

# 1 回限りの社会的ジレンマ問題における限定合理的な 行動主体の協力行動選択についての分析

獨協大学 今 福 啓

## 1. はじめに

我々の社会における問題に、より望ましい社会状況があるにもかかわらず、行動主体が自己の利益のみを追求するとそれを実現できない状況におちいる「社会的ジレンマ問題」がある。よく知られた問題のひとつは、囚人のジレンマ問題である。この問題は、同じ相手と継続的に行動選択する関係を築く「くり返し囚人のジレンマ問題」となるならば、個人が自己の利益のみを追求する合理性を追求しても協力行動が選択されることが明らかとなっている<sup>[1][2]</sup>。しかし合理性を追求する行動主体は、その後2度と会わない相手と1回だけ行動を選択する「1 回限りの囚人のジレンマ問題」に直面すると、得られる利益を最大化するため非協力行動のみを選択するため、望ましい状況を実現できない。

このような違いが生じるのは、くり返し囚人のジレンマ問題は自分の相手となる行動主体が誰であるのかを認識できる限定交換に属するが、1 回限りの囚人のジレンマ問題は相手が誰であるのかを認識しない一般交換に属するためである<sup>[1]</sup>。限定交換では各行動主体が過去に選択した行動を参照できるため、協力行動を継続することが合理的となる。それに対して、一般交換では相手が誰であるのかを認識しないことから、非協力行動が合理性となる。ところが実社会では、1 回限りの囚人のジレンマ問題においても、協力行動を選択することが生じている<sup>[1]</sup>。なぜ一般交換において協力行動が選択されるのかについて明らかにするため、これまで多くの研究がおこなわれてきた。

実験社会心理学では、行動主体が合理性を追求するという観点からでは協力行動を選択することの説明が困難であるとされた<sup>[3]</sup>。そのため、相手の得る利益から自己が満足感を得るという動機的説明や、自己のおかれた状況にたいする認知の違いから生じるという相互依存性の認知による説明、そして自分の行動を変えることで相手の行動をコントロールできるとする信念といった、自己利益の追求ではない要因から説明がなされてきた。反対に実験経済学では、行動主体の合理性のレベルの違いを想定したモデル構築による説明がおこなわれている<sup>[4]</sup>。行動主体のモデルとして、合理性の程度を1つのパラメータで表現し、それを変更することでその違いを表現できる Quantal Response Equilibrium(QRE) や、合理性のレベルの異なる行動主体からなる階層構造から構成されたとしたモデルである K レベル思考ステップモデルが提案されている。

これらの研究は、被験者による実験から得られた行動分析が主体となっている。そのため、被験者の合理性を想定しての実験であっても、被験者の感情面が行動選択に影響している可能性を否定できない。そこで本研究では、感情の影響を完全に排除することのできるコンピュータプログラムで、自己の利得を最大化する行動を選択する合理性をもつ行動主体を構築する。

本研究と同様に、コンピュータ内部にモデル化した行動主体が相手を認識せずに得た過去の経験を使用し、1 回限りの囚人のジレンマ問題で協調行動を維持できることを示した研究はすでにおこなわれている。その一つは、Q 学習とよばれる強化学習手法で過去に得た利得を学習し、将来の行動選択では協力行動が維持される結果を示した研究である<sup>[8]</sup>。この研究では行動主体が協力を維持するための条件として、Q 学習の学習速度を決めるパラメータである学習率と、過去の経験をどの程度割り引くのかという割引率が満たすべき条件を示し

---

1 結果は省略するが、3章で述べる手順において利得を1通りに限定したシミュレーションをおこない、すべての行動主体が非協力行動を選択する結果となったことを確認した。

ている。本研究はQ学習のような特定の行動決定手法を想定していない点と、以下に述べるとおり、より実社会に近い社会状況での行動選択をモデルに取り入れている点が異なっている。

通常、合理性をもつ行動主体が1回限りの囚人のジレンマ問題で選択する行動は非協力となり、協力行動が引き出される余地はない。これは、利得が1通りに限定されている問題において、相手が誰であるのかを認識せず1回のみ合理的に行動選択する場合の結果としてよく知られている。しかし、実社会における社会的ジレンマ問題は、企業間の投資におけるホールドアップ問題<sup>[5]</sup>、交通問題<sup>[6]</sup>、社会的行動の分析<sup>[7]</sup>と、さまざまに存在する。また実社会では、1回のみ社会的ジレンマ問題での行動選択をおこなうわけではなく、日々さまざまな社会的ジレンマ問題が生じ、そのたびに行動選択をおこなっていると考えることが自然である。

