知能と複雑系　127－25
(2002.　1.　10)

# $n$ 人繰り返し囚人のジレンマゲーム戦略の GA による進化

## 技術研究報告形式

池田　隆文[†]　　伊庭　斉志[†]

† 東京大学大学院 新領域創成科学研究科 基盤情報学専攻
〒 113-0033 東京都文京区本郷 7-3-1

E-mail: †{ikeda,iba}@miv.t.u-tokyo.ac.jp

**あらまし**　囚人のジレンマ (Prisoner's Dilemma, PD) は経済学、数学、ゲーム理論、政治学、社会心理学、人工知能など、さまざまな分野で広く注目を集め、研究されてきた問題である。囚人のジレンマにはさまざまな拡張が存在する。その中に $n$ 人による繰り返し囚人のジレンマ ($n$-person Iterated Prisoner's Dilemma, $n$-IPD) がある。$n$-IPD は一般性と現実世界の問題への応用性が高く、エネルギーや資源の保護の問題、インフレと賃上げの自粛の問題、環境汚染問題、人口増加問題、軍縮問題、などその適用範囲は広い。本研究では $n$-IPD を行うエージェント集団を遺伝的アルゴリズムを用いて進化させ、その時のエージェント集団の振る舞いを観察する。そしてゲームのプレイヤーを増やしたことによって生じる効果を明らかにし、$n$-IPD の特質について考察する。

**キーワード**　$n$ 人繰り返し囚人のジレンマ、協調行動の創発、遺伝的アルゴリズム

# The Strategy Evolution in N-person Iterated Prisoner's Dilemma Games

## The format of Technical Report of IEICE

Takafumi IKEDA[†] and Hitoshi IBA[†]

† Department of Frontier Information, Graduate School of Frontier Sciences, University of Tokyo
Hongo 7–3–1, Bunkyo-ku, Tokyo, 113–0033 Japan

E-mail: †{ikeda,iba}@miv.t.u-tokyo.ac.jp

**Abstract**　Althought the normal 2 person Iterated Prisoner's Dilemma has widely been studied for explaination of the cooperative behaviour evolution in social and biological systems, it began to be recognized this model has some limitation. Compared with the Prisoner's Dilemma played by 2 persons, the n-person Iterated Prisoner's Dilemma (n-IPD) has greater generality and applicability to real-life situations. In addition to the problems of energy conservation, ecology, and overpopulation, many other real-life problems can be represented by the n-IPD paradigm. The n-IPD can model those problems which cannot be handled by the 2-IPD. This paper shows the population dynamics of the 3 and 4-IPD game, in which different strategies are evolved by Genetic Algorithm. We discuss the emergent properties of n-IPD by example.

**Key words**　$n$-IPD, cooperative behaviour, Genetic Algorithm

## 1. はじめに

囚人のジレンマ (Prisoner's Dilemma, PD) は 1950 年 Merrill Flood と Melvin Dresher が考案し、Albert Tucker に定式化されて以来、経済学、数学、ゲーム理論、政治学、社会心理学、人工知能など、さまざまな分野で広く注目を集め、研究されてきた問題である [4]。囚人のジレンマにはさまざまな拡張が存在する。その中に $n$ 人による繰り返し囚人のジレンマ ($n$-person Iterated Prisoner's Dilemma, $n$-IPD) がある。$n$-IPD は一般性と現実世界の問題への応用性が高く、エネルギーや資源の保護の問題、インフレと賃上げの自粛の問題、環境汚染問題、人口増加問題、軍縮問題、などその適用範囲は広い。本研究では $n$-IPD を行うエージェント集団を遺伝的アルゴリズムを用いて進化させ、その時のエージェント集団の振る舞いを観察する。そしてゲームのプレイヤーを増やしたことによって生じる効果を明らかにし、$n$-IPD の特質について考察する。

本稿ではまず第 2.章で囚人のジレンマ (PD)、第 3.章で繰り返し囚人のジレンマ (IPD)、第 4.章で $n$ 人繰り返し囚人のジレンマ ($n$-IPD) について説明する。そして、第 5.章で 3-IPD を行うエージェントの戦略を遺伝的アルゴリズムを用いて進化させた実験の模様を紹介し、第 7.章でまとめを行う。

