

社会的ジレンマと協力概念

フリーライドの視点による N 人ゲームの構造分析

武藤 正義
(東京工業大学)

【要旨】

本稿の目的は, Dawes (1980) の後, Taylor (1987) や Elster (1989) らによって拡張された種々の社会的ジレンマ概念を, フリーライドの視点から整理することにある. 具体的にはつぎの2つの課題にとりくむ. 第1に, 種々の社会的ジレンマを協力行為が定義される状況 (= 協力状況) によって包括する. 協力行為 (協力状況) は, 「協力行為のコスト」と「共通の利益の存在」という2条件によって定義される. 第2に, 多元的な個人的合理性にもとづいて, 協力状況の下位にある種々の社会的ジレンマを整理する. すなわち, N 人囚人のジレンマ, N 人円満ゲーム, N 人保証ゲーム, N 人チキンゲームなどの諸状況を「均衡の配置」から統一的に定義し, それらを個人的合理性の異なる社会的ジレンマとして位置づける. 結果として, つぎの4点が明らかになった. Dawes (1980) の「社会的ジレンマ」と「協力状況」は独立である. Taylor (1987=1995) の先手コミットメントやナッシュ均衡戦略などの多元的な個人的合理性を考慮することで社会的ジレンマ概念を拡張・精緻化すると, 社会的ジレンマは協力状況に含まれる. さらに, 進化的な合理性を個人的合理性にカウントすると, 社会的ジレンマは協力状況に一致する (N 人円満ゲームでさえ社会的ジレンマといえる). 協力行為は, 個人にとってリスクがあるだけでなく, 全体にとってもリスクがある「過剰協力」をひきおこすこともある. 重度の過剰協力がおこる状況では, 集団淘汰 (group selection) によっても協力状態は実現できない. 本稿は N 人ゲームという複雑な社会状況の理解を刷新するものとして位置づけられよう.

キーワード 協力のコスト, 共通の利益の存在, 多元的な個人的合理性, 過剰協力

1. はじめに

社会的ジレンマの数理構造の分析には, すでに多くの先行研究がある. しかしそれらはフリーライドという視点からみて, 十分に整理されているとはいえない (フリーライドの問題性は明らかだろう). そこで, 本稿はこの視点から古典的なゲーム理論を用いて, 社会的ジレンマの基礎概念の整理を行う.

本稿と進化ゲーム理論との関係を述べておこう. 進化ゲーム理論的研究では2人ゲームのランダムマッチングが多く, N 人ゲームの分析はあまりみられない. これには古典的なゲーム理論の範囲でも N 人ゲームの分析がすすんでいないということが一因にあるだろう. 本稿はこれを補うものである. つまり, 社会的ジレンマへの進化ゲーム理論的アプローチの理論的基礎として, 本稿は (基本的には) N 人ゲームを古典的なゲーム理論の範囲で精査する¹.

¹ 私見では, 日本の数理社会学における社会的ジレンマ研究は, 盛山・海野 (1991) において一応の区切りを迎え, その後, 共有知識, 社会的動機, 進化ゲーム理論, ネットワークなどが数理社会学におけるゲーム理論的研究に導入された. このとき, N 人ゲームは構造が難

用語上の注意として、本稿では単に「協力」と書く場合には行為を表し、行為の組である社会状態（協力状態）を表すものではないとする。また、本稿における均衡とは、純粋戦略ナッシュ均衡を意味する。

1.1 諸問題

社会的ジレンマ研究に用いられる諸概念には、すくなくとも3つの問題点がある。

第1に、**協力行為**概念の曖昧さがあげられる。社会的ジレンマ研究では、行為選択肢（戦略）に協力(C)/非協力(D)というラベルが使われる（Cは Cooperate, Dは Defect の略）。このラベルはもともと囚人のジレンマゲーム（Dawes 1975 など）で使われていたものだが（以下「ゲーム」を省略）、チキンゲーム（Taylor 1987=1995: 30）や保証ゲーム（山岸 1991: 240）でもこのラベルが使われるようになった（表 1, 2, 3）。そしてこのことはこれまで特に違和感なく受けとめられてきた。

表 1 囚人のジレンマ

	C	D
C	3, 3	1, 4
D	4, 1	2, 2

表 2 チキンゲーム

	C	D
C	3, 3	2, 4
D	4, 2	1, 1

表 3 保証ゲーム

	C	D
C	4, 4	1, 3
D	3, 1	2, 2

注：太字は均衡を示す。他表も同様。

しかし、ゲームが拡張されれば、協力や非協力の意味合いも元の囚人のジレンマでのものとは変わってくる。すくなくとも、囚人のジレンマ、チキンゲーム、保証ゲームの3つにみられる共通の構造に依拠した形で、協力/非協力の意味が考えられるべきである。そうでなければこのラベルは類推でつけられたままになってしまうだろう。

第2に、**協力状態**概念の曖昧さがあげられる。表 1, 2, 3 の3つの2人ゲームでは、両行為者が協力をとった社会状態(C, C)が、ふつう協力状態とよばれる。一方、非協力状態はふつう(D, D)とされるが、この意味が非・協力状態であるならば、(D, C)も(C, D)も含まれるべきだろう。

第3に、特にN人ゲームにおいて、社会的ジレンマの**分類カテゴリー**の曖昧さがあげられる。うえの表 1, 2, 3 は2人ゲームの利得表であったが、ふつう社会的ジレンマはN人ゲーム（すなわち3人以上のゲーム）で定義される。しかし、社会的ジレンマ研究で用いられるN人チキンゲームとかN人保証ゲームなどの諸ゲームのクラス概念は、2人ゲームからの類推で、全体として未整理な形で用いられている。くわえて、N人ゲームでは、なにを協力状態とよび、なにを非協力状態とよぶのかがさらに曖昧になってくる。

しいので、2人ゲームにおいてもっばらこれらの新しいデバイスが使われた。このために、N人ゲームの分析は、木村（2002）など少数の例外の例外を除くと、理論的にはいったん下火になったものと考えられる。とはいえ、ゴミ問題等、社会的ジレンマの経験例は2人ゲーム（囚人のジレンマ）ではなく、N人ゲームのほうがはるかに多いのであり、調査、実験等の実証的研究ではやはりN人ゲームが用いられている。近年では、理論的にも大浦（2002）が集団淘汰（group selection）の再検討により、N人ゲームを再び問題にしている。

以上、3点の問題を解消すべく本稿では、協力行為、協力状態、諸 N 人ゲームをシステムティックに定義する。

1.2 先行研究と本稿の目的

先行研究としては、Dawes (1975, 1980), Taylor (1987=1995), Raub (1988), 永田 (1988), Elster (1989=1997: Ch13), 瀧川 (1989), 海野 (1988), 三隅 (1993), 木村 (2002), 鈴木 (2003) などがある。しかしまず、Dawes (1975, 1980) の社会的ジレンマの定義は狭すぎる。たとえば、Taylor は N 人チキンゲームを社会的ジレンマとみなしており、これは Dawes の定義には入らない。とはいえ、Taylor (1987=1995: 56-60) も N 人チキンゲームの定義をきちんと数理化してはいない。数理化の曖昧さは、Elster (1989=1997: Ch13) にもいえる。一方、海野 (1988), 三隅 (1993), 木村 (2002) は、経験的な資源の性質を数理化するという方向性ゆえに、くわえて Dawes の社会的ジレンマ定義に依拠しているゆえに、社会的ジレンマ一般のゲーム構造そのものは十分分析してはいない。

