繰返しのない囚人のジレンマの解決と信頼感の役割¹

北海道大学 林直保子²

Emergence of cooperation in one-shot prisoner's dilemmas and the role of trust

Naoko Hayashi (Department of Behavioral Science, Faculty of Letters, Hokkaido University, Kita-10 Nishi-7, Kita-ku, Sapporo 060)

The main purpose of this study is to apply the "selective play paradigm" to explain how cooperation emerges in one-shot prisoner's dilemmas. A unique feature of the selective play paradigm is the option for not playing a PD game. For this purpose, a computer simulation of 100-actor groups was conducted. At the beginning of each replication, each simulated actor was randomly assigned to one of eleven levels of trust, which indicates the actor's estimate of the overall cooperation rate in the group. Each simulated actor, then, decided whether or not to interact with the previous partner based on the calculated expected gains from interacting with one of the other partners. Results of the simulation show that: (1) when substantial opportunity costs exist, having a high level of trust benefits the actor; (2) the above effect of trust depends on the actor's cooperativeness in PD games; (3) but does not depend on the overall cooperation rate in the group.

Key words: prisoner's dilemma, selective play, trust, social dilemma.

個人の合理的な利益追及行動が集団に不利益をもたらす、結局個人にとっても都合の悪い結果が起こることがある。これが社会的ジレンマである。個々の成員と集団全体の間の利益の葛藤関係を表す社会的ジレンマは、社会に実際に存在する状況のモデルとして広い分野で研究されてきた。そしてこの社会的ジレンマは、 N 人囚人のジレンマとも呼ばれるように、囚人のジレンマ (prisoner's dilemma: 以下PDとする) が3者以上の集団に存在する状況としてとらえられてきた。1対1の2者の間に利益の葛藤が存在するPDは、次のように定義できる。

①両者がそれぞれ“協力”“非協力”の選択を行う。

②いずれにとっても、相手の選択にかかわらず、“協力”よりは“非協力”の選択を行った結果として、より大きな利益を得ることができるが、双方が“非協力”の選択を行えば、双方が“協力”の選択をとった場合に比べて得られる利益が小さい。

このように、PDにおいては非協力の選択が協力の選択に対して優越している。従って、人々が合理的に行動すると仮定する限り、協力行動は達成されないはずである。しかしAxelrod (1984)に代表される反復2者PD研究は、合理的な2者の間にでも相互協力が達成されることを示してきた。応報戦略 (tit-for-tat: 以下TFTとする) 等の“戦略合理的な行動” (山岸, 1989)により、PDにおける相互協力の達成が可能だというのが、PD研究の主要な知見の一つである。

ここで、“戦略合理性によるPDの解決”とは何を意味しているのかということについて、少し考えてみたい。戦略合理性とは、自己の行動が他者に与える影響をも考慮に入れた形で自分の行動の結果を予測し、その予測に基づいて合理的に行動することである。山岸 (1993) は応報戦略の有効性に関して、応報戦略を採用することにより各試行において存在しているPDの利得構造が、超ゲームのレベルでは信頼ゲームの利得構造に変換されているのだと議論している。PD関係にある2者のうち、一方のプレイヤーが応報戦略を採用している場合には、もう一方のプレイヤーにとっての利得構造が、相手が協力している限り協力が得になる構造になり、一方的な搾取が不可能になるのである。つまり、応報戦略による反復PDの解決とは、長期的に見た場合の利得構造が変換されることにより協力行動が達成されていることを示しているものであり、

¹ 本研究は日本学術振興会特別研究員奨励費の補助を受けた。また、本研究は、原子力安全システム研究所/社会システム研究所主宰による信頼研究プロジェクトの一環として行われた6度にわたるワークショップでの議論に負うところが大きい。同研究所および三隅二不二所長、山田昭副所長、春名康宏主任研究員に感謝する。