このような実社会の状況での行動決定を模倣するため、本研究ではジレンマ的な状況は同一であるが、複数の利得構造を含む社会モデルを構築する。そして、一定期間過去までの囚人のジレンマ問題の利得表と選択された行動を記憶し、それを利用して合理的に行動選択する行動主体をモデル化する。行動主体は相手が誰であるのか、過去にどのような行動選択をおこなったのかを認識せずに、複数の利得構造のうちランダムに遭遇した囚人のジレンマ問題において行動を選択する。行動主体は、記憶する過去の行動から新しい行動の組み合わせを作成し、最も高い利得となったものを次の行動決定に利用するという、合理性にもとづいた行動選択をおこなう。ただし、作成可能な行動は総数が膨大となることから、そのすべてを試すことは困難である。そのため、記憶する行動を試行錯誤的に変更するという、限定合理性をもつ行動モデルを構築する。

本研究のモデルにより、限定合理性をもつ行動主体であっても一定の割合の協力行動を導くことができることを、コンピュータシミュレーションの結果から示す。

## 2. 行動主体の行動決定モデル

本研究で構築する社会モデルは複数の行動主体から構成され、その中から選択された2者が自己の記憶する過去の経験をもとにして、協力が非協力の一方を選択する。1回限りの囚人のジレンマ問題は、利得構造が1つに限定されるならば、行動主体が過去の行動を記憶して行動選択に利用しても非協力行動のみを選択する状況に収束する。ところが1章で述べたとおり、実社会で直面する囚人のジレンマ状況にはさまざまな問題が存在する。そこで、本研究では複数の利得表からなる囚人のジレンマ問題を用意し、選択された行動主体はその中から行動決定時に示された1つの問題において行動を選択する。

各行動主体は、囚人のジレンマ問題において複数回の行動選択をおこなう。その際、行動主体は相手が誰であるのか、そして過去にどのような行動を選択したのかは認識せず、自分および相手が過去にどの利得表でどの行動を選択したのかという経験を記憶し、将来の行動選択に利用する。このようにして、1回限りの囚人のジレンマ問題をくり返しおこなう。

また、本研究では各行動主体の合理性を仮定するが、保持する情報のすべてを利用して行動決定する完全合理性をもつのではなく、情報の一部を利用する限定合理性をもつとしてモデル化する。その理由の1つは、囚人のジレンマ問題のように選択できる行動が2つに限定されていても、将来的に高い利得を得るために過去の経験を変更して新たな行動を作成すると、その組み合わせ総数が膨大となるためである。たとえば10回前までの行動を記憶する状況を考える。1回あたりの自分および相手の行動の組み合わせ総数は、 $2 \times 2 = 4$ 通りである。そのため、10回の行動の組み合わせ総数は $4^{10} = 1,048,576$ となり、きわめて膨大な数となる。もし行動主体の完全合理性を仮定すると、この組み合わせをすべて試した上で最も良い行動を選択することになるが、とても現実的な状況とはいえない。

2つ目の理由は、認知科学で人の記憶システムに制約があることを明らかにしている点である<sup>[9]</sup>。人の記憶システムは、長期記憶と短期記憶から構成される。長期記憶は過去の経験を保持し、蓄積された内容は一時的に引き出して置くための7つ前後の要素からなる短期記憶に想起するとされる。短期記憶のような限定的な部分の存在が明らかとなっていることから、行動主体の処理能力に制約を設けることは不自然でないと考えられ

る。

3つ目の理由として、実験経済学では限定合理性を支持する理論が提案されている点をあげることができる<sup>[4]</sup>。行動主体のモデルとして、1つのパラメータで合理性の度をあらわし、それを変更することで行動主体の合理性を調整できる QRE や、合理性のレベルが階層構造となっていることを示す K レベル思考ステップモデルが提案されている。これらは人の意思決定時の合理性に制約があることをあらわし、行動モデルに制約を設けることの妥当性を示している。

以上の点から、本研究では複数の利得構造から構成される 1 回限りの囚人のジレンマ問題で、限定合理性により行動を決定する行動主体が過去の経験を利用して行動決定をおこない、協力行動を導きだせるのかをコンピュータシミュレーションの結果から明らかにする。

### 3. シミュレーション

構築したモデルを用いたコンピュータシミュレーションの手順と、結果および考察を述べる。

#### 3.1 シミュレーションの手順

シミュレーションの手順は、次のとおりである。各手順の詳細は後述する。

手順 1. 総数  $N_r$  の利得表を用意する。利得表  $R_i$  ( $i$  は利得表の番号で、 $i=1, \dots, N_r$ ) において、行動主体は協力か非協力のいずれかの行動を選択し、表 1 の利得を得る。表中のカッコ内の値は、左側が行動主体 1 の得る利得、右側が行動主体 2 の得る利得である。

表 1 利得表  $R_i$

		行動主体 2	
		協力	非協力
行動主体 1	協力	$\left(3\frac{i}{N_r}, 3\frac{i}{N_r}\right)$	$\left(0, 5\frac{i}{N_r}\right)$
	非協力	$\left(5\frac{i}{N_r}, 0\right)$	$\left(\frac{i}{N_r}, \frac{i}{N_r}\right)$

