日本オペレーションズ・リサーチ学会和文論文誌
Transactions of the Operations Research Society of Japan
Vol. 60, 2017, pp. 100–112

（all D, all C）が一方のプレイヤーにとって最善のナッシュ均衡となる非対称繰り返し囚人のジレンマゲームについて

河野 敬雄

(受理 2017 年 3 月 21 日; 再受理 2017 年 10 月 3 日)

和文概要　従来の繰り返し囚人のジレンマゲームはフォーク定理に代表されるように，非支配戦略である協力行動 C が如何にして合理的選択の結果として実現し得るか，という文脈で定式化されてきた．

これに対して本論文においては，プレイヤー 2 が支配戦略である D 戦略を選択した場合，直ちにゲームを中止する，というサンクションを課すルールを設定した非対称な繰り返し囚人のジレンマゲームにおいて，割引率 δ が十分 1 に近い時，プレイヤー 1 が常に D 戦略を，プレイヤー 2 が常に C 戦略を選択する戦略セット $(\mathrm{all}\,D, \mathrm{all}\,C)$ がナッシュ均衡となり，プレイヤー 1 にとって最善のナッシュ均衡戦略となることを証明する．

キーワード: ゲーム理論，非対称繰り返し囚人のジレンマゲーム，ナッシュ均衡

1. はじめに

従来の標準的ゲーム理論において，ワンショットの標準形ゲーム (stage game) を無限回繰り返すゲーム[1]は展開形ゲームの 1 種とみなされ，展開形ゲームの定式化が用いられてきたが[2]，その場合，繰り返しゲームにおける行動戦略を含めた各プレイヤーの「戦略」の概念が必ずしも明確ではない．その結果として，ある戦略セットが繰り返しゲームにおけるナッシュ均衡であることの数学的証明に厳密さが欠けるうらみがあった．本論文においては，既に河野 [4, Section 6] あるいは [5] において試みたように，確率変数列，すなわち離散パラメータを持つ確率過程を用いて定式化することによって，相手が D 戦略を選択した場合，直ちにゲームを中止する，というサンクションを課すルールを設定した非対称な繰り返し囚人のジレンマゲームにおいて，割引率 δ が十分 1 に近い時，戦略セット $(\mathrm{all}\,D, \mathrm{all}\,C)$ がナッシュ均衡となり，$\mathrm{all}\,D$ プレイヤーにとって最善の戦略となることを数学的に厳密に証明することができることを示す．ただし，本論文は繰り返しゲーム一般を定式化することが目的ではないのでワンショットゲームが囚人のジレンマゲームの場合，すなわち，純戦略集合として $S = \{C, D\}$ に限定した必要最小限度の定式化を行う．先行研究と補足的議論については最後の 4 節を参照されたい．

2. 囚人のジレンマゲーム

次のような 2×2 の標準形ゲームを考える．なお，通常の囚人のジレンマゲームには 2 人のプレイヤーの役割に区別がないが本論文で考察する繰り返し囚人のジレンマゲームにおいては 2 人のプレイヤーの役割が異なることを仮定するから，彼らのワンショットゲームの利

[1]厳密にいうと有限回で終わることが確実ではあるがいつ終了するかが不確実なゲームのこと．詳しくは 4 節の (1) を参照のこと．

[2]例えば岡田 [17] の Section 6.2 を参照されたい．なお，河野 [9,10] が定義した展開形ゲームには本論文で考察する繰り返し囚人のジレンマゲームは含まれないことに注意．

表 1: 非対称囚人のジレンマゲームの利得表

		プレイヤー 2	
		C	D
プレイヤー 1	C	(b_1, b_2)	(d_1, a_2)
	D	(a_1, d_2)	(c_1, c_2)

得についても異なっていることを認めることにする．そのメリットは割引率 δ に関する条件がいずれのプレイヤーの利得に依存しているかが明らかになることである．ただし，囚人のジレンマ構造は両者ともに持っているとして各利得に対しては通常の 囚人のジレンマゲームの仮定と同様に純戦略 D が共に双方のプレイヤーにとって真に優越戦略となるように次の仮定をおく [3].

仮定 2.1. $a_i > b_i > c_i > d_i > 0\,;\, i = 1, 2.$

後の数学的定式化のために純戦略の組に対するプレイヤー $i;\, i = 1, 2$ の利得を $S \times S$ 上で定義された効用関数 f_i で表す．すなわち，

$$f_1(C,C) = b_1,\ f_1(C,D) = d_1,\ f_1(D,C) = a_1,\ f_1(D,D) = c_1,$$

$$f_2(C,C) = b_2,\ f_2(C,D) = a_2,\ f_2(D,C) = d_2,\ f_2(D,D) = c_2.$$

ただし，カッコ内の記号は左がプレイヤー 1 の，右がプレイヤー 2 の純戦略を表す.

次に繰り返しゲームを離散確率過程として定式化するためにワンショットゲーム自体も確率変数を用いて定式化し直しておく [4]．つまり，各プレイヤーの混合戦略とは S 値確率変数のことに他ならない [5]．ここで，$S \times S$ 値確率変数の全体を $\mathcal{R}(S \times S)$ と記す．プレイヤー 1 の戦略を X_1，プレイヤー 2 の戦略を X_2 とするとき，戦略の組 $(X_1, X_2) \equiv Z \in \mathcal{R}(S \times S)$ を戦略セットと呼ぶ．従来の標準的ゲーム理論のように戦略を分布で表現した場合，戦略セットは単に 2 つの分布のペアとして表現されるが，確率変数で表現した場合，単に 2 つの確率変数のペアで表現しただけでは不十分である．何故ならば，2 変数の確率変数の確率構造が定式化されていないからである．非協力ゲームであることを表現するためには X_1 と X_2 は独立な確率変数であることを大前提としなければならない [6]．改めて定義しておく.

定義 2.1. $Z = (X_1, X_2) \in \mathcal{R}(S \times S)$ がワンショットゲームの戦略セットであるとは，X_1 と X_2 が独立であるときをいう． ワンショットゲームの戦略セットの全体を $\mathcal{R}(G)$ とする.

---

[3]通常の繰り返し囚人のジレンマゲームにおいては，さらに $2b_i > a_i + d_i$ を仮定するが，本論文では必要としない．逆に通常は利得の正値性は仮定しないが本論文では必要である.

[4]河野 [4] ですでにそのように定式化している.

