

外生的資源制約を課した環境ジレンマゲームに関する研究

脇山 宗也[†] , 谷本 潤[‡] , 相良 博喜[†]

プレイヤー間の相互作用を 2x2 ジレンマゲームとし , これに外生的資源制約を課すことで一般的な環境ジレンマゲームのモデルを構築した . ジレンマが弱いゲームでは , 環境の状態に応じて行動を切り替えるサステナブルな戦略が進化するが , ジレンマが強くなると常に環境から資源を収奪しようとする戦略が優勢となる .

A research on Environmental Dilemma Game with exogenous resource restriction based on 2x2 game.

MOTOYA WAKIYAMA, JUN TANIMOTO and HIROKI SAGARA

A general Environmental Dilemma Game with exogenous resource restriction based on 2x2 game is established. A series of simulations reveals that a sustainable strategy allowing behavior-shifting to adjust environmental capability has been evolved. Whereas, only a set of dangerous strategies always trying to obtain resources has been prevailed.

1 . 緒言

社会ジレンマは一般に n 人型ゲームでモデル化され , その構造上 , 供給型と保存型ジレンマに大別される ([1],[2]) . 前者は社会科学では Olson 問題[3]と云われ本質的には n 人囚人ジレンマゲーム (n -Prisoner's Dilemma, n -PD) である Public Goods Game (PGG) で , 後者は本質的には n 人チキンゲーム (n -Chicken) である Tragedy of Commons (TOC) でモデル化される . 所謂 , 環境問題は後者の n -Chicken の性質を有する . このことは , 他者を出し抜くことが自己にとっての最良の帰結である一方 , 出し抜かれることより互いに裏切り合って環境そのものが崩壊して共倒れの状況になることが最悪の帰結である , とのゲーム構造が Chicken のジレンマ[4]により表されることから理解されよう .

TOC では , 環境には定率の資源回復があり , 社会総和がその範囲内であれば各プレイヤーは自己の要求通りの利得を得られるが , その閾値を超えると制約を受ける , との力学が組み込まれている . しかし , 環境の資源回復能力が時々刻々変容する環境の状態量に応じて変化 (例えば , 環境破壊が進むと回復能力も大き減退する) する一般的な環境問題のジレンマをモデル化するには不十分である . これに対し , Akiyama& Kaneko の Timberjacker's Dilemma [5], Tanimoto の Environmental Dilemma Model [6]では , 回復能力を資源状態量の関数で付与することで , 環境の資源制約を外生化したモデルを構築している . これらは具体的適用対象の比喩が強き意識されているため , エージェント間で繰り広げられる本体のゲーム構造の抽象化が不徹底で , 例え

[†] 九州大学大学院総合理工学府環境エネルギー工学専攻・修士課程
Graduate Student, Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

[‡] 九州大学大学院総合理工学府環境エネルギー工学専攻・教授・工博
Prof., Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Dr.Eng.

ば, Prisoner's Dilemma (PD) や Chicken など主として 2x2 ジレンマゲームで蓄積されてきた既往知見 (ジレンマ解消のための様々なゲームオプションに関する知見) が適用し難いとの問題点がある. 本論では, 資源制約を外生化しつつゲーム本体構造に 2x2 ゲームを適用した一般的環境ジレンマゲームのモデルを提示し, 併せてその基本的特性を報告するものである.

2. モデル

Player エージェントが互いに 2x2 ゲームを繰り返しながら, 自己の利得最大化を目指して, 有限の資源を環境から奪い合う, 或いはサステナブルな範囲で資源を獲得する (環境が危機に瀕すると環境護持的行動をとる) と云ったエージェントの行動戦略を進化させるエピソードをモデル化する. 環境は有限なストックを有し, 環境の状態量 (環境の資源保持量) の関数で表される回復能力をもつ.

