

囚人のジレンマゲームにおける協調行動とプレイヤーの合理性

小樽商科大学 商学部 社会情報学科
行方 常幸 (Tsuneyuki Namekata)

1. はじめに

本稿では個人的な目的と集団的な目的が一致しない社会的ジレンマの最も基本的な場合である 2 人囚人のジレンマゲームを取り上げる。この囚人のジレンマゲームは純粋に理論的な観点、社会心理学的な観点、社会学的な観点、等、様々な観点から研究されている。本稿では理論的な観点到重点をおき、他の観点を取り入れながら、協調行動を正当化する方法を合理性の多様な表現として捉え鳥瞰することを目標とする。

2. 囚人のジレンマゲームとは?

囚人のジレンマゲームは戦略形が表 1 で与えられるゲームである。自分の利得をより多くしたい各プレイヤーは、相手の取る戦略が C (協調) であれ D (裏切) であれ、自分は D (裏切) を取った方が自分の利得が大きくなるので、

表 1. 囚人のジレンマゲーム

		プレイヤー2	
		C (協調)	D (裏切)
プレイヤー1	C (協調)	1,1	b,a
	D (裏切)	a,b	$0,0$

$$a > 1, b < 0, a + b < 2$$

(支配戦略である) D を取り、結果として唯一のナッシュ均衡である(D,D)が実現する。しかし、この結果は両方のプレイヤーにとってお互いが C を取った時の結果よりも悪い状況であり、これがジレンマと呼ばれる理由である。

すなわち、通常のゲーム理論によると、このゲームの起こりうる唯一の結果として(D,D)が正当化され、協調行動(C,C)は正当化されない。

3. 協調行動(C,C)を正当化する種々の方法

この囚人のジレンマゲームに対して協調行動(C,C)の実現を正当化する試みがなされている。まず、ナッシュ均衡として協調行動を説明する試みがある。(1 回限り、有限回繰返し) 囚人のジレンマゲームはナッシュ均衡として協調行動を含まないので、この方法は元のゲームを新しいゲームに変換することを意味

する。さらに、新しく得られたゲームは囚人のジレンマゲームではないことに注意する。すなわち、プレイヤーは意識的に異なった観点から元の囚人のジレンマゲームを捉えることになる。このナッシュ均衡として協調行動を説明する試みとして、次の4つを取り上げる。

- (1) 無限回繰返しゲーム
各プレイヤーは割引率 δ (十分 1 に近い) で割引かれた無限期間にわたる自分の利得の総和を最大化する。
- (2) 有限回繰返しゲーム (ε -均衡 ; Radner (1978))
各プレイヤーは (十分に長い) 有限期間にわたる自分の利得の総和を最大値に十分近く (ε 以内に) する。
- (3) 有限回繰返しゲーム (p -TFT タイプ ; Kreps, Milgrom, Roberts, Wilson (1982))
一方のプレイヤーは 2 つのタイプからなる ; 確率 p (十分 1 に近い) で TFT を利用するタイプであり、確率 $1-p$ で通常の自分の利得の総和を最大化するタイプである。
- (4) 有限回繰返しゲーム (Bounded Complexity ; Neyman (1985))
各プレイヤーの利用できる戦略は有限オートマトン (サイズが 2 以上かつ期間数未満) に限定される。

これら 4 つを以下に簡単に説明する。

(1)の概略は表 2 のように新しくゲームを見ていることになる。この表では、1

表 2. 無限回繰返しゲーム

	TFT	ALT	AIID
TFT	1,1	$\frac{b+\delta a}{1+\delta}, \frac{a+\delta b}{1+\delta}$	$(1-\delta)b, (1-\delta)a$
ALT	$\frac{a+\delta b}{1+\delta}, \frac{b+\delta a}{1+\delta}$	$\frac{\delta}{1+\delta}, \frac{\delta}{1+\delta}$	$\frac{\delta b}{1+\delta}, \frac{\delta a}{1+\delta}$
AIID	$(1-\delta)a, (1-\delta)b$	$\frac{\delta a}{1+\delta}, \frac{\delta b}{1+\delta}$	0,0