[5]正確には，確率変数の定める分布と混合戦略が 1 対 1 に対応している．従って，分布が等しい 2 つの確率変数は戦略としては同一視する．ここで確率変数が定義されている抽象確率空間 $(\Omega, \mathcal{F}, P)$ は本論文で必要な確率構造を表現できるほどに十分「リッチ」な確率空間を設定しておく（存在定理はよく知られている）.

[6]戦略セットを 2 つの分布のペアで表現した途端に確率論的にはそれらは独立なランダム事象の分布であることを暗黙の内に仮定していると理解せざるを得ない．しかし，確率論的には周辺分布を与えただけでは両プレイヤーを含めた全体の確率構造は定まらない．実際，たとえば Luce and Raiffa [11, pp.115–116] に述べられている preplay communication の概念を導入した枠組みでの correlated equilibrium の概念を定式化するためには確率変数 X_1 と X_2 の独立性に代えて X_1 と X_2 を含む全体の確率構造を仮定する必要がある．詳しくは河野 [6, 7]，[8, Section 4] を参照されたい．戦略としての 2 つの確率変数の独立性の仮定はゲーム理論的には両プレイヤー間にコミュニケーションがない（出来ない）ということを表現している．しかし，ワンショットゲームである標準形ゲームだけを考察している間はわざわざ戦略を確率変数で表現することのメリットは少ないために多くの標準的ゲーム理論の教科書では分布を用いて戦略を定義している.

いま，ワンショットゲームの戦略セットを $Z=(X_1,X_2)$ とするとき，プレイヤー $i\,;\,i=1,2$ の期待利得 $u_i(X_1,X_2)$ は通常の確率論の記号を用いて次のように定義できる [7].

$$u_i(X_1,X_2)=E[f_i(X_1,X_2)],\ i=1,2.$$

このとき，ワンショット囚人のジレンマゲームのナッシュ均衡は次のように定式化される.

定義 2.2. 戦略セット $Z^*=(X_1^*,X_2^*)\in\mathcal{R}(G)$ がナッシュ均衡であるとは次の 2 つの不等式が満たされる時をいう.

(1) 任意の $(X_1,X_2^*)\in\mathcal{R}(G)$ に対して，$u_1(X_1^*,X_2^*)\geq u_1(X_1,X_2^*)$ が成り立つ.

(2) 任意の $(X_1^*,X_2)\in\mathcal{R}(G)$ に対して，$u_2(X_1^*,X_2^*)\geq u_2(X_1^*,X_2)$ が成り立つ.

なお，以上の定式化は標準形ゲーム一般に対して適用できる（河野 [4]）.

3. 繰り返し囚人のジレンマゲーム

さて，ここからいよいよ繰り返し囚人のジレンマゲームの定式化を行う．繰り返しゲームとは，ワンショットの同時手番ゲームを繰り返し行うことであるが，ここで大事なことは「戦略」とは何か，という定義の問題である．直感的には n 回目のワンショットゲームにおける各プレイヤーの戦略は $(n-1)$ 回までの両者の「手」（プレイ）に依存してよい．ということである．たとえばよく知られている TFT(tit for tat，しっぺ返し，オウム返し) 戦略や，Trigger 戦略等がそうである．そのために，繰り返しゲームとその戦略を確率変数列で表現しよう [8].
いま，$n=1,2,\ldots$ に対して，すべての $\{Z^{(k)}=(X_1^{(k)},X_2^{(k)})\in\mathcal{R}(S\times S)\,;\,k=1,\cdots,n\}$ を可測にする最小の σ 加法族を $\mathcal{F}^n$ とおく [9]．(なお，$\mathcal{F}^0=\{\emptyset,\Omega\}$，$\mathcal{F}^\infty=\vee_{n=1}^\infty\mathcal{F}^n$，つまりすべての $\mathcal{F}^n$ を含む最小の σ 加法族とする.)

定義 3.1. $\vec{Z}\equiv(Z^{(1)},Z^{(2)},\ldots,)\,;\,Z^{(n)}=(X_1^{(n)},X_2^{(n)})\in\mathcal{R}(S\times S)\,;\,n=1,2,\ldots$ が繰り返し囚人のジレンマゲームの戦略セットであるとは，$Z^{(n)}=(X_1^{(n)},X_2^{(n)})\,;\,n=1,2,\ldots$ が次の 2 つの条件を満たすときをいう.

(1) $Z^{(n)}=(X_1^{(n)},X_2^{(n)})$ は $\mathcal{F}^n-$ 可測 $(n=1,2,...)$.

(2) $X_1^{(1)}$ と $X_2^{(1)}$ は独立．$n=2,3,...$ に対して $X_1^{(n)}$ と $X_2^{(n)}$ は $\mathcal{F}^{n-1}$ に関して条件付き独立 [10]，すなわち，次の等式が成り立つ.

$$P(X_1^{(n)}=i,X_2^{(n)}=j/\mathcal{F}^{n-1})=P(X_1^{(n)}=i/\mathcal{F}^{n-1})P(X_2^{(n)}=j/\mathcal{F}^{n-1})\,;\,\forall i,j\in S.$$

---

[7] ここで，$E[*]$ は期待値，つまり Ω 上の抽象ルベーグ積分を表す．なお，後で出てくる記号 $E[*;A]$ は事象 $A\in\mathcal{F}$ 上の積分を表す．つまり，$E[*]$ は $E[*;\Omega]$ のことである.

[8] 河野 [4, Section 8] あるいは [5] を参照されたい．従来の標準的ゲーム理論の教科書における定式化との関係については最後の 4 節を参照されたい.

[9] 2^{2n} 個の排反事象による Ω の分割から生成される σ 加法族に他ならない．従来の標準的ゲーム理論に従ってゲームの木と情報集合で考えるより，公理的確率論の演算ルールに沿って推論する方が見通しがよくかつ数学的に厳密な推論が可能であるために敢えて抽象化した表現を用いている．なお，集合の包含関係の意味で $\mathcal{F}^n\subset F^{n+1}\,;\,n=0,1,2,\ldots$ を満たしているがこのような性質を持つ部分 σ 加法族のことを工学の分野では増大情報系 (filtration) と呼んでいる (津野 [20, 定義 1.2.1]). Rosenberg, Solan and Vieille [19, p.434] では "Let $(\Omega,\mathcal{A},P)$ be a probability space, and $\{\mathcal{F}_n\}$ be a filtration over $(\Omega,\mathcal{A},P)$(the information available at stage n") と最初に設定しているが，本論文でもこのように定式化しても構わない．彼らはこの確率空間上で定義された確率過程を用いて戦略を定式化している．なお，本論文および彼らの Ω は Mascher, Solan and Zamir [13, p.440] や Mertens, Sorin and Zamir [14, p.123] のいう "the Universal Belief Space Ω とは全く関係がないから注意されたい.