Dawes とは対照的に、Raub (1988) および永田 (1988) の社会的ジレンマ (問題状況) の定義は広すぎる (永田と Raub は独立にほぼ同じ内容の定義をしている)。Raub は Harsanyi (1977) を参考に「問題状況」を包括的に定義しているが、特にこれは $m \times n$ 行列の一般のゲームについてのものであって、実験・調査研究でよく用いられる 2 選択枝ゲームとしての社会的ジレンマに特有の性質を明らかにしてはいない。一方、瀧川 (1989) は 2 選択枝ゲームについて構造分析を試みているが、線形の場合のみであるにもかかわらず分析が複雑ゆえ実用上使いにくい (これは利得和の概念を用いているためでもある)。また、鈴木 (2003) は比較的明確に社会的ジレンマを定式化してはいるが、ステップ財に限ったものであり、限定的過ぎる。

このように、先行研究では N 人 2 選択枝ゲームとしての社会的ジレンマ一般の構造はきちんと分析してはいない。よって本稿の目的は、協力行為と協力状態を定義し、そして社会的ジレンマのサブカテゴリーにあたる N 人チキンゲーム、 N 人保証ゲームなどをシステムティックに定義することによって、社会的ジレンマ研究の基盤となる諸概念を整理することにある。ただし、本稿における協力概念は、フリーライドの視点から構成する。だから本稿の協力概念が唯一絶対というわけではない。

本稿の構成を述べる。まず、第 2 節では、協力 / 非協力という行為ラベルが使われるにふさわしいゲームのクラス (協力状況) を定義する (つまり協力行為を定義する)。また、この状況において、協力状態、非協力状態なる用語を定義する。つぎに第 3 節では、社会的ジレンマと協力状況の関係を多元的な個人的合理性の観点から議論する。さらに、 N 人チキンゲームや N 人保証ゲームを定義する。第 4 節はまとめである。

2. 協力状況と協力状態

この節では、協力行為が定義される状況として、「協力状況」なるものを定義する。協力という行為の意味は、ゲーム理論的には利得構造 (利得関数の組) によってきまる。

2.1 協力行為の定義

Elster (1989=1997: 152) は、協力行為 (cooperate) を「何人かが (あるいは全員が) ある行為をすればすべての人が便益を受けるような行為を、自分個人の利害に逆らって行うこと」として定義している。同著のほかの部分と照らし合わせると、「自分個人の利害に逆らう」は、明らかに被支配戦略を意味している。しかし、Taylor (1987 = 1995: 26, 30) は、支配戦略のない N 人チキンゲームにおいても協力行為 (cooperate) という用語を用いる。そこで本稿では基本的に、Elster の協力行為定義の拡張を試みる。

付録にあるように、「cooperate」の意味を英語辞書で調べると、「To work or act together toward a common end or purpose.」とか “To form an association for common, usually economic, benefit などとある (太字は引用者による)。このように多くの辞書が“cooperate”の意味に、“a common purpose”, “end”, “aim”, “benefit”すなわち「共通の目的」や、より弱い意味での**共通の利益**を認めている。これは“cooperate”が、たんに “work together” 「いっしょに働く」だけでなく、共通の利益にむかって方向づけられた行為であることを意味する²。

また、work や form an association には、明らかに労働や資金等のコストがかかる。「もっと家事に協力して」とか「募金にご協力ください」などというように、日本語の協力という言葉にもそのコストが前提にある。そこで、つぎのように「協力」を定義する。

定義 1 協力とは、(他者ととともに) コストを払うことで共通の利益をもたらさうる行為のことをいう。

「他者ととともに」に括弧をつけたのは、他者の行為を選択することはできないからである。この定義のポイントを述べよう。

1. 「コストを払う」ことは、それじたいとしては自分自身にとって不利益である。つまり、協力行為は、個人的な不利益によって、共通の利益を生みうる行為である。
2. 「共通の利益」について、たとえ自分だけしか協力しなくても、それは共通の利益を促進しうるから、協力者たる自分が利益を受け取る時はかならず、非協力者も同様の利益を受け取る。つまり、協力はこの行為者たちにとっての「公共財」を生みうる行為である。

2.2 協力行為のための必要条件

協力行為をゲーム理論的に定式化してみよう。行為者を均質的とする。このときゲームは

² 共通の利益は、社会にとって根幹的である。社会とは、人と人との関わりあいの総体であるが、この関わりあいは、一般形として、互いに正の外部性 (利益) と負の外部性 (害) をもたらさう (いうまでもなくそのモデルがゲームである)。その特殊形として、互いに正の外部性をもたらさうことが、典型的な共通の利益である。ただし、意味を広くとれば、共通の利益は、害がなくなることをも含む。つまり、共通の利益の本質は、「パレート改善」である。

対称である．また行為選択肢は，協力，非協力の2つのみとする．つぎのように記号を定める． N, m は整数とする．

- N 行為者の人数
- m 社会状態における自分自身を含めた協力者の人数，ただし $m \in [0, N]$
- $v(m)$ m 人協力しているときの，1人あたりの共通の利益
- $c(m)$ m 人協力しているときの，1人あたりの協力行為のコスト， $c(m) > 0$

注意点が2つある．第1に，協力の利益 $v(m)$ も協力のコスト $c(m)$ もともに協力者の人数 m によって変化しうる．第2に， $v(m)$ は負の値，つまり共通の不利益（害）をも表すことができる．利益，不利益の絶対量ではなく，相対的な利益の差だけが問題である．

うへの記号を用いて，協力行為を表現する式が満たすべき必要条件を考えよう．

第1に，協力者だけがコスト $c(m)$ を払い，協力者も非協力者も等しく共通の利益 $v(m)$ を受け取るから，

$$\text{協力者の利得} = v(m) - c(m), \quad \text{非協力者の利得} = v(m)$$

が成立する．したがって，非協力者は協力者の貢献につねに「フリーライド」する，

） 非協力者の利得 > 協力者の利得

なる関係（協力行為のコスト）が成立する（支配戦略関係ではない）．

第2に，共通の利益について考えよう．協力行為は，コストを払っているにもかかわらず，つねに共通の利益を生むとは一般にはいえない．コストを払って協力しても，かえってだれにも不利益を与えることもある．たとえば，Elster（1989=1997: 156-8）は，

- 「一方的な軍縮は，軍事バランスを崩すのでかえって緊張を高める」
（「過少協力」状態のパレート非効率性），
- 「パーティのあとの皿洗いは，全員でやるとむしろ非効率」
（「過剰協力」状態のパレート非効率性）

という例をあげている．このような状況では，協力行為は単にひとりだけが損をするというリスクだけでなく，だれもが損をするというリスクをもつ．

しかし，協力行為がまったく共通の利益を生む可能性がなければ，それはもはや協力行為とは定義上よばれない．したがって，協力行為が共通の利益を生むためには，すくなくとも全く協力者のいない社会状態よりはだれにとってもよいような，協力者がいる社会状態が存在しなければならない．つまり，協力行為には，

） 全員非協力の状態 $<_p$ 協力者がいるある状態（ $<_p$ は狭義パレート優越関係）

という条件（共通の利益の存在）が必要である．結局，この条件は「全員非協力がパレート非効率」ということである．そしてこれは，

$$v(0) < v(m_1) - c(m_1) \quad m_1 \in [0, N]$$

という式と同値である．というのは， $v(m_1) - c(m_1)$ は非協力者の利得である一方，協力者の利得は $v(m_1)$ であり， $c(m_1) > 0$ より， $v(0) < v(m_1)$ も成り立つので，このとき協力者，非協力者ともに，だれもが非協力者である状態 $m = 0$ よりよいからだ．

2.3 協利行為のための十分条件

） ）は，ある 2 つの行為選択肢が協力 / 非協力であるための必要条件であった．しかしじつは，これは十分条件でもある．このことを示そう．ある 2 つの行為選択肢 A, B において，

） B 選択者の利得 $>$ A 選択者の利得

） 全員 B を選択したときの結果 $<_p$ A 選択者が存在するある結果

という 2 つの関係が成立していれば， ） から A 選択者の人数を m として，

$v(m)$: = B 選択者の利得

$c(m)$: = B 選択者の利得 - A 選択者の利得 (> 0)