² 本論文の作成にあたり、山岸俊男教授 (北海道大学) に御指導頂きました。

純粋なPDの利得構造に対する協力的行動の達成を説明するものではない。合理計算による協力的行動が引き出されるためには、このような利得構造の変換が不可欠であり、利得構造が変換されるのは同一相手との関係が無限回繰り返される場合のみである。この“無限回の繰返し”がない場合には、将来の利益はあてにならないものとなり、現在の利害のみが重要視されるようになるであろう。極端な場合、繰返し回数が2回であれば、第1試行で協力的であった相手に対して第2試行で協力を返しても、それによって第3試行以降に相手が協力的になることによって自分の利益が保証されることにはならず、従って第2試行では、第1試行での相手の選択にかかわらず、裏切った方が協力するよりも有利な結果をもたらすことになる。

本研究の第1の目的は、従来のPD研究とは全く異なる方向から、1回限りのPDにおいても協力的行動が達成され得ることを示すことにある。既に述べたように、PDの利得構造では非協力的選択が協力的選択に優越しており、協力的動機ないし愛他動機といった社会的動機を仮定しない限りそこでの協力的行動を説明することは一見不可能のように思われる。本研究で扱う“選択的プレイパラダイム”は、このように一見不可能に思われる1回限りのPDにおける協力的説明が可能であることを示すものである。

繰返しのないPDにおける協力の達成

選択的プレイパラダイム 1990年以降の研究の中に比較的頻繁に見られるようになった“選択的プレイパラダイム”(e.g., 林, 1993; 林・神・山岸, 1993; 神・林・篠塚, 1993; Orbell & Dawes, 1991, 1993)は、各人が特定の相手とのジレンマ関係、あるいはジレンマ集団に参加するかどうかを決定するものである。従来のPD研究が特定の2者PD関係を取り出す形で行われてきたのに対し、選択的プレイパラダイムによるPD研究は、より大きな集団の中に成立する複数のPD関係を扱っている。選択的プレイパラダイムでは、各成員が1回ごとにプレイするかどうかを決定するため、特定の他成員との間に継続的關係は存在しない。従って、上述の応報戦略による構造変換は不可能であり、各々のPDは基本的に1回限りのPDと同じである。

out-for-tatによるアプローチ 選択的プレイ状況を扱った研究はこれまでに、特定の相手との間に関係の継続性が保証されない状況において、協力的行動の達成が可能であることを示してきた。例えばSchuessler(1989)は、解消可能な2者PD状況を扱ったコンピュータ・シミュレーションを行い、2者間に永続的な関係が存在しなくても協調関係が達成できる可能性を示している。また林(1993)は、選択的プレイ状況における各種戦略の有効性をシミュレーションにより比較

し、パートナー選択レベルでの応報戦略とも言えるout-for-tat(以下OFTとする)戦略が最も有効であることを示している。OFTとは、相手が裏切らない限りその相手との関係を維持する協力的戦略である。相手がOFT戦略を採用している場合には、非協力的行動をとった次の回にはプレイを拒否されることになるため、一方的な搾取は不可能となる。

Schuessler(1989)と林(1993)の研究結果は基本的に同じことを意味している。それはつまり、非協力的行動をとることにより相手にプレイを拒否された場合には、取引コストがかかる。このコストの存在により、PDでの非協力が協力よりも得にならない構造が生み出されているということである。既に述べたように、選択的プレイは特定の相手との関係として見ると、基本的には1回限りのPDと同じである。長期的な意味での構造変換を目的とした戦略的行動には意味がなく、PDの利得構造に対する合理的選択はやはり非協力である。しかし集団の中で1回限りのPDを繰り返す選択的プレイパラダイムにおいては、次回以降のPDをプレイするための取引コストの存在が協力的行動を達成させる。つまり、ここでは関係の継続性ではなく、試行の継続が協力的行動達成の条件となっているのである。

