

## One-shot 2x2ゲームにおけるコミュニケーションに基づく協調の創発

谷本 潤

2層の状態遷移機械をもつ知的エージェントがOne-shotの2x2ゲームを行う過程で、ライトのオン・オフを切り替えることで対戦相手とコミュニケーションをはかり、それに基づく協調行動を創発させることをモデル実験により示した。このコミュニケーションに基づく協調行動の創発は、互いに協調と裏切りを交互に出し合うことが公平なパレート解となるLeaderやHeroゲームで発生する。

### Emergence of cooperation supported by communication on a One-shot 2x2 game.

Jun Tanimoto<sup>†</sup>

This paper reports that an intelligent agent equipped with 2-layer Finite State Machines (FSM) can emerge a communication by lighting on and off, which leads to a social cooperation that solve the dilemma situation modeled by a One-shot 2x2 game. This communication amid two gaming agents can be particularly observed in Hero and Leader type dilemma games where alternating reciprocity by repeating Cooperation (C)-Defeat (D) after D-C is the Equal Pareto Optimum instead of sequence of mutual cooperation that is the Equal Pareto Optimum for Prisoner's Dilemma (PD) game.

キーワード： One-shot 2x2ゲーム, 状態遷移機械, コミュニケーションの創発  
Key words: One-shot 2x2 game, Finite State Machine, Emergence of communication

### 1. 緒言

コミュニケーションは様々な動物種で観察される機構である。その最も進化した形態が人間の言語であると言えよう。コミュニケーション、さらには人間の言語は如何なる契機で、どのように進化してきたのだろうか？この問題は学際的にみてもきわめて興味深いチャレンジであり、これまで様々な分野で議論されて来、未だ全容を理解するに至っていない（例えば1）。例えば、人類の言語獲得と頭蓋底の屈曲形状の発達とを関連付けて論じる試みがなされてきたが、コミュニケーションが過去に遡及して物的証拠を得ることがそもそも難しいソフトウェアに過ぎないことから、そのような生物学的実証アプローチには困難がある。このような中で、Aritaら（例えば2）、3)の複雑系に基づく人工生命のアプローチを適用して言語の創発過程を解明しようとの試みは、前記を補完する手法として有効である。

Dorigoら<sup>4)</sup>はフェロモンによる蟻エージェント間のコミュニケーションをAnt Systemとしてモデル化した、循環セールスマン問題（traveling salesman problem）に適用した。以降、Ant Systemは多くの最適化問題に適用されている（例えば5）。これらは適合度を最適化するためにフェロモンによるコミュニケーションが創発してきたとの前提に立っている。

コミュニケーションそのものの創発機構を扱ったものとして、Grimら<sup>6)</sup>は捕食者の襲来や食糧資源に関する情報を他個体と共有するために、原初は意味を持たない行動（gesture）が次第にコミュニケーションとして意味を持つように進化していくプロセスをマルチエージェントシミュレーションにより示した。

Lenaersら<sup>7)</sup>は言語及び語彙の進化プロセスを模擬する進化ゲームの数学的枠組みを提示している。

van Baalenら<sup>8)</sup>は、鳥のAlarm callにおけるコミュニケーション戦略を空間型のChickenゲームで検討

している。裏切りへのインセンティブが小さければ集団内で単一のシングルが使われる。しかしゲーム構造上裏切りインセンティブが大きくなると、偽情報を出す裏切り戦略によりwarm callの信頼性は低下する。これに対して協調戦略を採る個体は信号のチャンネルを換えてコミュニケーションの確保に努めようと試みるが、裏切り戦略との化かし合いに陥り、系は不安定な振動（warm callの信頼性が維持される期間と信号の意味が失われる期間が交互に表れる）を繰り返す。さらに裏切りインセンティブが大きくなるとwarm callによるコミュニケーションシステムは崩壊し、Tower of Babel（がやがや声が溢れる状況の英語比喩）に至ることを明らかにした。

