

96-J-17

有限反復囚人のジレンマにおける協調行動の進化

清水 剛

東京大学大学院経済学研究科大学院生

高橋伸夫

東京大学大学院経済学研究科助教授

1996年9月

このディスカッション・ペーパーは、内部での討論に資するための未定稿の段階にある論文草稿である。著者の承諾なしに引用・複写することは差し控えられる。

Evolution of Cooperation in the Finitely Repeated Prisoner's Dilemma

by

TAKASHI SHIMIZU

Division of Economics, Graduate School of the University of Tokyo

and

NOBUO TAKAHASHI

Faculty of Economics, the University of Tokyo

Abstract

This paper treats the emergence and preservation of cooperation in the finitely repeated prisoner's dilemma.

Cooperative behavior in the prisoner's dilemma has often been observed, even if the number of repetition is definite. But it can not be rationalized from the theory. On this subject, Shimizu (1996) reveals the emerging process of cooperation from the follow-up study of Axelrod's simulation.

In this paper, we discuss this, and examine the "gradual decline" of cooperation from our new simulation. We will show that when the nature of the interaction is long enough term, the gradual decline takes very long time. Therefore cooperation can be regarded as "pseudo stable".

有限反復囚人のジレンマにおける協調行動の進化

清水 剛

(東京大学大学院 経済学研究科 大学院生)

高橋 伸夫

(東京大学大学院 経済学研究科 助教授)

要約

有限反復囚人のジレンマにおいて、理論的には協調行動は成立し得ないにもかかわらず、実際には協調行動が観察される。これについて清水(1996)は、Axelrod のシミュレーションの追試から、有限反復囚人のジレンマにおける協調行動の出現の過程を明らかにした。本稿ではこれを見た上で、さらに有限反復囚人のジレンマにおいて起こるとされる協調行動の「段階的減少」がシミュレーションから検討され、繰り返す回数が多い場合、協調行動の「段階的減少」には非常に長い時間がかかり、むしろ協調行動の方が一般的な状態であることが示される。

1. はじめに

本稿の目的は、有限反復囚人のジレンマ、すなわち「囚人のジレンマ」がある決められた回数繰り返される場合における協調行動の出現と維持を、進化的なアプローチから検討することである。

「囚人のジレンマ」は、典型的には表 1 のような利得表によって表される 2 人ゲームである。各プレイヤーはそれぞれ「協調(C)」か「裏切り(D)」かを選択し、その組み合わせによって各プレイヤーの得る利得が決まる。もし自分も相手も協調を選択するならば二人とも R(表 1 の例で言えば 3) という利得を得る。もし片方が「裏切り」、片方が「協調」を選んだら裏切った方は T(同じく 5)、裏切られた方は S(0)の利得を得る。両方とも「裏切り」ならば利得は二人とも P(1)の利得になる。一般に、これらの利得について

$$T > R > P > S \quad (1.1)$$

の関係が成立するものが「囚人のジレンマ」と呼ばれる。また、通常

$$R > (T+S)/2 \quad (1.2)$$

という関係を付け加えることになっている(Rapoport and Chammah, 1965)。

この「囚人のジレンマ」においては、相手がどちらを選択しても「裏切り(D)」を選択する方が高い利得が得られる。しかし、両方とも「裏切り」を選択した場合に得られる利得は両方が「協調」を選んだ場合に比べ、二人とも低くなってしまふ。これが「ジレンマ」と呼ばれる理由である。

相手の選択に関わらず「裏切り」を取る方が有利であることから、この「囚人のジレンマ」が一回だけ行われた場合、理論的には「協調」という選択肢は取られない。また実際にも、協調行動はあまり取られないとされている。しかし、この「囚人のジレンマ」がある決められた回数だけ繰り返される場合、すなわち有限反復囚人のジレンマにおいては状況が異なってくる。この有限反復囚人のジレンマにおいては、理論的には協調行動は成立し得ないにもかかわらず、実際には協調行動が観察されているのである(Roth, 1988)。

この矛盾を説明するため、今までにいくつかの研究がなされてきた。例えば Kreps らは、対戦相手が何らかの理由により非合理的な戦略である協調的な戦略をとる可能性がある場合に協調行動が導かれることを示している(Kreps et al., 1982)。また Selten and Stoeker (1986)は学習理論による説明を試みているし、その他にもプレイヤーの行動がオートマトンとして定式化される場合を検討した Neyman (1985)、あるいはプレイヤーが協調行動から特別な効用を得られるとする Selten (1978)、Andreoni and Miller (1993) などが挙

げられるだろう。

しかし、進化的なアプローチによれば、このように特別な仮定を置かなくてもこの協調行動の成立が説明できることが分かってきた。ここで言う進化的アプローチとはゲーム理論に進化、淘汰といった進化論的な考え方を取り入れたもので(Taylor & Jonker, 1978; Zeeman, 1981; Maynard Smith, 1982)、社会の中の最適な行動パターンが、(通常ゲーム理論において考えられているように)個人の計算能力によってではなく淘汰や変異の過程の結果として達成されるとする考え方である。後で述べる Axelrod (1980b) はこの進化的アプローチの先駆的な業績の一つと言えるだろう。

Schussler (1989) は、この進化的なアプローチから有限反復囚人のジレンマを検討し、最初のうちは協調行動が支配的になるが、協調行動は徐々に減少し、最終的に合理的な解である全面的な裏切りに到達する「段階的減少(gradual decline)」が起こるとした。また Nachbar (1992) はこれとは別の観点から、同様の結論を導き出している。

このような中で清水(1996)は、Axelrod のシミュレーションについて追試を行い、有限反復囚人のジレンマにおいて協調行動が出現する過程を明らかにした。そこで本稿ではこのシミュレーションについて見た上で、Schussler や Nachbar の主張をシミュレーションから検討し、繰り返す回数が多い場合、協調行動の「段階的減少」には非常に長い時間がかかり、むしろ協調行動の方が一般的な状態であることを示したいと思う。

まず 2 節で反復囚人のジレンマについて見た上で、3 節で清水のシミュレーションを検討する。そして 4 節で協調行動の「段階的減少」についてのシミュレーションを行い、その結果を検討し、最後に結論として協調行動の「疑似安定性」について述べる。

2. 反復囚人のジレンマ

先に述べたように、一回限りの囚人のジレンマにおいては協調行動は成立しないとされている。しかし、これが繰り返されて「反復囚人のジレンマ」になると協調的な行動が取られるようになる、ということが「囚人のジレンマ」の実験の中から明らかになってきた(例えば Rapoport and Chammah, 1965)。

この反復囚人のジレンマの中で無限反復のケース、つまりある回が終わった後次の回が来るかどうか確率的に決められるような場合、協調行動が成立することが理論的にも確かめられている。ゲーム理論においてはトリガー戦略による協調的な部分ゲーム完全 Nash 均衡の存在が証明されている(例えば Gibbons, 1992, pp.91-96)、「フォーク定理」