[10] $P(*/\mathcal{F}^{n-1})$ は確率変数であることに注意されたい．なお，条件付き独立の定義については，たとえば Chung [1, p.284] を参照のこと.

繰り返し囚人のジレンマゲームの戦略セットの全体を $\mathcal{R}(G_\infty)$ とする．

注意 3.1. 条件 (1) は，従来のゲーム理論における説明において，たとえば，岡田 [17, p.232] では,「繰り返しゲームにおけるプレイヤーの戦略は，過去のプレイに依存して毎回の行動を決定する行動プロセスである」) あるいは Mailath and Samuelson [12, p.19] の "At the end of each period, all players observe the action profile chosen. In other words, the actions of every players are perfectly monitored by all other players." と文章表現された内容を数学的に定式化したものである．なお，工学の分野のことばで条件 (1) を表現すれば，確率過程 $\{Z^{(n)}\}$ は情報増大系 $\{\mathcal{F}^n\}$ に関して適合過程 (adapted process [11]) である，という仮定と同値である (津野 [20, 定義 1.2.2])．確率変数の観測可能性は，ある部分 σ 加法族に関するこの確率変数の可測性で定義される (同，まえがき iv 頁)．

条件 (2) は従来のゲーム理論では明確には意識されていないようで多分に直感的に理解されている気配がある．しかし，確率過程を用いて各プレイヤーの時刻 n における戦略としての分布（行動戦略）を求めようとしてみるとわかるが，この仮定なしには各プレイヤーが時刻 n において，それ以前の過去のすべての手を観察し考慮にいれながら時刻 n におけるワンショットゲームの戦略を相手とは独立に選ぶ，ということを数学的に定式化することは出来ない．実際に戦略セット $\vec{Z} = \{Z^{(n)} = (X_1^{(n)}, X_2^{(n)})\} \in \mathcal{R}(G_\infty)$ において $Z^{(2)}$ の分布 $P(X_1^{(2)} = i, X_2^{(2)} = j)\,;\, i, j \in S$ をベイズの公式と条件 (2) を用いて計算してみよう [12]．

$$
\begin{aligned}
&P(X_1^{(2)} = i_2, X_2^{(2)} = j_2, X_1^{(1)} = i_1, X_2^{(1)} = j_1) \\
&\quad = P(X_1^{(2)} = i_2, X_2^{(2)} = j_2 / X_1^{(1)} = i_1, X_2^{(1)} = j_1) \times P(X_1^{(1)} = i_1, X_2^{(1)} = j_1) \\
&\quad = P(X_1^{(2)} = i_2 / X_1^{(1)} = i_1, X_2^{(1)} = j_1) P(X_2^{(2)} = j_2 / X_1^{(1)} = i_1, X_2^{(1)} = j_1) \\
&\qquad \times P(X_1^{(1)} = i_1, X_2^{(1)} = j_1) \\
&\quad = P(X_1^{(2)} = i_2 / X_1^{(1)} = i_1, X_2^{(1)} = j_1) P(X_2^{(2)} = j_2 / X_1^{(1)} = i_1, X_2^{(1)} = j_1) \\
&\qquad \times P(X_1^{(1)} = i_1) P(X_2^{(1)} = j_1). \qquad (3.1)
\end{aligned}
$$

従って，$(X_1^{(2)}, X_2^{(2)})$ の分布は確率の加法性

$$
P(X_1^{(2)} = i_2, X_2^{(2)} = j_2) = \sum_{(i_1, j_1) \in S \times S} P(X_1^{(2)} = i_2, X_2^{(2)} = j_2, X_1^{(1)} = i_1, X_2^{(1)} = j_1)
$$

に上記の結果を代入することによって求められる．以下，同様にして $Z^{(n)} = (X_1^{(n)}, X_2^{(n)})\,;\, n = 3, 4, \ldots$ の分布はプレイヤー 1 の行動戦略（条件付き確率）$P(X_1^{(1)} = i_1), P(X_1^{(k)} = i_k / \mathcal{F}^{k-1})\,;\, k = 2, \ldots, n$ とプレイヤー 2 の行動戦略（条件付き確率）$P(X_2^{(1)} = j_1), P(X_2^{(k)} = j_k / \mathcal{F}^{k-1})\,;\, k = 2, \ldots, n$ から条件 (1) と条件 (2)（具体的な計算に必要な条件は (2) の方）と条件付き確率の公式 (ベイズの公式) に従って計算できる．ただし，後述する Trigger 戦略のように初めて D を選択した時刻の情報に基づいて容易にこれらの結合分布が計算できるような規則を持っている戦略もある．重要なことは確率変数を用いて厳密に定式化され，定義された戦略一般を対象にして，数学的に許されたルーティンな式の変形と推論に従うことによって初めて，得られた定理の数学的証明が可能になるということである [13]．

[11] 確率過程論でもよく知られた概念である．たとえば，Nevue [16, p.18].

[12] $P(B/A)$ は事象 A が与えられたときの事象 B の条件付き確率である．なお，$A = \emptyset$（空集合）に対しては定義しない．

[13] 特に定義 3.1 の条件 (2)（条件付独立性）の仮定が決定的に重要である．従来の標準的ゲーム理論の教科書

以上の考察から明らかなように，戦略セット $\vec{Z} = \{Z^{(n)} = (X_1^{(n)}, X_2^{(n)})\}$ に対する仮定からプレイヤー i は相手とは独立に自己の戦略（条件付き確率分布の無限列）$\{P(X_i^{(1)} = i_1), \ldots, P(X_i^{(n)} = i_n/\mathcal{F}^{n-1}), \ldots\}$ を自由に選択，決定することが出来る．なお，これらの条件付き確率分布は繰り返しゲームを展開形ゲームとみなした時の行動戦略に他ならない．このような意味で戦略セット $\vec{Z} = \{(X_1^{(n)}, X_2^{(n)})\} \in \mathcal{R}(G_\infty)$ において，プレイヤー i; $i = 1,2$ の戦略は $\vec{X}_i = (X_i^{(1)}, \ldots,, X_i^{(n)}, \ldots)$ で表現できるから，以後戦略セットを $\vec{Z} = (\vec{X}_1, \vec{X}_2)$ で表す．

以下によく知られた戦略の名称と定義を例示しておく．

(i) all C : $P(X_i^{(1)} = C) = 1, P(X_i^{(n)} = C/A) = 1\,;\, \forall A \in \mathcal{F}^{n-1}$, $n = 2, 3, \ldots$.