以上の既往研究では、個体の適合度を上げる機能としてコミュニケーションや言語を捉えている。人間を含む動物が繰り返す生存競争は、上記のモデル研究が前提としているように、自らの適合度を上げるためのジレンマゲームと看做すことが出来るだろう。ジレンマゲームは進化ゲーム理論を基礎とし、ここ10年非常に多くの研究が積み重ねられてきた。主にはOne-shotゲームでは裏切り（Defect, D）合いのNash均衡に陥ってしまうPrisoner's Dilemma（PD）が取り上げられ、どのような付加機構を入れるとジレンマが解消・緩和する（互いに協調Cooperation, C出来るか）か、そのジレンマ解消プロトコルを明らかにする営為が、ジレンマゲーム研究の中心的な課題であった。繰り返しゲーム（例えば9）、ネットワーク（例えば10）、Tag（例えば11）、Image Score（例えば12）等々は、いずれも対戦相手を識別、ないしは制限し、協調するに足る相手とだけ選択的にゲームをすることで協調を創発させる仕組みである。原初のコミュニケーションも同様の機能として進化したのではないだろうか？

佐藤ら<sup>13)</sup>は縄張り争いゲームにおいて、ライトの点灯とその認知機構を持つエージェントが同一の相

<sup>†</sup> 九州大学大学院総理工学研究院・教授・工博  
Prof., Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Dr.Eng.

表 1 2x2 ゲームの利得表  
Table 1. Payoff matrix for a 2x2 game.

Ego		Opponent		
		C	R	D
C	D			
	S			

手と繰り返し対戦をするプロセスを通じて、衝突を避けて交互に陣地を譲り、譲られる行動を創発させ、その行動がライトの点減と同期することを数値実験により示している。このゲームは、非対称な状況に対戦相手と平均的に取り合う状況がパレート最適になっているという点で、Chicken 型のジレンマをもつ Leader や Hero<sup>14)</sup>と同様の構造をもつゲームを繰り返し対戦している状況と考えられる。C と D を手とする 2x2 ゲームの利得表を表 1 のように定義すると、Leader は  $T>R$  &  $S>P$  か  $T+S>2R$  &  $T>S$  (一方が指導者 (D) 他方が服従者 (C) なる役割分担が出来ればよいが両者とも指導者になろうとするインセンティブがあるために生じるジレンマ)、Hero は  $T>R$  &  $S>P$  か  $T+S>2R$  &  $S>T$  (一方が他利をするヒーロー (C) 他方がそれを受け入れる者 (D) なる役割分担が出来ればよいが両者ともヒーローになろうとするインセンティブがあるために生じるジレンマ) なるゲームである。Leader と Hero は同一の相手と繰り返し対戦するならば、互いに C と D を交互に出し、S と T の利得を平均的に分け合うこと (S・T 互恵) が公平なパレート解となっており、互いに C を出し合うこと (R 互恵) が公平なパレート解となる PD ( $T>R>S>P$ ) とは異なる。ここで佐藤らの言説から敷衍して思量されるのは、コミュニケーションと S・T 互恵との関連性 (翻って R 互恵を前提にする PD ではコミュニケーションは創発しないのだろうか)、さらには相手の特性を正確に識別するプロトコルとしてコミュニケーションが機能するのなら、匿名の相手と対戦する状況でも互恵が創発する (間接互恵) ではないかと言う疑問である。

本稿では、2 層の情報入力-行動出力系で表される状態遷移機械 (Finite State Machine, FSM) と遺伝的アルゴリズム (Genetic Algorithm, GA) を組み合わせた学習エージェントが様々な構造の One-shot の 2x2 ゲームを行う過程で、原始的コミュニケーションを創発させ、これにより社会的な互恵関係が進化していくエピソードをモデル化し、上記の点について考

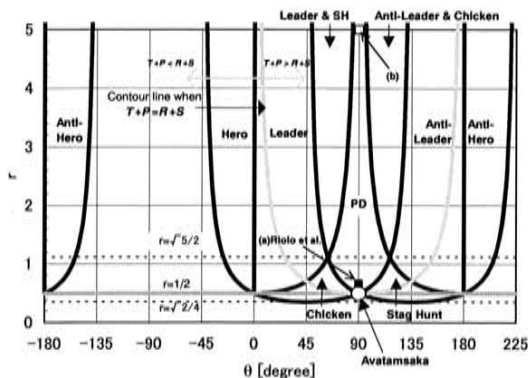


図 1  $\theta$  と  $r$  による 2x2 ゲームの表記.  
Fig.1 The scene of 2x 2 game world.