手順 2. 総数  $N_a$  の行動主体を用意する。各行動主体は、 $N_m$  回前までに直面した社会的ジレンマ問題の利得表の番号 ( $1, \dots, N_r$ ) と、その際に自分および相手が選択した行動を記憶する。ただし、どの相手であったのかは記憶しない。初期状態では、ランダムに発生させた利得表の番号と、協力と非協力を半々の確率で発生させた行動を記憶する。

手順 3.  $N_a$  の行動主体の中から、2つの行動主体をランダムに選択する。また、 $N_r$  の利得表から 1つをランダムに選択する。

手順 4. 選択された各行動主体は、より高い利得を獲得するため、記憶している過去の行動のそれぞれを確率  $P_a$  で変更した行動列を  $N_t$  個作成する。そして各行動列から得られる総利得を計算する。

手順 5. 作成した  $N_t$  個の行動列から、総利得が最大のものを選択する。そして記憶する利得表ごとに協力、非協力が選択された回数を数え、非協力の回数を協力と非協力の回数の合計で割り、行動確率を計算する。なお、過去に 1度も行動選択していない利得表での行動は存在しないため、行動確率は計算しない。

手順 6. 各行動主体が計算した行動確率を使い、手順 3 で選択された利得表での行動を確率的に選択する。この利得表における行動確率を手順 5 で計算しなかった場合は、協力か非協力を半々の確率で選択する。

手順 7. 手順 3 で選択された行動主体が、双方の選択した行動と利得表の番号を記憶する。そして、最も古い

記憶内容を削除する。

手順 8. 手順 3～7 を 1 回のシミュレーションの過程として、 $L \leq N_l$  ( $L$  はシミュレーションの実行回数、 $N_l$  は設定するくり返しの総数) の間くり返す。

行動主体が記憶する内容の例を表 2 に示す。表 2 では  $N_m=3$ 、 $N_r=10$  としている。表 2 で 1 回前となっている列は、この行動主体と相手 (どの相手であったのかは記憶していない) がシミュレーションを 2 回実行した時点 ( $L=2$ ) で選択され、利得表番号 7 を提示されて、行動として自分が協力、相手が非協力を選択したことを示している。獲得した利得は、1 回前は表 1 の利得表  $R_7$  であることから 0、2 回前は利得表  $R_3$  のため  $3 \cdot 3/10=10.9$ 、3 回前は利得表  $R_5$  のため  $5 \cdot 5/10=2.5$  で、総利得は  $0+0.9+2.5=3.4$  となる。なお、記憶するのは手順 3 で自分が選択されたときのみで、選択されなかった場合の結果は記憶しない。

手順 4 で作成する行動列について述べる。もし各行動主体が記憶する  $N_m$  回前までの結果をもとに、自分および相手が取りうるすべての行動の組み合わせを評価するならば、その総数は  $2^{2N_m}$  となる。総数は、 $N_m=10$  で 1048576、 $N_m=100$  ならば  $1.6 \times 10^{60}$  ときわめて膨大な値となり、行動選択のたびにすべての総利得を計算することは困難である。そのため、本研究では行動主体が限定合理的であるとして、一定の回数だけ行動列を作成する。行動主体は、1 回前から  $N_m$  回前まで過去の自分および相手のそれぞれの行動を、確率  $P_a$  で協力なら非協力に、非協力ならば協力に変更する。表 2 において、 $P_a=0.1$ 、つまり 1 割の確率で行動を変更するとして変更手順の例を説明する。最初に、0 から 1 の範囲でランダムな小数を発生させる。その値が 0.1 以下であるならば、1 回前の自分の行動を協力から非協力に変更する。2 回前の自分の行動でも同様にランダムな小数を発生させ、自分の行動を変更するかを決定する。この手順を、記憶する自分および相手のすべての行動にたいしておこなう。その後、作成した新しい行動列を用いて自分の獲得できる総利得を計算する。表 3 では、1 回前の相手の行動と 2 回前の自分の行動が変更され、総利得は 6.1 となっている。同様にして  $N_l$  個の行動列を作成する。

表 2 行動主体が記憶する内容の例  
( $N_m=3, N_r=10$  の例)

	1 回前	2 回前	3 回前
$L$	2	5	9
利得表の番号	7	3	5
自分の行動	協力	協力	非協力
相手の行動	非協力	協力	協力
自分の利得	0	0.9	2.5
総 利 得	3.4		

表 3 行動主体が作成する行動列の例  
(塗りつぶし箇所の行動を変更)

	1 回前	2 回前	3 回前
$L$	2	5	9
利得表の番号	7	3	5
自分の行動	協力	非協力	非協力
相手の行動	協力	協力	協力
自分の利得	2.1	1.5	2.5
総 利 得	6.1		