が定義できる．このとき， A 選択者の利得 = $v(m) - c(m)$ ， B 選択者の利得 = $v(m)$ となり，かつ ） から選択 A が共通の利益をもたらさうから， $c(m)$ は選択 A のコスト， $v(m)$ は共通の利益として解釈できる（ B 選択者の利得を共通の利益とよべるのは， A 選択者のコスト c が操作的に定義されているからである）．よって， A が「協力」であり， B が「非協力」であることがわかる．よって， ） ） は協利行為が成立するための十分条件である．

2.4 協利行為を定義できる状況としての協利状況

結局， ） ） はうえで定義した協利行為概念の必要十分条件と考えられる．もちろん，本稿の協利行為の定義はフリーライドを念頭においたものなので，協利行為の定義がこれではなくてはいけないということではない．しかし，うえの協利行為の定義は，協利のコスト $c(m)$ が協力者 m の関数であるためにかなり広いものである．以上の議論から，つぎの「協利状況」を定義する．

定義 2 ） ） をみたく対称ゲームを協利状況とよぶ．

協力状況では協力行為、非協力行為が定義できる。協力状況は当該の行為者たちにとって、彼らに公共財の供給が社会的に求められる状況である。

2.5 2人協力状況および協力の過剰と過少

状況の理解のために、2人の場合を考えておく（表4）。

表4 2人ゲームでの協力状況

この2×2対称ゲームが協力状況であるためには、

協力
非協力

	協力	非協力
協力	$v(2) - c(2), v(2) - c(2)$	$v(1) - c(1), v(1)$
非協力	$v(1), v(1) - c(1)$	$v(0), v(0)$

$$v(2) - c(2) > v(0) \quad \text{または} \quad v(1) - c(1) > v(0)$$

が成り立つことが必要十分である。ここで簡単のため、

表5 表4の略記

$$x = v(2) - c(2) \quad y = v(1) - c(1)$$

$$z = v(1) \quad 0 = v(0)$$

	協力	非協力
協力	x, x	y, z
非協力	z, y	$0, 0$

とする（表5）。このときうへの不等式は、 $(x > 0 \text{ または } y > 0)$ かつ $(z > y)$ とかける。このような対称ゲームを選好順序だけでかいてみると、つぎようになる。

表6 2人ゲームでの具体的な協力状況

	C	D	円満	保証	調整
$x > z > y > 0$	x, x	y, z	<u>3, 3</u>	<u>2, 2</u>	<u>1, 1</u>
$x > z > 0 > y$	x, x	y, z	1, 2	-1, 1	-2, -1
$x > 0 > z > y$	z, y	$0, 0$	2, 1	1, -1	-1, -2
$z > x > y > 0$	z, y	$0, 0$	0, 0	<u>0, 0</u>	<u>0, 0</u>
$z > y > x > 0$	<u>2, 2</u>	<u>1, 3</u>	1, 1	1, -1	-1, -1
$z > x > 0 > y$	<u>2, 2</u>	<u>1, 3</u>	1, 1	-1, 2	<u>1, 2</u>
$z > y > 0 > x$	<u>3, 1</u>	0, 0	<u>3, 2</u>	2, -1	<u>2, 1</u>
			0, 0	<u>0, 0</u>	0, 0

このような7つのゲームを表6にまとめておく。2×2対称ゲームは12種類あるので（武藤 2004），その半数以上が協力状況ということになる。表6の太字は均衡，下線部は後に定義する「協力状態」である。利得の数字は，順序のみに意味があり，基数的な意味はない。いずれも第1の選択Cが協力であり，第2の選択Dが非協力である。協力状況は，これら有名なゲームに共通の構造なのである。以下の4点について簡単に分析しておく。

- 個人的に最善な状態とナッシュ均衡について。 $(x > 0 \text{ または } y > 0)$ は全員協力(C, C)が個人的に最善の状態である（全員協力はナッシュ均衡でもある）。 $(z > y)$ では，自分は非協力で相手から一方的に協力してもらうのが個人的に最善の状態である（よって全員協力はナッシュ均衡でない）。このうち $(x > 0 \text{ または } y > 0)$ は多型状態(C, D), (D, C)が均衡であり，これら

は広義のチキンゲームといえる（チキンゲームの厳密な定義は後述）。

- 支配戦略について． C では協力が支配戦略となっており， D では非協力が支配戦略になっている．ほかのゲームでは支配戦略はない．
- 協力の個人的なリスクについて． C では協力が個人的に最悪の結果を起こしうる．ここでは，非協力がマクスミニ戦略でもある．
- 協力の集団的なリスクについて．協力はいつも望ましいとは限らず，集団的なリスクを含む場合がある．これは，協力の過少・過剰の問題である．
 - C のある種の「調整ゲーム」では，一方的な協力が，かえって全員非協力に対してパレート劣位になる**重度の過少協力**状態をひきおこす．
 - 「英雄ゲーム」では，全員協力が，かえって全員非協力に対してパレート劣位になる**重度の過剰協力**状態をひきおこす³．

当然，重度の過少協力・過剰協力状態はパレート非効率である．一方，軽度の協力の過少，過剰をつぎのように考える．

- C では一方的な協力が，全員非協力に対してパレート劣位になるわけではないが，パレート非効率的になる**軽度の過少協力**状態をひきおこす．
- D では全員協力が，全員非協力に対してパレート劣位になるわけではないが，パレート非効率的になる**軽度の過剰協力**をひきおこす．
- D では過少協力，過剰協力といえる社会状態はない．

2.6 協力状態の定義

全員非協力が望ましくないのは，パレート非効率だからである．その反対概念である協力状態は，当然，パレート効率的であるべきだろう．これは協力状況にみられる共通の利益の一側面である．

しかし，社会状態が単にパレート効率的であるだけでは，協力状態とはふつうよばれない．たとえば，表6の囚人のジレンマでは， (C, C) だけが協力状態とよばれ， (D, C) と (C, D) はパレート効率的であるにもかかわらず，協力状態とはよばれない．つまり，協力状態は，全員非協力に対してパレート改善されていることが必要だろう．これは共通の利益のもうひとつの側面である．以上から，協力状態をつぎのように定義する．

³ たとえば，だれもが協力的だと息苦しいということはよくある．適度にだれかがサボっているときのほうがうまくいくことは多い．だれもがあまりに協力的ゆえすすんで仕事をやろうとして，仕事の奪い合いや路線の対立などが起こり，いわゆる「内ゲバ」が起こって集団そのものが崩壊してしまうというのが英雄ゲームである．

定義 3 協力状態とは、全員非協力に対してパレート優位、かつパレート効率的な社会状態のことをいう。

協力状態以外の社会状態を「非協力状態」とよぶ。したがって全員非協力以外の非協力状態も一般に存在する。表 6 において協力状態を下線で示した（それ以外が非協力状態）。

2人協力状況の (C, C) は、全員協力 (C, C) がパレート効率的かつ (D, D) にパレート優位であり、協力状態といえる。一方、 (D, D) では、 (C, C) がパレート非効率である。特に、 (D, D) の英雄ゲームは重度の過剰協力が起こりうる（全員が英雄であろうとして争いが起こる）。結局、 (D, D) の場合には、 (D, C) と (C, D) のみを協力状態とよぶべきだろう。

ただし、 (D, D) のチキンゲームでは、やや直感に反し、 (D, C) と (C, D) はパレート効率的かつパレート改善されたものであり、これらも協力状態とよぶことになる。しかし、 (D, D) や (C, C) において (D, C) と (C, D) が協力状態とよばれるのであれば、チキンゲームにおいてもこれらを協力状態とよんでもおかしくはない。

N 人ゲームにおける協力状態の基本的な性質を述べよう。

定理（協力状態の存在） 協力状況では、協力状態が \forall に存在する。

証明) (C, C) より、協力状況では、全員非協力に対してパレート優位な社会状態 (C, C) が存在する。

がもしもパレート非効率であれば、その (C, C) に対してパレート優位な社会状態 (C, C) が存在する。行為者は N 人であり、社会状態は 2^N 個で有限であるから、この操作は有限回で終わり、パレート効率的な社会状態 (C, C) が存在して、全員非協力に対してパレート優位となる。この (C, C) が協力状態である（証明終）。