察するものである。

## 2. モデル

### 2.1. 2x2 ゲーム

全て 2x2 ゲームのゲーム構造は、

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

$$R = x_0 + 0.5 \cdot r_1 \cdot \cos(45) \quad \dots(1-2)$$

$$S = x_0 + r_2 \cdot \cos(45 + \theta) \quad \dots(1-3)$$

$$T = x_0 + r_2 \cdot \sin(45 + \theta) \quad \dots(1-4)$$

のようにパラメータ  $\theta$  [deg] と  $r = r_2 / r_1$  で表すことが出来る<sup>14)</sup> ( $x_0$  は解可能域がシフトするだけで  $P, R, S, T$  の相対関係上意味がないので、実質的には 2 パラメータとなる)。これら 2 パラメータで表されるゲーム空間を表すと図 1 のようになる。太線で囲われ、ジレンマゲーム名が冠されている領域以外が、ジレンマのない Trivial ゲームである。

### 2.2. ゲームのフレームと状態遷移機械

$n$  エージェントからなる完全混合集団を考える。エージェントは各自ライトを持ち、そのオン/オフ、およびゲームにおける C/D かの 2 つの行動を各ステップで行う。また、自分と対戦相手のライトのオン/オフ、前回の自分が行ったゲーム対戦の結果 ( $P, R, S, T$ ) を認知する受容器を持ち、これらは図 2 に示す 2 層の FSM を構成しているものとする。FSM #1 は前回ゲームにおける  $P, R, S, T$  が情報入力、ライト点減が行動出力の 4bit システム。FSM #2 は自他のライト点減が情報入力、ゲームにおける C か D が行動出力の 4bit システム。従って、各エージェントの遺伝子は合計 8bit である。各対戦ステップで、各エージェントは前回のゲーム結果に基づき自分のライトの点減を行い、自他のライトの点減状況を認識して自らのゲームにおける手 (C か D か) を意志決定する。

各エージェントは 1 世代中に集団中のランダムな相手と  $M$  回の One-shot ゲームを行い、獲得利得を適合度とする。各世代の期末に、各エージェントの FSM のビット列 (FSM #1 と FSM #2 を重ねた 8bit) は全集団をプールに GA を適用して進化操作を施す。その際、交叉確率は 1 で、FSM #1 を表す前半 4bit 部分と FSM #2 を表す広範 4bit 部分で各々 1 点交叉を行う。また、1% の確率で全 8bit 中の任意の文字列が反転する突然変異を考慮する。シミュレーションエピソードの初期には各エージェントの FSM はランダムにシャッフルしておく。

## 3. 実験

実験における共通のパラメータ設定は、 $n = 100$ ,  $M = 10$ ,  $x_0 = 0$ ,  $r_1 = 1.273$  とした。また、ゲームの手 (C か D か) 出し、ライト点減における行動過誤、認識過誤はともに 0 とした。1 エピソードは 20000 世代まで計算し、10000-20000 世代の平均値を取って擬似的な均衡の値とした (後述するように世代推移は安定的均衡を示すものではなく mimic の侵入により頻りに振動を示す相になるため)。また、次章で示す図 1 同様の全ゲーム領域のコンターで描いた結果 (図 5) は、5 試行のアンサンブル平均を取っている。

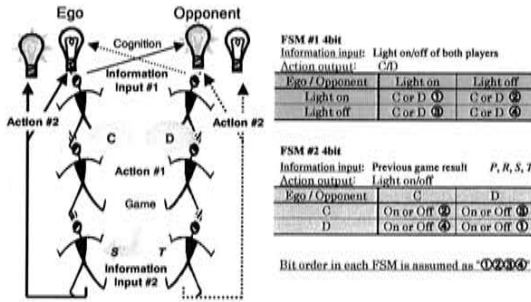


図2 FSMの構成とゲームの流れ。

Fig.2 Structure of assumed FSMs and game flow.

#### 4. 結果及び考察

##### 4.1. Hero と PD における世代推移比較

図3, 図4はあるエピソードにおける  $\theta=10deg$ ,  $r=1.0$  の Hero,  $\theta=90deg$ ,  $r=1.0$  の PD の 10000 世代までの世代推移である。夫々, (A)平均利得, 協調率 ( $P_C$ ), ライト点灯率 ( $P_{on}$ ), (B)P, R, S (T) の生起確率, (C)FSM #1 と FSM #2 の情報量 (注記1), (D)FSM #1 の各ビットにおいて“1”が占める割合, (E)FSM #2 の各ビットにおいて“1”が占める割合, (F)連続するゲームシーケンスにおいて S かつ ライト点灯 (on) → S かつ on, S かつ ライト消灯 (off) → S かつ off, T かつ on → T かつ on, T かつ off → T かつ off, S (もしくは T) かつ on (もしくは off) → T (もしくは S) かつ off (もしくは on) の生起確率, (G)連続するゲームシーケンスにおいて R かつ on → R かつ on, R かつ off → R かつ off の生起確率である。