第2のアプローチ——協力者を見分ける Frank, Gilovich, & Regan (1993)は、繰返しのないPDにおいて協力が達成される可能性について、上とは別の説明を試みている。このアプローチは、人々はわずかな情報から相手の協力傾向をある程度正確に推測することができるのだというものである。この場合の協力傾向とは、裏切ったことが相手にばれないような状況においても協力する傾向のことであり、善良さと言い換えてもよい。Frank et al.(1993)は、初対面の被験者間で30分の会話をさせることにより、その後の1回限りのPDにおいて、被験者は相手が協力するかどうかをかなり正確に予測できることを示している。Frank et al.(1993)はこの結果により、人々は相手とのインタラクションから得た情報をもとに、相手の協力傾向をある程度予測することができるため、協力的な人々は互いに見出し合い、栄えるのだと議論している。

繰返しのないPDへの第3のアプローチ 繰返しのないPDにおいて協力が達成されるもう一つの可能性は、Orbell & Dawes(1991, 1993)により示されている。以下にOrbell & Dawes(1991, 1993)の議論を紹介する。

Orbell & Dawes(1991, 1993)は相手が協力しなければ自己利益がマイナスになるようなPDの利得構造を仮定している。そこでは、集団の各成員は全他成員に対して、その相手とプレイするかどうかを決定する。つまりこの状況では、人々はプレイをする前の段階で、他の人々がどの程度協力的かを見積る必要があ

る。このとき、他の人々が協力するだろうと見積る人々は、他の人々が非協力的であると思う人々に比べて頻繁に参加すると考えられる。

社会的ジレンマに関するこれまでの研究で、協力的な人々は他の人々も協力的だと予測し、非協力的な人々は他の人々も非協力的だと予測する傾向があることが明らかにされている(e.g., Dawes, McTavish, & Shaklee, 1977; Yamagishi, 1986). Orbell & Dawes (1991, 1993)は、期待と行動の間のこの正の相関をもとに、選択的プレイ状況では協力的な人々で構成されるペアが成立する確率が高く、相互協力が達成されやすいのだと論じている。Orbell & Dawes (1993)は選択的プレイのPDを用いた実験において、協力者が非協力者よりも頻繁にプレイに参加していたことを示すことにより、彼らの議論を支持する証拠としている。

このOrbell & Dawes (1991, 1993)の議論には、一つの限界がある。Orbell & Dawes (1991, 1993)の議論では、協力者が非協力者よりも有利になるかどうかは、全体の協力率に依存している。彼らの議論で協力者が非協力者よりも有利になるのは、初期状態においてたまたまある程度の協力率が実現された場合のみであり、それ以外の場合には、協力者は搾取され衰退してしまう。

第4のアプローチ

本研究の第2の目的は、上記3種類のアプローチのうち、人々の協力者を見極める能力に依存しない二つのアプローチを統合することにより、Orbell & Dawes (1991, 1993)の議論で残されている問題を克服することにある。本研究で示す第4のアプローチは、以下に紹介するYamagishi & Yamagishi (1994)の信頼の理論に基づいている。

社会的な不確実性への対処法としての信頼
Yamagishi & Yamagishi (1994)は、相手の友好的行動に対する期待のうち相手の自己利益に根差した部分を“安心(assurance)”として定義し、“信頼(trust)”は、安心の部分を取り去った残りの部分であるとしている³。Yamagishi & Yamagishi (1994)は、このような客観的な行動の予測を超えた期待を信頼として定義し、状況の客観的認知からのずれとしての信頼は一種の認知的バイアスであると論じている。つまり信頼は、不確実な情報に基づいて、相手が裏切るか裏切らないかを予測する際に生じる認知的バイアスである。従ってそれは相手の意図を予測する必要のある場面でのみ存

³ 相手の友好的行動に対する期待のうち相手の自己利益に根差した部分(安心)とは、例えば自分の売る中古車に保証をつけたディーラーが欠陥のある車を売ることはないだろうと予測するといったように、信頼を裏切ると裏切った本人の自己利益が損なわれてしまうという知識に基づいて相手に期待する部分を意味する。

在し得るものである。相手の意図を予測する必要がある状況とは、相手に裏切りが得になるような状況で、かつ相手に裏切られるとこちらが損害をこうむる状況である(Yamagishi & Yamagishi (1994)の理論では、このような状況を社会的な不確実状況と呼ぶ)。

