

ネットワーク型囚人のジレンマにおける信頼度に基づく 戦略のシミュレーション

井 寄 幸 平

(受付 2009年6月1日)

1. は じ め に

本稿ではネットワーク型囚人のジレンマ [4, 5, 7, 8] のモデルを用いて「信頼度」に基づく意思決定シミュレーションをおこない、行動選択における種々の戦略の特性と有効性を分析する。

ネットワーク型囚人のジレンマは、従来特定の個人又は集団との継続的かつ強制的な対戦関係を仮定してきた囚人のジレンマモデルに対し、対戦相手の自発的選択による関係の構築と解消を可能にしたものである。ネットワーク型囚人のジレンマでは、各プレイヤーは集団の中から自由に自分の希望する対戦相手を指名することができ、プレイヤー同士の指名に基づいて実際の対戦をおこなう相手（ペア）を決定する。実社会で観察される継続的關係、例えば消費者間の情報のやり取りや企業間の取引關係の多くは当事者間の合意に基づいて形成されたものであり、必ずしも強制力を持たず、むしろ互いが自由に選択することが可能である。ネットワーク型囚人のジレンマにおける継続的な協調關係の発生や、各プレイヤーが使用する戦略とその有効性を分析することは、実社会における消費者や企業の行動分析に対しても有益な示唆を与えることができると考えられる。

本稿では、コンピュータシミュレーションを用いて戦略の有効性を分析する。囚人のジレンマゲームにおける戦略のシミュレーションには、Axelrodによる著名な戦略トーナメント [1, 2] や、神らによるネットワーク型囚人のジレンマの戦略シミュレーション [5, 8] などがある。これらの多くは事前に用意された固定的戦略を用いるものであるが、ゲーム中の細かな戦略の変化や、わずかな戦略パラメータの相違による結果の差異には言及していない。そこで本稿では、ゲームにおけるプレイヤーの行動分析のための手法として、「信頼度」に基づく意思決定シミュレーションを用いる。「信頼度」は、相手の行動に対する主観的な期待値として定義され、相手の事前情報がない場合の一般的信頼度を初期値とし、対戦を重ねるごとに逐次更新される。信頼度の計算に用いるパラメータを変化させることによって、通常用いられる多くの戦略を再現できるほか、中間的な特性を持つ微妙な戦略の変化を分析することが可能となる。以上より、本稿ではネットワーク型囚人のジレンマにおける戦略について

てコンピュータシミュレーションを行い、それらの特性と有効性を分析することを目的とする。

2. ネットワーク型囚人のジレンマのモデル

山岸らが提案したネットワーク型囚人のジレンマモデルの特徴は以下のように表される。

- 各プレイヤーは集団内から特定の相手に対戦候補として指名する。
- ペア形成規則に基づいて対戦ペアを決定し、原則としてゲームは1対1でおこなう。

ここで集団の人数やペア形成規則については限定されておらず、状況に応じて様々な数、方式を選択しうる。そこで本研究では以下のようなモデルを定義し、シミュレーションの構築をおこなう。

- (1) 4人1組の集団を形成し、1グループとする。
- (2) 各プレイヤーはグループ内から対戦を希望する相手を1名選択し指名する。
- (3) 各プレイヤーの指名に基づいて対戦ペアを決定する。対戦ペアは互いに指名しあったプレイヤー同士で形成するものとする(図1)。
- (4) ペアが成立したプレイヤー同士で1回の囚人のジレンマゲームの対戦をおこなう。プレイヤーは相手に対し「協力」か「非協力」を選択する。
- (5) ペアが成立しなかったプレイヤーはその回のゲームに参加しない。
- (6) 各プレイヤーの行動の組み合わせにしたがって、利得を決定する。

以上の手順を1ラウンドとし、同じ集団、同じ規則のもとでゲーム終了までおこなわれる複数回のラウンドを1セッションと定義する。

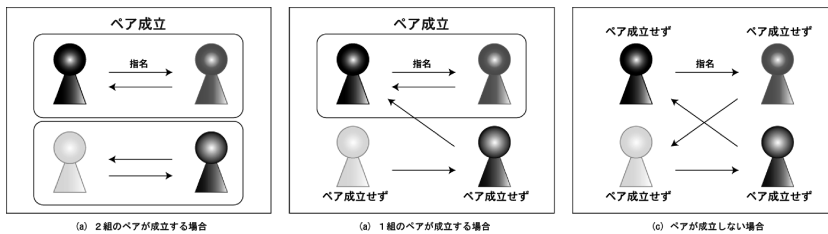


図1 ペア形成規則

また、各プレイヤーの利得は表1に示す利得表によって決定される。ゲームに参加できなかった(ペアを形成できなかった)プレイヤーの利得は0とする。よってゲームに参加した上で協力を選択し、相手が非協力を選んだ場合にはゲームに参加しなかった場合よりも低い利得を得ることになる。このことにより、ゲームに参加した場合には参加できなかった場合よりも