2-1 ゲーム構造およびエージェント

エージェントは 2x2 ゲームを行う. 従って, エージェント数は 2 以上の偶数 N とする. 協調 (Cooperation, 以下 C) もしくは裏切り (Defection, D) を手として出す 2x2 ゲーム構造は, 自他の手組み (D, D), (C, C), (C, D), (D, C) に対する利得をそれぞれ P (Punishment), R (Reward), S (Saint), T (Temptation) で表す. ここで, ゲーム構造は $\mathbf{M} = \begin{bmatrix} R & S \\ T & P \end{bmatrix} \equiv [R, S, T, P]$ なるマトリクスで, エージェント i の C か D かの手を $s_i = {}^T(1 \ 0)$ あるいは $s_i = {}^T(0 \ 1)$ なる縦ベクトルで表す (T は転置を表す). また, 本ゲームでは, ゲーム構造による定和性を保存するため, 以下の制約条件を付しておく.

$$P + R + S + T = 0 \quad \dots(1)$$

ゲーム構造 \mathbf{M} は, 谷本・相良[2] (図 1) に基づき, 2 つのパラメータ $r = r_2 / r_1$, $[\text{deg}]$ にて (2) 式で表す. 尚, $x_0 = (P+R)/2$ とする.

$$P = x_0 - 0.5 \cdot r_1 \cdot \cos(45^\circ) \quad \dots(2-1)$$

$$R = x_0 + 0.5 \cdot r_1 \cdot \cos(45^\circ) \quad \dots(2-2)$$

$$S = x_0 + r_2 \cdot \cos(45^\circ + \theta) \quad \dots(2-3)$$

$$T = x_0 + r_2 \cdot \sin(45^\circ + \theta) \quad \dots(2-4)$$

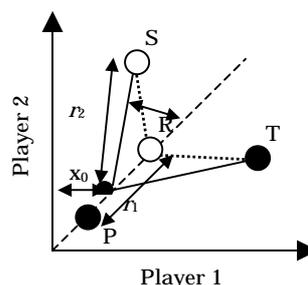


図 1 対称 2x2 ゲームの一般表記

Fig.1 General description for Symmetric 2x2 Game

エージェント i は k タイムステップにおいて, 環境から資源 $Ob_{i,k}$ を獲得する (獲得資源 $Ob_{i,k}$ の詳細は 2-3 節で述べる). $Ob_{i,k}$ は正值, 負値いずれも採り得る. 負値は各エージェントから環境への資源返還を意味する (植林その他の比喻).

2-2 エージェントの学習

各エージェントは戦略を有する. 戦略は自分 (focal) と対戦相手 (opponent) の前時間ステップの手組み (P, R, S, T のいずれか) と環境の状態 (現在の環境の資源保持量 Env_k がある閾値 $Env_{threshold}$ より大きい小さいか) により, C か D かの手を規定する行動ルールである. また, これとは別に対戦相手との初手を規定しておく. 従って, エージェントの振る舞いは, 前回対戦結果と環境の状態量を入力に, C か D かの行動を出力とする状態遷移機械 (Finite State Machine, 以下 FSM) で決定され, その戦略は $2^2 \cdot 2 + 1 = 9$ bit の文字列で表される (戦略の bit 表記は図 4 参照).

戦略の進化には遺伝的アルゴリズム (Genetic Algorithm, 以下GA) を適用する。各エージェントは個別に m_g 個の戦略で構成される遺伝子プールを有する。遺伝子の適合度は、自己を除く $N-1$ 人中からランダムに選択した相手と I_g 回繰り返しゲームを行い、これを1タームとして毎回相手をランダムに変えながら I_T ターム行った間に得た利得の総和、すなわち、 $I_g * I_T$ 回の繰り返しゲーム期間中の利得総和とする。GAの遺伝子操作はBoltzmann選択(温度係数 T_B とする)による自然淘汰、一点交叉(交叉確率 p_{cross})、突然変異(各遺伝子に対して突然変異確率 p_{mut})を適用する。