社会的な不確実性を前提とすると、個人が社会的関係をもつ唯一の手段は、相手を信頼することである。しかし、社会的な不確実性そのものを減らすことができれば、信頼の必要性は小さくなる。では社会的な不確実性そのものを減らす方法はあるだろうか。既に述べたように、反復2者PD研究は、応報戦略等の戦略的行動により両者にとって一方的な裏切りの誘因が小さくなることを示してきた。社会的な不確実性を客観的に減らそうとするときには、特定の相手と続けて付き合う関係を築けばよいのである。

他者を信頼するのはなぜか このように考えると、信頼と特定の相手とのコミットメントの形成は、社会的な不確実性に対する独立した対処法であることがわかる。そして、より単純で採用しやすいのはコミットメントの形成である⁴。しかし、コミットメントの形成は機会コストを生み出す。機会コストとはすなわち、ある相手とコミットした場合に、他の相手とのより実りある付き合いを逃してしまう場合のコストである。機会コストが増大するにつれ、コミットメント形成は不利になる。信頼の誘因は、機会コストが大きい状況において実際に不利なコミットメント関係から離脱する“推力”を提供する点にある。

信頼と選択的プレイ

選択的プレイパラダイムによる従来の研究は、複数の相手との間で同一の利得構造をもつPD関係を成立させる状況を扱っていた。これは、社会的な不確実性は存在するが機会コストは存在しない状況である。林 (1993)のシミュレーションでは、協力的な相手とコミットメントを形成するOFT戦略が最も有効であることが示されていた。しかし上述のYamagishi & Yamagishi (1994)の理論に基づいて考えると、選択的プレイ状況においても機会コストが存在すれば、信頼が意味をもってくるということになる。ここでの信頼の内容は、Orbell & Dawes (1991, 1993)の理論における“他者の協力への期待”と基本的に同じである。但し、Yamagishi & Yamagishi (1994)の理論では、信頼感の低い人々はコミットメントを形成するのだとされているため、Orbell & Dawes (1991, 1993)のような参加・不参加を選択するモデルではなく、関係の維持・

⁴ 社会的な不確実性が大きい状況では人々の間にコミットメントが形成されやすいことも実験により示されている(渡部・山岸・山岸・林・高橋, 1994; 山岸・山岸・林・高橋・渡部, 1994)。

方 法

コンピュータ・シミュレーション

このシミュレーションは、集団内の各個人が、ある行動決定規則に従って行動した場合にどれだけの利益を得ることができるかを調べるものである。ある行動決定規則とは、現在付き合っている相手との交換関係から得られる利益と現在の相手以外の成員との関係から得られる利益の期待値の比較を行い、同じか期待値の方が小さい場合には現在の相手との関係を維持するというものである。現在の相手以外の成員との関係から得られる利益の期待値の計算には、その成員のもつ信頼感が用いられる。ここで言う信頼感とは、自分以外の全他成員のうち、何パーセントが協力者であるかという見積りである。

シミュレーション状況 シミュレーションでは、100人集団のPDネットワーク状況を想定した。レプリケーションのはじめに、各成員の行動傾向が決定される。具体的には、協力者か非協力者のどちらかに割り当てられる。更に各成員は信頼の程度をランダムに割り振られた。信頼の程度は、0から100の範囲で10刻みの11段階とされた。この信頼と行動傾向の組合せがその成員の取る行動を決定する。信頼が0の成員は集団内の協力者の割合は0%であると予測し、100の成員は100%であると予測する。100試行を1レプリケーションとし、1条件につき300レプリケーションを行った。

シミュレーションの流れ 各レプリケーションの第1試行ではすべての成員の間でランダムにペアがつくられる。第2試行からは、各成員は前試行での自己利益と、前試行の相手以外の成員から得られる利益の期待値を比較して、期待値の方が大きい場合には前試行の相手との関係を断つ。そして前試行の相手との関係を断った成員と関係を断られた成員の間で新たにランダムにペアがつくられる。