と呼ばれる、より一般的な形での証明もされている(Friedman, 1971)。また、Axelrodにより「しっぺ返し」戦略(「しっぺ返し」戦略については後述)の集団安定性の証明もされている(Axelrod, 1984)。ただこの証明は不完全であるので、付録 A において補足しておくことにしたい。

しかし、これが有限反復のケース、つまり繰り返しの回数が決まっており、それをプレイヤー達が知っている場合には協調行動は理論的には成立しないのである。これは次のように説明されている。プレイヤーが合理的な選択をすると考えると、まず最後の回ではその次の回での相手の報復があり得ないため、いずれのプレイヤーによっても「裏切り」が選択される。次に最後から2番目の回でも、最後の回はいずれにしる(裏切り、裏切り)になると分かっているので、「裏切り」が選択される。さらに最後から3番目についても同様となり、最終的には最初から最後まで(裏切り、裏切り)という行動が取られることになるのである(Luce and Raiffa, 1957, p.98)。この全面的な裏切りという結果は、有限反復囚人のジレンマにおける唯一の Nash 均衡解なのである。

しかし、この結論は直感的な予想に反しているのと言えるのではないだろうか。事実、有限反復の状況で行われた囚人のジレンマについての実験でも、やはり協調行動は成立しているのである(Selten and Stoeker, 1986; Roth, 1988; Andreoni and Miller, 1993)。むしろ「有限反復囚人のジレンマにおいても協調行動は出現する」というのが定説となっていると言って良い。

先に述べた進化的なアプローチにより、このような有限反復囚人のジレンマに置ける協調行動の発生の道筋を示しているのが、清水(1996)による Axelrod のシミュレーションの追試である。そこで次節では、このシミュレーションを見ることにする。

3. 協調行動の出現

3.1 Axelrod のシミュレーション

反復囚人のジレンマにおける具体的な行動を検討するため、Axelrod はコンピュータによる「選手権」を開催した(1980a; 1980b; 1984)。すなわち、様々な分野の人からプログラムを募集し、そのプログラム同士で反復囚人のジレンマの試合を実際にやらせたのである。

第1回目の選手権は200回の繰り返しゲームで行われ、参加者から提出された14のプログラムと「でたらめ(random)」の15のプログラムの総当たり戦によって争われた。この結果をフィードバックした上で、63のプログラムによって第2回選手権が行われた。た

だし第2回では繰り返しの回数は予め決められてはおらず、各回ごとにある確率で終了するという形で行われた。なお、表1はこのとき用いられた利得表である。

結果は驚くべきものだった。「しっぺ返し(tit for tat; TFT)」というプログラムが2回とも優勝してしまったのである。この「しっぺ返し」はまず最初の回は協調し、次からはその前の回に相手が出したものと同じものを出すというプログラムであった。

さらに「しっぺ返し」の強さを示すため、Axelrodは第2回選手権について「生態学的」なシミュレーションを行ってみた。すなわち、多数の参加者によってこの選手権が何度も繰り返され、その中で高い得点を挙げたプログラム(=戦略)は次の選手権のなかで大きなシェアを占めるようになるのである。それぞれの選手権が「世代」、それぞれの戦略を取る参加者が生物だと考えると、これは進化の過程と同じであると考えられるだろう。

このシミュレーションは、具体的には以下のように整理される。

1. 様々な戦略(プログラム)を取るグループからなる集団を想定し、その中で世代ごとに選手権が行われる。
2. ある戦略がある世代において挙げた得点に比例して、次の世代におけるそのグループの成長の倍率が決まる。つまり来期の構成比率は(今期の構成比率)×(今期の得点)×(比例定数)となる。
3. 初期状態(第0世代)において、それぞれのグループは同数である。
4. ある戦略のある期における得点は、相手となる各戦略に対する得点の加重平均、つまり(相手となる戦略に対するその戦略の得点)×(相手の戦略の構成比率)の総和で表される。

このようなシミュレーションの結果、一番大きな比率を占めるようになったのはやはり「しっぺ返し」であった(Axelrod, 1980b)。しかも「しっぺ返し」は最初から最後までずっと1位を守り続け、他の戦略を大きく引き離したのである。また、その他に生き残った戦略も全て「上品な」—自分から裏切らない戦略であった。このようなことから、Axelrodは協調行動の優位性を結論づけたのである。

3.2 シミュレーションの追試

Axelrodの「生態学的」シミュレーションは、繰り返しの回数があらかじめ決められていない第2回選手権についてのものだった。そこで第1回選手権についてこのシミュレー

シミュレーションを行えば、有限反復囚人のジレンマにおいてどのような戦略が生き残るのか考えることができる。

このシミュレーションは次のように定式化される。戦略 $i=1, \dots, n$ について、 t 期における構成比率を $C_i(t)$ 、平均得点を $P_i(t)$ とし、戦略 i が戦略 j と対戦して得られる得点を r_{ij} とすると、 t 期における戦略 i の平均得点は

$$P_i(t) = \sum_{j=1}^n C_j(t) r_{ij} \quad (3.2.1)$$

となり、 $t+1$ 期の構成比率は

$$C_i(t+1) = \frac{P_i(t) C_i(t)}{\sum_{k=1}^n P_k(t) C_k(t)} \quad (3.2.2)$$

となるのである。ただし、定義より

$$\sum_{i=1}^n C_i(t) = 1 \quad (t=0, 1, \dots), \quad C_i(0) = \frac{1}{n} \quad (i=1, \dots, n) \quad (3.2.3)$$

である。

清水(1996)では、この第1回選手権のシミュレーションを Microsoft Excel を利用して行っている。詳細についてはそちらを参照していただくことにし、ここでは、その結果だけを見ることにしよう(図1, 2, 3)。

図1に見られるように、シミュレーションが始まってすぐに順調に成長する戦略と衰退していく戦略の2つのグループに分かれる。この2つのグループを分けているのは協調的な戦略かどうかであり、非協調的な(「上品」でない)戦略はすぐに衰退していく。この時点で1位は「しっぺ返し」である。

しかし、やがて「しっぺ返し」は衰退し、選手権で2位であったタイドマン&チェルτζジ(Tideman & Chierzzi)と6位のスタイン&ラパポート(Stein & Rapoport)という2つの戦略に完全に追い抜かれてしまう。図2を見ると、おなじ協調的な戦略でもこの2つの戦略以外はすべて減少傾向にあるのが分かるだろう。2000世代のシミュレーションが終わる頃には、このタイドマン&チェルτζジとスタイン&ラパポートの2つの戦略だけが生き残り、他の戦略はいずれ消滅してしまうことがほぼ明らかとなる(図3)。