回限りの利得表と比較可能なように、各プレイヤーの利得は δ で割引かれた無限回にわたる自分の利得の総和を $(1-\delta)$ 倍してある。TFT はしっぺ返し戦略 (後述) であり、ALT は C と D を交互に繰返す戦略であり、AIID はいつも D を取

る戦略である。この利得表から明らかなように、割引率が十分 1 に近ければ、すなわち、 $\delta \geq \max \left\{ \frac{a-1}{b}, \frac{a-1}{1-b} \right\}$ ならば (TFT, TFT) はナッシュ均衡となる。

(2) は表 3 のように利得表を見ていることになる。ここで、利得表の数値は 1

表 3. 有限回繰返しゲーム

	TFT	OPT
TFT	1, 1	$\frac{N-1+b}{N}, \frac{N-1+a}{N}$
OPT	$\frac{N-1+a}{N}, \frac{N-1+b}{N}$	$\frac{N-1}{N}, \frac{N-1}{N}$

期間あたりの平均利得であり、 N は期間数であり、OPT は TFT に対する最適反応戦略である (すなわち、始めの $N-1$ 期間は C を取り、最後の期は D を取る戦略)。この利得表により、 $N \geq \frac{a-1}{\varepsilon}$ ならば、(TFT, TFT) が ε -均衡となる。

(3) の結果だけを述べると、 $p \geq \max \left\{ \frac{a-1}{a}, \frac{-b}{1-b} \right\}$ ならば、残り数期間を残して協調行動を実現する均衡が存在する。

(4) においては、図 4 で与えられる TFT がそれ自身の最適反応となる。なぜなら、(2) で述べた OPT が TFT の最適反応戦略となるが、この OPT はサイズが期間数未満であるという (4) の仮定により、利用することが出来ないためである。

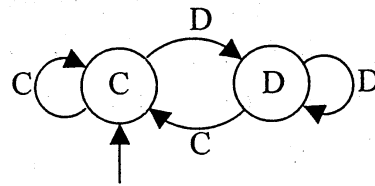


図 4. TFT のオートマトン表現

次に、ナッシュ均衡以外の概念を利用して 1 回限りの囚人のジレンマゲームにおいて協調行動を説明する試みがある。この (5)、(6)、(7) に関してはその考え方だけ述べる。詳細は元の文献を参照されたい。

(5) 有限回繰返し (Cooperating Equilibrium ; Rosenthal (1980))

自分が逸脱しない場合に得られると期待する最悪の利得よりも、自分が逸脱した場合に得られると期待する最悪の利得の方が大きくなければ、

逸脱しない。

(6) Farsighted Coalitional Stability (Chwe (1994))

各提携が、先を見越して自由に公表して意見交換する場合の安定な戦略の組の集合。

(7) ε -安定性 (Namekata (1997))

2人のプレイヤーが同時に戦略を変える確率 ε (十分1に近い) を考慮する。

以上の(1)から(4)において、プレイヤーの合理性は次のように想定されている：(1)では自分の利得の厳密な最大化を目指す。(2)では厳密な最大化ではなくて ε の誤差を許容する。(3)ではプレイヤーは確率 p で自分の利得の最大化を必ずしも目指さない(協調行動を促進する)TFTを利用するタイプである。(4)ではプレイヤーの利用できる戦略は有限オートマトンに限定される。(2)から(4)では利得の最大化を部分的に限定しているが、基本的には自分の利得の最大化を目指す。(5)では均衡という名がついているが、互いの戦略が他の最適反応になっているという意味の均衡ではなく、相手の取る戦略に対する揺らぎを考慮している。すなわち、自分の利得の単純な最大化ではない。(6)と(7)ではそれが提唱する解を求める過程で(C,C)と(D,D)の結果((1,1)と(0,0))の比較(共通の利益に対する考慮)を行っている。