条件 集団内の協力率に関する4条件(20%, 40%, 60%, 80%), 機会コストに関する2条件(機会コストなし・あり)の組合せによる8条件が行われた。

利得構造 PDの利得構造は、(協力, 協力)=1, (協力, 非協力)=-2, (非協力, 協力)=3, (非協力, 非協力)=-1とした(但し利得は括弧内の左側のプレイヤーの利得を示す)。

機会コスト 機会コストは、PDでの行動が同じ場合でも、誰との間にPD関係が成立するかによって得られる利益が異なるという形で導入された。具体的には、レプリケーションのはじめに、各成員が他成員それぞれに対してもつ“基準点”として0, 1, 2のいずれかがランダムに与えられた。従って、同一成員の基準点は一定ではなく、相手によって変化した。このよ

解消の選択を扱う必要がある。第4のアプローチでは、人々はある交換から得られる利益と別の交換から得られる利益の比較により行動を決定するとされるため、交換に伴う社会的不確実性のみが前提となり、利得の絶対量に関する制限は必要ない。また、関係の継続・解消を選択する状況では、OFT戦略を採用することができるため、高いレベルの信頼をもつ者が搾取される頻度が集団内の協力率により直接影響されることはない。

シミュレーションの目的

本研究の最終的な目的は、Yamagishi & Yamagishi (1994)の理論に基づく第4のアプローチにより、繰返しのないPDにおいて協力行動が達成される可能性を示すことにある。本研究ではこの目的のためにコンピュータ・シミュレーションを行った。このシミュレーションでは、協力行動が達成されるためには協力行動をとることが実際に有利になる必要があるという考えに基づき、機会コストの存在と信頼感、PDでの行動(協力・非協力)が取得点に及ぼす影響を検討する。具体的には以下の3点を明らかにする。

まず第1に、機会コストの存在する状況と存在しない状況では、利益に及ぼす信頼の影響力が異なることを示す。林(1993)が明らかにしてきたように、機会コストのない状況では、特定の相手との間に継続関係を形成して相互協力を維持するのが最上の策であるが、機会コストが存在する状況では、他人を信頼する傾向が有利に働くことが予測される。

第2に、上述の信頼の効果が個人の行動(協力・非協力)に影響されることを示す。人々が相手の行動そのものは考慮せず純粋な自己利益の比較により継続・解消の決定を行うと仮定しても、自己利益自体は相手の行動により影響されている。相手が非協力している場合の利益は相手が協力している場合の利益よりも確実に小さくなる。逆に言うと、非協力行動をとっている者は、相手から関係を解消される可能性が高いということになる。従って非協力的な人にとっては、信頼感を用いてこちらから相手を選ぶことの意味が相対的に小さくなると考えられる。

第3に、機会コストの存在する状況では他者を信頼することが有利に働くという上述の信頼の効果は集団内の協力率によっては影響されないことを示す。第4のアプローチでは、高いレベルの信頼をもつ者が有利になるのは、他の協力者とたまたま付き合うからではなく、非協力者とは付き合わずにより良い相手を探すからである。従って、上述の機会コストの存在する状況での信頼の利益に対する効果は、集団内の協力率が低い場合にも同様に見られると予測することができる。

うにして、全成員が全他成員に対して、0, 1, 2 のいずれかの基準点をもつ状況がつけられた。基準点は、PDでの協力・非協力に関係なくPD関係をもった相手の利益に加算される得点である。成員Aが成員Bに対して基準点2をもち、成員Bが成員Aに対して基準点1をもちている場合、相互協力の場合の利益は、Aにとっては1(相互協力の利益)+1(BのAに対する基準点)=2であるのに対して、Bにとっては1(相互協力の利益)+2(AのBに対する基準点)=3となる。この場合、AとBの間には相互協力関係が成立しているが、Aにとっては自分に対して2の基準点をもっている協力者との関係を形成することにより、より大きな利益が得られる状況が存在している。