PD (図4) では R 互恵 ( $P_C$  がほぼ 1) と裏切り合い ( $P_C$  がほぼ 0) の期間を交互に繰り返して、安定しない。これは、一旦、R 互恵が達成されても突然変異による mimic が侵入して、系全体は振動的な相を呈するからである。R 互恵の期間では互いにライトを点灯もしくは消灯することで C を出していることがわかる (図4(G))。 “合図”としてのライト点灯もしくは消灯がエージェント間で一度合意されると、ゲーム対戦のシーケンス内で mimic の侵入により崩壊するまで変化せず、点灯 (消灯) し放しとなって、点滅することはない。この点で、この現象は後に述べる Hero における S・T 互恵をコミュニケーションに準えたとするならば、表象と云うべきものである。換言すれば、この合図は世代推移に可塑性を有する Tag として機能し、これにより対戦相手を識別して R 互恵が達成されているのである。

では、如何にして一旦構築された R 互恵が mimic の侵入を受けて崩壊するのであろうか？ 図中に点線の囲みで示した 6000 世代後、8000 世代後の崩壊プロセスを詳しく見てみる。夫々、表象としての合図はライト点灯と消灯である (図4(G))。

前者では FSM #1 の 1 ビット目がほぼ 1, すなわち、自他共に点灯時に C 出す、と云う合図がエージェント内で合意されている (図4(D))。ところがこの期間末期に当初 FSM #1 の 4 ビット目はほぼ 0 であったのが 1 へと急激に変容している (図4(D))。つまり、自他共に消灯時でも C 出すエージェントが徐々に集団内で数を増やしている。自他共に点灯時に加えて消灯時にも C 出す戦略が多くを占めるようになると、それまで殆どが点灯されていたライトが

一部のエージェントが消灯するようになる。すると、点灯エージェントと消灯エージェント間の対戦が生じ、この場合は、互いに裏切り合いとなる。このようにして前者の R 互恵は崩壊に至ったと考えられる。

後者では FSM #1 の 4 ビット目がほぼ 1, すなわち、自他共に消灯時に C 出す、と云う合図がエージェント内で合意されている (図4(D))。この期間中、図4(E)をみると FSM #2 の 4 ビット目がほぼ 1, すなわち相手を食った T 獲得時に点灯するとのいわば“懺悔コール”が確立している。ところが、最終期に至って、この 4 ビット目が 0 の mimic が侵入して来、一気に拡散する。これにより何かの間違い (協調率は厳密には 1 ではない) で相手を食ってしまったときのセーフティネットが確信犯の出現により破られ、遂には R 互恵そのものが崩壊してしまう。

以上ここで例示して説明したのは 2 パターンだが、詳細に調べてみると R 互恵の崩壊過程には様々な形態がある。翻って云うと R 互恵では、FSM の様々なビットの mimic に対して互恵が崩壊する trail が有り得ることになり、これが協調率が安定せず振動的な振る舞いを示す因になっている。

これに対して図3の Hero ではどうだろうか？ 基本的には S・T 互恵と R 互恵 (図3(B)) がある割合で生起しており、利得、協調率は安定的に推移している (図3(A))。このことは定性的には、図4の PD より Chicken 型のギャンプル性ジレンマも弱く、かつ Stag Hunt 型のリスク回避性ジレンマも存在せず<sup>14)</sup>、総じてジレンマ構造そのものが弱いことから来していると推量される。ライト消灯を合図とする R 互恵 (図3(G)) には及ばないながら、図3(F)をみると交互に S と T を取り合う非常に効率的な S・T 互恵 (Crowley の云う CAD 型の Coordinate Alternating Reciprocity<sup>15)</sup>) が高率で発生していることがわかる。匿名対戦となる One-shot ゲームにおいてこのような間接互恵 (Indirect Reciprocity) を達成させる上では、ライトの点灯、消灯がもつ情報送受信が重要な役割を担っている。このことは、図4(C)の PD と較べたときの情報量の大きさからも諒解される (図3(C))。実際、図3(D)をみると、FSM #1 の 2 ビット目がほぼ 0 であることから自分点灯、相手消灯時に D を出し、3 ビット目がほぼ 1 であることから自分消灯、相手点灯時に C を出す、とのプロトコルが形成されていることが確認出来る。これにより、高率で CAD 型の S・T 互恵が可能となる。既述したように、対戦シーケンスで点灯続ける、あるいは消灯続けることが “合図” となって R 互恵を達成している PD では、ライトの点灯 (もしくは消灯) はいわば時間方向に静的な “表象” に過ぎない。それに対して、ここで指摘した現象は、原始的ながらコミュニケーションの創発とそれに基づく S・T 互恵の達成と云うことが出来ると考えられる。