## 2. 囚人のジレンマ

二人の囚人が別々の独房に入れられている。彼らは一度は信頼しあってチームを組み、銀行強盗を働いたのだが、運悪く捕まってしまった。二人はそれぞれ独房に収監されているので連絡しあう手段はない。彼らには司法取引による減刑のチャンスが与えられている。取調べで 2 人とも相棒に協調 (Cooperate, C) して黙秘すれば 2 人ともそこそこの減刑を得られる (減刑 3 年)。しかし自分が黙秘していても、相棒が裏切って (Defect, D) 自白すると自分の刑は極端に重くなり (減刑 0 年)、相棒は無罪放免 (減刑 5 年) となる。また両方が自白すればそれぞれ刑 (減刑 1 年) に処せられる。

囚人のジレンマゲームの得点 (減刑年数) 表の例を図 1に示す。図 1はプレイヤー 1 から見た得点 (減刑年数) を表している。プレイヤー 2 が得る得点はこれと対称的である。表中の 2 つのラベル「プレイヤー 1」、「プレイヤー 2」を入れ替えれば、プレイヤー 2 から見た得点の表になる。

2 人の囚人は黙秘するか自白するかでジレンマに陥ることになる。囚人各々にとって見れば、相棒が

| | | プレイヤー 2 | |
|---|---|---|---|
| | | $D$ | $C$ |
| プレイヤー 1 | $C$ | 0 | 3 |
| | $D$ | 1 | 5 |

図 1 プレイヤー 1 から見た囚人のジレンマの得点

協調して黙秘するにせよ、裏切って自白するにせよ、自分は裏切って自白した方が協調して黙秘した場合よりも得られる減刑は大きい。しかし両方ともそう考えてお互いを裏切り自白してしまった場合、両方とも協調して黙秘した場合よりも刑が重くなってしまう。相棒は裏切らないかもしれないが、その保証はない。相棒を信用して相棒に協調し黙秘したとき、相棒が裏切って自白すれば最悪の結果となる。その最悪の結果を避けるためには、自分も裏切って自白するしかない。もちろん相手も同じように考えて裏切るだろう。こうして両者は裏切りあって自白しあい、減刑 1 年を得ることになる。両者とも協調して黙秘を続け、減刑 3 年を勝ち取った方が両者にとって望ましいにもかかわらず。これが囚人のジレンマである。

個人として最良と考えられた、裏切って自白するという行動が、プレイヤー 1、プレイヤー 2 の集団としては最悪の結果を導くのである。逆に個人としては最悪の、協調して黙秘を続けるという行動に出ると、集団としては最良の結果になる。集団は個人の集合であり、個人を統合してできたものである。しかしながら、集団の合理性は個人の合理性を統合したものではなく、むしろ、それと真っ向から対立している [9]。また、両者がたとえ協調したとしても、その協力関係は均衡状態にない。この協力関係はそこから脱するものが利得する構造になっているからである。両者が協調しあっている状況では、自らの選択を一方的に変更したプレイヤーは得をする。つまり両者には協力関係から脱する動機がある。これに反して、個人の合理性に対応する両者が裏切りあうという状況は均衡状態にある。両者が裏切りあっている状況では、自らの選択を一方的に変更したプレイヤーは損をする。そのため両者にはもはやそこから離れる動機がない。

## 3. 繰り返し囚人のジレンマ

囚人のジレンマの対戦が1回きりではなく、何回も行われる可能性が高いとしたら、各人のとるべき行動は微妙な問題になるだろう [8] [11]。次の対戦での相手の報復を怖れたり、将来の相手の協力を引き出すため協調的な行動に出ることが最善であるかもしれない。この場合、合理的で利己的なプレイヤーにとって適切な戦略は、もはや1回限りの囚人のジレンマのように自明ではなくなる。1回きりの対戦ではなく不定回の対戦を繰り返す場合、繰り返し囚人のジレンマ (Iterated Prisoner's Dilemma, IPD) という。

IPD では、囚人のジレンマゲームを同じ相手と何回か繰り返して行い、得点の総和あるいは平均点をその囚人の得点とする。以下では1回ごとの対戦における囚人の選択を"手"と呼ぶことにする。具体的には、手は協調する（黙秘する）C か、裏切る（自白する）D かのいずれかである。この場合に考えられる戦略としては、例えば、

