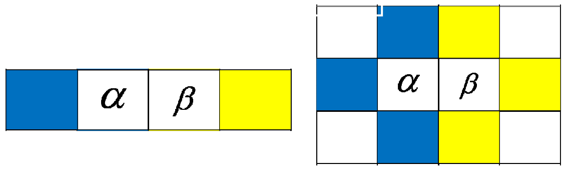

Fig.1 ゲームを行うセルの選択

プレーヤー(セル)同士の勝敗は m+2d 回ゲームを行った際に得た合計得点で決定される。1 つのゲームにつき、反復型囚人のジレンマゲームは N 回繰り返しゲームを行う。得点が低かった方のセルおよびパッチの戦略は、得点が高かった方のセルの戦略に塗り替えられる。もし、両者の得点が同一だった場合、1/2 の確率でどちらかのセルの戦略がもう一方の戦略に塗り替えられる。

3. m→∞の理論

パッチ(コロニー)に無限のプレーヤー(m→∞)がいる場合、勝者が AC であることを証明することができる。次は同戦略間における 1 ステップあたりの平均得点(A)である。

(1) AC vs. AC

両プレーヤーの戦略が AC であった場合、$2A$ を得る。1 ステップで考えられるケースは次の 3 つである。ⅰ)両者エラー有り、ⅱ)片方エラー有り、ⅲ)両者エラー無し

ケースⅰ)、ⅱ)およびⅲ)では、得点の合計は 6、5、および 2 がそれぞれ与えられる。従って、

$$2A = 6(1-x)^2 + 5\times 2(1-x)x + 2x^2$$

よって

$$A = 3 - x - x^2 \quad (2)$$

(2) AD vs. AD

両プレーヤーの戦略が AD であった場合、

$$A = 1 + 3x - x^2 \quad (3)$$

(3) PAV vs. PAV

両プレーヤーの戦略が PAV であった場合、反復型囚人のジレンマゲームを N 回行う時、$n(1 \leqq n \leqq N)$ 回目では相手プレーヤーは C あるいは D のいずれかを示す。相手プレーヤーがオプション $j(k)$ (j, k=CorD)を選択する確率 P は以下のように与えられる。

$$P_{cc}(n+1) = (1-x)^2 P_{cc}(n) + x^2 P_{cd}(n) + x^2 P_{DC}(n) + (1-x)^2 P_{DD}(n) \quad (4a)$$

この推移確率はシャノン図(Fig2.)から与えられる。jk によって示された円は、プレーヤー(相手)が $j(k)$ の状態を示すことを意味する。同様の方法で Fig2. から以下の式を得る。

$$P_{CD}(n+1) = x(1-x)P_{CC}(n) + x(1-x)P_{CD}(n) + x(1-x)P_{DC}(n) + x(1-x)P_{DD}(n) \quad (4b)$$

$$P_{DC}(n+1) = x(1-x)P_{CC}(n) + x(1-x)P_{CD}(n) + x(1-x)P_{DC}(n) + x(1-x)P_{DD}(n) \quad (4c)$$

$$P_{DD}(n+1) = x^2 P_{CC}(n) + (1-x)^2 P_{CD}(n) + (1-x)^2 P_{DC}(n) + x^2 P_{DD}(n) \quad (4d)$$

式(4a-d)から以下の式が導き出せる。

$$A = 3 - 5x + 11x^2 - 8x^3 \quad (5)$$

(4) TFT vs. TFT

エラーが生じなければ(x=0)、A=3 である。エラーが生じる場合、PAV と同様の方法で平均得点 A を算出できる。

$P_{cc}(n) - P_{DD}(n) = Q(n)$ と置いて、

$$Q(n+1) = (1-2x)Q(n) \quad (6)$$

この回帰方程式を解いて

$$Q(n) = (1-2x)^{n-1} \tag{7}$$

$n\to\infty$の時、$Q(n)\to 0$　これは $P_{CC}(n) = P_{DD}(n)$を意味する。同様にして、$P_{jk}(n)$が任意の j あるいは k でも同じ値になることを証明できる。したがって、

$$A = 2.25 \qquad \text{for } x\neq 0 \tag{8}$$

上記のケース(1)-(4)の結果により、$m\to\infty$の場合の勝者を決定することができる。Fig.2 から得点が最も高くなる戦略は AC であることがわかる。AC と比較すると TFT と PAV の両方が小さい値を取るが、それは報復主義によるものである。x=0.1 の時の理論値を Table1 に示す。

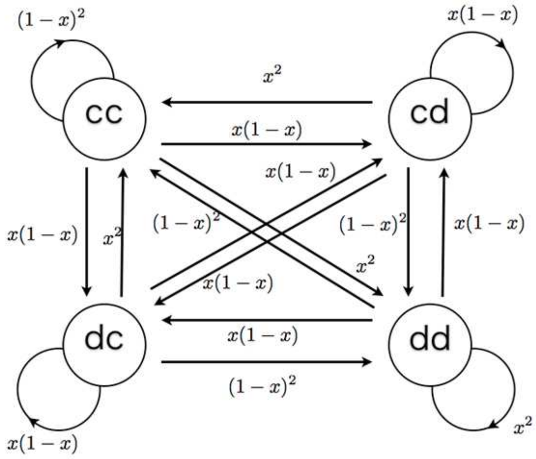

Fig.2 シャノン図と確率推移

Table.1

	Opponents			
	AC	PAV	TFT	AD
AC	$3-x-x^2$ (2.89)	1.5+1.5x (1.65)	$3-4x+4x^2+2x^3$ (2.642)	$4x+x^2$ (0.41)
PAV	4-3.5x (3.65)	$3-5x+11x^2-8x^3$ (2.602)	2.25 (2.25)	0.5+3.5x (0.85)
TFT	$3+x-6x^2+2x^3$ (3.402)	2.25 (2.25)	2.25 (2.25)	$1+2x+2x^3$ (1.202)
AD	$5-6x+x^2$ (4.41)	3-1.5x (2.85)	$1+7x-10x^2+2x^3$ (1.602)	$1+3x-x^2$ (1.29)