図中に点線の囲みで示した 7000 世代後にコミュニケーションに基づく S・T 互恵がやや低調になる期間が観察される (図3(F))。この期間は図3(G)から R 互恵も低調となる。これは、図3(E)から読めるように、FSM #2 の 3 ビット目が 1 に近い値から、mimic の侵入で小さい値になることが擾乱として作用しているためと考えられる。つまり、通常は食られて S の利得となったときに点灯する筈がこの期間はこの

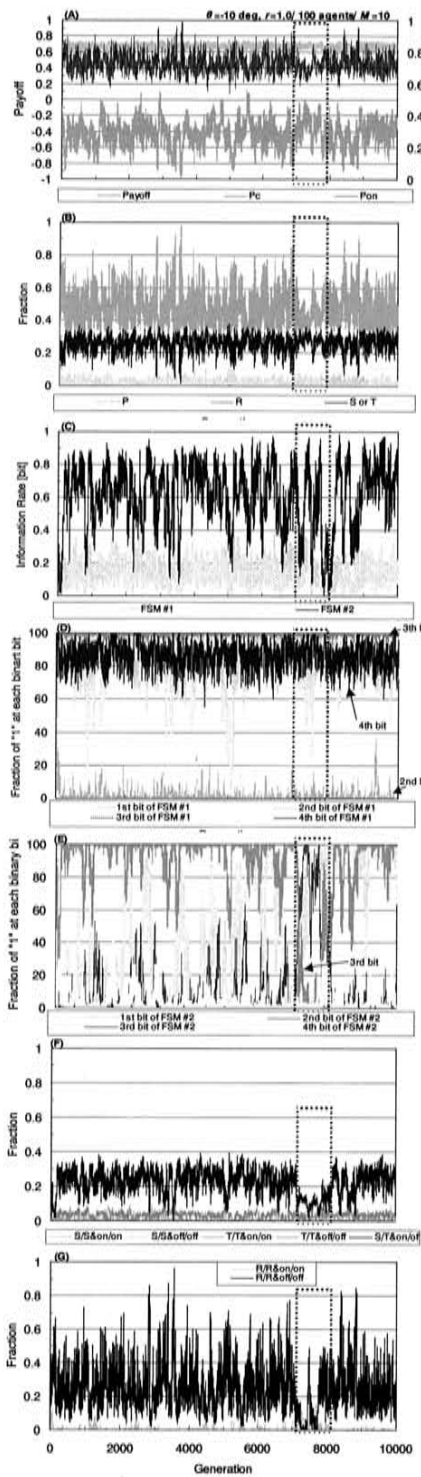


図3  $\theta=-10\text{deg}$ ,  $r=1.0$ の結果  
Fig.3 Result in case of  $\theta=-10\text{deg}$ ,  $r=1.0$ .