(1) まったくでたらめにその回の手を決める
(2) 相手の手にかかわらず常に協調する（黙秘する）
(3) 相手の手にかかわらず常に裏切る（自白する）
(4) 協調と裏切りを繰り返す

などが考えられる。

IPD では1度きりの PD ゲームでは見られなかった解が生じることが知られている。有名なのはしっぺ返し戦略 (Tit for Tat, TFT) と呼ばれるものである。この戦略は、

- 1回目はランダムに手を出す
- 2回目以降は相手の1回前の手をまねる

という戦略である。この戦略は"目には目を、歯には歯を"を体現する。つまり、この戦略は、前回相手が裏切って自白した場合には自分も今回は裏切って自白し、逆に前回相手が協調して黙秘した場合には自分も今回は協調して黙秘する、というものである。

IPD に世間の関心が集中したのは国際政治学者 Robert Axelrod の一連の研究が 80 年代の初期に発表されてからである。 Axelrod はゲーム理論家たちを招いてコンピュータ・プログラムによる IPD ゲームコンテストを2度開催した。そして全てのプログラムを他のプログラム（自分自身と、ランダムに協調か裏切りかを選択するプログラムも含む）と数百回対戦させた。このようなゲームでは、どの相手と対戦しても最高の得点を稼ぐという意味で"最強"の戦略が存在しないことはすぐに分かる。もし相手戦略が次の手を選択する時に今までの歴史をまったく考慮しないような戦略だったら、無条件に裏切る戦略がベストである。もし相手戦略が最初は協調するが、一度でも裏切ったらその後は"報復"としてずっと裏切りつづけるような戦略だったら、無条件に強調する戦略が有効だろう。にもかかわらず、どんな相手ともうまくやっていけるような特徴をもった戦略があった。 Axelrod の最初の大会で優勝した戦略は TFT であった。そしてより重要な事実として挙げるべきは、第2回の大会では 63 人の参加者を数え、その全員に前回の大会の結果を知らせていたにもかかわらず、再び TFT が優勝してしまったことであろう [10]。Axelrod の IPD コンテスト以後、TFT のような上品で (自分からは裏切らない)、すぐに反応し、(相手の裏切りに即座に反応し)、寛容な (相手にしつこく報復しない) 戦略が IPD ゲームでは強いと考えられている。

## 4. $n$-IPD への拡張

### 4.1 $n$-IPD とは?

2-IPD は数十年にわたって広範囲に研究されてきたが、実世界の問題、特に社会学や経済学の問題には 2-IPD ではモデル化できないものが数多くある [1] [3]。$n$ 人繰り返し囚人のジレンマ ($n$-person Iterated Prisoner's Dilemma, $n$-IPD) はそのような問題をモデル化できる、より現実的で一般的なゲームである。簡単な例をあげよう。

1976 年夏、イギリスは深刻な干ばつに見舞われた。ほとんどの地方で、水は貴重な資源になった。マスコミは全国民に水の節約を強く勧めた。この時、各市民に突きつけられたジレンマは次のようにまとめられる。各個人が節水により利益を得られるのは、その他多くの人々が同じように節水を実行した場合である。しかし、個々人にとって見れば、明らかに節水の必要はない。この人が節水に協力しなかったとしても水は使えるし、罰を受けることもない。一方、もしほとんどの人が節水の呼びかけを無視したら、節水をした人の努力は無駄になってしまう。個人の立場で見れば、節水はどうころんでも無益である。明らかに個人の利己心の中には他の人がどうしようと節水の呼びかけを無視したいという気持ちがある。しかし、もしみんなが個人的な合理性を追求し水を節約しなかったら、みんなが集団の合理性にに基づいて節水を実行した場合に比べ事態は悪くな

る。1976 年夏の間、イギリス国民のうち節水の呼びかけを応じたものは少数であった。約 40%の国民が節水しただけで、50%の国民はまったく気にもかけなかったという。

上記の例は限りある資源の保護に関する問題であった。$n$-IPD はこのような問題だけでなく、環境汚染問題、人口増加問題、軍縮問題、インフレーションと労働組合の賃上げ自粛問題など広範囲な現実世界の問題をモデル化できる。