手順 5 では、手順 4 で利得が最大となった行動列の中から、手順 3 で選択された利得表での行動を取り出し、行動確率を計算する。例として、記憶している利得表番号 2 の数が 10 存在し、そのうち非協力となっている行動数が 3 とする。この場合、行動確率は  $3/10=0.3$  と計算される。この値が 0 に近いほど、協力を選擇する確率が高くなる。これをすべての利得表ごとに計算する。

手順 6 の行動選択時には、0 から 1 の範囲でランダムな小数を発生させ、それが手順 5 で計算した行動確率を上回る場合には協力、そうでなければ非協力と確率的に選択する。なお、行動主体の記憶数が限定されているため、利得表の番号と行動結果が記憶する結果に含まれない場合がある。この場合には行動確率を計算できないため、協力が非協力を半々の確率で選択する。

### 3.2 シミュレーションのパラメータ

シミュレーションのパラメータ設定を以下に示す。

表 4 シミュレーションのパラメータ

パラメータ	値
利得表の数 $N_r$	10, 100
行動主体の数 $N_a$	10
過去の記憶数 $N_m$	100
行動の変更確率 $P_a$	0.1
行動列の作成数 $N_t$	10000
シミュレーション回数 $N_i$	30000

利得表の数  $N_r$  は、行動主体が直面する問題の複雑さが行動決定にどのような結果をおよぼすのかを明らかにするため、10と100の2通りを設定した。

行動主体の数が多いと、各行動主体の選択される回数が減少し、行動確率の更新頻度が下がり、シミュレーション回数を増やす必要がある。計算時間の短縮のため、行動主体の数は  $N_a=10$  と小さい値に設定した。

過去の記憶数は  $N_m=100$  とした。脳の記憶システムには短期記憶と長期記憶があり、長期記憶の内容を一時的に活性化して短期記憶に保持している<sup>[9]</sup>。長期記憶の大きさがどの程度であるのかは不明だが、短期記憶の容量制限がたかだか7であるとされることから、それと比較して大きい値を設定した。記憶数を限定しているため、利得表の数が  $N_r=10$  の場合と比較して、 $N_r=100$  の場合には限定合理性が強くなる。

個人が完全合理的に行動する場合、過去の自分および相手の協力、非協力の行動の組み合わせ総数は  $1.6 \times 10^{60}$  と膨大な値となる。本研究は行動主体が限定合理性であることを想定しているため、行動の作成数は、組み合わせ総数と比較して現実的な時間で計算できる  $N_t=10000$  とした。

行動主体が過去の行動を変更する確率  $P_a$  は、直面する問題にたいする個人の選好をあらわすといえる。もちろん、問題の選好が高い場合には積極的に試行錯誤するよう、選好の違いを考慮して行動主体ごと、利得表ごとに異なる  $P_a$  を設定することもできる。しかし、ここではすべての行動主体が同じ問題にたいしては同じ程度の選好をもつと想定して、 $P_a=0.1$  に統一した。記憶する自分および相手の行動の総数は  $100 \times 2 = 200$  であり、1回の行動列の作成時に20程度の行動が変更される。

### 3.3 シミュレーション結果

コンピュータシミュレーションを実行した結果を示す。プログラムはJava 8で作成し、CPUがIntel Core i7 2.5GHz、メモリ 16GByteのMacBook Proで実行した。シミュレーション1回あたりの実行時間は、約1時間15分である。シミュレーションは、利得表の数  $N_r=10$  および  $N_r=100$  のそれぞれで10回実行した。シミュレーションでは、行動主体が行動決定の際にランダムな数値を使用している。ランダムな数値はシードとよばれる値を基準にして計算されるため、同じシードを使用すれば同じ結果を得られる。本研究では得られる結果の傾向に類似性があるかを確認するため、シミュレーションごとにシードをランダムに変更して実行した。そのため、結果はすべて異なるものとなったが、得られた結果の傾向はいずれも以下に述べる傾向をもつことが確認できた。

利得表の数を  $N_r=10$  とした場合のシミュレーション結果を示す。図1は、表1の  $i/N_i$  が最小、中程度、最大となる利得表  $R_1, R_5, R_{10}$  で、各時刻において全行動主体が持つ行動確率の平均値を計算し、それをプロットしたものである。図では非協力を選択する確率を表しているため、値が1に近いほど行動主体が非協力を選択する確率が高くなる。また、表5は各シミュレーションのそれぞれで得たすべての行動確率の平均値を計算し、その中で最小、最大となった値と、計算した平均値の平均および標準偏差を示している。