利得の高い場合、同じ場合、低い場合がそれぞれ含まれることになり、状況に応じて行動選択をおこなう幅が増加する。

表1 利得表

	協 力	非協力
協 力	3, 3	-1, 4
非協力	4, -1	0, 0
ゲームに参加できなかった場合：0		

以上のモデルを用いてコンピュータシミュレーションをおこなう。

3. コンピュータシミュレーション

3.1 信頼度を用いた意思決定モデル

本研究では、各プレイヤーは同じグループに所属する他のプレイヤーに対する「信頼度」を用いて行動を決定するモデルを用いる。繰り返し囚人のジレンマのシミュレーションを扱った研究では、事前に特定の戦略をプログラムしておく方法 [1, 2, 5, 8] や、強化学習などの適応的手法を用いるもの [3] などがある。しかし特定の戦略プログラムでは行動戦略の微妙な差異の表現ができず、また強化学習等の得られた利得をもとに戦略を絞り込んでいく手法では、利得の高い非協力行動に偏る傾向が見られ、いずれも現実に見られる多様な戦略表現にふさわしいとはいえない。さらにネットワーク型囚人のジレンマでは、協力・非協力以外に「相手の選択」という別の戦略も必要となるため、さらに多様な戦略を考慮しなければならない。しかしながら相手の選択と行動決定は1つの戦略として連動しているはずであり、何らかの基準に基づく包括的な戦略表現が必要である。

そこで本研究では、対戦相手の過去の行動履歴から相手の行動に対する主観的な期待値を「信頼度」としてモデル化し、その値をもとに相手の指名と行動選択の双方をおこなうアルゴリズムを使用する。ただしゲーム開始時点や対戦経験のない相手に対しては過去の行動履歴が存在しないため、初期値としての信頼度を設定する。信頼度は実際に対戦した相手の行動によって逐次更新されて行くものとし、1回の対戦で更新される信頼度の割合と信頼度の初期値によって多様な戦略を表現することができる。本アルゴリズムの特徴は以下のように表される。

- 信頼度という単一の基準値によって相手指名と行動選択をおこなうことが可能である。
- 信頼度の逐次更新により、行動戦略の細かな変化や繰り返しにとまらぬ収束を表現できる。

- 信頼度の初期値と更新割合の2つのパラメータを様々に変化させることにより、従来の戦略組込型シミュレーションで用いられてきた多くの戦略を再現できる。
- 相手の行動履歴を基準とするため、強化学習に見られる非協力への極端な偏向を回避することができる。

次に、信頼度が満たすべき条件について考える。信頼度は相手の行動履歴から、将来相手がどのような行動を取るかについての主観的な期待を数値化したものであり、一定の基準値内で変化することが必要である。また、相手が特定の行動を取り続ける場合、その行動に対する期待値は一定値に収束すると考えるのが自然である。さらに、過去の行動履歴については通常新しいものほど重視されることが多いと思われる。以上より、信頼度の性質について以下のような仮定を満たすものと定義する。

- (1) 信頼度の値は相手が協力した場合に増加し、非協力の場合は減少する。
- (2) 値は0から1の間とし、相手が常に協力の場合は1に、常に非協力の場合は0に収束する。
- (3) 過去の行動履歴には一定の割引率を乗じ、新しいものほど大きく評価する。
- (4) 指名しなかったプレイヤーに対する信頼度は変化しない。
- (5) 指名したにも関わらず相手から指名されなかった場合、信頼度は減少する。

以上の仮定を満たすものとして、ラウンド t におけるプレイヤー i のプレイヤー j に対する信頼度を以下のように定義する。

$$T_{ij}(t) = Y_{ij}(t) \left(\frac{1}{1+r_i} \right) \left\{ r_i X_{ji}(t) + T_{ij}(t-1) \right\} + \{1 - Y_{ij}(t)\} T_{ij}(t-1) \quad (1)$$

ただし信頼度の初期値、いわば見知らぬ相手に対する一般的信頼度は $0 \leq T_{ij}(0) \leq 1$ の範囲で所与とする。ここで r_i は割引率であり、 $0 \leq r_i \leq 1$ とする。 r_i の値が大きいくほど相手の直前の行動に対する信頼度の変化は大きくなり、 $r_i = 0$ であれば相手の行動に関わらず信頼度の値が変化しない（つまり一定の行動を取り続ける）ことになる。 $X_{ji}(t)$ はラウンド t におけるプレイヤー i のプレイヤー j に対する行動であり、以下のように定義する。

$$X_{ij}(t) = \begin{cases} 1 & (\text{協力}) \\ 0 & (\text{非協力または指名せず}) \end{cases} \quad (2)$$

また $Y_{ij}(t)$ はラウンド t においてプレイヤー i がプレイヤー j を指名したかどうかを示す変数であり、以下のように定義する。

$$Y_{ij}(t) = \begin{cases} 1 & (\text{指名}) \\ 0 & (\text{指名せず}) \end{cases} \quad (3)$$

各プレイヤーは式 (1) ~ (3) によって定義される信頼度をもとに相手の指名と行動選択をおこなう。指名戦略として、各プレイヤーは自分を除くグループ内で信頼度の最も高いプレイ

ヤを指名するものとする。複数のプレイヤーが条件に合致する場合にはその中からランダムに1名を選択し指名する。これは信頼度が相手の協力行動に対する主観的な期待値であり、かつ相手が自分を指名するかどうかを包含しているために、信頼度の低い相手を指名するメリットがないためである。協力するにしろ非協力を選ぶにしろ、相手が自分を指名した上で協力を選択する方が利得は高くなる。よって信頼度が最も高いプレイヤーを指名することが妥当と考えられる。