$n$-IPD の最も重要な特質は、全員が個人の利益を追求すると、全員が個人の利益を追求しなかった場合よりも自体が悪くなるが、誰にも自分から率先して私利追求をやめる動機がない、というものである。

### 4.2 $n$-IPD の構成

$n$-IPD をもう少しきちんと定義してみる。

$n$-IPD ゲームでは、毎回各プレイヤーは協調する (Cooperate, C) か、裏切る (Defect, D) かの選択を迫られる。$n$-IPD ゲームの得点表を図 2 に示す。図 2 はプレイヤー 1 が得る得点を表している。プレイヤー 1 の得る得点は自分の出した手と、残りの $n-1$ 人のプレイヤーの中で協調的な手 (C) を出したプレイヤーの数で決まる。他のプレイヤーが得る得点はこれと対称的である。表中のラベル「プレイヤー 1」を「プレイヤー $k$」に替えれば、プレイヤー $k$ が得る得点の表になる。

**残りのプレイヤーの中で C を出したプレイヤーの数**

| | | 0 | 1 | 2 | | $n-1$ |
|---|---|---|---|---|---|---|
| **プレイヤー 1** | $C$ | $C_0$ | $C_1$ | $C_2$ | ... | $C_{n-1}$ |
| | $D$ | $D_0$ | $D_1$ | $D_2$ | ... | $D_{n-1}$ |

図 2 $n$ 人囚人のジレンマの得点表

この得点表は、次のような性質を満たすべきであると考えられている [6]。

1. $D$ 選択のの優位性

$$D_i > C_i \qquad i = 1, 2, \ldots, n-1$$

この条件は、他のプレイヤーがどのような行動をとるにせよ、自分は常に裏切った方が得られる得点が大きくなることを表している。この条件のため、個人の利益を優先するプレイヤーは裏切りを選択することになる。

2. 単調性

$$C_i > C_{i-1}$$
$$D_i > D_{i-1} \qquad i = 1, 2, \ldots, n-1$$

この条件は、他のプレイヤーがより多く協調するほど得られる得点が大きくなることを表している。そしてそれは自分の手が協調であろうと、裏切りであろうと同じである。

3. 協調行動の効率性

$$(i+1)\, C_i + (n-i-1)\, D_{i+1} > i\, C_{i-1} + (n-i)\, D_i$$
$$C_{n-1} > D_0 \qquad i = 1, 2, \ldots, n-1$$

第 1 の不等式は、集団全体の得点の総和は、協調するプレイヤーが増える (あるいは、裏切るプレイヤーが減る) につれて単調に増加することを表している。また第 2 の不等式は、全てのプレイヤーが協調した場合の得点 $C_{n-1}$ は、すべてのプレイヤーが裏切った場合の得点 $D_0$ よりも大きいことを表している。

条件 1 のため、プレイヤーは裏切りの誘惑にかられることになる。しかし、その誘惑に負けて裏切るプレイヤーが増加すると条件 2 のため自分の得点も減少する。事態の悪化を防ぐためにはまず自分から協調行動をとるべきだが、それは抜け駆けするプレイヤーに利益を与えることにもなる。また条件 3 は、協調するプレイヤーが増加することのみが集団の利益につながることを保証するものである。もしこの条件が満たされない $i$ が存在するとしたら、協調したプレイヤーの数を $i$ 人に保つように $n$ 人のプレイヤーが交互に協調と裏切りを繰り返し利益を譲渡しあう、といった解が生じてしまう可能性があるからである。

以上の 3 つの条件を満たす得点表はいくらでも考えられる。例えば、図 3の得点表がそうである ([7] より)。

## 5. 実　　験

$n$-IPD ゲームを行うエージェントの戦略を遺伝的アルゴリズム (Genetic Algorithm, GA) で進化させる実験を行った。以下ではその実験のやり方や結果について述べる。

### 5.1 遺伝子の構成

$n$-IPD を行うエージェントの戦略を GA で進化させるためには、まずエージェントの戦略をなんらか

の計算可能なモデルで表し、それを遺伝子とみなして遺伝子オペレータで遺伝子を操作できなくてはならない。まず、エージェントの遺伝子 (戦略) の構成の仕方を考える。

**残りのプレイヤーの中でCを出したプレイヤーの数**

| | | 0 | 1 | 2 | | $n-1$ |
|---|---|---|---|---|---|---|
| プレイヤー1 | $C$ | 0 | 2 | 4 | ... | $2(n-1)$ |
| | $D$ | 1 | 3 | 5 | ... | $2(n-1)+1$ |