2-3 環境

環境の資源保持量 Env_k は $[0, Env|_{max}]$ なる連続量を取る。環境はエージェントからの収奪 Rob_k を受け一方、自らの回復能力 Rc_k により保持する資源量を回復させる。タイムステップ k における環境が保持する資源の量を Env_k とすると、(3)式で更新される。

$$Env_{k+1} = Env_k - Rob_k + Rc_{k+1} \quad \dots(3)$$

ただし、初期資源量は上記のタームごとに $Env|_{max}$ にリセットされる Env_{k+1} が0を下回る、或いは $Env|_{max}$ を上回る場合は、それぞれ $Env_{k+1}=0$ 、 $Env_{k+1}=Env|_{max}$ とする。 Rob_k はエージェントの収奪(あるいは返還)する資源の総量であり、(4)で与えられる。

$$Rob_k = \sum_i^N Ob_{i,k} \quad \dots(4)$$

Rc_k は現在の環境の資源保持量 Env_k に依存し、(5)式で与える。ただし、 α はモデル定数である。

$$Rc_k = \alpha \times Env|_{max} \times \left(\frac{Env_k}{Env|_{threshold}} \right) \quad \text{if } Env_k < Env|_{threshold}$$

$$= \alpha \times Env|_{max} \times \left(1 - \frac{Env_k - Env|_{threshold}}{Env|_{threshold}} \right) \quad \text{else}$$

...(5)

この資源回復関数は、 $Env = Env|_{threshold}$ で最大となり、その点から環境が豊穡になるほど、もしくは枯渇するほど資源の回復は緩やかになることを示す。

ここで、エージェント j と対戦するエージェント i の獲得資源量 $Ob_{i,k}$ は(6)式のように与える。

$$Ob_{i,j} = 0 \quad \text{if } Env_k - Rob_k < 0 \ \& \ Ob_{i,k} > 0$$

$$= {}^T \mathbf{s}_j \cdot \mathbf{M} \mathbf{s}_i \quad \text{else}$$

...(6)

(6)式の意味するところは、以下の通りである。当初、環境が豊穡な状態では、エージェントは2x2ゲームの結果に応じて環境との間に資源をやり取りすることができる。しかし、環境が危機的状態に窮すると、資源を収奪することを強制的に停止され、資源の返還のみ許される(以下、環境破壊とする)。

3. 数値実験

3-1 計算条件

仮定したパラメータを表1に示す。結果は1試行1000世代の5試行アンサンブル平均である。

表1 計算パラメータ
Table1 Assumed parameters

$Env _{max}$	$5 * N * R (=100)$	N	10
$Env _{threshold}$	$Env _{max} / 2 (=50)$	p_{cross}	0.5
I_g	100	p_{cross}	0.5
I_T	10	T_B	0.0001
m_g	2000		
			$\frac{3}{4} * N * R / Env _{max} (=0.15)$

3-2 結果及び考察

特定の2条件のMについて論及した後、MをPDの領域内で連続的に変化させながら行った結果について述べる。

3-2-1 M=[2.0,-4.0,4.0,-2.0]における結果

図2, 図3は夫々M=[2.0, -4.0, 4.0, -2.0] (以下条件1)とした場合の協調率 P_c の世代推移と, 環境破壊確率及び社会平均利得(1ゲーム・1エージェントあたりの利得, 最大値はTに相当)の世代推移である. 尚条件1のMはPrisoner's Dilemma (PD)の構造を持つ(図9中で+マークしたゲーム構造). エージェントは環境が豊穡な場合($Env_k > Env|_{threshold}$), 環境が貧少な場合($Env_k < Env|_{threshold}$)の両状況において協調率が高くなるように学習は推移する(図2). 1000世代終了時の環境の破壊確率は0.5と高く, 1ゲームあたりの利得はRの1/4の0.5程度である(図3).

協調率が高いにも拘らず, 獲得利得が公平なPareto解であるRに比べ低いのは, 資源制約が利いて(6式), 環境破壊が生じている状況下ではそれ以上の資源収奪が行い得ないことに起因する. つまり, 通常の2記憶長メモリFSMでPDを行った場合, エージェントたちは互いにCを出してRの利得を得るように学習が推移するわけだが, この様な状況が長く続くと環境破壊は不可避免的に生じ, 環境の回復能力がそれに耐え得なければ環境破壊は常態化する. 条件1では $S+T=0$ となるので, エージェント間で優勝劣敗が生じている限りにおいては, 環境へのインパクトはない. むしろ, エージェント内に協調(Rなるゲームの結果)が創発し, それが継続されることが問題なのであり, 条件1ではまさにそのような状況が生起しているものと考えられる.