協力状態は、かならずしも均衡ではない（たとえば表 6 の「囚人のジレンマ」）。

協力状態は、複数存在しうる（たとえば表 6 の「チキンゲーム」）。

2.7 他の協力行為定義の可能性

他の協力行為（協力状況）の定義の可能性を考えておこう。まず、 (C, C) に対して、協力行為にコストがかからない場合とコストがむしろマイナスである場合が考えられる。この場合、非協力者の利得より協力者の利得のほうが大きくなりうる。これは、他者を出し抜けば全体的にもよくなりうる、という状況である。ここでは、個人的な行為と社会との間に緊張関係はない。つまり、ここではフリーライドが存在しない可能性がある。この状況を数理モデル化することもできるが、本稿では、協力という概念をフリーライドと関連させて定義したいゆえ (C, C) という条件を課した。

なお、もしも (C, C) を仮定しないならば、2つの選択を協力と非協別に区別するメルクマールには、どのようなものがあるだろうか。たとえば、

’) 全員が選択 1 をとった状態と、全員が選択 2 をとった状態という両極を考えたとき、パレート優位になるほうの選択を協力、もう一方を非協力とよぶ

というものがあろう。このとき) の条件は自然に満たされる。しかし ’) だけでは、両極のうちのパレート優位な「全員協力」状態でもパレート非効率である可能性がある（つまり軽度の過剰協力）。

そこで、いま、選択 A を全員がとっている状態が、選択 B を全員がとっている状態に対してパレート優位であるとし（すなわち、選択 A = 協力、選択 B = 非協力）、かつ、どちらの状態もパレート非効率であるとしよう。この状況では、選択 A と選択 B が混在する「協力状態」が存在して両極に対してパレート優位であり、かつパレート効率的である。このとき、この協力状態において 2 つの選択肢の一方を協力とよび、もう一方を非協力とよぶべき積極的な理由はない。なぜなら、どちらの選択肢もこの協力状態には欠かせず、どちらかを非協力とはいえないからである（このことはじつは)) によって定義される協力状況においてもいえるのだが、)) では非協力のほうが協力よりつねに利得が高いという非対称性があるから問題はない）。

よって、’) の仮定のもとでは、’) 全員協力がパレート効率的である、という第 2 の仮定が必要になる。しかし、この場合には Elster (1989=1997: 156-8) が問題にした「過剰協力」という興味深い問題が扱えなくなる。具体的には、表 6 の 指導者ゲームと 英雄ゲームを協力状況から排除することになる。しかし本稿では、これら 2 つのゲームも協力状況(社会的ジレンマ)として扱いたい。後述するように、これが制度の存立と関係するからである。

なお、協力ゲーム理論における協力概念は、「提携」coalition を意味する。協力ゲームでは、協力的行為とそれに伴うコストは考えていない。ただし提携概念は、本稿における協力状態(パレート効率性とパレート改善)を前提にしている。

3. 社会的ジレンマと協力状況

この節では、先行研究において考えられてきた多くの社会的ジレンマの定式化が、協力状況であることを示す。なお、この節における選択 C, D は前節において定義した協力、非協力とはさしあたり別物とする。

3.1 既存の社会的ジレンマ定義の問題点

社会的ジレンマの本質は、「個人的合理性と社会的合理性の乖離」にあるとされる（木村 2002: 20）。社会的合理性は、平等性などもありえるが、ふつうはパレート効率性とされる。一方、個人的合理性には幅がある。たとえば、Dawes (1980) は個人的合理性を支配戦略としてとらえ、社会的ジレンマをわかりやすい形で定式化している⁴。

⁴ Dawes (1975) では、(1)(2)以外に「 $C(m), D(m)$ は単調増加」、「協力者が増加すると、全行為者の総利得も増加する」という 2 つの条件をつけて社会的ジレンマをさらに狭く定義している。海野 (1988) に詳しい。

定義 4 Dawes による社会的ジレンマとは、次の条件(1)(2)をみたす行為選択肢 C, D をもつ N 人 2 選択肢対称ゲーム（ $= 2^N$ 対称ゲーム）である。

- (1) $D(m-1) > C(m) \quad m \in \{1, \dots, N\}$
 (2) $D(0) < C(N)$

ここで、 m は自分自身を含む C 選択者の人数であり、 $D(m)$ は D 選択者の利得、 $C(m)$ は C 選択者の利得である。よって $C(m)$ の定義域は $\{1, \dots, N\}$ であり、 $D(m)$ の定義域は $\{0, \dots, N-1\}$ である。なお、 $D(m), C(m)$ は減少しうるから、この状況は意外に一般的ではある。

(1)(2)を満たす状況は、「 N 人囚人のジレンマ」とよばれる（ $N=2$ の場合に通常の囚人のジレンマに一致する）。(1)は D が**強支配戦略**であることを意味し、(2)は全員 D ($m=0$) が全員 C ($m=N$) に対してパレート劣位であることを意味する。つまり(1)(2)は、だれもが個人的に合理的な強支配戦略（ D ）を選択すれば、社会的に非合理的なパレート非効率状態を帰結してしまう、ということをも主張する。

このような状況の構造から、ふつう C は「協力」、 D は「非協力」とよばれるが、これらははたして前節の協力行為の条件（ ）を満たすのだろうか。このことの吟味も含めて、既存の社会的ジレンマの定義に対して4つの問題点（パレート劣位条件(2)の限定性、支配戦略条件(1)の限定性、パレート非効率均衡の限定性、逆転現象）をあげる。

問題点 1 パレート劣位条件(2)の限定性 (2)の本質は、個人的に合理的な選択 D をだれもがとるとパレート非効率になる点にあるので、特に全員 C が全員 D に対してパレート優位である必要はない。したがってこの条件は、

$$(2)' \quad D(0) < C(m_1), D(m_1) \quad m_1 \in \{1, \dots, N\} \quad \text{ただし } D(N) \text{ は定義できない}$$

におきかえられるべきである⁵。じっさい、Elster（1989=1997: 152）は社会的ジレンマを条件(1)と(2)'で定式化している。条件(2)'は、 $D(0) > C(N)$ がありうること、すなわち過剰協力の可能性を暗示している。

問題点 2 支配戦略条件(1)の限定性 個人的な合理性は、(1)が意味する強支配戦略に還元されるものではない。それはナッシュ均衡として表現されてもよい⁶。したがって支配戦略が存

⁵ ただし、 $C(m)$ が非減少であること（= 選択 C が負の外部性を他の C 選択者に与えない）を仮定すれば、(2) (2)'がいえる。証明) () は明らかなので、() を証明しておこう（これもほぼ自明だが）。 $C(m) > C(m+1)$ とする。したがって任意の m に対して、 $C(m) > C(N)$ が成立する。よって、(2)' からある m_1 に対して $D(0) < C(m_1)$ が成立すれば、(2) $D(0) < C(N)$ が成立する（証明終）。

⁶ 支配戦略が他者の戦略を顧みない「個別的」合理性を意味するのに対して、ナッシュ均衡戦略は他者の戦略と相関的にとられるから個別的ではない（丸田 1991）。しかし、それは自

在しない状況でも、「すべてのナッシュ均衡がパレート非効率」〔Raub および永田による社会的ジレンマ定義〕であれば社会的ジレンマとして考えられる（Raub 1988, 永田 1988）。(1)は、社会的ジレンマ一般の定義には狭すぎるのである。

問題点 3 パレート非効率均衡の限定性 そうかといって(1)を「すべてのナッシュ均衡がパレート非効率」におきかえるだけでは社会的ジレンマを包括できない。ナッシュ均衡がパレート効率的でありうる N 人チキンゲームも社会的ジレンマに含めるべきという議論があるからだ（Taylor 1987=1995: 2章, 木村 2002: 22）。このとき条件(1)をどのような条件におきかえるべきかは次節で議論する。