図3　$n$ 人囚人のジレンマの得点表の例

各プレイヤーは、$m$ 回過去の対戦の歴史を参照して自分が次に出す手を決定する。対戦の歴史は、自分の手と協調行動をとった自分以外のプレイヤーの数とで表現される。対戦の歴史はバイナリ列で表される。例えば、次のバイナリ列は4-IPDゲーム2回分の歴史を表現している。

... 1 1 0 0 0 0

左側の3つのビット 1 1 0は前々回の対戦の歴史を表し、右側の3つのビット 0 0 0は前回の対戦の歴史を表している。左側の3つのビット 1 1 0の中で一番左のビット 1は、前々回自分は1 (=C、協調) の手を出したことを表している。その右にある2つのビット 1 0は、前々回 $10_{(2)} = 2_{(10)}$ 人の他プレイヤーがCの手ををとったことを表している。同様にして右側の3つのビット 0 0 0は、前回の対戦では、自分は0 (=D、裏切り) の手を出し、他のプレイヤーでCを出したのは $00_{(2)} = 0_{(10)}$ 人であったことを表現している。

プレイヤーの戦略とは、$m$ 回過去までの対戦において生じえるあらゆる歴史に対して、次に出すべき手を対応づける参照表である。戦略もバイナリ列で表される。例えば、次のバイナリ列は4-IPDの戦略を表現している。

[ 1 1 0 0 0 0 1 1 ]

このプレイヤーは1回過去の対戦まで記憶している。例えば前回の対戦を記述した歴史が 1 1 0 であったとすると、このプレイヤーは自分の戦略の左から $110_{(2)} = 6_{(10)}$ 番目の値を参照する。そして今回出す手を1 (=C) と決める。

### 5.2　遺伝子オペレータの構成

5.1で定義したビット列を遺伝子として、突然変異、遺伝子重複、遺伝子分割という遺伝子操作を行い、この遺伝子集団を進化させる。

突然変異によって遺伝子のビットは1文字反転する。突然変異は遺伝子の各ビットに対してある一定の確率 $p_p$ で生じる。

例:　[01] → [00]

遺伝子重複によって遺伝子自信のコピーが $2n-1$ 個その遺伝子の後ろに付け加えられる。例えば、4-IPDでは、遺伝子重複により遺伝子長は8倍に伸びる。遺伝子重複は各エージェントに対してある一定の確率 $p_d$ で生じる。遺伝子重複はエージェントの記憶長の1増加と同時に起こる。遺伝子重複は中立変異であることに注意してほしい。つまり遺伝子の表現としては異なったものになるが、手として現れる戦略は遺伝子重複前とまったく同じである。次の例は、4-IPDにおいて記憶長が $m=1$ から $m=2$ に増加する時の遺伝子重複の様子である。

例:　[11000011] →
[11000011110000111100001111000011
11000011110000111100001111000011]

また遺伝子分割によって、遺伝子のある部分を残してその他の部分は取り除かれる。遺伝子分割の際、もとの遺伝子はまず $2n$ 個の等しい長さのブロックに分割される。例えば、4-IPDでは、遺伝子分割により遺伝子は8個のブロックに分割される。そして、そのうちランダムに選ばれた1ブロックのみが残される。遺伝子分割は遺伝子重複と同様、各エージェントに対してある一定の確率 $p_s$ で生じる。遺伝子分割はエージェントの記憶長の1減少と同時に起こる。次の例は、4-IPDにおいて記憶長が $m=2$ から $m=1$ に減少する時の遺伝子分割の様子である。

例:　[11000110110100101101001110100011
11011011110000111110000110100111]
→ [11010011]
(前から3番目のブロックが残された)

### 5.3　集団のダイナミクス

各世代での対戦を終えると、各エージェント集団は以下の式にしたがい、その成績に応じて全体に占める割合 $x_i$ を増減させる。ここでは個別のエージェントごとではなく、戦略レベルで同じ戦略を持った