図1の左側と右側は、それぞれ表5で非協力確率が最小、最大となった結果である。利得表の番号 $i$ が大きく、 $i/N_i$ が大きい結果ほど非協力の割合が増えることがわかる。また、行動確率はどの利得表の結果でも協力、非協力のいずれかに収束するのではなく、時間とともに変化する結果となっている。表5の平均値をみると、非協力を選択する確率は0.5以上である。ただ、 $R_1, R_5, R_{10}$ のいずれにおいても行動確率は時間とともに変化しており、常に非協力が高い確率で選択されるわけではないことがわかる。

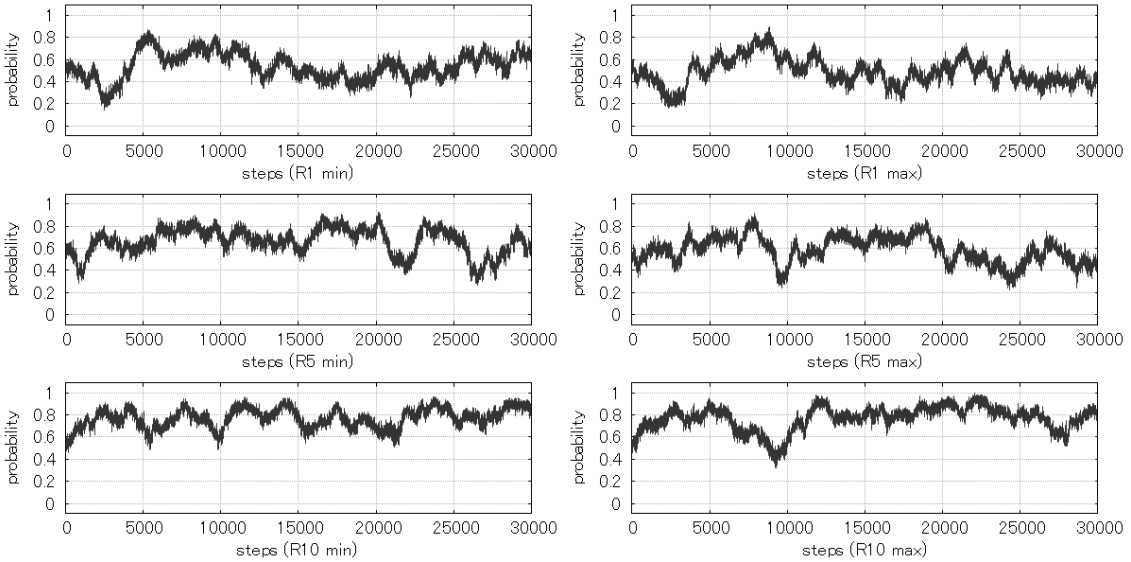


図1  $N_r=10$  の行動確率の平均値

表5  $N_r=10$  で非協力を選択する割合

	$R_1$	$R_5$	$R_{10}$
最 小	0.481	0.587	0.743
最 大	0.560	0.676	0.779
平 均	0.509	0.631	0.757
標準偏差	0.026	0.025	0.017

利得表の数が $N_r=100$ の場合の結果を示す。図2は、表1の $i/N_i$ が最小、中程度、最大となる利得表 $R_1, R_{50}, R_{100}$ で、各時刻に全行動主体が持つ行動確率の平均値を計算してプロットしたものである。表6は、表5と同様にして計算した全行動確率の平均値のうち最小、最大の結果と、10回のシミュレーションで得られた平均値の平均、標準偏差を示している。図2の左側が表6で最小となった結果、右側が最大となった結果である。図1と同様に行動確率は時間とともに変化しているが、利得表 $R_1$ の18000ステップ付近や $R_{50}$ の17000ステップ付近で非協力の確率が0となり、一定の期間にわたって協力が選択される状態を実現できている点に違いがみられる。また、行動確率が0から1の範囲で大きく動いている。図2の左右の利得表 $R_1$ における結果を比べると、左では12000~20000回付近と25000回以降で確率が0.5を下回っているが、右では5000回以降のほとんどで0.5を上回っている。このようにシミュレーションごとに違いが生じた結果、表5と比較して表6では標準偏差の値が大きくなったといえる。