条件1における優勢な戦略を見てみる. 9bit戦略のうち以下の3bitに注目して, 全戦略を8クラスにカテゴリ化した(図4). すなわち図4に示すように, (1) $Env_k > Env|_{threshold}$ で前ステップのゲームがPであるときの次手(第2bit),

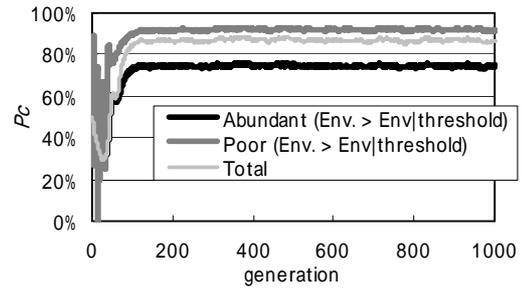


図2 協調率の世代推移
Fig.2 Transition at generation of P_c

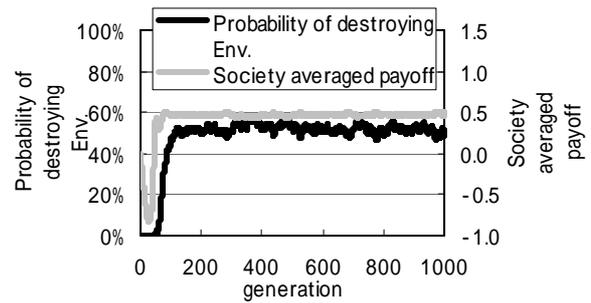


図3 環境破壊確率及び社会平均利得の世代推移
Fig.3 Transition at generation of Probability of destroying Env. and Society averaged payoff.

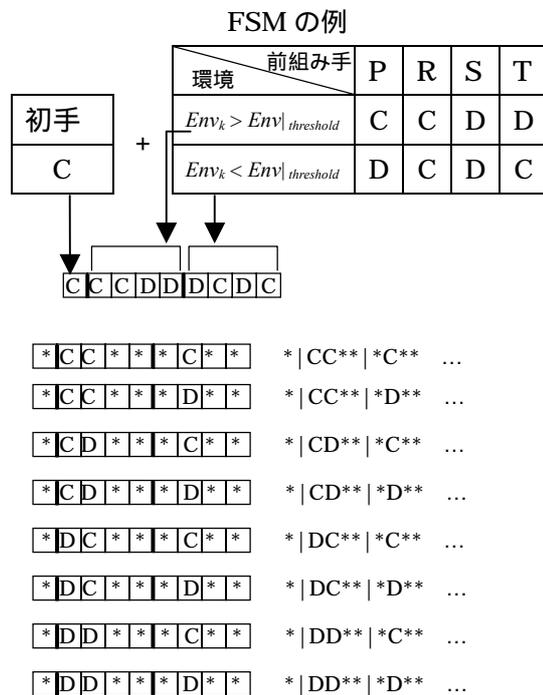


図4 戦略の類別パターン

(2) $Env_k > Env|_{threshold}$ で前ステップのゲームがRであるときの次手(第3bit), (3) $Env_k < Env|_{threshold}$ で前ステップのゲームがRであるときの次手(第

7bit), により類別した 8 クラスである. 図 4 の*はワイルドカード(C or D)とする.

図 5 は 8 クラスの戦略比率の世代推移である. 条件 1 では “*|DC**|*C**” (以下戦略) 及び “*|CC**|*C**” (以下戦略) に属する戦略がマジョリティを占めている. これらの両クラスは環境が豊穡な場合にも, 貧少な場合にも, R の次手には C を出す点で共通している (第 3bit と第 7bit). 戦略 及び は, 環境が悪化しようとも収奪を止めようとしない, 環境に対する “貪り派” の戦略であり, 彼らが社会のマジョリティを占めたため, 条件 1 では高い環境破壊確率と低利得となった.