次にペアが成立したプレイヤーの行動選択のための戦略（行動戦略）を定義する。行動戦略は様々なものが考えられる。例えば信頼度の値に応じて協力・非協力の選択を確率的におこなうものや、信頼度の値に一定の閾値を設定して行動を変化させるものなどが考えられる。行動自体も信頼度の高い相手に協力する場合と、逆に信頼度の高い相手に対して非協力を選択する場合などが考えられる。そこで本研究では以下のような6種類の戦略を用意し、それぞれを混在させてシミュレーションをおこなうこととする。

● 戦略1： $P(t) = T_{ij}(t-1)$

● 戦略2： $P(t) = 1 - T_{ij}(t-1)$

● 戦略3： $P(t) = \begin{cases} 1 & (T_{ij}(t-1) > 0.5) \\ 0 & (T_{ij}(t-1) \leq 0.5) \end{cases}$

● 戦略4： $P(t) = \begin{cases} 1 & (T_{ij}(t-1) < 0.5) \\ 0 & (T_{ij}(t-1) \geq 0.5) \end{cases}$

● 戦略5： $P(t) = \begin{cases} 1 & (0.3 < T_{ij}(t-1) < 0.7) \\ 0 & (T_{ij}(t-1) \leq 0.3, 0.7 \leq T_{ij}(t-1)) \end{cases}$

● 戦略6：ランダムに協力または非協力を選択

ここで $P(t)$ はラウンド t においてプレイヤー i がプレイヤー j とペアが成立していた場合に協力を選択する確率として定義する。戦略1, 2はそれぞれ信頼度に応じて確率的に行動を選択するものであり、戦略1は信頼度の高い相手に協力的、戦略2は信頼度の高い相手には非協力的な行動をとる。戦略3, 4, 5は閾値を設定した上で信頼度に応じて行動を決定するもので、戦略3は信頼度の高い相手に協力的、戦略4はその逆、戦略5は信頼度が中間の相手に対して協力を選択する。戦略6は比較のための戦略であり、信頼度に関わらずランダムに協力か非協力を選択する。また、行動戦略はセッションを通して変化しないものとする。

以上の意思決定モデルと戦略を用い、コンピュータシミュレーションによりそれぞれの戦略の有効性について分析をおこなう。

3.2 シミュレーション手順

上述のモデルにより、コンピュータシミュレーションをおこなう。シミュレーションの手順は以下の通りである。

- (1) 4人のプレイヤーによる初期集団を生成し、 r_i 、 $T_{ij}(0)$ および $P(t)$ を各プレイヤーにランダムに与える。
- (2) 各プレイヤーは指名戦略に従ってグループ内から1名を指名する。
- (3) 各プレイヤーの指名に基づき、互いに指名しあっているプレイヤーをペアとする。ペアが成立しなかったプレイヤーの利得は0とする。
- (4) ペアが成立したプレイヤーは行動戦略に従って協力または非協力を決定し、表1に示す利得を得る。
- (5) 各プレイヤーは自分および指名した相手の行動に応じて信頼度の値を更新する。

以上の手順を1ラウンドとし、同様の手順を繰り返す。また、1セッションは50ラウンドでおこなわれるものとする。

3.3 シミュレーション結果

以下にシミュレーションの結果を示す。シミュレーション結果は500,000セッションの平均値をとって示している。

図2は r_i 、 $T_{ij}(0)$ と平均利得との関係、および形成されたペア数とそのうち相互協力を達成しているペア数を示す。全体的に r_i が大きいほど利得が高く、またわずかであるが $T_{ij}(0)$ が0.5の付近にピークが見て取れる。また、 $r_i=1$ 、 $T_{ij}(0)=0.1$ の付近にも高い値が見られる。形成されたペア数はラウンドが進むごとに徐々に増加しており、信頼度を更新していくことによって適切な対戦相手を指名できていることを示している。ただし相互協力を達成しているペア数の平均はわずかに増加するものの0.25程度にとどまっている。

次に各戦略が得た利得を個別に示す。図3は戦略1～6の平均利得と r_i 、 $T_{ij}(0)$ の関係を示したグラフである。いずれの戦略も r_i が低いときの平均利得は低くとどまっており、 r_i が大きい場合の方が平均利得も高くなっている。特に戦略2、4でこの傾向が顕著に現れている。戦略2、4は信頼度の高い相手に対して非協力的な戦略であり、この場合は $T_{ij}(0)$ が低い方が平均利得が高くなるという結果になっている。一方戦略1、3のように信頼度の高い相手に対して協力的な戦略では r_i の増加に対して戦略2、4ほど顕著な傾向はないが、 r_i が0～0.2程度のときに比べてそれ以上の場合には高い利得を得ていることが分かる。

また、いずれの戦略についても平均利得のカーブが $T_{ij}(0)$ の軸方向において行動戦略の関数 $P(t)$ の形に沿っていることが分かる。戦略1、2では $T_{ij}(0)$ が増加するにしたがってそれぞれ増加・減少しており、戦略3、4、5のように閾値を持つ行動戦略の関数ではその閾値