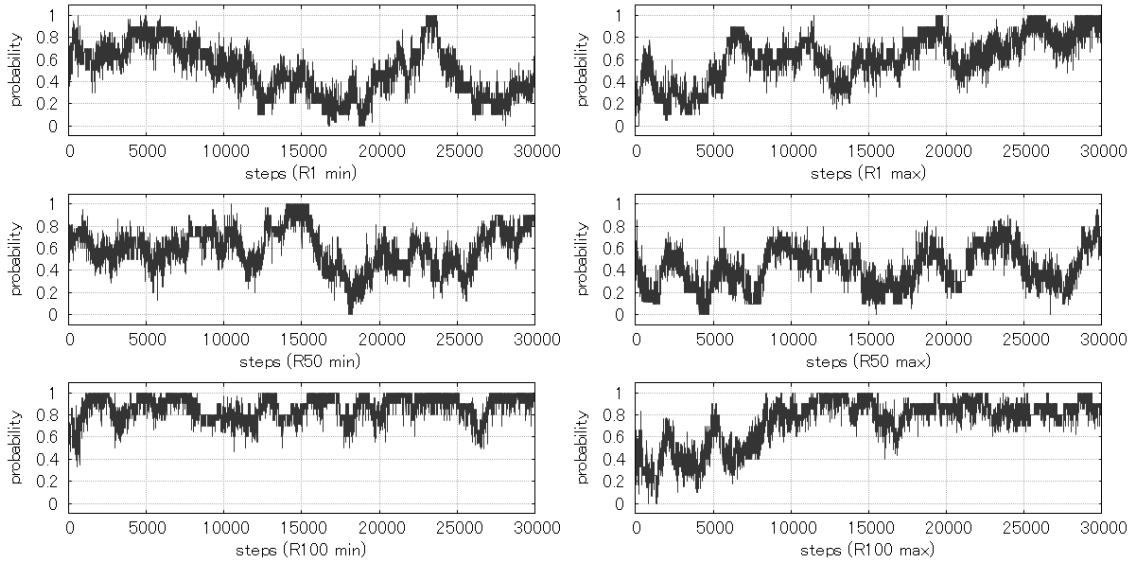


図2  $N_r=100$  の行動確率の平均値

表6  $N_r=100$  で非協力を選択する割合

	$R_1$	$R_{50}$	$R_{100}$
最 小	0.419	0.412	0.616
最 大	0.703	0.755	0.880
平 均	0.526	0.606	0.725
標準偏差	0.104	0.094	0.066

行動主体の行動確率と選択した行動を示す。図3、図4は、それぞれ利得表の数  $N_r=10, 100$  において、1つの行動主体に着目し、その行動主体がもつ行動確率の時間変化を図示したものである。図中の×印は、この行動主体が3.1節の手順3で選ばれた際に選択した行動である。0が協力、1が非協力を示している。図中の左側と右側は、それぞれ表5、表6で最小、最大となった結果である。また表7と表8に、それぞれ図3、図4において協力、非協力が選択された回数と割合をまとめた。

図3をみると、 $N_r=10$  では行動確率が特定の値に収束することなく、0から1の範囲で連続的にゆるやかに変化している状況が見てとれる。それにたいして  $N_r=100$  の結果である図4では、行動確率が多くの場合0または1のいずれかの値となり、ステップ状に急激に変化していることがわかる。表7と表8から、ともに利得表  $R_1$  では協力を選択した割合が高くなっているが、利得表の番号  $i$  が大きくなるにつれて、非協力が選択される割合が高くなっていくことがわかる。