(ii) all D : $P(X_i^{(1)} = D) = 1, P(X_i^{(n)} = D/A) = 1\,;\, \forall A \in \mathcal{F}^{n-1}$, $n = 2, 3, \ldots$.

(iii) TFT : $P(X_i^{(1)} = C) = 1, P(X_i^{(n)} = C/\{X_j^{(n-1)} = C\} \cap A) = 1\,;\, \forall A \in \mathcal{F}^{n-1}$, $P(X_i^{(n)} = D/\{X_j^{(n-1)} = D\} \cap A) = 1\,;\, \forall A \in \mathcal{F}^{n-1}$, $n = 2, 3, \ldots, i \neq j$.

(iv) Trigger : $P(X_i^{(1)} = C) = 1$. $P(X_i^{(n)} = C/\{H_{jD} > n-1\} \cap A) = 1\,;\, \forall A \in \mathcal{F}^{n-1}$, $P(X_i^{(n)} = D/\{H_{jD} \leq n-1\} \cap A) = 1\,;\, \forall A \in \mathcal{F}^{n-1}$, $n = 2, 3, \ldots,\ i \neq j$. ここで，$H_{jD} \equiv \min\{k\,;\, X_j^{(k)} = D\}$, ただし，$\{*\} = \emptyset$（空集合）の場合は $H_{jD} = \infty$ と定める [14]

いよいよ本論文で考察する非対称繰り返し囚人のジレンマゲームを定式化する．非対称という意味は上記の Markov time H_{2D} を用いるからである．以下ではこのように定式化した繰り返し囚人のジレンマゲームを G_∞^D で表すこととする．すなわち，

定義 3.2. 繰り返し囚人のジレンマゲーム G_∞^D とは，戦略セット $\vec{Z} = (\vec{X}_1, \vec{X}_2) \in \mathcal{R}(G_\infty)$ に対するプレイヤー i; $i = 1,2$ の期待利得 $u_i(\vec{X}_1, \vec{X}_2)$ を次のように定義した繰り返しゲームのことである [15].

$$u_i(\vec{X}_1, \vec{X}_2) \equiv E\left[\sum_{n=1}^{H_D} \delta^{n-1} f_i(X_1^{(n)}, X_2^{(n)})\right]\,;\, i = 1, 2.$$

ここで，$H_D \equiv \min\{k\,;\, X_2^{(k)} = D\}$．ただし，$\{*\} = \emptyset$（空集合）の場合は $H_D = \infty$ と定める (Trigger 戦略を定義したところで導入した H_{2D} と同じものである)．ただし，δ $(0 < \delta < 1)$ は従来の標準的繰り返しゲームにおける定式化と同様，割引率ないし時間選好率として踏襲した [16].

この期待利得の定義は，プレイヤー 2 が D を選択したらその回でゲームを終了する，ということを意味している．社会学的にはプレイヤー 1 はプレイヤー 2 が裏切ったら即座にサンクション [17] としてゲームを中止する，ということである．

次に繰り返しゲーム G_∞^D のナッシュ均衡を定義する．

定義 3.3. 戦略セット $\vec{Z}^* = (\vec{X}_1^*, \vec{X}_2^*) \in \mathcal{R}(G_\infty)$ がナッシュ均衡であるとは次の 2 つの不等式が満たされるときをいう．

（と多くの論文）では分布を用いて定義しているために条件付独立性を結果的に暗黙の内に仮定している，ないし最初からそう理解されている．そのような曖昧さを避けるために本論文ではワンショットゲームの段階から確率変数を用いて定式化したのである．より詳しい議論は 4 節を参照されたい

[14] H_{jD} はプレイヤー j が初めて D を選択する時刻である．確率過程論でよく知られた D への first hitting time であり，Markov time（確率変数）となる．すなわち，任意の $n = 1, 2, \ldots$ に対して，$\{H_{jD} \leq n\} \in \mathcal{F}^n$ が成り立つ．

[15] 2 人の期待利得がこのように定義されていることが両者にとっての「共有知識」となっていることは，ゲーム G_∞^D の定式化において大前提である．

[16] この δ の解釈に関する先行研究については 4 節の (1) を参照されたい．

[17] そのために利得はすべて正値であることを仮定している．

(1) 任意の $(\vec{X}_1, \vec{X}_2^*) \in \mathcal{R}(G_\infty)$ に対して，$u_1(\vec{X}_1^*, \vec{X}_2^*) \geq u_1(\vec{X}_1, \vec{X}_2^*)$ が成り立つ，

(2) 任意の $(\vec{X}_1^*, \vec{X}_2) \in \mathcal{R}(G_\infty)$ に対して，$u_2(\vec{X}_1^*, \vec{X}_2^*) \geq u_2(\vec{X}_1^*, \vec{X}_2)$ が成り立つ．

以上の定式化の下で次の定理が得られる．

定理 3.1. $(c_2 - d_2)/c_2 \leq \delta < 1$ であるとき，$\vec{Z}^* = (\mathrm{all}\,D, \mathrm{all}\,C)$ は繰り返しゲーム G_∞^D のナッシュ均衡であり，かつ，すべての戦略セット $(\vec{X}_1, \vec{X}_2) \in \mathcal{R}(G_\infty)$ に対して $u_1(\mathrm{all}\,D, \mathrm{all}\,C) \geq u_1(\vec{X}_1, \vec{X}_2)$ である．

証明　$\mathrm{all}\,D = \vec{X}_1^*, \mathrm{all}\,C = \vec{X}_2^*$ とおく．$\vec{X}_2^*$ に対しては $P(H_D = \infty) = 1$ だから，

$$u_1(\vec{X}_1^*, \vec{X}_2^*) = \sum_{n=1}^{\infty} f_1(D, C)\delta^{n-1} = \frac{a_1}{1-\delta},\ u_2(\vec{X}_1^*, \vec{X}_2^*) = \sum_{n=1}^{\infty} f_2(D, C)\delta^{n-1} = \frac{d_2}{1-\delta}.$$