問題点 4 逆転現象 以上の 3 つの問題点が基本的には社会的ジレンマの Dawes 定義が狭すぎることを問題にする一方、第 4 の問題点は、それが広すぎることを問題にする。この問題点は本稿独自のものである。Dawes 定義の問題条件(1)(2)だけだと、 $D(m)$ と $C(m)$ がと

表 7 Dawes 定義での逆転現象

C の人数 m	0	1	2	3
$D(m)$	0	4	<u>2</u>	-
$C(m)$	-	-1	<u>3</u>	1

注：太線部は均衡を示す。他表も同様。

ともに減少することがある場合に $D(m) < C(m)$ となることがある⁷。したがって)により、この状況は協力状況ではない。すなわち、条件(1)(2)で定まる C を「協力」、 D を「非協力」とよぶことは一般にはできない。たとえば、 $D(0) = 0, C(1) = -1, D(1) = 4, C(2) = 3, D(2) = 2, C(3) = 1$ となる 3 人ゲームを考えよう（表 7）。この表では条件(1)として、任意の $D(m-1)$ より、その右下のセルの $C(m)$ は小さくなる。また条件(2)として、最も左上の $D(0)$ より最も右下の $C(3)$ のほうが大きい。よって、このゲームは(1)(2)をみたく。その一方、 $m = 2$ では $D(m) < C(m)$ である。したがって Dawes の社会的ジレンマは協力状況に包含されない。

その一方で、協力状況では支配戦略の組がパレート効率的となる場合があり（たとえば円満ゲーム（表 6, 表 8））、この場合は古典的なゲーム理論の範囲では明らかに社会的ジレンマではない。よって、協力状況と Dawes の社会的ジレンマは、論理的には独立なのである。しかし多くの社会的ジレンマ研究では、

$$(3) \quad D(m) > C(m) \quad m \in \{1, \dots, N-1\}$$

を事実上仮定している。この式の必要性を次節で考える。

3.2 社会的ジレンマは協力状況に含まれる

問題点 2 「支配戦略条件(1)の限定性」に対する Raub および永田の代替案は、行為選択肢が 3 個以上のゲームをも対象にしていた。しかし本稿ではこの方向は考えず、簡潔のため最

身の利得を最大化するという点で「個人的」合理性にかなう。

⁷ この点、Dawes (1975) では、このようなことは起こらない。 $C(m)$ が非減少関数であれば、 $m = 1, 2, \dots, N$ に対して $C(m) - C(m+1) < D(m)$ だからつねに $D(m) > C(m)$ が成り立つ。

も基本的な 2 選択枝の場合のみを考える⁸。

問題点 3 「パレート非効率的均衡の限定性」は、 N 人チキンゲームをも社会的ジレンマに含めようという議論から出てきた。 N 人チキンゲームの本質は、Taylor (1987=1995: 56-60) によれば**先手コミットメントのインセンティブの存在にある**。これは**全員 C がナッシュ均衡ではなく、かつ $D(m^*) > C(m^*) (> D(0))$ をみたす多型ナッシュ均衡が存在すること**を意味する⁹ [Taylor による N 人チキンゲームの定義]。この状況は、多型均衡の実現を意図して「先に」 D を選択すれば他者に C を強制できるけれども (= 先手コミットメント)、だれもがそのように考えて行為選択すれば全員 D (パレート非効率) が帰結してしまう、というものである。先手コミットメントを個人的合理性として考えることで、この状況もやはり社会的ジレンマ (個人的合理性と社会的合理性の乖離) として考えられるだろう。このときいうまでもなく、先手コミットメントは、フリーライダー問題のキー概念である。

問題点 4 「逆転現象」の観点から、さらに N 人チキンゲームを考えてみよう。Taylor は、 D 選択者が有利な多型均衡と全員 D (均衡でなくてもよい) のみを実現可能な社会状態として考えている。しかし、現実にはもちろん理論上でさえかならずしもそれらの社会状態が実現するとは予測できない。なぜなら、支配戦略の場合とは異なり、ゲーム理論の個人合理的な推論は、「ひとたびナッシュ均衡が実現すればだれもそこから動かない」と予測するだけで、「**ナッシュ均衡戦略を行為者が選択する**」とまでは予測しないからだ。よって、 $D(m) < C(m)$ なる均衡外の社会状態がありえれば、 D が個人合理的な選択であるとは断定しにくくなる。

すくなくとも進化的過程を考えれば、 $D(m) < C(m)$ なるある状態が一定期間実現すれば D は C に淘汰される (利得単調性)。あるいはこのことを合理的に考えられるとすれば、行為者は C をとりうるので D の個人的な合理性は弱まる。このような**自分の利得^{マイナス} - 他者の利得**という相対的利得を最大化する原理を**進化的合理性**とよぶ¹⁰。進化的合理性は個人的合理性の一種としてカウントできよう。したがって、 N 人チキンゲームにおいて D が確実に個人合理的な選択であるためには、条件(3)が必要になる。

まったく同様にして、個人的合理性としての進化的合理性を考えると、2 選択枝社会的ジレンマ一般においても(3)は必要になる。じっさい、山岸 (1989) や木村 (2002) に登場するほとんどの社会的ジレンマの具体例では、この条件が仮定されている。結局、先手コミットメントは、進化的過程の考慮によって拡張される。このときもちろん、(3)はフリーライド

⁸ もっと積極的に 2 選択枝社会的ジレンマを位置づけることもできる。2 選択枝はなにかをする / しない、という根本的な二項対立にもとづくことで「ゲームの外部がない」ということと「相互行為のルールから逸脱する / しない」という特別な意味をもつからだ (武藤 2005a)。

⁹ 多型とは異なる行為の選択者がいる状態のこと。

¹⁰ 相対的利得の最大化は、社会的動機としては競争主義とよばれる (土場 1995)。競争主義あるいは進化的合理性は、いわゆる嫌がらせ (スパイト) 行動に象徴的である。戦略や社会的動機や遺伝子を抽象した進化の単位をタイプとよぼう。進化するタイプ (= シェアを拡大するタイプ) は、他者に対して相対的に利得が高いタイプである (利得単調性)。したがって、スパイト行動や競争主義はあらゆる N 人ゲームの協力状況 [(), ()] で進化する。一方、状況が 2 人ゲームのランダムマッチングの場合、これらの進化はゲームによって条件づけられている。たとえばチキンゲームでは競争主義 (非協力戦略) は一定の範囲までしかシェアを拡大できない。なお、利得単調性については Weibull (1995=1998: 182) 参照。

が実行されるための十分条件である。

翻って問題点 1「パレート劣位条件(1)の限定性」にかんして、前述したように Elser (1989=1997)は社会的ジレンマを(1)と(2)'で定式化しているが、条件(3)を明示してはいない。そこで(3)を仮定すれば、条件(2)'は、さらに簡潔に、

$$[2] \quad D(0) < C(m_1) \quad m_1 \in \{1, \dots, N\}$$

とかくことができる。

結局、ナッシュ均衡戦略、先手コミットメント、進化的合理性という多元的な個人的合理性を考えていくと、条件(1)「支配戦略条件」は社会的ジレンマには不要であり、かわりに条件(3)が必要であることがわかった。そこで(1)を条件(3)におきかえ、これを条件[1]と改めよう。これと上の[2]を併せると、社会的ジレンマはつぎの2条件をみたす必要がある。