3-2-2 M=[2.0,-2.4,2.4,-2.0]における研究

図 6, 図 7 は M=[2.0, -2.4, 2.4, -2.0] (以下条件 2)とした場合の図 2, 図 3 に当たる. 尚条件 2 の M も条件 1 同様 PD である(図 9 中で・マークしたゲーム構造)が, 条件 1 より r が小さな, よりジレンマが弱いゲーム構造である. 条件 2 の協調率をみると, 豊穡な場合は条件 1 と同様高いが, 貧少な場合は低いことが特徴的である(図 6). 1000 世代終時の環境破壊確率は条件 1 と対照的にほぼ 0 で, 社会平均利得は約 1.2 程度であることが分かる(図 7).

図には示していないが条件 2 では, 環境は豊穡な範囲(下限に近く)と貧少な範囲(上限近く)を時間ステップごとに交互に繰り返すような推移を示す. これは, 後段述べるように条件 2 では優勢な戦略 により, 豊穡な範囲で資源の収奪, 貧少な範囲で資源の返還を切り替えて行うので, 結果的に環境は豊穡と貧少の切り替え点 $Env|_{threshold}$ 近くを移ろうことになる. 環境が, 資源回復能力が最大となる $Env|_{threshold}$ 近くあるため, エージェントたちは持続可能且つ効率的な資源収奪が可能とな

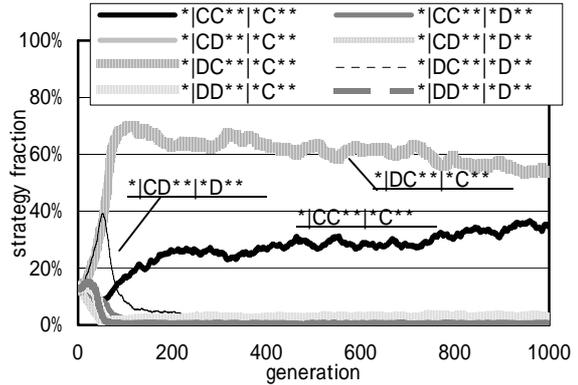


図5 戦略比率の世代推移
Fig.5 Transition at generation of strategy type fraction

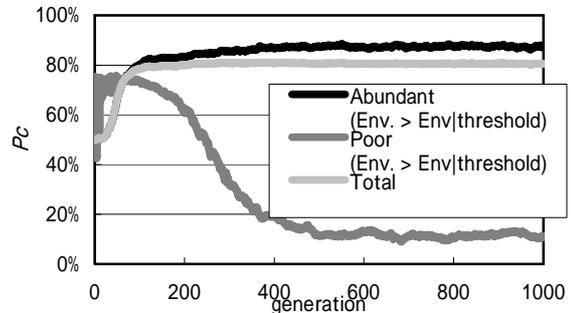


図6 協調率の世代推移
Fig.6 Transition at generation of Pc

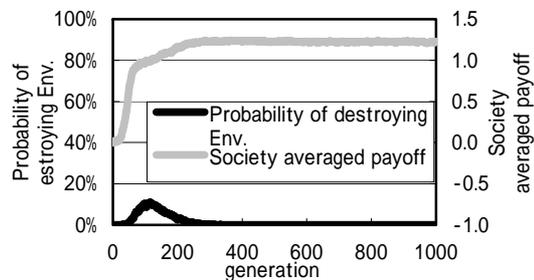


図7 環境破壊確率及び社会平均利得の世代推移
Fig.7 Transition at generation of Probability of destroying Env. and Society averaged payoff