プレイヤーが囚人のジレンマゲームに面した場合、多くは、1回限りのゲームではないであろう。もし、無限回繰返しゲームであれば、(1)のように協調行動を正当化できる。また、有限回であっても、利得の最大化を厳密には求めない場合、相手に協調を目指す可能性を想定できる場合、相手が利用できる戦略が限定される場合は、それぞれ(2)から(4)のように、協調行動を正当化できる。実際われわれが他の人と出会う場面において、1回限りの場合は珍しいし、厳密な最大化など望みはしないし、相手に協調を目指す可能性を想定するだろうし、相手が利用できる戦略は限定されているとみなす場合が多い。従って、(1)から(4)において、モデル内で想定されているプレイヤーの合理性は普通のわれわれの限定された合理性を反映するようになっている。詳しくは述べなかったが、(5)から(7)の考え方もわれわれの合理性を表現したものになっている。しかし、(1)から(4)の方法では、1回限りのゲームにおいて、協調行動を正当化できないことに注意を要する。

また、そのモデルによる協調行動の正当化の(計算)過程を理解するためには読者にかなりの合理性が要求される。従って、より容易に理解可能な正当化の過程が望まれる。この線に沿ったものとして以下が考えられる。

(8) 利得行列の主観的変換 (Pruitt & Kimmel の目標/期待理論; 1977) 囚人のジレンマ的状况を何度も経験するうちに、主観的に保証ゲーム (表 5) に変換する。

表 5. 囚人のジレンマゲームの保証ゲームへの主観的変換

		プレイヤー2	
		C (TFT)	D (AllD)
プレイヤー1	C (TFT)	<u>1,1</u>	y,x
	D (AllD)	x,y	<u>0,0</u>

$$x=(1-\delta)a < 1, y=(1-\delta)b < 0$$

(1) に関しては括弧内を読む。

(1)のゲームの利得表 (表 2) の ALT の行と列を削除すれば、表 5 の利得表と一致する。すなわち、(1)においては自分の利得の総和という客観的操作を用いてゲームを変換したが、(8)においては主観的に行っている。(6)と(7)と同様に(8)においても共通の利益に対する考慮がモデルの中にすでに組み込まれている。保証ゲームには純粹戦略におけるナッシュ均衡(C,C)と(D,D)があるので、利得支配している(C,C)が選ばれる可能性がある。この協調行動の説明は読者にも容易に納得できるものである。

利得行列の主観的な変換に関しては、例えば、次のようなものがある。

(A) Social Orientation (Schulz, May (1989))

自分の利得と相手の利得から自分の効用を構成する場合にいくつかのタイプ (C:協調的、K:競争的、I:個人的、等) が存在する：

$$\text{自分の効用} : u(p_o, p_y) = ap_o + bp_y$$

p_o :自分の利得, p_y :相手の利得

$$a = b = 1 \Rightarrow C$$

$$a = 1, b = -1 \Rightarrow K$$

$$a = 1, b = 0 \Rightarrow I$$

協調的なタイプのプレイヤーが見るゲームは $a+b < 0$ ならば表 6 のように保証ゲームなる。

(B) 利他主義のモデル (Andreoni, Miller (1993))

表 6. 囚人のジレンマゲームの変換 1

互惠的利他主義: 両方のプレイヤーが協調を取ったときだけ α の効用を得

		プレイヤー2	
		C(協調)	D(裏切)
プレイヤー1	C(協調)	2,2	<u>$a+b, a+b$</u>
	D(裏切)	<u>$a+b, a+b$</u>	<u>0,0</u>

$$a > 1, b < 0, a+b < 0$$

る。この互惠的利他主義者が見るゲームは $\alpha > a-1$ ならば表 7 のように保

証ゲームになる。

通常、ゲーム理論では利得行列に記入されている数値は効用である。すなわち、上記の(A)や(B)による変換後の数値が記入されていると想定されている。しかし、これは余りにも安易な想定である。現

実をより良くモデル化するためには、この変換もモデル化の重要な部分であるので、この変換の過程もゲーム理論の範囲内で考察すべきである。この変換が客観的に行われる場合が(1)であり、主観的に行われる場合が上記の(A)や(B)である。このように考える状況に応じて様々な変換があり得る。

以上、プレイヤーは囚人のジレンマゲームを主観的に保証ゲームとみなしている可能性があることが分かる。従って、1回限りでも協調行動が実現する可能性がある。このように1回限りの囚人のジレンマゲームにおいて協調行動を正当化できる点が主観的変換の特徴である。自分がDを取るのは自分の利得を大きくするためではなく、相手がDを取る可能性を高く見積もっているために、ナッシュ均衡(D,D)の実現を目指しているからである。