社会的ジレンマの必要条件

$$[1] \quad D(m) > C(m) \quad m \in \{1, \dots, N-1\} \quad \text{協利行為のコスト}$$

$$[2] \quad D(0) < C(m_1) \quad m_1 \in \{1, \dots, N\} \quad \text{共通の利益の存在}$$

明らかに、[1] = (), [2] = () がいえる。つまり、個人的合理性が多様な意味を容れることで、社会的ジレンマは協力状況（公共財供給状況）に包含される（社会的ジレンマ = 協力状況）。よって、[1][2]の成立をもって C を協利行為、 D を非協利行為とよぶ。第2節の記号を用いれば、共通の利益 $v(m)$ とコスト $c(m) > 0$ を用いて、協力者の利得 $C(m)$ と非協力者の利得 $D(m)$ をつぎのようにかける。

$$C(m) = v(m) - c(m), \quad D(m) = v(m)$$

これまで公共財供給状況は社会的ジレンマの一種として考えられてきたから（木村 2002: 20-30）、これはすこし奇妙にもみえる。しかし、もともと公共財供給状況は表6 円満ゲームの例にみたように、個人的合理性と社会的合理性の「乖離」のない状況をも含むのである¹¹。

結局、Dawes 定義においては社会的ジレンマと協力状況は独立だったが、個人の多元的合理性に照らして社会的ジレンマを拡張かつ精緻化していくと、社会的ジレンマは、協力状況に包含される。

ただし、[1][2]はこのままでは社会的ジレンマ（二水準の合理性の「乖離」）に対する十分条件ではない。しかし、次節にみるように、進化的視点ではこれは異なってくる。

¹¹ 「乖離のない状況」は公共財の価値 $v(m)$ の傾き $v(m+1) - v(m)$ が大きい場合、すなわち $D(m) < C(m+1) \iff v(m) < v(m+1) - c(m+1) \iff v(m+1) - v(m) > c(m+1)$ が任意の m で成り立つ場合である。つまり1人の協利のコストよりもそれによる1人あたりの利益のほうが大きい場合には、「乖離」はないのである。しかしこのような状況は人数が多いほど稀だろう（大集団のジレンマ）。

4. 協力状況のサブカテゴリー

この節では、協力状況の下位のゲームのクラスである N 人囚人のジレンマ、 N 人チキンゲーム、 N 人保障ゲーム、 N 人円満ゲームを「均衡の配置」という視点から統一的に定義する。そしてこの4つのゲームがそれぞれ4つの個人的合理性すなわち、支配戦略、先手コミットメント、マクスミニ戦略、進化的合理性に対応した社会的ジレンマであることを明らかにする。

4.1 種々の N 人ゲームの定義

ゲーム理論的な分析軸は、当然ながら均衡による。ここで、条件[1][2]を満たす状況には均衡が存在することが証明できる（武藤 2005b）。したがって、均衡が存在すればそこでの協力者数 m^* が $0 \sim N$ をとるので、図 1 が得られる。協力状況の条件[1][2]の下で、純粹型として、つぎの N 人ゲームのクラスを均衡の配置によって操作的に定義する。

定義 5

N 人囚人のジレンマ	均衡が全員非協力 ($m^* = 0$) のみからなる
N 人円満ゲーム	均衡が全員協力 ($m^* = N$) のみからなる
N 人チキンゲーム	均衡が多型状態 ($m^* \in [1, N - 1]$) のみからなる

言い換えれば、チキンゲームとは、全員協力および全員非協力がいずれも均衡になっていないゲームとして定義される。すなわちチキンゲームとは、すべての均衡が協力者と非協力者が混在する多型状態からなるゲームのことである。

定義 5 の 3 つのゲームは典型的なもので、これらだけで協力状況を埋め尽くすわけではなく、たとえば N 人囚人のジレンマと N 人円満ゲームの「あいのこ」のような、これらの中間的な状況が存在する。さらに各ゲームのクラスについていくつかの注意点がある。

- N 人囚人のジレンマでは**非協力が強支配戦略**であることが証明できる (ibid)。したがって、この N 人囚人のジレンマ定義は、Elster (1989=1997: 152) の社会的ジレンマ定義 = 条件(1)(2)' に一致する。
- N 人円満ゲームでは**協力が強支配戦略**であることが証明できる。また、その組である**全員協力 $m = N$ が唯一のパレート効率状態**であることが証明できる (武藤 2005b)。つまり N 人円満ゲームでは、全員協力が唯一の協力状態でありかつ唯一の均衡である。なお、 N 人円満ゲームは、非協力者のほうが協力者より利得が高いが (条件[1])、だれかが協力したほうがよいとだれもが思っており (条件[2])、しかもだれもが協力をしたいと思っている (協力が支配戦略)、という状況である。
- N 人チキンゲームの定義では、Taylor (1987) のものと異なり、 $m = 0$ が均衡ではないことが要請される。しかしこの定義は Taylor の先手コミットメントの考え方を継承し

ている．これは純粋型という都合である．また，協力者数 m を異にする複数の均衡が存在する場合もある．いずれにしても，**全員非協力に対してパレート優位な多型均衡が存在**することが証明できる（武藤 2005b）．

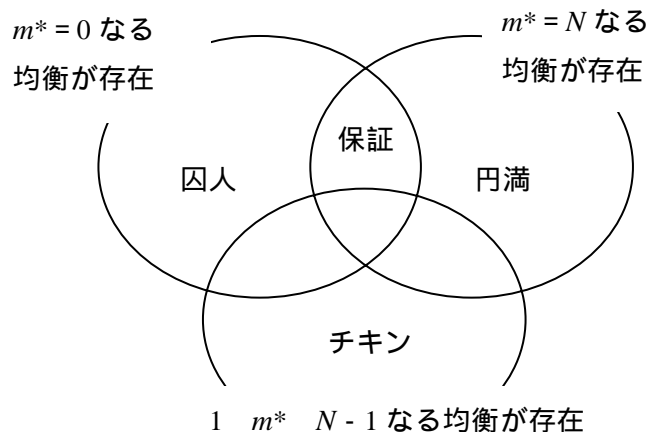


図1 協力状況のサブカテゴリー

N 人保証ゲームを N 人囚人のジレンマと N 人円満ゲームの「あいのこ」として定義する．

定義 6 N 人保証ゲーム 全員非協力 ($m^* = 0$) と全員協力 ($m^* = N$) が均衡で，かつ均衡はこれらのみである．

- N 人保証ゲームでは，非協力が**マクスミニ戦略**であることが証明できる (ibid)．そして，マクスミニ戦略を個人的合理性として考えることもできるので， N 人保証ゲームも社会的ジレンマとして考えられる (次項)．なお， N 人保証ゲームでは，よく知られた**分水嶺** (critical mass) が存在することが証明できる (ibid)．

N 人囚人のジレンマ， N 人円満ゲーム， N 人保証ゲームがかなり特殊な協力状況である一方， N 人チキンゲームは，かなり一般的な協力状況である．このような違いは，先行研究ではほとんど意識されていなかった． N 人チキンゲームの内訳は多様で複雑である．

4.2 N 人保証ゲームとマクスミニ戦略

Taylor (1987) は N 人保証ゲームそのものを社会的ジレンマとは考えていない． N 人保証ゲームでは全員 D だけでなく，全員 C も均衡になるからだ．一方，森村 (1987) は，社会的ジレンマというわけではないが，ホップズの自然状態が囚人のジレンマだけでなく，保証ゲームとしても考えられると論じている (ただし 2 人ゲームでの議論だが)．なぜなら表 3 にみるように，2 人保証ゲームでは，相手が協力しないときに自分が協力すれば，その結果は最悪となるからだ (したがってこのとき非協力 D はマクスミニ戦略である)．

このようにチキンゲームとは異なり，保証ゲームでは協力 C がリスクを伴う．この点を敷

衍して、 N 人保証ゲームを社会的ジレンマとみなすことができる。このとき、個人的合理性にあたるものが「マクスミニ戦略」である。**マクスミニ戦略**とは、それをとったときに起こりうる最悪の結果が最もましになるような戦略のことである。