エージェント集団をまとめて評価している点に注意してほしい

戦略 $i$ を持つ個体が 1 回のラウンドで得る得点の平均 $s_i$ は、$h^i_{kl}$ を戦略 $i$ を持つ個体が $k$ 回目の対戦、$l$ 回目のラウンドで得た得点として、

$$s_i = \frac{1}{fT}\sum_{k=1}^{f}\sum_{l=1}^{T} h^i_{kl}$$

で表せる。ただし $T$ は 1 回の対戦で行われるラウンド数、$f$ は戦略 $i$ を持つ個体が 1 世代で行う対戦の数である。厳密には、$F$ を集団全体で 1 世代に行われる対戦の総数として、

$$f = Fx_i$$

とすべきだが、シェアの小さい戦略にも十分な対戦の機会を与えるため、$species\_num$ をその世代で存在する戦略の数として

$$f = \frac{F}{species\_num}$$

とする。

また、全個体の 1 回のラウンド当たりの得点の平均は、

$$s = \sum_i s_i x_i$$

である。

戦略 $i$ の適合度 $w_i$ は、戦略 $i$ を持つ個体の 1 回のラウンド当たりの平均得点と、全個体の 1 回のラウンド当たりの平均得点との差で、

$$w_i = s_i - s$$

である。戦略の適合度に比例した値が次の世代への増殖率となるように個体集団は増減する。

$$x_i(t+1) - x_i(t) = dw_i x_i(t)$$

ただし、$d$ は定数である。この式は次のようにも書ける。

$$x_i(t+1) - x_i(t) = ds_i x_i(t)\left(1 - \sum_j \frac{s_j x_j(t)}{s_i}\right)$$

ここで、

$$\sum_i (x_i(t+1) - x_i(t)) = 0$$

であることに注意してほしい。また、$x_i$ が $1/N$ 以下になったら（$N$ はエージェントの総数である）$x_i = 0$ とし、戦略 $i$ を持つ個体は絶滅したものとされる。

### 5.4 その他

初期世代はランダムに生成した記憶長が 1 の 20 の戦略で構成し、各戦略の集団全体に対する割合は等しくする。つまり各戦略の全集団内のシェアはそれぞれ 1/20 である。各戦略集団は各世代でランダムに $n-1$ 個の戦略を選び対戦する。集団全体で 1 世代当たり $F$ 回の対戦が行われる。各対戦は $T$ ラウンド続く。

また、ゲームを行う際、通信路にはノイズが入る。ある一定の確率 $p$ でプレイヤーの手は誤って相手のプレイヤーに伝えられる。

## 6. 結果と考察

### 6.1 プレイヤー数の増加による協調率の低下

各パラメータを図 4のように設定し、3-IPD の実験を行った。以下に 3-IPD での実験結果を示す。図5は戦略集団の進化の様子である。横軸は世代数、縦軸は全集団に各戦略が占める割合である。また図 6に全集団の平均点の推移を示す。

| parameter | value | description |
|---|---|---|
| $N$ | 1000 | population size |
| $F$ | 1000 | number of fights |
| $p$ | 0.01 | error probability |
| $p_p$ | $2 \times 10^{-5}$ | point mutation rate |
| $p_d$ | $10^{-5}$ | gene duplication rate |
| $p_s$ | $10^{-5}$ | split mutation rate |
| $d$ | 0.1 | growth rate |
| $T$ | 1000 | fight length |
| $m_{max}$ | 3 | maximum memory size |

図 4　パラメータの設定

実験の結果、3-IPD でも 2-IPD のような進化の推移が観測された。比較的長い期間、優生種の独占時代が続く。その間にもさまざまな突然変異体が発生と消滅を繰り返す。ある時点で独占時代は突然終わりを告げ、速やかな優生種の交替が起こる。いままで首位の座にあった種は絶滅し、新たな突然変異体が集団を支配し始める。これは、Kristian Lindgren が 2-IPD で見出した断続平衡進化 [5] の様相そのものである。進化の推移の様子では 2-IPD と 3-IPD の間には違いはなさそうである。

図 6の平均点の推移をみると、5000 世代以降、平均点は 2.8 近辺で安定している。これは 3-IPD の全対戦を通して、3 人中平均 1.8 人が協調した（平均 1.2 人が裏切った）ことを示している。これについて少し説明しよう。