以上の結果から見られるように、有限反復囚人のジレンマにおいても協調行動は出現するのである。しかし、それは単純な「しっぺ返し」によるものではない。この点について次節で検討してみよう。

3.3 「最後に裏切る」戦略の優位性

まずはじめに、より合理的な戦略であるはずの非協調的な戦略が協調的な戦略に敗れた理由を考えてみよう。その理由は、協調的な戦略はお互いに協調することで高い得点を挙げることができるのに対し、非協調的な戦略はお互い相手を出し抜こうとする結果、逆に高い得点を得ることができないということである。

つまり、先の Axelrod の選手権で言えば、「しっぺ返し」同士なら最初から最後まで(協調、協調)が選択されて $3 \times 200 = 600$ 点になるのに対し、最初から最後まで裏切る戦略(「全面裏切り」)同士が挙げる得点は $1 \times 200 = 200$ 点にしかない。もちろん非協調的な戦略は協調的な戦略を出し抜くことで少しは高い得点を挙げることができるが、それは「しっぺ返し」同士が挙げる得点よりはるかに小さいのである。このシミュレーションには「全面裏切り」は参加していないが、もし参加していたとしてもすぐに衰退してしまっただろう。

それでは、協調的な戦略の中で発展と衰退を分けた要因は何であろうか。実は、生き残った2つの戦略は、最後の2回までは「しっぺ返し」的に振る舞うのだが、最後の2回は必ず「裏切り」を選択するという特徴を持っているのである。

この2つの戦略がその他の戦略と対戦するなら、その結果は表3のようになり、この2つの戦略の得点は600点、一方相手の得点595点になる。そして、もしこの2つの戦略が自分自身もしくはもう一方と対戦するときには、表4のようになり、点数はお互い596点にしかない。しかし、それでも「しっぺ返し」がこの戦略を相手にしたよりも1点多いのである。

結果として、先の2つの戦略は自分たちに対して596点、「しっぺ返し」やその他の戦略に対しては600点の利得を挙げられるのに対し、「しっぺ返し」やその他の戦略は自分たちに対しては同じく600点を挙げられるのだが、2つの戦略に対しては595点しか挙げられない。ゆえに、この2パターン戦略しか考慮に入れる必要が無い状況では、「しっぺ返し」やその他の戦略の得点は、必ずこの2つの戦略の得点を下回るのである。このような状況では「しっぺ返し」やその他の戦略はこの2つの戦略と同じか低い成長率にしかならず(3.2.2参照)、最終的にはこの2つの戦略だけが生き残ることになる。

以上のようなことから、このシミュレーションにおいては「全面裏切り」でもただの「しっぺ返し」でもなく、「最後に裏切る」しっぺ返しが勝利を収めるのである。

4. 協調行動の維持

4.1 協調行動の「段階的減少」

このシミュレーションの、有限反復囚人のジレンマにおいても協調行動が出現するという「非合理的な」結果は、ある意味でむしろ当然のことと言えるだろう。というのは、先にも述べたとおり協調的な戦略はお互いに協調して高い得点を上げることが出来るのに対し、非協調的な戦略はお互いに点数を下げってしまうからである。

むしろここで注目されることは、「単純な」しっぺ返し「最後に裏切る」しっぺ返しに打ち負かされてしまうという点である。別な言い方をすると、途中まで協調しつつ、最後に相手を出し抜いて高い得点をあげるという戦略が生き残るということである。このような「最後に裏切る」行動は実験においても観察されている(Selten and Stoeker, 1986; Andreoni and Miller, 1993)。

Schussler (1989) は、この「最後に裏切る」行動が拡大していく結果、最終的に「全面裏切り」に到達すると考えた。つまり、最初は協調行動が支配的になるのだが、この「最後に裏切る」行動がどんどん広まり、エスカレートしていく結果、協調行動は減少していき、最終的に「合理的な」解である全面的な裏切り状態になるというのである。これが協調行動の「段階的減少(gradual decline)」である。

この現象を考えるために、まず「しっぺ返し」を TFT、最後の 1 回を裏切る「しっぺ返し」を E1、最後の 2 回を裏切るものを E3、以下 E4、E5…と表すことにしよう。また、繰り返しの回数を k とする。利得は表 1 のものを使うと考えると、これらの戦略間の得点は表 2 のように表される。

この表を見ると、いかなる戦略に対しても TFT は E1 と同じかそれより低い得点しか挙げられないことが分かる。これを E1 が TFT を支配(dominate)していると言うが、先のシミュレーションと同様、このような場合 TFT は E1 より同じか低い成長率しかもち得ない。ゆえに初期状態において F がどのような比率であっても、最終的には F は 0 に収束することになる。F が 0 に収束すると今度は E が E1 をほぼ支配することとなる。そして E1 がほとんど 0 になってしまうと今度は E2 が E3 に支配されることになり…というように、最終的には E_k 、すなわち「全面裏切り」に収束するのである。

Nachbar (1992) は、(3.2.1)-(3.2.3)式で示されるような動態分析において全ての戦略を考慮した場合、最終的には Nash 均衡解に収束する事を示しており、その Nash 均衡解への収束の経路として「段階的減少」を主張している。また彼は 6 回繰り返しゲームについてシミュレーションを行っている。

ただし注意すべき点として、この「段階的減少」が起こるためには、E1, E2, …, E k といった戦略が存在していなくてはならない。この点に関しては後ほどまた検討する。

4.2 「段階的減少」のシミュレーション

それでは、この「段階的減少」についてシミュレーションによって検討してみよう。

ここでは TFT, E1-E10 の 11 の戦略について、Axelrod のシミュレーションと同様に、初期状態における構成比率を全て同じ 1/11 とし、(3.2.1)-(3.2.3)によって示されるシミュレーションを Microsoft Excel を使用して行っている。

$k=10$ とした時の結果が図 4 である。これを見ると、最初の頃は $F \rightarrow E1 \rightarrow E2$ と目まぐるしく首位が移り変わっているのだが、そのうちに E5, E6 となると構成比率がほとんど 1 になるようになり、E7, E8 あたりになると 1 でしばらく安定してしまう。そしてしばらくたつと次の戦略が成長してきて首位が交替し、またその戦略がしばらく 1 近傍で安定しているがそのうちにまた次の戦略と交替してしまうというパターンを取るわけである。この、ある戦略での擬似的な安定を Nachbar は「擬似収束(pseudo convergence)」と呼んでいる。

この「擬似収束」の状態では現在支配的である戦略(例えば E7)をその次の戦略(E8)が近似的に支配しており、得点は「次の戦略」の方が高い。ただ、「次の戦略」の構成比率はかなり小さくなっている(ここでの例では 10^{-30} 程度)ので、逆転するまでに時間を必要とする。その間擬似的に「収束」しているのである。