マクスミニ戦略は、ゼロサムゲームのときのみ有効で、非ゼロサムゲームでは意味がないと一般には考えられている。しかし、**現実の相互行為状況では、かならずしも他者が合理的であるとは限らない**。そのような相手がナッシュ均衡戦略をとる保証はどこにもない。このような場合には、マクスミニ戦略を用いて最悪の結果を回避するという手は、ひとつの個人的合理性の表現である。

もっとも、相手に合理性を仮定できないときには、マクスミニ戦略だけではなく、似た構造をもつ「ラプラス戦略」と「マクシマクス戦略」にも個人的合理性があるといえよう。**ラプラス戦略**とは、それをとったときに起こりうる結果の平均が最もよくなるような戦略のことである。**マクシマクス戦略**とは、それをとったときに起こりうる最高の結果が最もよくなるような戦略のことである。

しかし、この3つのなかではマクスミニ戦略だけが「保証水準」を明確にすることができる。**保証水準**とは、当該の状況で相手がどんなことをしてきても、マクスミニ戦略をとることによって最低限確保できる利得である。このように、マクスミニ戦略は、 N 人ゲームという幾重にも不確定な状況に、保証水準という確定的な利得を保証する。この「確定性」、簡単にいえば「確実さ」という点で、マクスミニ戦略は、ラプラス戦略やマクシマクス戦略に比べて個人合理的であると考えるべきであろう。

つまり、非ゼロサムゲームでは無視されてきたマクスミニ戦略も、他者がかならずしも合理的に振舞うとは限らない現実の相互行為では、確実さという点で考慮に値する個人的合理性なのである。よって、**マクスミニ戦略の組がパレート非効率になる**状況を、社会的ジレンマの一種として考えることができる。

結局、保障ゲームでは、協利行為は個人的に最悪の結果を招きうる。そのため、だれもがそれを回避しようとするれば全員非協力 = パレート非効率が帰結するのである¹²（表8参照）。

表8 N 人保証ゲーム

C の人数 m	0	1	2	3
$D(m)$	0	0	0	-
$C(m)$	-	-10	-10	1

¹² なお、保証ゲームの議論でよく用いられるリスク支配戦略は、その前提にナッシュ均衡戦略をとることを想定しているため、すでに他者の合理性を前提にしている（Harsanyi and Selten 1988, 岡田 1995: 377）。そのうえで立場を交換して期待利得計算を比較することが必要であり、それじたいとしては、社会学的には合理性を要求しすぎる（本稿では、経済学におけるゲーム理論とは異なり、確率や利得和を行為者が計算するという仮定は入れない）。ただし、Kandori et al. (1993) のように、リスク支配戦略を進化論的に限定合理性の下で解釈する議論もあるので、リスク支配戦略にかんする議論すべてを棄却することはできない。しかし、本稿のような古典的なゲーム理論の枠組みでは、このリスク支配戦略は合理的すぎるゆえ、考えない。

4.3 N人円満ゲームと進化的合理性

繰り返しになるが、N人円満ゲーム（表9）は、社会的ジレンマとは通常考えられていない。ところが、進化的合理性を考慮すると、N人円満ゲームも一種の社会的ジレンマとして考えられる。なぜなら、進化的合理性の下では、協力が強支配戦略であるにもかかわらず、全員協力は均衡ではあっても安定ではないからだ（進化ゲームの解概念は均衡ではなく安定性である）。すなわち、だれかが全員協力から逸脱すれば、すべての者の利得は下がるけれども、あらゆる社会状態において非協力をとる逸脱者は協力者より利得が高いために、利得単調性（模倣や学習や教育）にもとづいて増加していく。このとき、協力者は淘汰され、全員非協力が実現する。これはN人円満ゲームと2人円満ゲームのランダムマッチングとの大きな違いである¹³。

このことは、N人円満ゲームだけでなく、すべての協力状況においていえる。したがって、進化的合理性を考慮するならば、条件[1][2]は、社会的ジレンマであるための十分条件でもあることになる（4つの個人的合理性：進化的合理性、マクスミニ戦略、強支配戦略、先手コミットメントは互いに打ち消しあうものではない）。つまり、進化的にみれば、[1][2]は社会的ジレンマの必要十分条件なのである。

表9 N人円満ゲーム

Cの人数 m	0	1	2	3
$D(m)$	0	2	4	-
$C(m)$	-	1	3	5

4.4 過剰協力・集団淘汰・制度

2.5節で述べたように、協力はある状況において全員非協力 $m^*=0$ よりも悪い結果をもたらすことがある。たとえば、野外パーティのあと、全員で芝生のうえのゴミを拾ったならば、芝生がダメになってしまい、だれにとっても最悪の結果となる（Elster 1989=1997: 157）。これは重度の過剰協力である。重度の過剰協力は、N人チキンゲームの一種であるN人英雄ゲーム（表10）、N人囚人のジレンマ（表11）、N人保証ゲーム（表12）など円満ゲーム以外の協力状況で起こりうる。

表10 過剰協力（N人英雄ゲーム）

Cの人数 m	0	1	2	3
$D(m)$	0	2	1	-
$C(m)$	-	1	0	-1

表11 過剰協力（N人囚人のジレンマ）

Cの人数 m	0	1	2	3
$D(m)$	0	2	4	-
$C(m)$	-	-1	1	-1

表12 過剰協力（N人保証ゲーム）

Cの人数 m	0	1	2	3	4
$D(m)$	0	2	2	-2	-
$C(m)$	-	-1	1	-3	-1

¹³ 強支配戦略の組は、ランダムマッチング下のレプリケーターダイナミクスにおいて漸近安定である。なぜなら、強支配戦略以外の強支配される戦略はすべて淘汰されるからである。Weibull（1995=1998: 103）参照のこと。

進化的合理性を考えると、このような重度の過剰協力のある社会的ジレンマでは、集団淘汰（group selection）による協力状態の実現も不可能になる¹⁴。なぜなら、集団淘汰では、当の集団が協力者のみからなる場合にだけ進化する（シェアを拡大する）が、重度の過剰協力がある場合では、協力者のみの集団は非協力者のみの集団に淘汰されるからである。

したがって、重度の過剰協力が存在する場合には、協力者と非協力者が順番に交代するような制度や協力者に報奨を与えるような制度を導入しなければ、社会的ジレンマは解決できない。しかし、これは上位下位あるいは指導者を要する組織や民主主義のような制度を考えていくうえできわめて示唆に富むのではないだろうか。これは非対称性としての権力の問題とももちろん関係してくる。また、集団は国家や企業やNPOなど、社会学の中心的概念でもある¹⁵。従来の研究では、進化は慣習と関係づけられてきた一方、集団や制度にはあまり接点がなかった。この重度の「過剰協力」を介して、進化・集団・制度の関係を議論していくことができるだろう。

5. おわりに

本稿の内容をまとめておこう。本稿では、まず、フリーライダー問題に準拠して、「協力的行為」が定義される状況としての「協力状況」を、（ ）協力のコストと（ ）共通の利益の存在（ $= [1][2]$ ）によって定義した。つぎに、これを起点として、協力状態を定義し、さらに協力状況のサブカテゴリーとして、 N 人囚人のジレンマ、 N 人円満ゲーム、 N 人保証ゲーム、 N 人チキンゲームを均衡の配置によって操作的に定義した。これらはそれぞれ異質な個人的合理性によって特徴づけられることにより、社会的ジレンマとして考えることができた（表13にまとめた）。

表13 協力状況のサブカテゴリーと個人的合理性

ジレンマ分類	操作的定義	個人的合理性
N 人囚人のジレンマ	均衡が $m^* = 0$ のみ	強支配戦略
N 人円満ゲーム	均衡が $m^* = N$ のみ	進化的合理性
N 人チキンゲーム	均衡が $0, N$	先手コミットメント
N 人保証ゲーム	均衡が $m^* = 0$ と N のみ	マクスミニ戦略