任意の戦略セット $\vec{Z} = (\vec{X}_1, \vec{X}_2^*)$ に対しては，

$$u_1(\vec{X}_1, \vec{X}_2^*) = \sum_{n=1}^{\infty} E[f_1(X_1^{(n)}, C)]\delta^{n-1} \leq \sum_{n=1}^{\infty} a_1\delta^{n-1} = \frac{a_1}{1-\delta} = u_1(\vec{X}_1^*, \vec{X}_2^*)$$

が得られるから定義 3.3 の (1) が成り立つ．また，任意の戦略セット $\vec{Z} = (\vec{X}_1, \vec{X}_2)$ に対して，$f_1(X_1^{(n)}, X_2^{(n)}) \leq a_1$ だから $u_1(\vec{X}_1^*, \vec{X}_2^*) \geq u_1(\vec{X}_1, \vec{X}_2)$ が成り立つことも明らかである．

次に任意の戦略セット $\vec{Z} = (\vec{X}_1^*, \vec{X}_2)$ を考える．H_D とゲーム G_∞^D の定義から，

$$\begin{aligned}
u_2(\vec{X}_1^*, \vec{X}_2) &= \frac{d_2}{1-\delta}P(H_D = \infty) + E\left[\sum_{n=1}^{H_D} \delta^{n-1} f_2(D, X_2^{(n)})\,;\, H_D < \infty\right] \\
&= \frac{d_2}{1-\delta}P(H_D = \infty) + E\left[c_2\delta^{H_D-1} + d_2\sum_{n=1}^{H_D-1}\delta^{n-1}\,;\, H_D < \infty\right] \\
&= \frac{d_2}{1-\delta} + \frac{c_2 - d_2 - c_2\delta}{1-\delta}E\left[\delta^{H_D-1}\,;\, H_D < \infty\right].
\end{aligned}$$

従って，

$$u_2(\vec{X}_1^*, \vec{X}_2^*) - u_2(\vec{X}_1^*, \vec{X}_2) = \frac{\delta - (c_2 - d_2)/c_2}{(1-\delta)/c_2}E[\delta^{H_D-1}\,;\, H_D < \infty] \geq 0$$

を得る．　□

定理 3.2. δ が十分 1 に近い時（その範囲は証明中に与える），戦略セット $\vec{Z}^* = (\vec{X}_1^*, \vec{X}_2^*)$（$\vec{X}_1^*$ =TFT or Trigger，$\vec{X}_2^*$ =TFT or Trigger）は繰り返しゲーム G_∞^D のナッシュ均衡である．

証明　$\vec{Z}^* = (\vec{X}_1^*, \vec{X}_2^*)$（$\vec{X}_1^*$ =TFT or Trigger，$\vec{X}_2^*$ =TFT or Trigger）のとき，$P(H_D = \infty) = 1$ だから，

$$u_1(\vec{X}_1^*, \vec{X}_2^*) = \sum_{n=1}^{\infty} f_1(C, C)\delta^{n-1} = \frac{b_1}{1-\delta},\ u_2(\vec{X}_1^*, \vec{X}_2^*) = \sum_{n=1}^{\infty} f_2(C, C)\delta^{n-1} = \frac{b_2}{1-\delta}.$$

$\vec{X}_1 \neq$ TFT or Trigger，$\vec{X}_2^*$ =TFT or Trigger とする．$H_{1D} \equiv \min\{k\,;\, X_1^{(k)} = D\}$ とおくと，$\vec{X}_2^*$ が TFT または Trigger 戦略であるから，$H_{1D} = \infty \Longrightarrow H_D = \infty$ であり，$H_{1D} < \infty$

ならば $H_D = H_{1D}+1$ が成り立つから，

$$\begin{aligned}
u_1(\vec{X}_1,\vec{X}_2^*) &= \frac{b_1}{1-\delta}P(H_{1D}=\infty) \\
&\quad + E\left[\sum_{n=1}^{H_{1D}-1}\delta^{n-1}f_1(C,C)+\delta^{H_{1D}-1}f_1(D,C)+\delta^{H_{1D}}f_1(X_1^{(H_{1D}+1)},D)\,;\,H_{1D}<\infty\right] \\
&= \frac{b_1}{1-\delta}P(H_{1D}=\infty) \\
&\quad + E\left[\frac{b_1(1-\delta^{H_{1D}-1})}{1-\delta}+\delta^{H_{1D}-1}a_1+\delta^{H_{1D}}f_1(X_1^{(H_{1D}+1)},D)\,;\,H_{1D}<\infty\right] \\
&\le \frac{b_1}{1-\delta}+\left(a_1+c_1\delta-\frac{b_1}{1-\delta}\right)E\left[\delta^{H_{1D}-1}\,;\,H_{1D}<\infty\right].
\end{aligned}$$

故に，

$$u_1(\vec{X}_1^*,\vec{X}_2^*)-u_1(\vec{X}_1,\vec{X}_2^*) \ge \frac{g_1(\delta)}{1-\delta}E\left[\delta^{H_{1D}-1}\,;\,H_{1D}<\infty\right];\; g_1(\delta)\equiv c_1\delta^2+(a_1-c_1)\delta-(a_1-b_1).$$

仮定2.1から，$g_1(0)<0$, $g_1(1)>0$ だから，$g_1(\delta)=0$ の正根を δ_1 とするとき，$0<\delta_1\le\delta<1$ に対して

$$u_1(\vec{X}_1^*,\vec{X}_2^*)-u_1(\vec{X}_1,\vec{X}_2^*)\ge 0$$

を得る．

つぎに，$\vec{X}_1^* =$ TFT or Trigger, $\vec{X}_2 \neq$ TFT or Trigger とすると，

$$\begin{aligned}
u_2(\vec{X}_1^*,\vec{X}_2) &= \frac{b_2}{1-\delta}P(H_D=\infty) \\
&\quad + E\left[\sum_{n=1}^{H_D-1}\delta^{n-1}f_2(C,C)+\delta^{H_D-1}f_2(C,D)\,;\,H_D<\infty\right] \\
&= \frac{b_2}{1-\delta}P(H_D=\infty) \\
&\quad + E\left[\frac{b_2(1-\delta^{H_D-1})}{1-\delta}+\delta^{H_D-1}a_2\,;\,H_D<\infty\right] \\
&= \frac{b_2}{1-\delta}+\frac{a_2-b_2-a_2\delta}{1-\delta}E\left[\delta^{H_D-1}\,;\,H_D<\infty\right].
\end{aligned}$$