この場合、最終的には 4500 世代程度で「全面裏切り」である E10 に収束することになる。一旦「全面裏切り」に到達してしまえば、「全面裏切り」と対戦して「全面裏切り」より高い得点を得ることはできないので、ここで安定することになる。

今度は $k=200$ としてみよう(図 5)。ここでも基本的なパターンは同一であるが、交替に必要な世代数は大きく異なる。このシミュレーションでは E4 が首位を占めるようになるのが 3000 世代程度、5000 世代より少し前でようやく E5 が逆転し、そこから E6 へは更に 5000 世代ぐらいかかっている。そして 20000 世代でたどり着くのは E7 までである。このシミュレーションで E10 が支配的になるには 50000 世代以上かかってしまうのである。

4.3 繰り返し回数の増大の効果

今まで見てきたように、繰り返しの回数が長くなると、収束するまでの時間は長くなる。この理由は、繰り返し回数の増加につれて世代ごとの変化が遅くなることにある。すなわ

ち(3.2.2)式から、

$$\frac{C_i(t+1) - C_i(t)}{C_i(t)} = \frac{P_i(t) - \sum_{k=1}^n P_k(t)C_k(t)}{\sum_{k=1}^n P_k(t)C_k(t)} \quad (4.3.1)$$

となり、 t 期での変化率は戦略の得点と全体の平均得点の差に比例する。そしてこのシミュレーションにおいては得点は試合回数が長くなるほど大きくなるのである。

繰り返しの回数が多いことは、もう一つ別の意味を持っている。つまり、繰り返し回数が多くなると、一回の交替が全体としての協調行動の比率に与える影響力が小さくなるのである。例えば、 $k=10$ の場合と $k=200$ の場合を比べると、 $k=10$ の場合 E1 から E2 への交替は協調行動の比率を 0.9 から 0.8 へ下げるのに対し、 $k=200$ の場合 0.995 から 0.99 に下げることにしかならないのである。

結果として、繰り返しの回数が多くなると、まず変化率が小さくなり、また一つの交替の影響が小さくなるため、協調行動が長期にわたって存続するようになる。

この点を示すために、試合の回数と協調行動の関係を表したのが図 6 である。グラフは k 回繰り返しゲームでの世代と協調行動比率(ゲームの中で協調行動がとられる割合)を示している。これを見ると $k=10$ の時、協調行動比率が 2000 世代で 0.2 に落ちているのに対し、 $k=20$ なら 0.7 程度、 $k=200$ ならほとんど 1 で推移している。

「段階的減少」においても、 k が大きいならば協調行動は長期にわたって維持されるのである。

4.4 突然変異

先にも述べた通り、この「段階的減少」が成立するには、E1, E2, E3…の全ての戦略の構成比率が、この「段階的減少」が起こっている期間中正でなくてはならない。これを満たすために、Nachbar は「初期時点において全ての戦略が存在する(正の構成比率を持つ)」、「構成比率は 0 にはならない」の 2 つの条件を置いている。

しかし、このシミュレーションの中では戦略の構成比率はかなり小さくなり、場合によっては 10^{-100} 、 10^{-200} というレベルにまでなる。このように小さい数ではむしろ消滅したと考える方が妥当であろう。また、例えば 2 節のシミュレーションのようにそもそもある戦略が最初から存在していないという可能性もある。

そう考えると上の条件は多少非現実的ではあるが、シミュレーションの中に突然変異を

入れるのが難しいところから突然変異の変わりに用いられたと考えられる。

しかし、本質的には戦略の生成は突然変異もしくは他の何らかの方法で新しい戦略が生成されると考えるべきものであろう。そこで突然変異を考慮した場合を少し検討してみよう。

勿論、突然変異の発生率や規模によってシミュレーションは大きな影響を受けることになり、このシミュレーションにおける結果は不確定なものになる。ただ、もし突然変異というのがそれほど起こるものではないというように考えるならば、「段階的減少」は遅くなると考えるのが自然であろう。

また、新しい戦略が取り入れられることにより「全面裏切り」の中から協調行動が出現する可能性もある。

戦略 A が B に対して挙げる得点を $V(A|B)$ とし、戦略 A の平均得点を U で表わすことにしよう。また、戦略 A の構成比率を C_A で表わす。これより

$$\begin{aligned} V(\text{TFT}|\text{TFT}) &= 3k \\ V(\text{TFT}|E_k) &= (k-1) \\ V(E_k|\text{TFT}) &= 5+(k-1) \\ V(E_k|E_k) &= k \end{aligned} \tag{4.4.1}$$

である。さて、「全面裏切り」まで行き着いた集団の中で、何らかの方法によって「しっぺ返し」が発生したとする。ここで

$$C_{\text{TFT}}=p, C_{E_k}=1-p, C_{E_1}=C_{E_2}=\dots=C_{E_{k-1}}=0 \tag{4.4.2}$$

とすると、

$$U(\text{TFT})=pV(\text{TFT}|\text{TFT})+(1-p)V(\text{TFT}|E_k)=3kp+(1-p)(k-1) \tag{4.4.3}$$

$$U(E_k)=pV(E_k|\text{TFT})+(1-p)V(E_k|E_k)=p\{5+(k-1)\}+(1-p)k \tag{4.4.4}$$

より、 $U(\text{TFT})>U(E_k)$ すなわち

$$p>1/(2k-3) \tag{4.4.5}$$

の時、「全面裏切り」の中から協調行動が出現することになる。これは $k=200$ の時 0.25% 程度であり、何らかの方法によってこれだけ「しっぺ返し」を取る個体が発生すれば良いことになる。Axelrod が指摘しているように、「内輪づきあい」を認めるならさらに小さい (Axelrod, 1984)。

まとめると次のようになる。もし突然変異を考慮に入れた場合はその規模や生起確率によって影響が異なるが、一般的には「段階的減少」が遅くなると考えられる。また場合によっては「全面裏切り」の中から協調行動が出現する可能性もある。

この4節での結論は、「段階的減少」という条件の下でも繰り返し回数が多ければ協調行動は長期にわたって維持されるし、また「段階的減少」によって考えられるより長く協調行動が維持される可能性もあるということである。

5. 結論

これまで見てきたように、進化的アプローチから考えれば有限反復囚人のジレンマにおいても協調行動は出現し、かつ長期にわたって維持されると考えられる。もちろん、厳密に言えばこの協調行動は最終的には消滅する。しかし、そこまでの期間は長期にわたると考えられるし、仮に協調行動が消滅したとしても、例えば協調行動の方が有利であることに気づく人がある程度いればまた協調行動が出現することになる。

このように考えると、有限反復囚人のジレンマは「消え去るまでの一時的現象」ではなく、むしろ「裏切り」の方が特殊な現象であると考えて良いであろう。有限反復囚人のジレンマにおける協調行動は擬似的に安定しているのである。