関係の維持・離脱 既に述べたように、各成員は現在の相手との関係から得られる利益と別の相手との関係から得られる利益の期待値の比較を行い、現在の相手との関係を維持するかどうか決定する。期待値の計算に際しては、各成員は自分に対して0, 1, 2の基準点をもつ成員がそれぞれ何人ずついるかという情報と自分の信頼のレベルを用いる。信頼感20の協力者の場合、自分に対する基準値0をもつ成員のうち、20%からは1の利得、80%からは-2の利得を得ると考える。同様に、自分に対する基準値1をもつ成員のうち、20%からは2, 80%からは-1の利得を得るというように考え、現在の相手以外の98人の相手から得られる利益を計算し、98で割る。その計算結果が、前試行の利益以下である場合には現在の相手との関係を維持する。

結果⁵

機会コストなし条件における集団内協力率の各条件のシミュレーション結果をFigure 1に、機会コストあり条件における集団内協力率の各条件のシミュレーション結果をFigure 2に示した。

機会コストの効果

平均得点に及ぼす機会コストと信頼の効果調べるために、機会コスト×集団内協力率×信頼感×行動傾向の分散分析を行った。この分析においては、機会コストと信頼の交互作用が有意であった($F(10, 720) = 13.55, p < .001$)。信頼の効果はいずれの条件においても有意であるが(機会コストなし条件: $F(10, 360) = 134.91, p < .001$; 機会コストあり条件: $F(10, 360) = 12.13, p < .001$)、機会コストなし条件に関してTUKEY法により多重比較した結果、信頼感に関しては0とそれ以外の間にのみ有意な差があった($p < .001$)。機会コストのない条件では信頼のレベルでは

Figure 1. 機会コストなし条件における平均取得点。

Figure 2. 機会コストあり条件における平均取得点。

なく、少しでも信頼をもっていることが平均取得点に影響を与えていることが、Figure 1からもわかる⁶。

⁶ 本シミュレーションの利得構造では、機会コストなし条件の場合、信頼レベル T をもつ協力者の離脱の決定は期待値 $3T-2$ と現在の利益(1または-2)の比較により行われる。従って、 $T=0$ 以外の成員は非協力者とは必ず関係を断つ。同様に、機会コストありの場合、信頼のレベル T をもつ協力者の離脱の決定は、期待値 $3T-1$ と現在の利益(-2から3)の比較により行われる。従って、現在の利益が2または3の場合には離脱せず、-2または-1の場合には必ず離脱する。現在の利益が1の場合には $T > .66$ の成員が、現在の利益が0の場合には $T > .33$ の成員が離脱することになる。Figure 1, Figure 2での平均取得点のパターンは、この点から解釈可能である。

⁵ 以下の分析では、シミュレーションのデータを30レプリケーションごとに合算したものをを用いた。

行動戦略と信頼の効果

機会コストあり条件に関する上述の分散分析において、信頼感と行動戦略の間に有為な交互作用が見られた($F(10, 360) = 26.42, p < .001$)。Figure 2からもわかるように、高い信頼をもつことが有利に働くのは、その個人がとっている行動が協力的の場合に限られる。

集団内協力率の影響

機会コストあり条件に関する上述の分散分析において、集団内協力率と信頼の交互作用は有意ではなかった($F(30, 360) = 1.24, ns$)。この結果から、機会コストあり条件においては、平均得点に及ぼす信頼の効果が集団内協力率には影響されていないことが示された。

考 察

本研究ではまず、特定の2者関係を取り出す形で行われてきた従来のPD研究においてPDが解決されてきたのは、実際には関係の継続を前提とした戦略的行動により長期的に見た場合の利得構造が変換されていることを示していた点について論じた。そして、2成員の合意のもとにPD関係が成立する選択的プレイパラダイムを用いることにより、従来の2者PD研究では説明することのできなかった問題——特定の相手との間に関係の継続性が存在しないPDにおける相互協力の達成——が解決され得ることを示した。この点に関する本研究の議論をまとめると、①他成員がOFT戦略を採用している場合には、一方的な搾取が不可能となり、純粋なPDの利得構造においても協力的行動が達成される、②人々が相手の協力傾向を見極められると仮定すると、協力者同士のペアが成立する、③他者への期待と自己の行動の間には相関が存在するため、ある特定の場合には相互協力が達成される、④信頼とコミットメントの形成を社会的な不確実性への対処法であると考え、機会コストが存在する状況では他者を信頼し協力的に振る舞う傾向が実際に有利に働く、ということになる。