図 7に 1 回の 3-IPD ゲームおけるプレイヤーの得点の平均を示す。横軸は C の手を出したプレイヤー

の数、縦軸は得点の平均である。Cの手を出したプレイヤーが0人のときは、全てのプレイヤーはみなそれぞれ1点を得る。よって得点の平均も1点である。Cの手を出したプレイヤーが1人のとき、Cを出したプレイヤーは0点、Dを出した2人のプレイヤーは3点を得る。よってこの時、得点の平均は

$$\frac{0\text{点} \times 1 + 3\text{点} \times 2}{3} = 2\text{点}$$

である。 同様にCを出したプレイヤーが2人の時

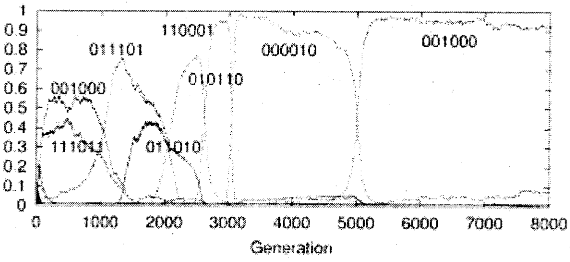

図5　戦略集団の推移

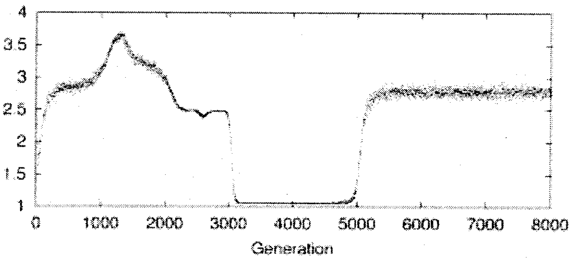

図6　全集団の平均点の推移

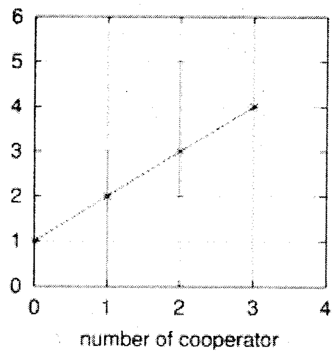

図7　3-IPDゲームの平均点

の得点の平均は

$$\frac{2\text{点} \times 2 + 5\text{点} \times 1}{3} = 3\text{点}$$

であり、Cを出したプレイヤーが3人の時は、みなが4点を得るので、得点の平均は4点である。よって次の式が成り立つ。

$$\bar{P} = C_{num} + 1$$

ただし、$\bar{P}$は得点の平均、$C_{num}$はCを出したプレイヤーの数である。この式の両辺の全ラウンド、全対戦についての平均をとると、

$$\frac{1}{F}\sum^{F}\frac{1}{T}\sum^{T}\bar{P} = \frac{1}{F}\sum^{F}\frac{1}{T}\sum^{T}(C_{num}+1)$$

$$\therefore\ avg(C_{num}) = s - 1$$

ただし、

$$avg(C_{num}) = \frac{1}{FT}\sum^{F}\sum^{T}C_{num}$$
$$s = \frac{1}{FT}\sum^{F}\sum^{T}\bar{P}$$

である。

図6で表されている平均点とは$avg(a)$のことである。よって、平均点が2.8近辺であることから、3-IPDの全ゲームを通して、3人中平均約1.8人が協調した（平均1.2人が裏切った）と結論づけることができる。

$n$-IPD（$n > 2$）では2-IPDに比べて協調行動が生じにくいことが知られている。2-IPDでは相手の裏切りに対して裏切り返すという行動にそれなりの効果があった。しかし$n$-IPD（$n > 2$）ではその報復効果が$n-1$人に分散してしまうため、裏切り行為のリスクが減少する。実験の結果はこの事実を反映しているものと思われる。

### 6.2　対戦数の限定による協調率の低下

次に3-IPDで全集団の平均点の推移に関する2つの結果を示す。図8は総当たり対戦をした時の推移を表し、図9は対戦数$F$を500にした時の推移を表している。対戦数を限定した場合（図9）、対戦数を限定しなかった場合（図8）に比べて平均点が低いことが分かる。これは次のように考えることができる。