故に，$0<\delta_2\equiv(a_2-b_2)/a_2\le\delta<1$ に対して，

$$u_2(\vec{X}_1^*,\vec{X}_2^*)-u_2(\vec{X}_1^*,\vec{X}_2)=\frac{\delta-(a_2-b_2)/a_2}{(1-\delta)/a_2}E\left[\delta^{H_D-1}\,;\,H_D<\infty\right]\ge 0$$

を得る．結局，$\max\{\delta_1,\delta_2\}\le\delta<1$ であれば，戦略セット $\vec{Z}^*=(\vec{X}_1^*,\vec{X}_2^*)$ はナッシュ均衡であることが証明された． □

定理3.1の意味するところは，たとえば，大企業が系列下請けに対して要求を呑まなければ取引を中止する，という「脅し」に対して大企業の要求を受け入れても取引を継続して貰わざるを得ない系列下請けの状態をゲーム理論を用いて定式化していると解釈することが可能である．その他，近年のブラック企業と従業員の関係等社会的に必ずしも好ましくない現

象を表現している．多くの社会学者が密かに期待している好ましい社会選択 (TFT,TFT) あるいは (Trigger,Trigger) は δ が十分 1 に近い時，ナッシュ均衡ではあるが，$u_1(\mathrm{all}\,D, \mathrm{all}\,C) = a_1/(1-\delta) > u_1(\mathrm{TFT}, \mathrm{TFT}) = b_1/(1-\delta)$ であるからプレイヤー 1 にとって all D は ベストな選択なのである．さらにいえば，ナッシュ均衡である (TFT,TFT) の状態にあってもプレイヤー 1 が戦略を all D に変更した場合，それに対するプレイヤー 2 の最適応答は all C であり，これに対するプレイヤー 1 の最適応答はもちろん all D であるから結局ナッシュ均衡 (all D, all C) に落ち着いてしまう，という意味でもナッシュ均衡 (all D, all C) は極めて安定な均衡であることがわかる．

4. ディスカッション

なお，最後に本論文における繰り返しゲームの定式化と従来の標準的ゲーム理論におけるそれとの関係について若干コメントをしておく．

(1) 岡田 [17] では δ について割引率と呼んで説明なしにアドホックに導入しているが，たとえば，Myerson [15, p.308] では ($\delta = 0.99$ として)

> For example, suppose that the number of times that the game will be played is a random variable[18], unknown to the players until the game stops, and that this random stopping time is a geometric distribution with expected value 100.

と説明している．しかし，以降の説明に “random stopping time” という概念を一般化はしていない[19]．なお，random stopping time であると理解した場合，δ は各回でゲームが続く確率であり，丁度 n 回でゲームが終了する確率は $(1-\delta)\delta^{n-1}$ である．さらに，$\sum_{n=1}^{\infty}(1-\delta)\delta^{n-1} = 1$ であるから，公理的確率論の帰結としてゲームは確率 1 で有限回で終了する．しかし，同じ頁には “a repeated game has an infinite time horizon, unlike the finite game trees, . . . ” と説明してあって，あたかもゲームが本当に無限回行われるかのごとく説明されているが，このような理解は正しくない[20]．いずれにしろ，無限，任意有限，(確定的に) 有限の 3 通りの概念的違いを峻別しなければならない．確率変数を用いて説明すると分かり易い．いま random stopping time T を ∞ または自然数に値を取る確率変数とすると，$P(T = \infty) > 0$ となる場合は正の確率でゲームが無限に続き，$P(T < \infty) = 1$ であるが，$\forall n$(自然数)，$P(T > n) > 0$(つまり分布の台が無限集合) の場合，ゲームは任意に長く続く可能性はあるが必ず有限時間で終了するのに対して，$\exists n$(自然数)，$P(T \leq n) = 1$ ならば，ゲームは確定時刻 n までに必ず終了するということである[21]．なお，たとえば random stopping time の分布が幾何分布の場合，ラスムセン [18, pp.198–199] は，ゲームがどの時刻で終了するかが不確実である状況

[18]正しくは，さらに全ての混合戦略とは独立な確率変数であることを仮定しなければならない．詳しくは河野 [4, pp.50–51] を参照されたい．

[19]河野 [4, 定理 8.1] では自然数に値を取る独立な確率変数としての “stopping time” を導入し，幾何分布以外の分布に対しても一定の条件の下で Trigger 戦略がナッシュ均衡になることを証明した．本論文においてもそのような定式化は可能であるが，幾何分布であることを仮定する限り，得られる結果に違いは生じないので深入りしなかった．

[20]繰り返しゲームにおいて δ を割引率と呼んでアドホックに導入している従来の標準的ゲーム理論の教科書は，大抵真に無限回ゲームが繰り返されるかのごとき説明がなされている．Mailath and Samuelson [12, p.106] には “the game has a finite horizon with probability 1. However, once the players have reached any peirod t, they assign probability δ to play continuing, allowing intertemporal incentives to be maintained despite the certainty of a finite termination.” と正しく指摘してある．

[21]確定時刻 T まで繰り返すゲームを Maschler, Solan and Zamir [13, p.524] は T-stage game と呼んでいる．彼らは (p.524 13.3) “Our goal is to characterize the limit set of equilibrium payoffs as T goes to infinity.” と述べているが，確定している T までの繰り返しゲームの構造を $T \to \infty$ とした繰り返しゲームと T が確率変数で幾何分布を持つ stopping time として定式化した繰り返しゲームの構造は数学的には異なる．

と確かにある時刻までに終了する可能性が高くとも，現実にその時刻まで続けば，ゲームは出発時点と同じ状況に置かれてしまうことは互いに異なる，と説明しているが，可算加法性を認める公理的確率論の下ではこの二つの状況に関する認識は完全に一致するのであって矛盾はしない．何故ならば，最後に終了したラウンドにおいて，ゲームが続く正の確率 $\delta > 0$ を持っているではないか，という疑問を持つかもしれないが，実はゲームが終了した時刻はランダムではあるが脚注 14 で説明した Markov time ではない．従って，公理的確率論の立場からは確率を論じてはならない（論じることが出来ない）のである．