本研究の最終的な目的は、第4の説明に基づき、他者を信頼し、協力的に振る舞うことが実際に有利に働くことを示すことにあった。本研究ではこの目的のために、コンピュータ・シミュレーションを行った。このシミュレーションにより、機会コストが存在する状況では他者を信頼する成員が有利になることが示された。これに対して機会コストが存在しない状況では、信頼は、相手に裏切られている成員が別の成員に乗り換えようとする場合のみ意味をもつ。つまり非協力的な相手との関係を断つためには“協力してくれる人だっているはずだ”と思うだけでよい。このシミュレーションでは更に、高いレベルの信頼感をもつことが

有利に働くのは、自らが協力的である場合のみであることが示された。これは自らが非協力的である成員は、そもそも相手から取引の合意を得ることができない可能性が高いためである。このことは逆に言うと、高い信頼感をもつ者でも、非協力的な相手と避けることができることを意味している。シミュレーション結果も、信頼の効果が集団内の非協力者の割合には影響されないことを示している。

本シミュレーションでは、現在の相手の行動にはかわりなく、純粋に現在の利益と現在の関係外部からの期待利益の比較により行動を決定する状況が設定されていた。従って当然のことながら、PDそのものの利得構造(相手の行動と自分の行動によって決定される利得)と、基準点(関係性により決定される利得)の相対的大小関係によって結果は変わってくるが予測される。例えば、行動にかかわらず決定される利得が相対的に小さくなると、協力的であることの有利さが小さくなり、不利な関係から離脱できることのみが重要になってくるだろう。

社会的ジレンマへの拡張は可能か

最後に、本研究の議論により、利得構造が変換され得ない1回限りの社会的ジレンマにおける協力的行動の達成をも説明することができるかという点を考えてみたい。

選択的プレイ状況におけるOFTは基本的に、他者の行動に影響を与える戦略的行動であった。しかし応報戦略の場合と同様に、3者以上の社会的ジレンマ集団においては、特定の成員の行動に影響を与えることを目的とした戦略的行動が、それ以外の成員に対しては影響力をもってしまうという問題がある。従って、他者に影響を与えて協力的行動を引き出す方法では、社会的ジレンマの解決を説明することはできないということになる。

では第2の“見極め”アプローチではどうであろうか。このアプローチにより社会的ジレンマの解決を試みるためには、社会的ジレンマ集団に参入する際の“入会審査”を設けなければならない。つまり、第1のアプローチの場合と同様に、見極め能力に意味があるのは、非協力的な成員を関係から排除できる場合に限られるのである。

第3のアプローチであるOrbell & Dawes(1991, 1993)の議論は、戦略合理性によるアプローチでは説明不可能だった社会的ジレンマの解決が、選択的プレイパラダイムでは可能となることを示唆している。協力的な人は他人も協力的だと考えるため、社会的ジレンマ集団に占める協力者の割合が高くなるためである。

第4のアプローチでは、社会的ジレンマ状況に関しても、“信頼をもつことでより有利になることができ

る状況”が存在することを示すという方向から取り組むことになる。Yamagishi & Yamagishi(1994)の理論に基づいて考えると、信頼をもつことで有利になれるのは、客観的社会的な不確実性の低い集団においては得られない利益が、客観的社会的な不確実性の高い集団(社会的ジレンマ集団)では得られる場合である。ここでも信頼と協力傾向の間の相関が存在すれば、社会的ジレンマ集団に占める協力者の割合は高くなる。