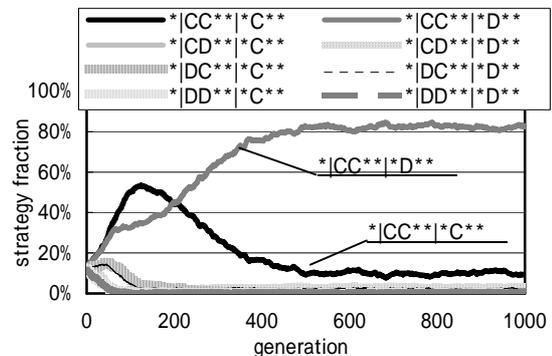


図8 戦略比率の世代推移
Fig.8 Transition at generation of strategy type fraction

っている。

図8は図5同様に示した戦略クラスの世代推移である。条件2では図4で示した戦略が優勢となっている。この戦略の特徴は、(1)第7bitがDであること、(2)第2bitと第3bitにCがあること、である。(1)により、環境が豊穡から貧少なエリアに切り替わっても、即座にそれまでの協調し合い(Cの出し合い)を破棄して、一転、Dを出し合い、互いに環境に資源を返還することで、環境を豊穡な範囲に復さしめる行動をとる。(2)により、環境が貧少から豊穡に復旧した直後、それまでのDの出し合いからCの出し合いに転じる(第2bitのC)し、環境に余剰があって暫くRを続け得る状況にあるときにはやはりCを出し続ける(第3bitのC)行動をとる。

図8を再び見ると、世代が進んでも、条件2では、条件1のように戦略と戦略以外は殆ど存在しなくなるような社会構成ではなく、戦略が優勢を保持しながらも戦略が10%程マイノリティーとして残っている。戦略は、貧少な状況に切り替わってもCを出さず(第7bitがC)。従って、対戦ペアの組み合わせが戦略同士であれば、ほかの多数ペア(戦略同士)が環境危機に際して資源を返還するのに、その狭間をついて資源収奪を続けることが可能になる。すなわち free rider である。無論、戦略が戦略とペアを組んだ場合は、環境が貧少な状況に切り替わった際には戦略は戦略に貪られることになる。戦略あつての戦略であるから、マジョリティになることはない。このとき戦略が戦略同様に残れない理由は、豊穡な状況で資源回復直後に戦略同士のペアでは裏切り合いになる(第2bitがD)ためである。戦略がマジョリティとなる条件2では既述したように豊

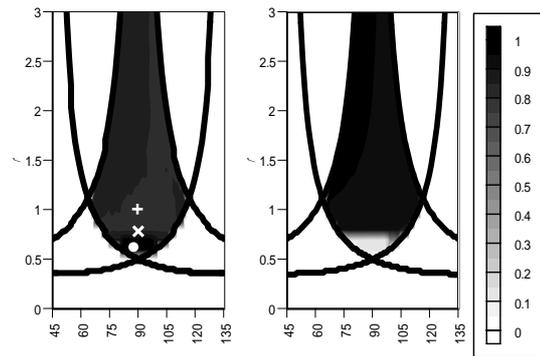


図9 豊穡状態における協調率
Fig.9 Pc in the case of Abundant Env.

図10 貧少状態における協調率
Fig.10 Pc in the case of Poor Env.

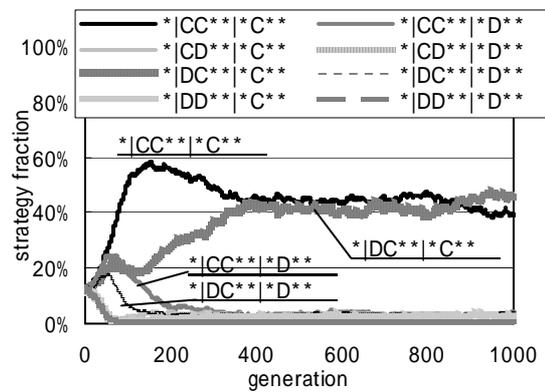


図11 戦略比率の世代推移

Fig.11 Transition at generation of strategy type fraction

穡、貧少の切り替えが頻繁に生起するので、戦略は早い段階で淘汰されてしまう。

以上述べたように戦略は環境が悪化した場合は資源の収奪を止め資源を保護しようとし、結果的に、環境との共生を達成して、自らも公平な Pareto 解に近い高い利得を維持している。すなわち “sustainable 派” の戦略である。

3-2-3 Mを連続的に変化させた場合の実験・考察

MをPDの領域内で連続的に変化させて、同様の実験を行った。MはR=2.0に固定してP, S, Tの値を連続的に変えることでゲーム構造を変化させた。データはrの値を0.2刻み、の値を2deg刻みで数値計算を行い、結果は5試行のアンサンブル平均である。尚先程の条件1及び条件2は(2)式より夫々r=1.0、