ネットワーク型囚人のジレンマにおける信頼度に基づく戦略のシミュレーション

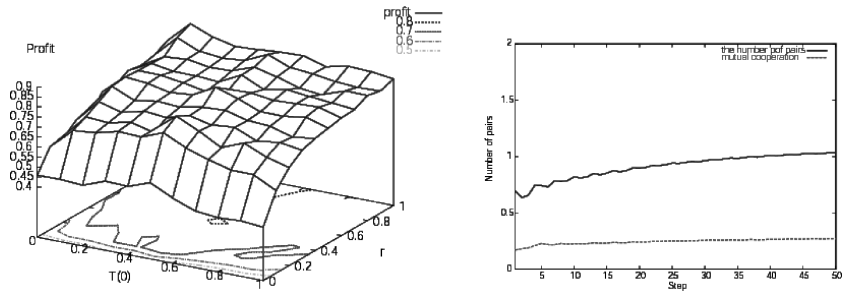


図2 平均利得とペア形成数

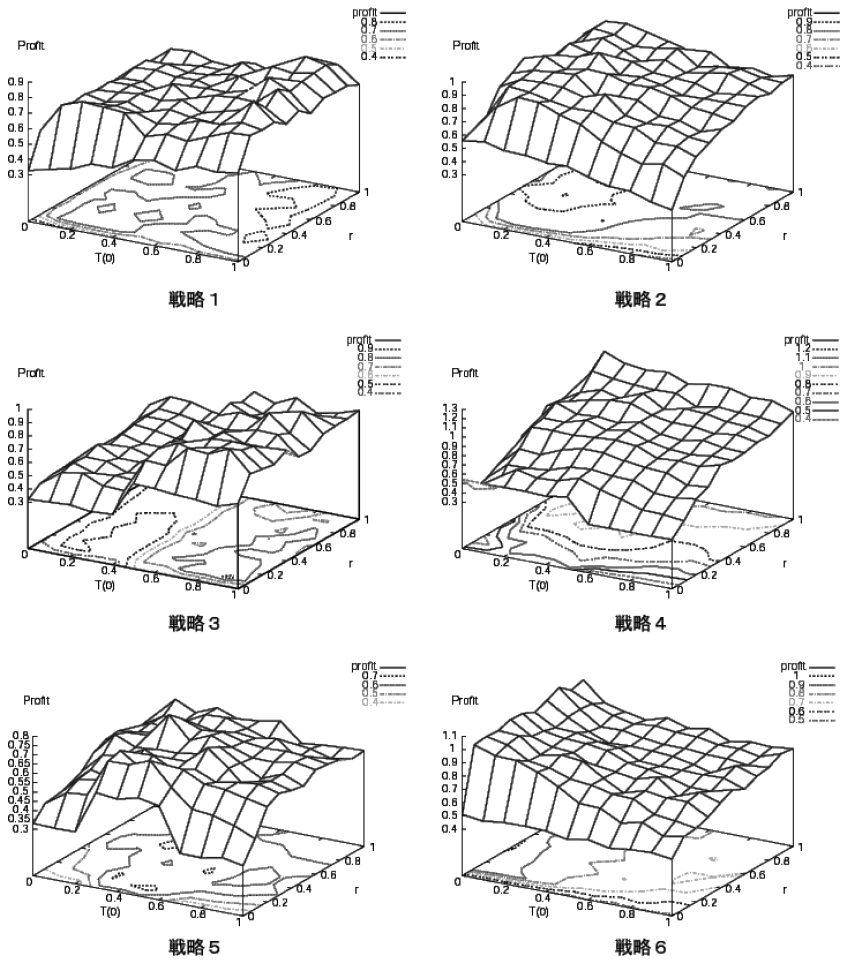


図3 各行動戦略の平均利得

を境にして平均利得が大きく変化している。これは信頼度の高い相手に対して協力的な場合にははっきりと現われており、特に戦略3では $T_{ij}(0)$ が0.5のときを境にして平均利得が階段状に大きく増加している。戦略4で r_i が大きい場合はそれほど明確な変化は現われていないが、 r_i が小さい場合にやはり $T_{ij}(0)$ が0.5のときを境にして平均利得が減少していることが分かる。戦略5は2つの閾値を持つ行動戦略であるが、やはり閾値部分を境に平均利得が増加・減少している。これらに対しランダムに行動を選択する戦略6では、 r_i が0の場合には $T_{ij}(0)$ の値に関わらずほぼ同じ平均利得となっている。しかしながら r_i が増加していくにつれ、 $T_{ij}(0)$ が小さい場合の平均利得が高くなり、 $T_{ij}(0)$ が増加するにしたがってゆるやかに平均利得が減少する様子が見られる。他の戦略では $T_{ij}(0)$ が小さいときの平均利得の傾向がほぼ他のパラメータの場合でも踏襲されるのに対し、戦略6では $T_{ij}(0)$ が小さい場合のみ別の傾向を示すこととなった。