4. 一次元格子上の結果

4-1. シミュレーション結果

m(パッチ数)と x(ノイズ)の値を変えてシミュレーションを行った。今回のシミュレーションでは、ゲームに負けたセルは勝ったセルの戦略に塗り替えられる。従って、最終的には 1 つの戦略が他の戦略に打ち勝つ。Fig.3 で代表的なパターンダイナミクスを示す。横軸と縦軸は空間と時間を表す。m=0、x=0 の時: Fig3.(a)より TFT が勝利していることが分かる。エラーが無い(x=0)限り、TFT は AD に勝つ。その一方で、AD は PAV と AC に勝つ。「ギャンブラーの破産問題」(Epstein 1995)によって TFT が勝利者となる。m=0、x=0.1 の時: Fig3.(b)より TFT、PAV、AD の「じゃんけんゲーム」を示す。格子サイズが有限の場合、1 つの戦略が残存する(Itoh and Tainaka 1994)。勝利者は、TFT、PAV および AD の中からランダムに決定される。m≠0、x≠0 の典型的なダイナミクスは Fig.4 である。m=1 から m=5 までは PAV が占領する。m=1 の場合は、ほとんどのセルが PAV によって占領されるが、1 個あるいは 2 個の AD セルは残存する。m=1 において AD が 2 個を超えるセルを占領する場合、AD は PAV に打ち負かされるが、AD が 1 つのセルを占領する場合、AD は PAV を打ち負かすからである。m=6 を超えると AC が占領するようになる。m=6 でも m=1 の時と同様の現象が見られる。

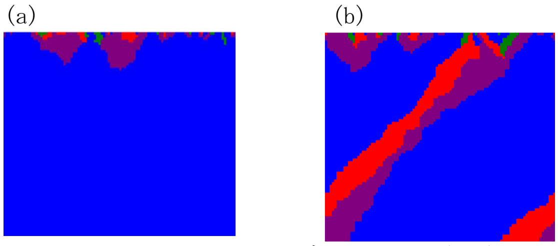

Fig.3 パターンダイナミクス(m=0)

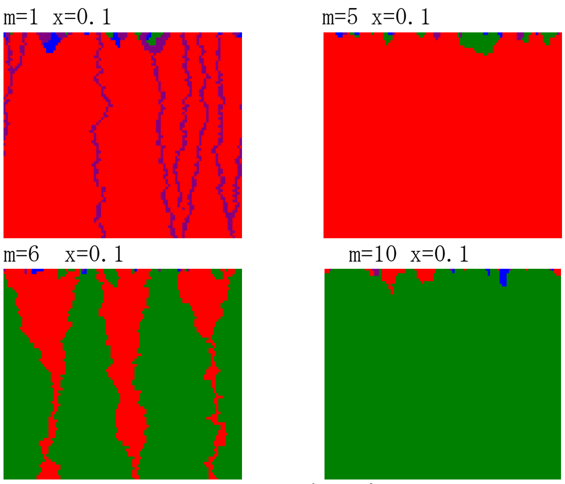

Fig.4 m≠0 の場合のダイナミクス

5. 考察

前セクションで述べた理論はシミュレーション結果を利用して全て説明することができる。まず、m=0 および x=0.1 の場合のじゃんけん状態について説明する。例えば、Fig.5(a)のような TFT と PAV のクラスタ間でのゲームを考える。Table1 によれば、セル A の平均得点は 2.25+2.25=4.5 で与えられる。セル B の場合は 2.602+2.25=4.852 である。したがって、PAV はゲームに勝つ。同様に考えると、AD は PAV に対して勝利するが、TFT に対しては負ける。従って、TFT、PAV、および AD は単にじゃんけんの関係と同じである。次に、m=1 および x=0.1 になぜ AD セルが 1,2 個残存するのかを考察する。例えば、Fig5(b)のような PAV の中に AD が存在する場合を考える。セル A の平均得点は 2×2.85+1.29=6.99 から与えられる。一方、セル B は 2×2.602+0.85=6.054 を得る。つまり、AD は PAV に勝利する。しかしながら、AD2 個の近隣のセルが PAV だった場合、勝利者は逆になる。AD のスコアは 2×1.29+2.85=5.43 であるが、PAV のスコアは 2×2.602+0.85=6.054 である。従って、AD の単独セルは PAV の海の中をランダムウォークする。格子が有限の場合、1 つの AD セルだけが最終平衡で残る。同様の方法で、m=6 および x=0.1 の AC の海の中を PAV のセルがランダムウォークすることを証明することができる。

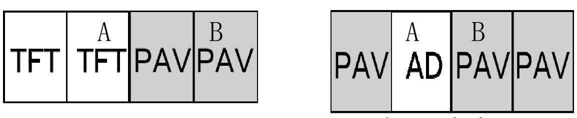

Fig.5 シミュレーション結果の考察

人間社会では、人々は皆友人を持つ。同じ戦略を持つ友人のコミュニティは一種のコロニー (パッチ) とみなすことができる。多くの動物の中で形成される「家族」は尚更コロニーとみなすことができるだろう。社会性昆虫は多くの個人(m>>1)からなるコロニーを形成する。「類似理論」(Hamilton 1964)を要求しない我々のモデルは、アリまたはスズメバチのような利他主義の発展について説明することが可能である。