$\theta=90\text{deg}$ 及び $r=0.6$, $\theta=90\text{deg}$ に当たる。図 9 は豊穡状態における協調率, 図 10 は貧少状態における協調率を 951 世代から 1000 世代までで平均したものである。豊穡な場合において協調率は高く, ゲーム構造による違いはあまり見られない(図 9)。しかし, 環境が貧少な場合の協調率は, ジレンマ大となる(r が大きな) 領域では高いが, 相対的にジレンマ小さい(r が小さい) 領域では低くなっている(図 10)。このとき豊穡か貧少に拘らず協調率が高い部分は, 条件 1 同様に貪り派優勢に, 豊穡では協調率が高いが貧少では協調率が低いゲーム領域では条件 2 のように sustainable 派優勢の社会を形成する。

これらの 2 領域は r が 0.6~0.8 で切り替わる。ただし, この領域では戦略, 戦略, 戦略の併存平衡になるのではなく, 戦略優勢(条件 2) から戦略 および戦略 優勢(条件 1) が不連続に相変化する。

図 11 はこの遷移域に位置する $r=0.7$, $\theta=90\text{deg}$, すなわち $M=[2.0,-2.8,2.8,-2.0]$ なるゲーム構造(図 9 中で \times マークしたゲーム構造; 以下, 条件 3)における戦略クラスの世代推移である。

マーク・ \times + の順, すなわち, r を小から大へ(ジレンマを小から大へ) 変異させる状況を考える。

マーク・ では, ジレンマは大きくない。このような状況では, sustainable 派である戦略が優勢となる健全な社会が構成される。

ジレンマが大きくなると(マーク \times), 次第に S と T の差異が大きくなるので, 環境が貧少から豊穡に切り替わった直後に C を出す戦略 は第 2bit が D である戦略に貪りを受けるようになり, 遂には優勢たり得なくなる。替わってマジョリティを担うようになるのは戦

略 であるが, これも戦略 同様第 2bit は C であるから, 第 2bit が D の戦略 からの貪りを受ける。従って, 図 11 と図 8 の比較から明らかなように, この遷移域では, 戦略 は淘汰され, 戦略 と戦略 の拮抗状態の相を呈する。

4. 結論

既往の 2x2 ジレンマゲームに関する知見を有機的に適用するために, エージェント間のゲームは 2x2 とし外生的資源制約を課した一般的な環境ジレンマモデルを構築した。

エージェント間のゲームを PD 領域で様々変化させたところ, ジレンマが弱い領域では, 環境が豊穡な状況では資源を収奪し, 貧少の状況になると返還する, すなわち閾値前後で行動を切り替える sustainable な戦略が進化した。この場合の社会利得は公平な Pareto 解に近い高いものであった。しかるにジレンマが強くなると, 閾値の前後で常に資源を収奪しようとする貪り戦略が進化し, 社会の利得は前者に比べ低調な状況から脱することが出来ない,

【参考文献】

- [1] 神事直人, 1996, 「共有地のジレンマ」モデル再考-海野モデルの再検討と一般モデルの構築, 理論と方法 Vol.11 (2), 145-156.
- [2] 神事直人, 1997, 「保存」型ジレンマモデルの再検討, 理論と方法 Vol.12 (1), 15-30.
- [3] Olson, M., 1965, The Logic of Collective Action: Public Goods and the Theory of Groups, Cambridge (MA), Harvard University Press.
- [4] 谷本潤, 相良博喜, 2005, 環境問題への適用のためのジレンマゲームの基本構造に関する考察, 日本建築学会環境系論文集 No.595, 135-140.

- [5] Akiyama,E., Kaneko, K, 2002, Dynamical systems game theory 2 A new approach to the problem of the social dilemma, Physica D 167, 36-71.
- [6] Tanimoto,J, 2005, Environmental dilemma game to establish a sustainable society dealing with an emergent value system, Physica D 200, 1-24.