また平均利得の大きさを見るとそれぞれ0.3~1.3程度の間分布していることが分かる。信頼度の高い相手に対して協力的な戦略1や戦略3の $T_{ij}(0)$ が0.5以上の部分については、 r_i が特に小さい場合を除いて比較的平坦なグラフとなっており、平均利得の最小値と最大値の間にそれほどの差は見られない。しかし戦略2、4のように信頼度の高い相手に対して非協力的な戦略では r_i が大きいほど、また r_i が大きく $T_{ij}(0)$ が小さいほど高い利得を得ていることが分かる。ただし $T_{ij}(0)$ が0の付近では、全体的に見れば高い利得ではあるが $T_{ij}(0)$ がわずかに高い場合よりも利得が下がっている。全体では r_i が1、 $T_{ij}(0)$ が0.1の付近で戦略2、4とも高い利得を得ており、特に戦略4で観察された平均利得約1.3は戦略全体で見た場合の最高得点となっている。

3.4 被験者実験との比較

3.3節において、戦略による平均利得の違いと r_i 、 $T_{ij}(0)$ に応じた傾向を知ることができた。本節ではこれまでに起こった被験者実験の結果 [6] をもとに、シミュレーションで得られた平均利得の傾向を照らし合わせて実際に被験者が用いている戦略がどのようなものか分析する。

図4は被験者実験における協力率と平均利得の関係を示したものである。ここで協力率はセッションを通してペアが成立し、ゲームの対戦をおこなった上で協力行動を選択した割合を指す。図より協力率の高い被験者ほど高い利得を得ていることが分かる。ここでの協力率はセッション全体を通しての協力の割合であるため、シミュレーションでの信頼度の初期値 $T_{ij}(0)$ とは完全には一致しないが、被験者実験では協力的な被験者は初回から協力的であることが多く、その結果相互協力を達成して高い平均利得を得ることが確認されている [6]。また、協力率が高いにも関わらず平均利得が特に低い被験者が存在するが、これはセッショ

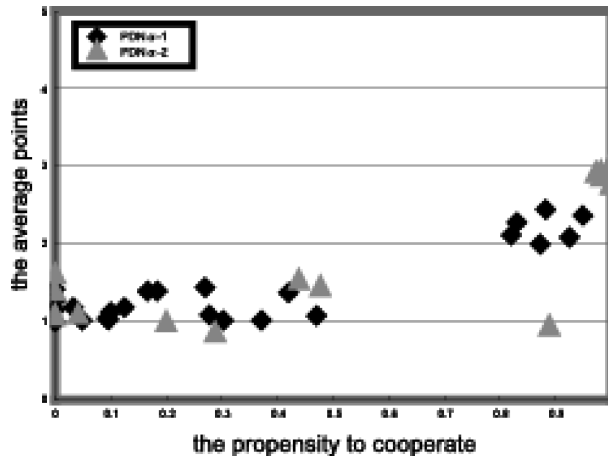


図4 被験者実験における協力率と平均利得

ン初期のラウンドでペアになった相手に対し協力を選択したが相手が非協力を選択し、その後も相互協力を達成できなかったためである。つまり事前情報のない相手に対して協力的ではあるものの、相手の行動との兼ね合いで高い利得が得られなかったことになる。これらの事柄から、被験者実験で観察された戦略は3.1節の戦略1のように信頼できる相手（協力を選択しそうな相手）に対して協力的であり、相互協力を通じて高い利得を得るものであると推測される。実際に戦略1と同様のアルゴリズムを用いておこなったシミュレーションでは図5のように被験者実験の特徴をほぼ再現できており、被験者が実際に用いている戦略として戦略1を採用することが妥当と考えられる。ただし被験者実験で用いた利得表は本研究での利得表と異なり（協力-非協力）の組み合わせで利得がそれぞれ（0, 5）、相互非協力で（1, 1）、ペアが成立しなかった場合の利得が1となっている。図5のシミュレーションは被験者実験と同じ利得表を用いたものである。

以上にもとづき、プレイヤーの行動戦略を戦略1に限定してシミュレーションをおこなう。ここでの利得表は表1のものを使用することとし、3.2節の手順と同様に50ラウンドの対戦を500,000セッションおこなうものとする。

図6は戦略1のみでシミュレーションをおこなった場合の平均利得と r_i 、 $T_{ij}(0)$ の関係、および形成されたペアと相互協力状態にあるペアの数を示したグラフである。

まず平均利得を見ると3.3節の結果と同様、平均利得は r_i 、 $T_{ij}(0)$ が大きいほど高い値を示しており、全体の形は行動戦略 $P(t)$ に沿うような形になっている。しかし r_i が0のときを除き、 $T_{ij}(0)$ が1のときには $T_{ij}(0)$ が1未満のときと比べて利得がやや低くなっている。また、この $T_{ij}(0)$ が大きいときの利得の減少幅は r_i の値が高いほど大きくなっていることが分かる。全体の利得の大きさは0.3~0.9程度の間分布しており、図3の戦略1が得た利得と平