ここまで、囚人のジレンマゲームにおける協調行動の正当化をプレイヤーの様々な合理性の表現として捉えてきた。その道筋は、まず、繰返しゲームにより利得表を保証ゲームへ客観的に変換し、次に、従来の厳密な合理性を弱め限定的な合理性に限り、そして、最後に利得表を主観的に変換した。この最後の主観的変換によって、1回限りの囚人のジレンマゲームにおいても協調行動を正当化できた。

この道筋を逆に辿ってみる。まず、われわれは囚人のジレンマゲームに面した時、利得表を主観的に変換してしまう。その後、その変換したゲームを繰返し行うことを頭に描く。この繰返しゲームを考える時に、通常のゲーム理論が扱うように利得の総和の最大化を目指し、過去の履歴に依存した戦略を考慮するということはしない。すなわち、プレイヤーに

表 7. 囚人のジレンマゲームの変換 2

		プレイヤー2	
		C(協調)	D(裏切)
プレイヤー1	C(協調)	$1+\alpha, 1+\alpha$	b, a
	D(裏切)	a, b	$0, 0$

$$\alpha > a-1, a > 1, b < 0, a+b < 2$$

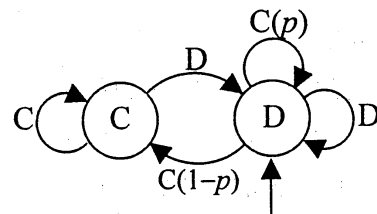


図 8. Suspicious Tit-for-Tat (SuTFT(p) (0 ≤ p < 1))

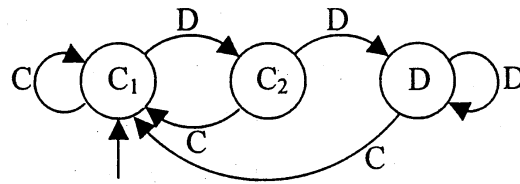


図 9. Tit-for-Two-Tats
(TF2T)

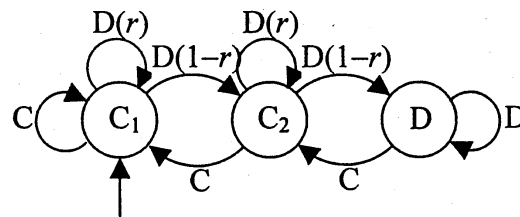


図 10. Sticky Tit-for-Two-Tats
(STF2T(r)) ($0 \leq r < 1$)

そこまでの合理性を要求しない。ただ、プレイヤーは保証ゲームである表 5 の利得表における 2 つのナッシュ均衡のどちらかの実現を望む。このような戦略として図 8 から図 10 で与えられる 3 つの戦略を考察する。これら 3 つの戦略は自分と相手が共に C (または D) をとれば、C (または D) を取り続けるという意味で、ナッシュ均衡である(C,C)または(D,D)の実現を望んでいる。しかしながら、初期にどちらを狙うか、また、その実現をどれくらい強く望むかの点で違っている。図 8 の Suspicious Tit-for-Tat (SuTFT) は TFT とよく似ているが、初期に(D,D)を狙う点が大きく違う。また、自分が D をとっている時に相手が C をとっても、それを(C,C)実現の相手のサインと単純にみなさない警戒心があり、それを確率 p で表している。図 9 の Tit-for-Two-Tats (TF2T) と図 10 の Sticky Tit-for-Two-Tats (STF2T) の 2 つの戦略は初期的には(C,C)の実現を望む。しかしながら、(C,C)の実現を望む方法が若干違う。STF2T において(C,C)を望む意思の強さを確率 r で表している。

これら 3 つの戦略が繰り返しゲームを行ったらどのような結果になるだろうか？ これらすべての戦略が(C,C)または(D,D)の実現を望んでいるのだから、長い時間がたてばどちらかが実現することが期待される。TF2T と STF2T が対戦すれば(C,C)が実現することは明らかである。そこで、SuTFT と TF2T または STF2T を対戦させてみた (行方)。解析的に求めることは困難なので、 $p=r=0$ から 0.01 刻みでそれぞれ 100 個、総数 10,000 個の組み合わせについて最終的に何が実現