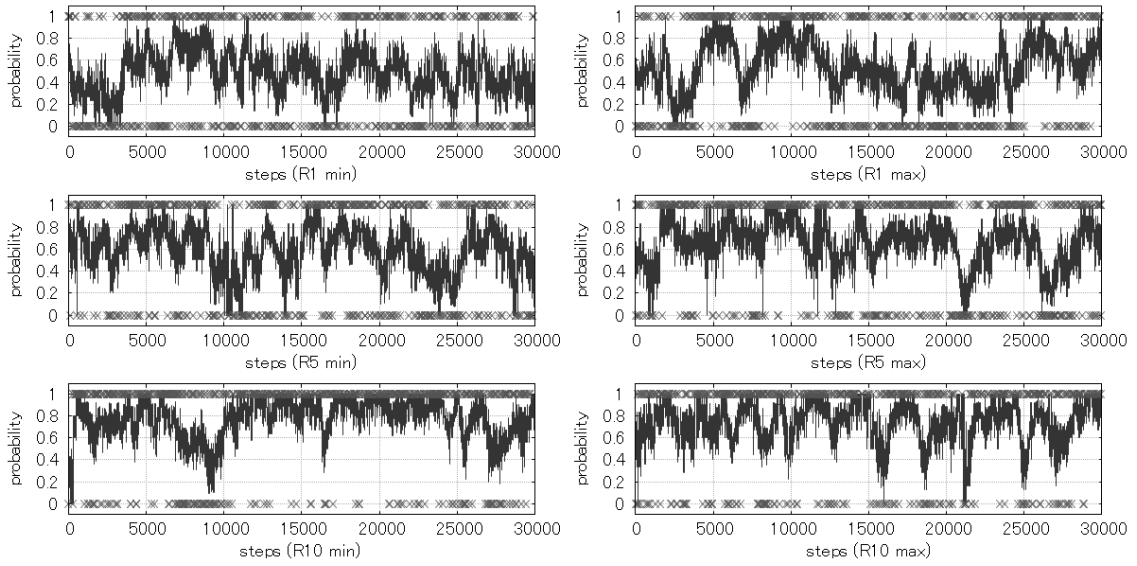


図3  $N_r=10$  における行動主体の行動確率

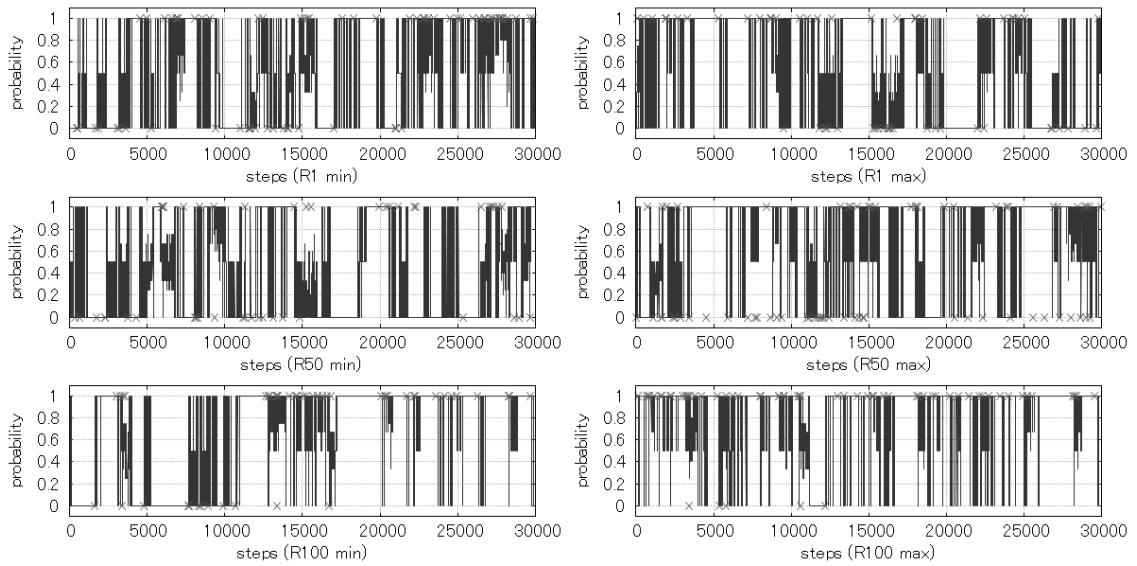


図4  $N_r=100$  における1つの行動主体の行動確率



表7  $N_r=10$  における行動回数  
(値は行動選択回数で、カッコ内はその割合)

		最 小	最 大
$R_1$	協 力	285 (0.484)	301 (0.470)
	非協力	304 (0.516)	340 (0.530)
	合 計	589	641
$R_5$	協 力	221 (0.386)	213 (0.347)
	非協力	338 (0.614)	400 (0.653)
	合 計	573	613
$R_{50}$	協 力	147 (0.229)	147 (0.248)
	非協力	495 (0.771)	445 (0.752)
	合 計	642	592

表8  $N_r=100$  における行動回数  
(値は行動選択回数で、カッコ内はその割合)

		最 小	最 大
$R_1$	協 力	34 (0.596)	24 (0.444)
	非協力	23 (0.404)	30 (0.556)
	合 計	57	54
$R_{50}$	協 力	28 (0.519)	16 (0.276)
	非協力	26 (0.481)	42 (0.724)
	合 計	54	58
$R_{100}$	協 力	14 (0.304)	5 (0.079)
	非協力	32 (0.696)	58 (0.921)
	合 計	46	63

### 3.4 考察

図1、図2では、利得表の数が  $N_r=10,100$  のいずれの場合でも、行動確率は時間とともに変化している。また、得られる利得の大きい結果ほど非協力が高い割合で推移している。このような行動の時間変化は、行動主体が3.1節の手順6で記憶する、過去の行動を確率的に変更して作成した新しい行動列から、総利得が最大のものを選択して行動決定に使用することから生じる。どの利得表でも、獲得できる利得は表1の  $i/N_r$  により決定される。この値が大きい場合、自分が協力を選択しても相手が非協力を選択して利得を得られなければ、総利得の減少分として大きく影響する。そのため、 $i$  が大きければ確実に利得を得られる非協力を選択することが合理的となる。逆に  $i$  が小さいならば、作成した行動列で自分が協力、相手が非協力を選択して利得が得られなくとも、行動列が保持する他の利得表での行動選択で獲得できる利得により、減少分を補うことが容易となる。そのため、利得が小さく設定されている利得表での行動決定ほど、行動主体からは協力行動を引き出しやすくなる。あまり生じる状況ではないが、 $i$  が大きい利得表でも、行動列に同じ利得表が存在して、その行動の多くが協力となる結果のときに総利得が最大となれば、行動選択時に協力が選択される。それが、 $i$  が大きい利得表である  $R_{10}$  や  $R_{100}$  でも協力が選択されている理由である。