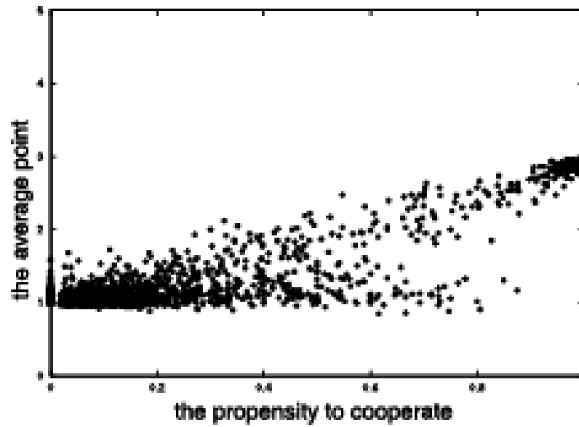


図5 戦略1を用いたシミュレーションにおける協力率と平均利得

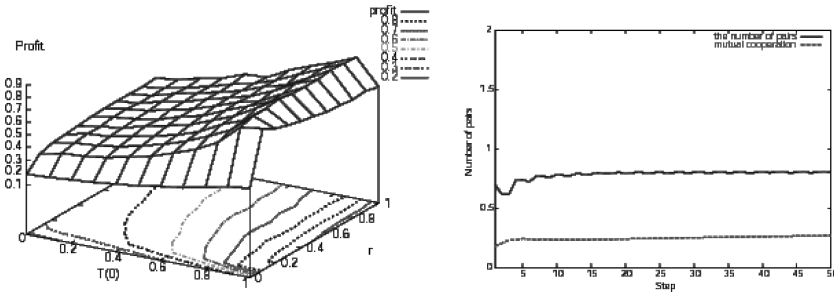


図6 戦略1の平均利得とペア形成数

均的にはほぼ同じ値となっている。

図6の右側は形成されたペア数および相互協力状態にあるペア数である。ラウンドが進むにつれて徐々に増加しているが10ラウンドを過ぎたあたりからほぼ横ばいとなっており、以降それほど変化は見られない。図2で見られるように複数の行動戦略を混在させた場合は形成されるペア数はラウンドを通じて増加傾向にあったのに対し、戦略1のみでシミュレーションをおこなった場合には早い段階で一定のペア数に収束することが分かる。

3.5 進化的手法によるパラメータの変化

次に戦略1を用いたシミュレーションにおいて、プレイヤーの意思決定を左右する r_i , $T_{ij}(0)$ の値を進化的手法を用いて変化させる。本研究では遺伝的アルゴリズムを用いて r_i , $T_{ij}(0)$ の値を進化させ、どのようなパラメータ値が生成されるかを分析する。

シミュレーションは以下の手順でおこなう。

- (1) 初期集団としてランダムな r_i , $T_{ij}(0)$ の値を持つプレイヤーを10,000人作成する。

- (2) 初期集団をランダムに4人ずつのプレイヤーからなるグループに分割する。
- (3) 各グループごとに50ラウンド（1セッション）のシミュレーションをおこなう。
- (4) 全プレイヤーのうち利得の上位10%のプレイヤーを複製し、利得の下位10%のプレイヤーを削除して入れ替える。
- (5) 以上の手順を1世代とし、得られたプレイヤー集団を新たな初期集団として(2)以降の手順を繰り返す。

ただし r_i , $T_{ij}(0)$ の値は0～1の間で0.1刻みでランダムに生成する。また遺伝的アルゴリズムの形式には様々なものがあるが、本研究では上記のように非常にシンプルなものを用いる。また、一定の割合でパラメータを変化させる突然変異は用いず、利得の大小によって増加・減少させる淘汰のみを使用する。

以上のシミュレーションを1,000世代おこない、 r_i , $T_{ij}(0)$ の変化とそのときに各プレイヤーが得た平均利得を分析する。

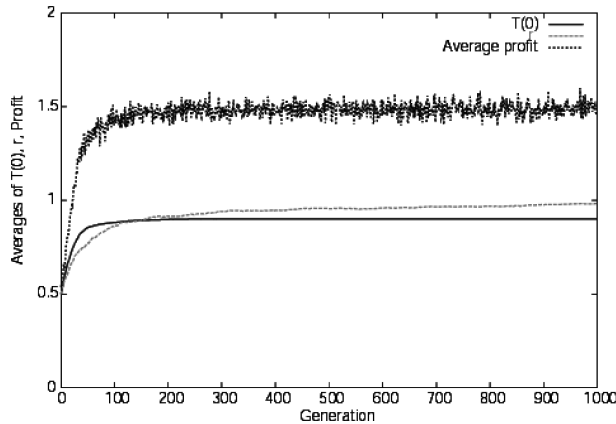


図7 戦略パラメータおよび平均利得の変化

図7は戦略パラメータである r_i , $T_{ij}(0)$ およびその世代における全プレイヤーの平均利得を示したものである。図より最終的に r_i は1, $T_{ij}(0)$ は0.9付近に収束していることが分かる。これは図6で見られた平均利得の最も高い部分であり、進化的手法により同じパラメータを持つプレイヤー同士の集団が形成されても高い利得を得られることが分かる。また、 $T_{ij}(0)$ は比較的早い世代（50世代付近）で収束が終わっているが、 $T_{ij}(0)$ は少しずつ増加を続けている。また平均利得は100世代付近ではほぼ収束しているが、その値は1.5付近まで上昇しており、図6で見られた利得の最大値よりもかなり高い値となっている。