第4のアプローチが第3のアプローチと異なる点の一つは、選択的プレイ状況の設定にある。第3のアプローチでは成員はある社会的ジレンマにおける協力者の見積りにより参加・不参加を決定するのに対して第4のアプローチでは、複数の集団の中での期待利益の比較により、どの集団に参加するかを決定する。つまり、これら二つのアプローチは基本的には同じであり、どちらを採用するかという問題は、社会的ジレンマをどのような問題ととらえるかという問題と同じである。

これら二つのアプローチのもう一つの違いは、信頼と協力傾向の相関が存在する理由である。Orbell & Dawes(1991, 1993)はこの相関の存在をフォールス・コンセンサスにより説明している。しかし第4のアプローチでは、他者に対して協力を期待するのは、そうすることにより有利になる場合があるからであるとされるため、同じ説明は用いられない。Yamagishi & Yamagishi(1994)の理論に基づいて考えると、この相関が生まれている間接的な原因としては、“評判”が考えられる。Yamagishi & Yamagishi(1994)の理論において、信頼は相手の意図に関する不完全な情報を処理する際の認知的バイアスであるとされている。そして評判は相手の意図に関する、間接的な、そして不完全な情報源である。彼らの理論において、評判は二つの異なる役割を果たす。一つはコミットメント関係にある人々に対して、コミットメント外部の人との取引に関する情報を提供する“情報的役割”であり、もう一つは人々に不正直な振舞いを控えさせる“サンクションの役割”である。評判のもつ後者の役割は、選択的プレイ状況で非協力行動が得にならない構造をつくりだすものである。

ここでは本研究でPD解決のために用いたアプローチが、戦略合理性によって解決が困難だった社会的ジレンマにも適用され得ることを示すにとどまる。今後は実際に選択的プレイの社会的ジレンマ状況を扱った研究が求められる。

引用文献

- Axelrod, R. 1984 *The evolution of cooperation*. New York: Basic Books.
- Dawes, R., McTavish, J., & Shaklee, H. 1977 Behavior, communication, and assumptions about other peoples' behavior in a commons dilemma situation. *Journal of Personality and Social Psychology*, **35**, 1-11.
- Frank, R. H., Gilovich, T., & Regan, D.T. 1993 The evolution of one-shot cooperation: An experiment. *Ethology and Sociobiology*, **14**, 247-256.
- 林 直保子 1993 TIT-FOR-TAT から OUT-FOR-TAT へ——ネットワーク型囚人のジレンマにおける戦略選手権—— 理論と方法, **8**, 19-32.
- 林 直保子・神 信人・山岸俊男 1993 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, **8**, 33-43.
- 神 信人・林 直保子・篠塚寛美 1993 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—— 実験社会心理学研究, **33**, 21-30.
- Orbell, J., & Dawes, R. 1991 A cognitive miser theory of cooperators' advantage. *American Political Science Review*, **85**, 515-528.
- Orbell, J., & Dawes, R. 1993 Social welfare, cooperator's advantage, and the option of not playing the game. *American Sociological Review*, **58**, 787-800.
- Schuessler, R. 1989 Exit threats and cooperation under anonymity. *Journal of Conflict Resolution*, **33**, 728-749.
- 渡部 幹・山岸俊男・山岸みどり・林 直保子・高橋伸幸 1994 社会的な不確実性のもとでの一般的信頼とコミットメント 日本グループ・ダイナミクス学会第42回大会発表論文集, 104-105.
- Yamagishi, T. 1986 The provision of a sanctioning system as a public good. *Journal of Personality and Social psychology*, **51**, 110-116.
- 山岸俊男 1989 社会的ジレンマの主要な理論的アプローチ 心理学評論, **32**, 262-294.
- 山岸俊男 1993 利得構造の変換と規範の生成 理論と方法, **8**, 51-68.
- Yamagishi, T., & Yamagishi, M. 1994 Trust and commitment in the United States and Japan. *Motivation and Emotion*, **18**, 129-166.
- 山岸俊男・山岸みどり・林 直保子・高橋伸幸・渡部幹 1995 信頼とコミットメント形成——実験研究—— 実験社会心理学研究, **35**, 23-34.

——1994. 9. 6 受稿, 1995. 5. 6 受理——