(2) 本論文で考察したゲーム G_∞^D の期待利得は定義 3.2 であたえたが，本論文は基本的に不確実性（ランダムネス）を確率変数で表現してきた原則に従えば，(1) で説明してきたように割引率 δ も random stopping time T を用いて表現すべきである．その場合，定義 3.2 は次のように書き直せる．

定義 4.1. 繰り返し囚人のジレンマゲーム G_∞^D とは，戦略セット $\vec{Z} = (\vec{X}_1, \vec{X}_2) \in \mathcal{R}(G_\infty)$ に対するプレイヤー i ; $i = 1, 2$ の期待利得 $u_i(\vec{X}_1, \vec{X}_2)$ を次のように定義した繰り返しゲームのことである．

$$u_i(\vec{X}_1, \vec{X}_2) \equiv E\left[\sum_{n=1}^{T\wedge H_D} f_i(X_1^{(n)}, X_2^{(n)})\right];\ i = 1, 2.\ \text{（ここで，} a \wedge b = \min\{a, b\}\text{）}$$

ただし，T は $\mathcal{F}^\infty$ とは独立で幾何分布を持つ確率変数であるとする．

幸い，定義 3.2 と定義 4.1 は同値である[22]．本論文では取りあえず繰り返し囚人のジレンマゲームにおいて，従来知られていない非対称なサンクション，プレイヤー 2 が D を選択すれば直ちにゲームが終了する，を仮定することによって $(\text{all}\, D, \text{all}\, C)$ の組がナッシュ均衡になる，という定理を確率過程を用いて厳密に証明することを目標としたために random stopping time は導入せずにあえて割引率とした．ただ，将来の可能性として定義 4.1 の T に対する幾何分布を持つという仮定や独立性の仮定を緩めることによって繰り返しゲームをより厳密に定式化した上でさらに一般化して考察できる可能性があることを指摘しておきたい．

(3) 繰り返しゲームの戦略についての本論文の定義 3.1 に関する若干のコメント：

繰り返しゲームの定式化は 1990 年頃には一通り完成して現在に至っているように思われるが繰り返しゲームの行動戦略一般についての定義や説明が多くの教科書において不十分かつ不正確であるように思われる．たとえば，Fudenberg and Tirole [2, Section 5, Repeated Games], Myerson [15, Section 7, Repeated Games], Mailath and Samuelson [12, Section 2, The Basic Structure of Repeated Games with Perfect Monitoring], Kandori [3, p.98, repeated games], 岡田 [17, Section 6 繰り返しゲーム], M. Maschler, E. Solan and S. Zamir [13, Section 13, Repeated games], Mertens, Sorin and Zamir [14, Chapter IV, General Model of Repeated Games] 等と本論文における定義 3.1 を比較検討されたい．

これらの教科書では本来不確実な状況の時間的経過を確率過程を用いて定式化していないために戦略とは何かが必ずしも明確に定式化されていないのである．もう少し具体的に述べ

[22] ついでに指摘しておくと，従来の割引率 δ を用いた通常の繰り返し囚人のジレンマゲームの期待利得は $u_i(\vec{X}_1, \vec{X}_2) \equiv E[\sum_{n=1}^{T} f_i(X_1^{(n)}, X_2^{(n)})]$; $i = 1, 2.$ と表現することが出来て（ただし，T は $\mathcal{F}^\infty$ とは独立で幾何分布を持つ確率変数)，この式は $u_i(\vec{X}_1, \vec{X}_2) = E[\sum_{n=1}^{\infty} \delta^{n-1} f_i(X_1^{(n)}, X_2^{(n)})]$; $i = 1, 2.$ に等しい．normalization factor, average payoff あるいは weighted average of the daily payoff と称して右辺に $(1-\delta)$ を乗じている教科書（たとえば，Fudenberg and Tirole [2, p.147]，Mailath and Samuelson [12, p.21], Maschler, Solan and Zamir [13, p.543]) もあるが，いずれが「自然な」定義であるか自ずと明らかであろう．

ると，Fudenberg and Tirole [2, p.147] にはa player's strategy can depend only on the past values と述べており，これは本論文の定義3.1の条件(1)に対応している．しかし，数学的に最も重要な仮定である条件(2)については，"Note also that each period of play begins a proper subgame. Moreover, since moves are simultaneous in the stage game, these are the only proper subgames, ..." と説明してあるだけで，条件付き独立性を意味しているとは思われない．Myerson [15, pp.310–311] には "The interpretation of such a general repeated game is as follows." に続いて半頁ほど繰り返しゲームの構造を説明している部分が本論文の定義3.1に相当すると思われるが分布（行動戦略）の言葉で表現してあるために，本論文の定義3.1の(2)(条件付き独立性)に相当する部分は暗黙の内に仮定されていると理解せざるを得ない．Kandori [3, p.100] の場合は, "A repeated game strategy for player i is a complete contingent action plan, which specifies a current action after any history" となっており，文章表現ではあるが，本論文の定義3.1の条件(1)しか表現していない．岡田 [17, p.239] には純戦略は定義してあるが混合戦略については明示的には定義されていない．同工異曲な定義は最近の教科書に至るまで繰り返されている[23].

Maschler, Solan and Zamir [13, p.527] には全プレイヤーの "strategy vector τ" から定まる結合分布 $\mathbf{P}_\tau$ を次式

$$\mathbf{P}_\tau(s^1, s^2, \ldots, s^t) = \mathbf{P}_\tau(s^1, s^2, \ldots, s^{t-1}) \times \tau_1(s_1^t | s^1, s^2, \ldots, s^{t-1}) \\ \times \tau_2(s_2^t | s^1, s^2, \ldots, s^{t-1}) \times \cdots \times \tau_n(s_n^t | s^1, s^2, \ldots, s^{t-1}) \tag{4.1}$$

を用いて帰納的に求め，この公式について次のように説明している．

> This formula for $\mathbf{P}_\tau$ expresses the fact that the mixed action that a player implements in any given stage can depend on the actions that he or other players played in previous stages, but the random choices of the players made simultaneously in each stage are independent of each other.