現実において、確実に有限反復囚人のジレンマと考えられることはそれほど多くないかもしれない。しかし、ここで考察された有限反復囚人のジレンマにおける擬似安定性は、現実を説明する上で一つの示唆になると思われる。

有限反復囚人のジレンマにおいても協調行動は進化するのである。

付録 A 「しっぺ返し」の集団安定性の証明

この証明は、Axelrod (1981) において提出され、Axelrod (1984) の付録 B に再録された命題についてのものである。この命題の必要条件は Axelrod (1981) によって証明されているが、十分条件については証明が不完全である。ここで証明するのは十分条件についてである。

命題 2 いかなる戦略 A についても

$$V(\text{TFT} | \text{TFT}) \geq V(A | \text{TFT}) \Leftrightarrow w \geq \max\{(T-R)/(R-S); (T-R)/(T-P)\}$$

十分条件の証明：

(I) 「しっぺ返し」対「しっぺ返し」の対戦は表 1 に示されているような形になる。ここで表 6 に示されるような戦略 $D_1(k)$ をプレイヤー 2 が取ると考える。

表 5 と 6 の違いは太線で囲まれた k 回目と $k+1$ 回目だけである。ここで $U_2(k)$ を k 回目と $k+1$ 回目を除く合計点とすると、

$$V(\text{TFT} | \text{TFT}) = w^{k-1}(R+wR) + U_2(k)$$

$$V(D_1(k) | \text{TFT}) = w^{k-1}(T+wS) + U_2(k)$$

故に全ての $k=1, 2, 3, \dots$ について

$$V(\text{TFT} | \text{TFT}) \geq V(D_1(k) | \text{TFT}) \Leftrightarrow w \geq (T-R)/(R-S)$$

(II) 次に表 7 と 8 で示されるような $D_n(k)$ と $D_{n+1}(k)$ をプレイヤー 2 が取るとする。 $n=1$ の場合は (I) に示されたとおりである。故に $n \geq 2$ と仮定する。

表 7 と 8 の違いは太線で囲まれた $k+n$ 回目と $k+n+1$ 回目だけである。ここで $U_2(k+n)$ を $k+n$ 回目と $k+n+1$ 回目を除く合計点とすると、

$$V(D_n(k) | \text{TFT}) = w^{k+n-1}(S+wR) + U_2(k+n)$$

$$V(D_{n+1}(k) | \text{TFT}) = w^{k+n-1}(P+wS) + U_2(k+n)$$

故に全ての $k=1, 2, 3, \dots$ について

$$V(D_n(k) | \text{TFT}) \geq V(D_{n+1}(k) | \text{TFT}) \Leftrightarrow w \geq (P-S)/(R-S)$$

$$V(D_n(k) | \text{TFT}) \leq V(D_{n+1}(k) | \text{TFT}) \Leftrightarrow w \leq (P-S)/(R-S)$$

(a) $w \geq (P-S)/(R-S)$ かつ $w \geq (T-R)/(R-S)$ の場合、(I)より全ての k について

$$V(\text{TFT} | \text{TFT}) \geq V(D_1(k) | \text{TFT}) \geq V(D_2(k) | \text{TFT}) \geq \dots \geq V(D_n(k) | \text{TFT}) \geq \dots$$

(b) $w \leq (P-S)/(R-S)$ かつ $w \geq (T-R)/(T-P)$ の場合、全ての k について

$$\begin{aligned} & V(D_1(k) | \text{TFT}) \leq V(D_2(k) | \text{TFT}) \leq \dots \leq V(D_n(k) | \text{TFT}) \leq \dots \\ & \leq V(\text{TFT} | \text{TFT}) - w^{k-1} V(\text{TFT} | \text{TFT}) + w^{k-1} V(\text{ALL D} | \text{TFT}) \\ & = (1-w^{k-1}) V(\text{TFT} | \text{TFT}) + w^{k-1} V(\text{ALL D} | \text{TFT}) \\ & \leq \max\{V(\text{TFT} | \text{TFT}); V(\text{ALL D} | \text{TFT})\} \end{aligned}$$

ここで $w \geq (T-R)/(T-P)$ より

$$V(\text{TFT} | \text{TFT}) = R/(1-w) \geq T+wP/(1-w) = V(\text{ALL D} | \text{TFT})$$

故に、

$$V(D_1(k) | \text{TFT}) \leq V(D_2(k) | \text{TFT}) \leq \dots \leq V(\text{TFT} | \text{TFT})$$

(a)と(b)から、 $w \geq (T-R)/(R-S)$ かつ $w \geq (T-R)/(T-P)$ ならば、全ての k 及び $n(\geq 2)$ について

$$V(\text{TFT} | \text{TFT}) \geq V(D_n(k) | \text{TFT})$$

(III) (I)と(II)から、しっぺ返しに対してはいかなる「裏切り」も得点を増加させることができない。故に1回以上の長さを持つ「裏切り」を取る戦略は、しっぺ返しより高い得点を挙げることができない。

ここで $(k-1)$ 回目から $(k+n+1)$ 回目までの間協調をとり続ける任意の戦略 $D'_0(k)$ に対して、 k 回目から $k+n$ 回目までの間 $n+1$ 回は裏切り続けるが、他は同じ戦略を $D'_{n+1}(k)$ としよう。(I),(II)と同様の方法により、 $w \geq (T-R)/(R-S)$ かつ $w \geq (T-R)/(T-P)$ ならば全ての k, n に対し、

$$V(D'_0(k) | \text{TFT}) \geq V(D'_{n+1}(k) | \text{TFT})$$

である。

これで命題2は証明された。

付録B Microsoft Excel による「段階的減少」のシミュレーションの方法

ここでは「段階的減少」のシミュレーションの方法を簡単に述べておく。基本的なコンセプトは清水(1996)と同一であるので、わかりにくい点はそちらを参照していただきたい。なお、{}は入力すべきセル内の式や文字を表している(カッコ内を入力する)。