図8は1セッションのうちに形成されたペア数と、そのうち相互協力を達成しているペア数の平均値を第1, 25, 1000世代について示したものである。世代が進むにつれてペア数・

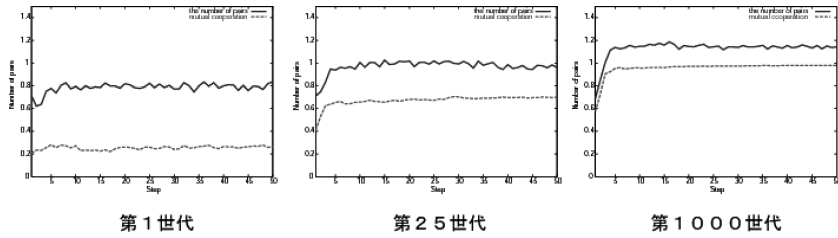


図8 形成されたペア数および相互協力ペア数の変化

相互協力ペア数とともに増加していることが分かるが、形成されるペア数の増加と比べて相互協力を達成するペア数の割合の増加が非常に大きくなっている。これは戦略パラメータの進化にともなって初期段階で協力をを選択するプレイヤーが増加し、結果的に継続的な相互協力状態を導くことができているためと思われる。しかし第1000世代においてもペア数、相互協力ペアともに最大値の2には到達しておらず、ペア数は1.2組程度、相互協力状態にあるペア数は1組程度にとどまっている。この状態では戦略パラメータの進化はほぼ収束しており、これ以降シミュレーションを継続しても大きく変わることはないと思われる。

4. 考 察

以下にシミュレーション結果についての考察を述べる。

まず6種類の行動戦略を用いたシミュレーションでは、全体として過去の相手の行動に対する割引率 r_i が大きいほど得られる平均利得が高いことが示された。つまり相手の行動に対して敏感に反応し、信頼度の値を大きく更新しているプレイヤーほど高い利得を得ていることになる。使用した行動戦略には協力的な相手に対して協力的なものや、逆に協力的な相手に対して非協力的なものなど様々なパターンがあるが、行動戦略の特性に関わらず、いずれの場合も相手の行動に応じて敏感に反応する方が利得につながる。これは相手の行動の変化に対して迅速に信頼度を更新することで、変化後の相手と関係が継続する場合にはより高い利得を得、対戦関係を解消して新たな相手を探す場合にはその判断を早め、さらに他の候補から適切な相手を選択することができるためと考えられる。 r_i が小さい場合には相手が行動を変化させても信頼度への反映が遅くなるため、例えば相互協力による継続的な協調関係を築けたかもしれない場合にも非協力を出し続けて相手に報復されたり、非協力的な相手との関係を解消できずに長期間にわたって搾取される状態に陥りやすくなると考えられる。特に行動を常に変化させない $r_i=0$ の場合には他のパラメータ値の場合と比べて平均利得が格段に低くなっており、相手の行動に対して信頼度を適切に更新することの重要性が分かる。

次に信頼度の初期値である $T_{ij}(0)$ と平均利得の関係を見ると、戦略6 (ランダム戦略) を

除いて行動戦略の関数 $P(t)$ の形に沿って変化している。協力的な戦略 1 や戦略 3 では $T_{ij}(0)$ が高いほど利得が大きく、非協力的な戦略 2 や戦略 4 では $T_{ij}(0)$ が低いほど利得が大きくなっている。これは協力的な戦略が相手との継続的な協調関係によって高い利得を得ているのに対し、非協力的な戦略はなるべく信頼度の高い（協力を選択しそうな）相手に対して非協力をを選択することで高い利得を得ているためと考えられる。ネットワーク型囚人のジレンマでは複数のプレイヤーによる集団内でゲームがおこなわれるため、非協力によって得点を稼いだ後に他のプレイヤーと新たな関係を構築し、さらに同様の手段で得点を重ねるという方法が可能になっている。特に非協力的な戦略では継続的な協調関係を築くことは困難であり、協力的な相手を見抜くことが自分の利得に大きく関わってくるため、相手の行動に対して迅速に反応しうる (r_i の大きい) プレイヤーほど高い得点を得る傾向が顕著に見られる。また平均利得の最大値を見ると非協力的な戦略の方が協力的な戦略よりも高くなっているが、これは表 1 の利得表が示すように相手の協力に対して非協力を選んだ場合の得点が大きく、また協力を出したプレイヤーの得点が負であることに起因するものと考えられる。しかし非協力を出すことは自分の信頼度を下げることにもつながるため、適切に相手の行動を見抜けない場合にはかえって利得を下げることになる。このため非協力的な戦略ではパラメータの違いに対して平均利得が大きく異なっていると思われる。これに対し継続的な協調関係を築きやすい協力的な戦略では、 $r_i=0$ の場合を除いて平均利得のグラフはなだらかで安定した形になっている。以上より、協力的な戦略は比較的多くのパラメータで安定した利得を得られるのに対し、非協力的な戦略は利得の最大値では上回るが、適切な行動を選択するためのパラメータは狭い範囲に限られると言える。またランダムに行動を選択する戦略 6 では $r_i=0$ の場合を除いて $T_{ij}(0)$ が小さい方が高い利得を得ているが、これは相手の協力度をあらかじめ低く見積もっておくことでより協力的な相手を探しやすくし、自分の行動がランダムに決定される場合でもなるべく高い得点を得られるようになるためと考えられる。