するかを計算したのが図 11 のグラフである。横軸が p 、縦軸が r で、OXY の左上が(C,C)、右下が(D,D)の実現する領域である。TS は直線 $p=0.63$ である。 $p=0.63$ の時は TF2T では(C,C)を実現することは出来なかったが、 $r=0.2$ の STF2T では(C,C)を実現することが出来る。また、 $p=0.9$ と SuTFT の警戒心がかなり強い場合でも、 $r=0.74$ で(C,C)を実現できる。すなわち、警戒心がかなり強い SuTFT を相手にする場合、協調行動を実現するという観点に立てば、TF2T よりも(C,C)実現の意思が強い STF2T の方が有効であることが分かった。

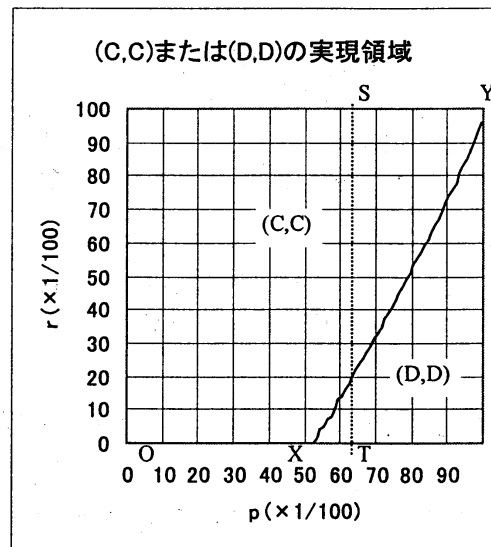


図 11. グラフ

4. 終わりに

本稿では囚人のジレンマゲームにおいて協調行動を正当化する様々な方法をプレイヤーの合理性という観点から調べた。ここでいう合理性とはプレイヤーがどのように考えるかという漠然とした分類であり、厳密に定義されたものではない。しかしながら、従来のゲーム理論において考察されてきた、厳密に定義された合理性から、いかなる方向へ歩みだすかを見出すためには、貴重な視点を提供している。われわれが協調行動を望むとき、どのように考えて行動を取るのだろうか？本稿で述べた方法の中にその答えがあれば、その方法を更に検討し充実していくことがゲーム理論のやるべき仕事である。

われわれが社会生活を行うときにどのように考えて（どのような合理性に基づいて）行動しているのかは複雑である。本稿でみたように簡単に割り切るこ

とは出来ないであろう。しかしながら、より包括的なゲーム理論を構築するために、様々な考え方を取り入れていくべきである。

参考文献

- Andreoni, J. & Miller, J.H. (1993) Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence, **The Economic Journal** 103, 570-585.
- Chwe, M.S.-Y. (1994) Farsighted Coalitional Stability, **Journal of Economic Theory** 63, 299-325.
- Kreps, D.M., Milgrom, P., Roberts, J. & Wilson, R. (1982) Rational Cooperation in the Finitely Repeated Prisoners' Dilemma, **Journal of Economic Theory** 27, 245-252.
- Namekata, T. (1997) Prisoner's Dilemma and the Relation between Players, **Mathematica Japonica** 45, 433-439.
- Neyman, A. (1985) Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoners' Dilemma, **Economic Letters** 19, 227-229.
- Pruitt, D.G. & Kimmel, M.J. (1977) Twenty Years of Experimental Gaming: Critique, Synthesis, and Suggestions for the Future, **Annual Review of Psychology** 28, 363—392.
- Radner, Roy. (1980) Collusive Behavior in Noncooperative Epsilon-Equilibria of Oligopolies with Long but Finite Lives, **Journal of Economic Theory** 22, 136-154.
- Rosenthal, R.W. (1980) New Equilibria for Noncooperative Two-Person Games, **Journal of Mathematical Sociology** 7, 15-26.
- Schulz, U. & May, T. (1989) The Recording of Social Orientations with Ranking and Pair Comparison Procedures, **European Journal of Social Psychology** 19, 41-59.
- 行方 常幸、行方 洋子 (1999)「プレイヤーの合理性と囚人のジレンマゲーム」(研究ノート)『理論と方法』26(2), 127-133。