1. B1 に繰り返しの回数を入力する。
2. A3 から L14 の間に表 2 の得点表を入力する。まず A4-A14 に{TFT},{E1},{E2}…と入力し、B3-L3 も同様にする。
3. B4 には $3k$ 、すなわち $\{=3*\$B\$1\}$ と入力する。そして B5 は $3k-3$ 、すなわち $\{=3*\$B\$1-3\}$ とする。以下他のセルも同様にして得点表を作成する。
4. 各世代における構成比率は、B 列から L 列の 17 行から下に列ごとに表示される。そのため、B17 から L17 にラベルとして各戦略の名前を入れ(その順番は得点表と同じ)、B18-L18 に初期構成比率として $\{=1/11\}$ を入力する
5. 各世代における平均得点は、N 列から X 列までの 17 行から下にやはり列ごとに表示される。そのため、N17-X17 にも各戦略の名前を入れておく。
6. N18-X18 に得点の計算式を入力する。N18 に TFT の得点として $\{=SUMPRODUCT(B18:L18,\$B\$4:\$L\$4)\}$ 、O18 に E1 の得点として $\{=SUMPRODUCT(B18:L18,\$B\$4:\$L\$4)\}$ とするように、N18 から X18 まで一つ右のセルになるごとに $\$B\$$ と $\$L\$$ の後の数字を一つずつ増やしながらか入力する。
7. Z 列には群全体の成長率(各戦略の得点の加重平均)が表示される。この成長率の計算式として Z18 に $\{=SUMPRODUCT(B18:L18,N18:X18)\}$ と入力する。
8. B19-P19 には新しい構成比率の計算式を入力する。B19 に $\{=B18*N18/\$Z18\}$ とし、C20-P20 にコピーすれば良い。
9. A 列は世代を表示する。このため、A18 から下に 0,1,2…と言うように数字を埋め込んでいく。 [編集]-[フィル]-[連続データ作成]を使用する。
10. 後は計算式を下にコピーして行けばシミュレーションが完成する。そのためまず N18-Z18 までを N19-Z19 にコピーし、次に B19-Z20 までを下の行にコピーすれば良い
11. シミュレーションが完成したら、 [グラフウィザード]を利用してグラフを描かせる。

参考文献

- Andreoni, James and John H. Miller (1993) "Rational cooperation in the finitely repeated prisoner's dilemma: Experimental analysis," *Economic Journal*, 103, 570-585.
- Axelrod, Robert (1980a) "Effective choice in the prisoner's dilemma," *Journal of Conflict Resolution*, 24, 3-25.
- Axelrod, Robert (1980b) "More effective choice in the prisoner's dilemma," *Journal of Conflict Resolution*, 24, 379-403.
- Axelrod, Robert (1981) "The emergence of cooperation among egoists," *American Political Science Review*, 75, 306-318.
- Axelrod, Robert (1984) *The Evolution of Cooperation*. Basic Books, New York. (松田裕之訳 『つきあい方の科学』 HBJ 出版局, 1987)
- Friedman, James W. (1971) "A non-cooperative equilibrium for supergames," *Review of Economic Studies*, 38, 1-12.
- Gibbons, Robert (1992) *Game Theory for Applied Economists*. Princeton University Press, Princeton, N. J. (木村憲二訳 『応用経済学のためのゲーム理論入門』 マグロウヒル 出版, 1994)
- Kreps, David M., Paul Milgrom, John Roberts & Robert Wilson (1982) "Rational cooperation in the finitely repeated prisoner's dilemma," *Journal of Economic Theory*, 27, 245-252.
- Luce, R. Duncan and Howard Raiffa (1957) *Games and decisions: Introduction and Critical Survey*. John Wiley & Sons, New York.
- Maynard Smith, John (1982) *Evolution and the Theory of Games*. Cambridge University Press, Cambridge. (寺本英・梯正之訳 『進化とゲーム理論』 産業図書, 1985)
- Nachbar, John H. (1992) "Evolution in the finitely repeated prisoner's dilemma," *Journal of Economic Behavior and Organization*, 19, 307-326.
- Neyman, Abraham (1985) "Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma," *Economic Letter*, 19, 227-229.
- Rapoport, Anatol & Albert M. Chammah (1965) *Prisoner's Dilemma: A Study in Conflict and Cooperation*. The University of Michigan Press, Ann Arbor, Mass. (広松毅・平山朝治・田中辰雄訳 『囚人のジレンマ』 啓明社, 1983)

- Roth, Alvin E. (1988) "Laboratory Experimentation in economics: A methodological overview," *Economic Journal*, 98, 974-1031.
- Schuessler, Rudolf (1989) "The gradual decline of cooperation: Endgame effects in evolutionary game theory," *Theory and Decision*, 26, 133-155.
- Selten, Reinhard (1978) "Chain store paradox," *Theory and Decision*, 9, 127-159.
- Selten, Reinhard and Rolf Stoeker (1986) "End behavior in sequence of finite prisoner's dilemma supergames," *Journal of Economic Behavior and Organization*, 7, 47-70.
- 清水剛 (1996) 「進化のシミュレーション」高橋伸夫編『未来傾斜原理－協調的な経営行動の進化』, 白桃書房(近刊)。
- Taylor, Peter D. and Leo B. Jonker (1978) "Evolutionarily stable strategies and game dynamics," *Mathematical Bioscience*, 40, 145-156.
- Zeeman, E. C. (1981) "Dynamics of the evolution of animal conflicts," *Journal of Theoretical Biology*, 89, 249-270.

表1 「囚人のジレンマ」の利得表

	C(協調)	D(裏切り)
C(協調)	R=3, R=3	S=0, T=5
D(裏切り)	T=5, S=0	P=1, P=1

(出典)Axelrod, 1980a, Table 1.

表2 「段階的減少」における戦略間の得点

	TFT	E1	E2	E3	E4	...	En
TFT	$3k$	$3k-3$	$3k-5$	$3k-7$	$3k-9$...	$3k-1-2n$
E1	$3k+2$	$3k-2$	$3k-5$	$3k-7$	$3k-9$...	$3k-1-2n$
E2	$3k$	$3k$	$3k-4$	$3k-7$	$3k-9$...	$3k-1-2n$
E3	$3k-2$	$3k-2$	$3k-2$	$3k-6$	$3k-9$...	$3k-1-2n$
E4	$3k-4$	$3k-4$	$3k-4$	$3k-4$	$3k-8$...	$3k-1-2n$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
En	$3k+4-2n$	$3k+4-2n$	$3k+4-2n$	$3k+4-2n$	$3k+4-2n$...	$3k-2n$

表 3 対戦の結果(TFT 対 T&C)

	1	2	……	198	199	200	計
TFT	C	C	……	C	C	D	
T&C(S&R)	C	C	……	C	D	D	
得点(TFT)	3	3	……	3	0	1	595
得点(T&C)	3	3	……	3	5	1	600

表 4 対戦の結果(T&C 対 T&C)

	1	2	……	198	199	200	計
T&C(S&R)	C	C	……	C	D	D	
T&C(S&R)	C	C	……	C	D	D	
得点(T&C)	3	3	……	3	1	1	596
得点(T&C)	3	3	……	3	1	1	596

表5 TFT 対 TFT

プレイヤー	戦略	1	$k-1$	k	$k+1$	$k+2$
1	TFT	C	C	C	C	C
2	TFT	C	C	C	C	C
プレイヤー2の利得		R	R	R	R	R

表6 TFT 対 $D_1(k)$

プレイヤー	戦略	1	$k-1$	k	$k+1$	$k+2$
1	TFT	C	C	C	D	C
2	$D_1(k)$	C	C	D	C	C
プレイヤー2の利得		R	R	T	S	R