協調しようという個体は単独であると、早晩死に絶える。しかし、どんなに小さくても群れをなすことができれば、協調者たちは敵意に満ちた環境の中でも生き残り、繁殖することができる。そして裏切り者どもがお互いにお互いの潰し合いをしている一方で、協調者たちは互恵集団をつくり、かれらはや

がて裏切り者どもを駆逐していく。つまり協調者が生き残るためには、自分と同じように協調してくれる仲間が必要なのである。3-IPD において総当たり

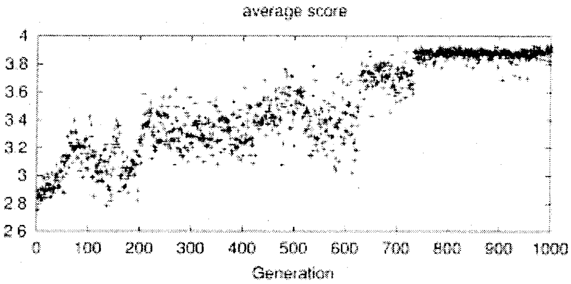


図 8　総当たり対戦時の全集団の平均点の推移

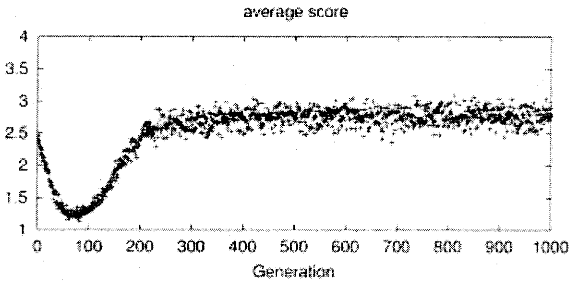


図 9　対戦を制限した時の全集団の平均点の推移

対戦をした場合、協調的なプレイヤーは必ず協調的なパートナーを見つけることができた。しかし、対戦数を限定した場合は、協調的なプレイヤーは必ずしも協調的なプレイヤーと対戦できるとは限らない。この事実が対戦数を限定した場合の協調率の低下につながったものと考えられる。

$n$-IPD の対戦数を限定することの効果は、素朴に考えると、総当たり対戦をした場合に 1 世代で起こっていた現象が数世代をかけて起こるようになり、結果的に進化のスピードが遅れる、という程度であるようにも思われる。しかし実はそれだけには止まらず、ゲームの質そのものを変え、その後の集団の推移に重大な影響を及ぼしかねない。

## 7. おわりに

本研究では、3-IPD においても 2-IPD のようなエージェント集団の進化の推移が見られること、3-IPD ですでにプレイヤー全員の協調行動が生じにくいこと、対戦数を限定すると協調行動が生じにくくなることを示した。今後は、$n$ をさらに大きくした場合の様子を調べるとともに、ゲームの推移をよりミクロな観点から分析していくつもりである。

## 文　献


[1] A. M. Colman. *"Game Theory and Experimental Games"*, pp. 156–159. Pergamon Press, Oxford, England, 1982.

[2] N. S. Glance and B. A. Huberman. "The dynamics of social dilemmas". *Scientific American*, pp. 58–63, March 1994.

[3] G. Hardin. "The tragedy of the commons". *Science*, 1968.

[4] Stenen Kuhn. "Prisoner's Dilemma". In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy (Fall 2000 Edition)*. URL =http://plato.stanford.edu/archives/fall2000/entries/prisoner-dilemma/.

[5] Kristian Lindgren. "Evolutinary Phenomena in Simple Dynamics". In *ARTIFICIAL LIFE-II*, pp. 295–312. Addison-Wesley, 1990.

[6] Per Molander. "The prevalence of free riding". *Journal of Conflict Resolution*, Vol. 36, , December 1992.

[7] Xin Yao and Paul J. Darwen. "An Experimental Study of N-Person Iterated Prisoner's Dilemma Games". *Informatica*, Vol. 18, pp. 435–450, 1994.

[8] 伊庭斉志. 「遺伝的アルゴリズムの基礎」. オーム社, 1994.

[9] 松原望. 「現代人の統計 4 新版 意思決定の基礎」. 朝倉書店, 1985.

[10] ロバート・アクセルロッド松田裕之訳. 「つきあい方の科学 バクテリアから国際関係まで」. CBS 出版, 1987.

[11] 星野力. 「進化論は計算しないとわからない」. 共立出版, 1998.