問題なのは，上の(4.1)のところで本論文の定義3.1の条件(2)の条件付き独立性を暗黙の内に用いていることである[24]．対応する説明として "the random choices of the players made simultaneously in each stage are independent of each other" とあるが，彼らの戦略に対する定義式(13.3)（Definition 13.3 p.525）"A behavior strategy for player i in a T-stage game is a function associating a mixed action with each history of length less than T" には条件付き独立性の仮定は含まれていない（暗黙の仮定としか理解できない）．本論文のように確率過程として定式化してみると容易に理解できるが，"strategy vector τ" の定義だけでは確率過程の確率構造を一意に定義したことにはならない．つまり，"the random choices of the players made simultaneously in each stage are independent of each other"（各stage gameの行動戦略の説明）のところで「あらかじめ，明示的に」かつ数学的に仮定しておかなければならない．そのキー概念が条件付き独立性なのである．さもなければ，後に得られた定理を数学的に厳密に証明することはできない．

要するに，繰り返しゲームは文字通り同じ標準形（同時手番）ゲームを繰り返すことを想定しているわけであるから，特に2回目以降のゲームが戦略自身は過去の履歴に依存してもよいが，ワンショットゲームとしては同時手番ゲームとゲーム構造が同一でなければならな

[23]たとえば，Mailath and Samuelson [12, pp.19–21], Maschler, Solan and Zamir [13, pp.525–527].
[24]本論文で例示した(3.1)を参照されたい．

い．そのことの数学的表現が条件付き独立という概念である．この概念を正確に表現するためには戦略自体を確率過程で表現する必要があったのである．さらにいえば，プレイヤー 2 がランダムに初めて D を選択する時刻 H_D が確率変数として明確に定義できるためにはプレイヤー自身の戦略自体を確率過程で定式化しておくことが不可欠である．このような定式化によって初めて確率過程論でよく知られた計算テクニックをルーティンに用ることによって，得られた定理を数学的に厳密に証明することが出来たのである．定理そのものは従来の標準的ゲーム理論の枠内での説明からでも予想すること自体は容易である．

謝辞

投稿前の論文原稿について，種々のコメントと有益なるアドバイス，文献をご教示して下さった村上範明氏に深く感謝します.

参考文献

[1] K.L. Chung: *A Course in Probability Theory 3rd ed.* (Academic Press, Sandiego, 2001).

[2] D. Fudenberg and J. Tirole: *Game Theory* (The MIT Press, Cambridge, 1991).

[3] M. Kandori: Repeated games. In S.N. Durlauf and L.E. Blume(eds.): *The New Palgrave Dictionary of Economics, Second Edition* Volume 7 (Macmillan, 2008).

[4] 河野敬雄: ゲーム理論アラカルト: 確率論の立場から. *Rokko Lectures in Mathematics*, **13** (2003), 1–89.
http://www.math.kobe-u.ac.jp/HOME/home-j/rokko.html.

[5] N. Kôno: Nash equilibria of randomly stopped repeated prisoner's dilemma. In T. Tokimasa et al.(eds.): *Applied Economic Informatics and Systems Sciences*, **31** (2005), 81–93.
http://kup.or.jp/en/booklist/ss/872.html.

[6] N. Kôno: Noncooperative game in cooperation: Reformulation of correlated equilibria. *The Kyoto Economic Review*, **77-2** (2008), 107–125.

[7] N. Kôno: Noncooperative game in cooperation: Reformulation of correlated equilibria (II). *The Kyoto Economic Review*, **78-1** (2009), 1–18.

[8] 河野敬雄: ゲーム理論アラカルト (続): 確率論の立場から. *Rokko Lectures in Mathematics*, **21** (2011), 1–124.
http://www.math.kobe-u.ac.jp/HOME/home-j/rokko.html.

[9] 河野敬雄: 展開形ゲームの新しい定義とその帰結. 理論と方法, **31-1** (2016), 138–149.
https://www.jstage.jst.go.jp/article/ojjams/31/1/31_138/_pdf.

[10] 河野敬雄: ゲーム理論の数学的基礎: 批判的再検討. *Rokko Lectures in Mathematics*, **25** (2017), 1–94.
http://www.math.kobe-u.ac.jp/HOME/home-j/rokko.html.

[11] R.D. Luce and H. Raiffa: *Games and Decisions -Introduction and Critical Survey-* (John Wiley & Sons, New York, 1957).

[12] G.J. Mailath and L. Samuelson: *Repeated Games and Reputations: Long-Run relationships* (Oxford University Press, Oxford, 2006).

[13] M. Maschler, E. Solan, and S. Zamir. *Game Theory* (Cambridge University Press, New York, 2013).

[14] J-F. Mertens, S. Sorin, and S. Zamir: *Repeated Games* (Cambridge University Press, New York, 2015).

[15] R.B. Myerson: *Game Theory: Analysis of Conflict* (Harvard University Press, Cambridge, 1991).

[16] J. Neveu: *Martingales à Temps Discret* (Masson et Cie, Paris, 1972) Translated by T.P. Speed, *Discrete-Parameter Martingales* (Norh-Holland Publishing Company, 1975).

[17] 岡田 章: ゲーム理論 (新版) (有斐閣, 2011).

[18] E. Rasmusen: *Games and Information: An Introduction to Game Theory* (Blackwell, Malden, 2007) 細江守紀他 訳: ゲームと情報の経済分析 [基礎編](九州大学出版会, 2010).

[19] D. Rosenberg, E. Solan, and N. Vieille: Stopping games with randomized strategies. *Probability Theory and Related Fields*, **119** (2001), 433–451.

[20] 津野義道: Kalman-Bucy のフィルター理論 (共立出版, 2006).

河野敬雄
〒520-0016 大津市比叡平 1-18-20
E-mail: `kono.norio.58x@st.kyoto-u.ac.jp`

ABSTRACT

ON A NON-SYMMETRIC REPEATED PRISONER'S DILEMMA GAME IN WHICH (all D, all C) STRATEGY IS THE BEST NASH EQUILIBRIUM FOR ONE PLAYER

Norio Kôno

In the well-known repeated prisoner's dilemma game, how to achieve cooperative behaviors among both players has been the main concern, like the fork theorem.

In contrast, in this paper we formulate, by using the theory of standard discrete time stochastic processes, a non-symmetric repeated prisoner's dilemma game in which player 1 can stop the game immediately when player 2 has chosen "D" strategy. In consequence, we can rigorously prove that player 1 and player 2's strategy set $(\mathrm{all}\ D, \mathrm{all}\ C)$ is the best Nash equilibrium for player 1 among any other strategy sets.