表7 TFT 対 $D_n(k)$

プレイヤー	戦略	1	$k-1$	k	$k+1$	$k+n-1$	$k+n$	$k+n+1$	$k+n+2$
1	TFT	C	C	C	D	D	D	C	C
2	$D_n(k)$	C	C	D	D	D	C	C	C
プレイヤー2の利得		R	R	T	P	P	S	R	R

表8 TFT 対 $D_1(k)$

プレイヤー	戦略	1	$k-1$	k	$k+1$	$k+n-1$	$k+n$	$k+n+1$	$k+n+2$
1	TFT	C	C	C	D	D	D	D	C
2	$D_{n+1}(k)$	C	C	D	D	D	D	C	C
プレイヤー2の利得		R	R	T	P	P	P	S	R

図1 第1回選手権の結果によるシミュレーション(20世代まで)

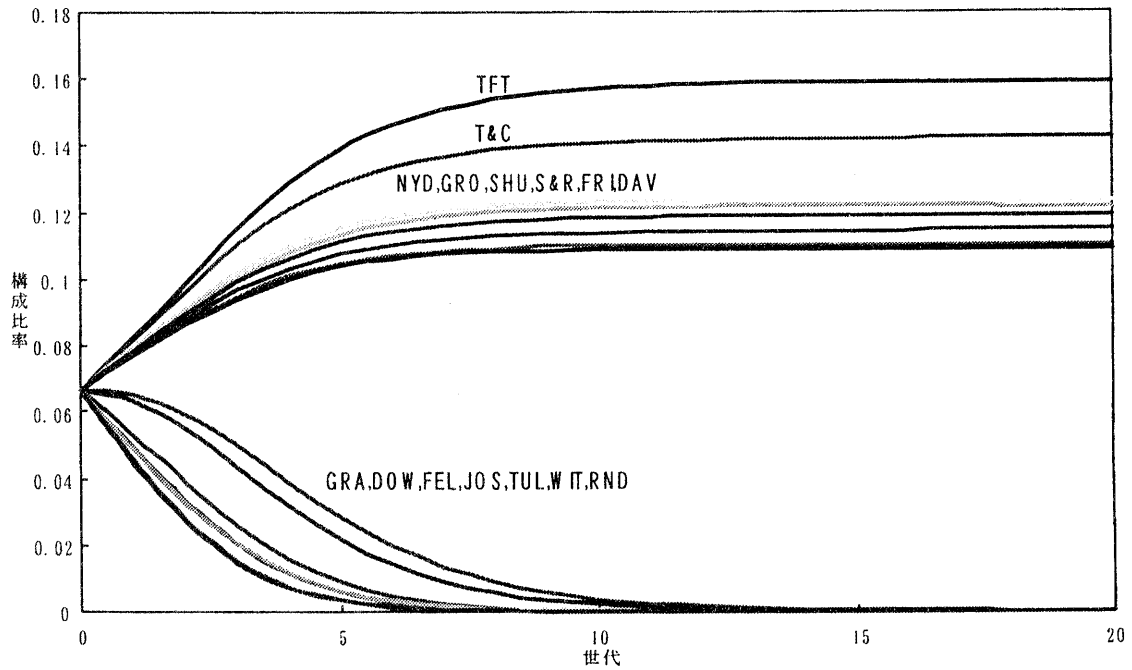


図2 第1回選手権の結果によるシミュレーション(200世代まで)

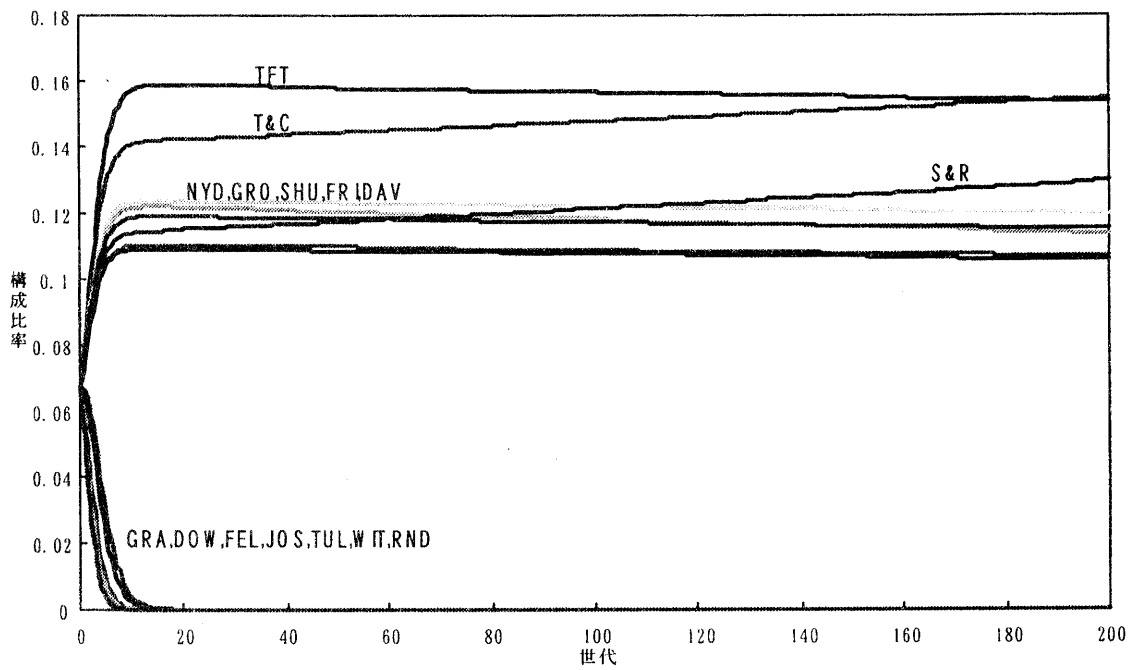


図3 第1回選手権の結果によるシミュレーション(2000世代まで)