次に被験者実験結果との比較から戦略 1 のみでシミュレーションをおこない、さらに進化的手法を用いてパラメータの変化を分析した。戦略 1 のみでシミュレーションをおこなった場合、複数の行動戦略が混在する場合と基本的な傾向は同じであるが、 $T_{ij}(0)$ に応じて平均利得が増加する傾向がよりはっきりと現われた。これは協力的な戦略である戦略 1 においては、初期段階から協力的に振る舞い、相手と継続的な協調関係を築くことが利得につながるためである。ただし、 $T_{ij}(0)$ の値によっては初期に非協力をを選択するプレイヤーも存在するため、 $T_{ij}(0)$ の値を最大の 1 にすると相手の非協力的行動によって利得を失ってしまう場合がある。このため $T_{ij}(0)=1$ の場合には $T_{ij}(0)$ が 0.9 前後のときよりも平均利得が低くなっていると思われる。これは遺伝的アルゴリズムを用いたシミュレーションで得られた結果からも確かめられる。図 7 より r_i は 1、 $T_{ij}(0)$ は 0.9 付近に収束しており、様々なパラメータ値が混在する

場合に有効な戦略は、プレイヤー集団が同種のパラメータ値に収束しても有効性を保っている。むしろ図8よりパラメータの収束によってより多くの協調関係を築くことができ、相互協力を達成するペアおよび平均利得もより高くなっていくことが確認できる。

ただし1000世代を経た後でもすべてのプレイヤーが相互協力状態になるわけではなく、全体の半数程度にとどまっている。これは集団内から自分で相手を選択し、さらに相互の指名によってペアが形成されることから、ペア形成自体が比較的困難なためである。それでも遺伝的アルゴリズムによって最終的に得られた相互協力ペアの数は被験者実験 [6] で観察された数 (0.5) を大幅に上回っており、パラメータの進化によりより効率的に高い利得を得るための相手と行動の選択が行われていることを示している。

5. お わ り に

本稿ではネットワーク型四人のジレンマにおける戦略について、信頼度を用いて意思決定をおこなうアルゴリズムを用いてシミュレーションをおこない、様々な戦略の特徴と有効性を分析した。

シミュレーション結果から行動選択に使用する戦略に関わらず、相手の行動に対して迅速に反応し信頼度を更新しうるプレイヤーほど高い利得を得ることが示された。また、信頼度の初期値、すなわち対戦経験のない相手に対する一般的信頼度については行動選択の戦略の形に応じて有効性が変化することが確認された。協力的な戦略を使用する場合には一般的信頼度が大きいプレイヤーほど高い利得を獲得し、逆に非協力的な戦略を用いる場合には一般的信頼度が低い方が高い利得を得られた。

さらに被験者実験との比較から、実際に被験者が使用している戦略が相手の信頼度に応じて確率的に協力を選択するものであると推定し、対戦シミュレーションおよび遺伝的アルゴリズムによるパラメータ進化により有効な条件を示した。この場合、相手の行動に迅速に対応し、かつ一般的信頼度が高いプレイヤーほど高い利得を得ることが分かった。ただし一般的信頼度が高すぎる場合には相手に搾取されて利得を失う場合があるため、最大値よりもわずかに低い場合に最大の利得を得られることを示した。

今回のシミュレーションでは相手の選択に1種類のみ戦略を使用したため、今後は相手選択の戦略を拡張した分析をおこなっていく必要がある。

参 考 文 献

- [1] Axelrod, R.: *The Evolution of cooperation*, Basic Books, 1984.
- [2] Axelrod, R.: *The Complexity of cooperation*, Princeton University Press, 1987.
- [3] Erev, I. and Roth, A.E.: *Simple reinforcement learning models and reciprocation in the prisoner's dilemma game*, *The Bounded Rationality: The Adaptive Toolbox*, pp. 215–231, MIT Press, 2002.
- [4] Hayashi, N. and Yamagishi, T.: *Selective play: Choosing partners in an certain world*, *Personality and Social Psychology Review*, Vol.2, pp. 276–289, 1998.
- [5] Yamagishi, T., Hayashi, N. and Jin, N.: *Prisoner's dilemma network: Selection strategy versus action strategy*, *Social Dilemmas and Cooperation*, Springer Verlag, pp. 233–250, 1994.
- [6] 井寄幸平, 鳩野逸生, 小田宗兵衛, 上田完次, 「ネットワーク型囚人のジレンマにおける協調行動の創発に関する研究」, システム制御情報学会論文誌, 第16巻第9号, pp. 36–43, 2003.
- [7] 神 信人, 林直保子, 篠塚寛美, 「ネットワーク型囚人のジレンマの実験的研究: PD 関係におけるコミットメントの形成」, 実験社会学研究, 第33巻第1号, pp. 21–30, 1993.
- [8] 林直保子, 神 信人, 山岸俊男, 「ネットワーク型囚人のジレンマ: 戦略のシミュレーション」, 社会心理学研究, 第8巻第1号, pp. 33–43, 1993.