注：うへの定義は協力状況を前提する。社会的合理性はパレート効率性である。

¹⁴ 社会学における集団淘汰のモデルには、大浦（2002）、金井（2003）などがある。

¹⁵ 佐藤（1995）は、ウェーバーのプロテスタンティズムの議論を発展させ、ゼクテという集団が会社の母体となることが、近代資本主義社会の成立にとって重要であったことを明らかにしている。「会社によって社会をつくる」のである。ゼクテ（会社）では、集団間を比較的自由に移動できる。本報告書に本稿とは別の論文として収められている諸々の実験的研究は、このような文脈から解釈できるだろう。

この枠組みから結果的に、社会的ジレンマと協力状況についてつぎの3点がわかった。

1. Dawes (1980) の「社会的ジレンマ」と「協力状況」は独立である。
2. Taylor (1987=1995) の先手コミットメントやナッシュ均衡戦略などの多元的な個人的合理性を考慮することで社会的ジレンマ概念を拡張・精緻化すると、社会的ジレンマは協力状況に含まれる。
3. さらに、進化的な合理性を個人の多元的な合理性にカウントすると、社会的ジレンマは協力状況に一致する (N 人円満ゲームでさえ社会的ジレンマといえる)。

また、過剰協力と集団淘汰の関係についてつぎの点がわかった。

4. 協力は、個人にとってリスクがあるだけでなく、全体にとってもリスクがある重度の「過剰協力」をひきおこすこともある。重度の過剰協力が起こる状況では、集団選択 (group selection) によっても協力状態は実現できない。

本稿は、これまで曖昧になっていた社会的ジレンマ研究における協力行為概念、協力状態概、諸 N 人ゲームをシステムティックに定義し、多元的な個人的合理性という視点からこれらの N 人ゲームを社会的ジレンマとして位置づけた。これにより、社会的ジレンマ研究の見通しがよくなれば幸いである。

今後の課題は、本稿の基礎概念と木村 (2002) などの現実の現象に注目した研究とを融合しながら、経験的な研究に生かしていくことである。なお、個々の協力状況のサブカテゴリーについての定理の証明は、武藤 (2005b) を参照してほしい。

付録 協力 (cooperate) の辞書的な意味

いくつかの英語辞典で、協力 (cooperate) の意味を調べた。共通の目的、共通の利益などを太字にしてある。

The American Heritage® Dictionary of the English Language

1. To work or act together toward **a common end or purpose**.
2. To acquiesce willingly; be compliant:
3. To form an association for **common, usually economic, benefit**.

Encarta® World English Dictionary, North American Edition

1. work together: to work or act together to achieve **a common aim**.
2. comply: to do what is asked or required

The Wordsmyth English Dictionary-Thesaurus

To work with others, esp. for **a common purpose**.

Infoplease Dictionary

1. To work or act together or jointly for **a common purpose or benefit**.
2. To work or act with another or other persons willingly and agreeably.
3. To practice economic cooperation.

Cambridge Dictionary of American English

To act or work together for **a shared purpose**, or to help willingly when asked

Rhymezone

Work together on **a common enterprise** or project

Cambridge International Dictionary of English

To act or work together for **a particular purpose**, or to help someone willingly when help is requested:

Longman Dictionary of Contemporary English

1. If two people or two groups cooperate, they work together in order to achieve **a result that is good for both of them**.
2. To help someone willingly when they ask for your help.

Merriam-Webster's Online Dictionary, 10th Edition

1. To act or work with another or others : act together
2. To associate with another or others for **mutual benefit**

また、「共通の利益」を明示しないほかの辞書の定義も参考にあげておく（下線は引用者）。

UltraLingua English Dictionary

To work together in harmony; to collaborate; “To succeed at this task, all members of the group must cooperate.”

Online Plain Text English Dictionary

To act or operate jointly with another or others; to concur in action, effort, or effect.

【文献】

Dawes, R. M., 1980, “Social Dilemmas”, *Annual Review of Psychology*, 31: 169-193.

土場学, 1995, 「社会的価値と合理性：社会的ジレンマに関する実験社会心理学的研究とゲーム理

- 論的研究の対話の試み」『理論と方法』10(2): 115 - 132 .
- Elster, J., 1989., *The Nuts and Bolts for the Social Sciences*, Cambridge University Press. (= 1997 , 海野道郎 訳 『社会科学の道具箱』 ハーベスト社 .)
- Harsanyi, J. C., 1977, *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge: Cambridge University Press.
- Harsanyi, John C. and Reinhard Selten, 1988, *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, Massachusetts.
- 金井雅之, 2003, 「進化ゲームにおける選択的相互作用モデルの意義と課題」『理論と方法』18(2): 153-167 .
- Kandori, Michihiro, George J. Mailath, and Rafael Rob, 1993. "Learning, mutation, and long-run equilibria in games", *Econometrica*, 61: 29 - 56.
- 木村邦博, 2002, 『大集団のジレンマ』 ミネルヴァ書房 .
- 永田えり子, 1988, 「自由と効率」『理論と方法』3(1): 43-56 .
- 丸田利昌, 1991, 「個人主義的社会理論と Nash 均衡概念」『理論と方法』6(1): 21-38 .
- 三隅一人, 1993, 「社会的ジレンマモデルの一般化」『理論と方法』8(1): 89-92 .
- 武藤正義, 2004. 「社会的動機による 2×2 対称ゲームの変換」数土直紀編 『社会理論の実践的可能性の探求』平成 13～15 年度科学研究費補助金研究成果報告書, 学習院大学, 103-130 .
- 武藤正義, 2005a, 「秩序問題としての社会的ジレンマの一般性：《二重の選択性》の視点から」大浦宏邦編 『秩序問題への進化ゲーム理論的アプローチ』平成 14～16 年度科学研究費補助金・研究成果報告書, 帝京大学 .
- 武藤正義, 2005b, 「社会的ジレンマの諸定理：フリーライドの視点による N 人ゲームの構造分析 II」大浦宏邦編 『秩序問題への進化ゲーム理論的アプローチ』平成 14～16 年度科学研究費補助金・研究成果報告書, 帝京大学 .
- 岡田章, 1996, 『ゲーム理論』有斐閣 .
- 大浦宏邦, 2002, 「農耕戦略の進化モデル：非ランダムモデルによるジレンマの回避」『帝京社会学』16: 97-120 .
- Raub, W. 1988. "Problematic Social Situation and the Large Number of Dilemma: A Game-theoretical Analysis." *Journal of Mathematical Sociology*. 13(4): 311-357.
- 佐藤俊樹, 1993, 『近代・組織・資本主義』 ミネルヴァ書房 .
- 盛山和夫・海野道朗編, 1991, 『秩序問題と社会的ジレンマ』 ハーベスト社 .
- 鈴木鉄忠, 2003, 「集合行為問題のゲーム理論的分析及び社会学的考察」平成 14 年度東京工業大学社会理工学研究科修士論文 .
- 瀧川哲夫, 1989, 「N 人ゲーム状況の構造解析」『北海道大学文学部紀要』37: 77-146 .
- Taylor, M. 1987. *Possibility of Cooperation*. Cambridge University Press. (= 1995 松原望訳 『協力の可能性』木鐸社 .)
- Weibull, J., 1995, *Evolutionary Game Theory*, M.I.T. Press. (大和瀬達治監訳 . 1998. 『進化ゲームの理論』文化書房博文社 .)
- 海野道郎, 1988, 「社会的ジレンマの数理モデル：共有物の性質とモデルの対応関係」『社会学研究』52 .
- 山岸俊男, 1989, 「社会的ジレンマ研究の主要な理論的アプローチ」『心理学評論』32(3):262 - 294 .

Social Dilemmas and Cooperation Concept:

N-person Game Structure Analysis from the viewpoint of Free Ride I

Masayoshi MUTO

Graduate School of Decision Science and Technology
Tokyo Institute of Technology
2 - 12 - 1 Oookayama, Meguro, Tokyo 152 - 8552, JAPAN