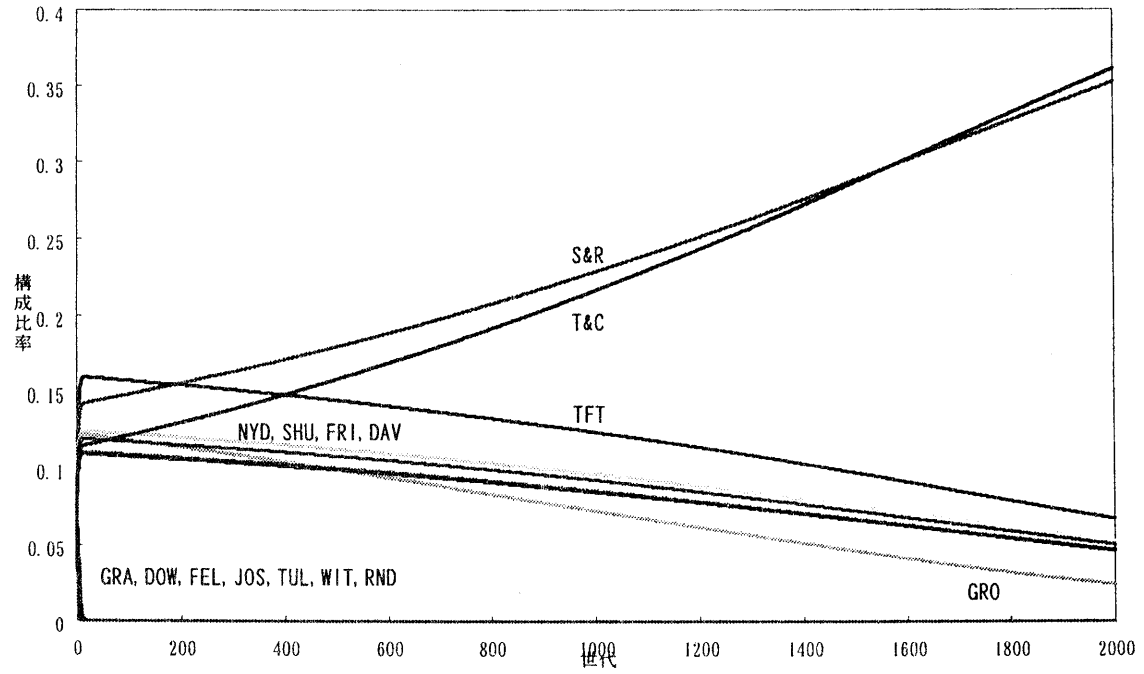


図4 「段階的減少」のシミュレーション($k=10$)

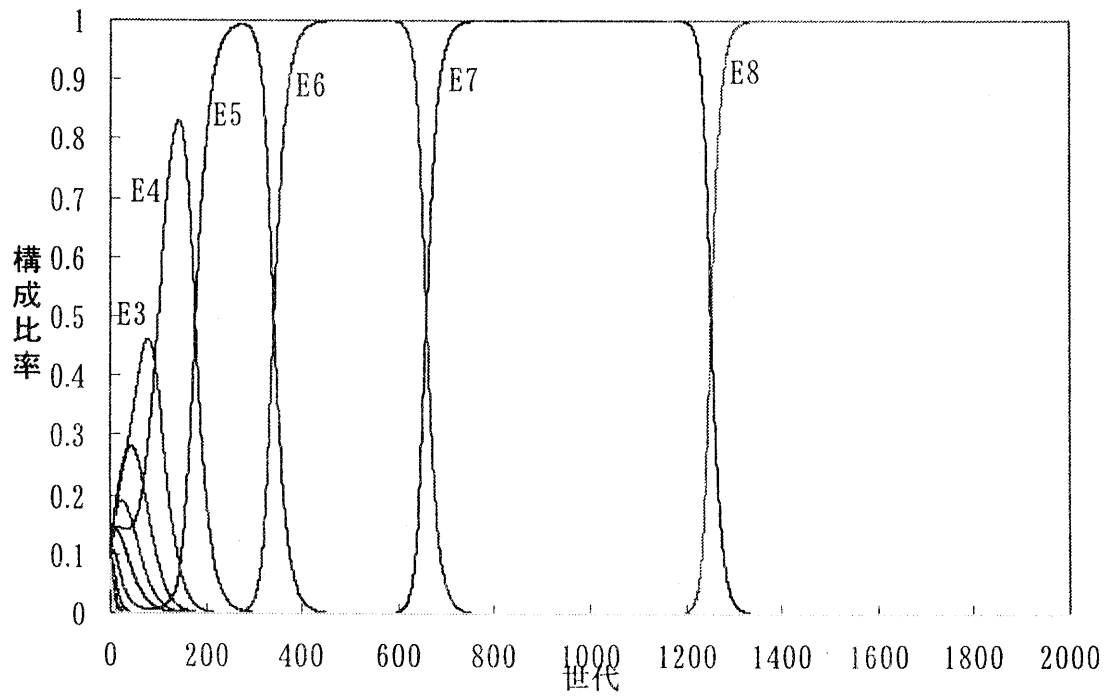


図5 「段階的減少」のシミュレーション($k=200$)

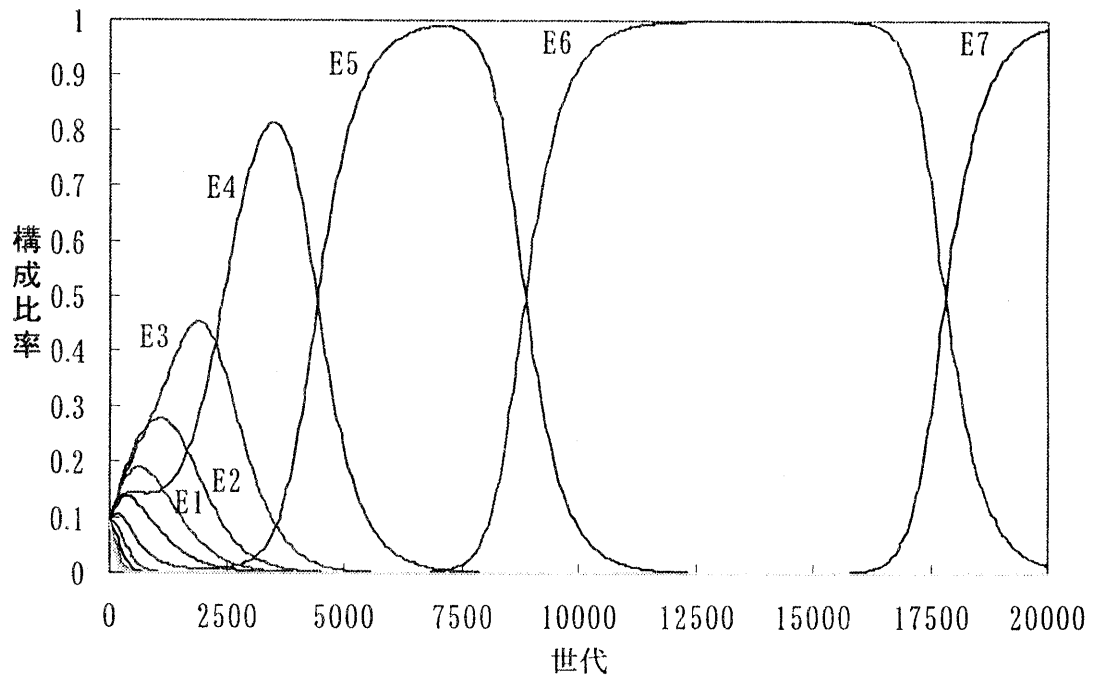


図6 「段階的減少」における協調行動比率