なお、被験者実験による結果<sup>[3]</sup>では、協力行動を実行した割合である協力率が0.556と示されている。表5、表6の利得表  $R_1$  における値から協力率を計算すると0.519, 0.581、表7と表8では0.48, 0.596であることから、利得の小さい状況での結果は被験者実験に近いといえる。

図1、図2では、全行動主体がもつ行動確率の平均値は連続的に変化しており、急激な変化はない。しかし1つの行動主体に着目した行動確率をみると、 $N_r=10$ の結果である図3では連続的に変化しているのに対し、 $N_r=100$ の結果である図4では0か1のいずれかとなっている状況が多い。 $N_r=10$ の場合、行動の記憶数 $N_m=100$ にたいして利得表の数が少ない。そのため、すべての利得表での行動結果が記憶されている可能性が高く、それを利用して行動確率を毎回計算できる。そのため、行動確率は極端に変化せず、図1の平均値の変化もゆるやかとなっている。しかし $N_r=100$ の場合、記憶にすべての利得表での結果が含まれない可能性が高く、少ない経験をもとにした行動選択となることが生じる。たとえば、1回分の結果しか記憶していない利得表では、必ず協力か非協力が選択される。結果として、図4では行動確率が極端になっている。

全行動主体がもつ行動確率の平均値を示した表5と表6を比較すると、両者で大きな違いはない。しかし、 $N_r=10$ では図3のように行動確率の変化がゆるやかなため、全行動主体の平均値を示した図1では、平均値から大きく外れる状況がすくない。それに対して $N_r=100$ の場合、図4のように行動が協力か非協力と極端になるため、全行動主体の平均値である図2では0と1の間でひんばんに変化している。これらの結果から、利得表の数 $N_r$ の違いが行動決定に影響することがわかる。

以上の考察から、限定合理性をもつ行動主体は影響の小さい、得られる利得が小さい状況では協力を選択する割合が高くなるが、得られる利得が大きく、選択をあやまると大きな影響が生じる状況では非協力を選択しがちであるといえる。また、社会的ジレンマ問題の総数が少なく多様性の低い社会では、協力を選択する割合は社会全体では一定程度に保たれ大きく変化しないものの、問題の総数が多く多様性の高い社会では、個人は協力か非協力のいずれかを極端に選択し、社会全体で協力する割合が激しく変動する状況になると結論づけられる。

#### 4. おわりに

本研究では、社会的ジレンマ問題の一つである1回限りの囚人のジレンマ問題で、限定合理的な行動主体から一定程度の協力を引き出せることを目的としてコンピュータシミュレーションを行った。シミュレーションでは、実社会にさまざまな囚人のジレンマ問題が存在することを考慮し、ジレンマ的な状況は同一だが利得構造が異なる複数の利得表からなる問題をモデル化した。また、複数存在する行動主体が過去の行動および利得表を記憶し、それを試行錯誤的に変更して得た行動の組み合わせの中で総利得が最大となるものを利用して行動選択する、限定合理性をもつ行動主体をモデル構築した。

シミュレーション結果から、行動選択時に得られる利得の大小が協力、非協力を決定する割合に影響をあたえることを示した。そして、複数ある利得表の数の違いから生じる社会の多様性が、協力、非協力を決定する行動確率に影響することを明らかにした。

#### 参考文献

- [1] 亀田達也, 村田光二『複雑さに挑む社会心理学 適応エージェントとしての人間』有斐閣アルマ, 2000
- [2] 今福啓「プログラム自動生成手法によるジレンマ的問題における社会的な公平性の創発についての分析」情報学研究(獨協大学情報学研究所), Vol.2, pp.5-18, 2013
- [3] 渡部幹, 寺井滋, 林直保子, 山岸俊男「互恵性の期待にもとづく1回限りの囚人のジレンマにおける協力行動」実験社会心理学研究, 第36巻, 第2号, pp.183-196, 1996
- [4] 川越敏司『実験経済学』東京大学出版会, 2007
- [5] 中山幹夫, 武藤滋夫, 船木由喜彦『ゲーム理論で解く』有斐閣ブックス, 2000
- [6] 藤井聡「TDMと社会的ジレンマ: 交通問題解消における公共心の役割」土木学会論文集, No.667, IV-50, pp.41-58, 2001

- [7] 伊藤正人『行動と学習の心理学 日常生活を理解する』昭和堂, 2005
- [8] 森山甲一「囚人のジレンマゲームにおけるQ学習による協調の維持」コンピュータソフトウェア, Vol.25, No.4, pp.145-153, 2008
- [9] 岩崎祥一『脳の情報処理 - 選択から見た行動制御 -』サイエンス社, 2008