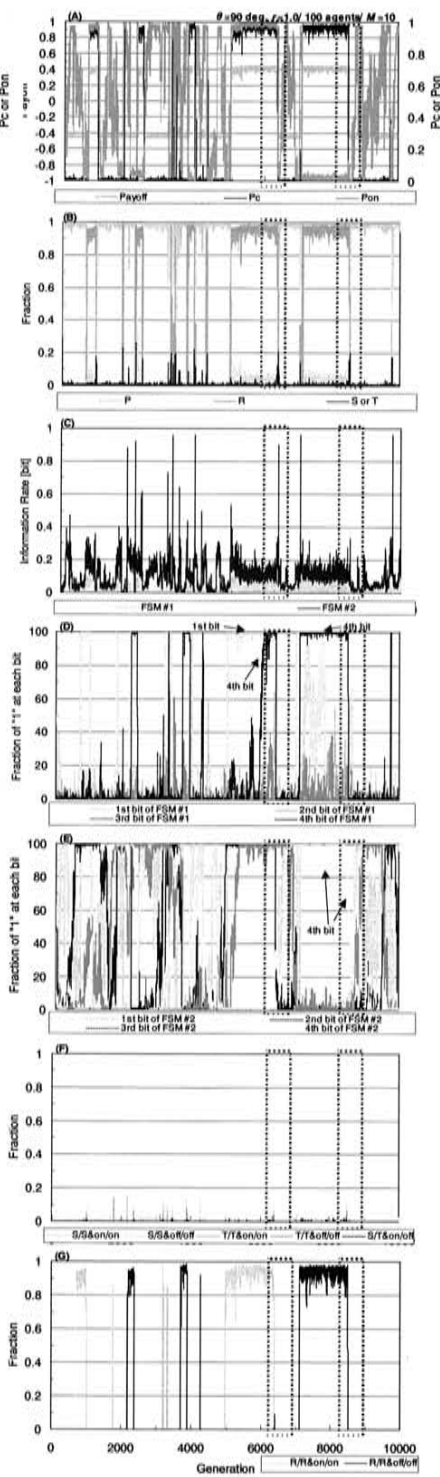


図4  $\theta=90\text{deg}$ ,  $r=1.0$ の結果  
Fig.4 Result in case of  $\theta=90\text{deg}$ ,  $r=1.0$ .

“ヘルプコール”が機能せずに、結果的に  $S \cdot T$  互恵も  $R$  互恵も低調になったと思われる。

#### 4.2. 全ゲーム領域における特性

図 5 は全ゲーム領域における (A)Replicator Dynamics による解析解との利得差, (B)FSM #1 の情報量, (C)FSM #2 の情報量, (D) 連続するゲームシーケンスにおいて  $S$  (もしくは  $T$ ) かつ on (もしくは off)  $\rightarrow T$  (もしくは  $S$ ) かつ off (もしくは on) の生起確率, (E) $S$  (もしくは  $T$ ) かつ on (もしくは off)  $\rightarrow S$  (もしくは  $T$ ) かつ on (もしくは off) の生起確率, (F)  $R$  かつ on (もしくは off)  $\rightarrow R$  かつ on (もしくは off) の生起確率である。

図 1 で云うと PD の SH 側, Chicken, Anti-Leader & Chicken (本質的には PD) の領域で本モデルはジレンマ緩解・解消プロトコルの一切ない解析解より高利得を上げている (図 5(A))。しかし, PD の Chicken 側, Leader & Hero (本質的には PD) と Leader の一部で強いジレンマの領域ではその効果は及ばない。が, 逆に, Leader の一部 (Hero 側) と Hero ゲームでジレンマ大 ( $r$  大) ほど解析解より高利得を上げており (図 5(A))。特に Leader より Chicken 型ジレンマが弱い Hero でその効果が顕著である<sup>14)</sup>。この領域は FSM の情報量, 特に FSM #2 の情報量が大きいエリアに重なる。また, ライト点滅と C 出すか D を出

すかが同期した, 換言するとコミュニケーションにより CAD 型  $S \cdot T$  互恵を創発させている領域 (図 5(D)) の中心部分と一致している。

FSM #1 の情報量大小は C/D をどう出すかの行動戦略に自他ライト点滅状況の情報はどう利いているか, FSM #2 のそれはライトの点滅行動に前回ゲーム帰結の情報 ( $P, R, S, T$ ) がどう利いているかを意味する。換言するなら, FSM #1 はゲームの手を直接規定する行動戦略の, FSM #2 はコミュニケーションのプロトコル (もっと言えば言語の種類) の秩序程度を意味している。

本論の結果は, コミュニケーションと同期した互恵と云う観点からは,  $R$  互恵を前提とする PD より CAD 型  $S \cdot T$  互恵を前提にする Hero と Leader の方が重要であることを示唆していると考えられる。これは, 既述したように PD ではライト点滅は時間方向に static な Tag 的表象として作用しているに過ぎないのに対して, CAD 型  $S \cdot T$  互恵では時間方向に変化させるライトの点灯, 消灯に大きな情報が包含されているからである。また, 図 5(F)の  $\theta=90\text{deg}$  付近を見ると,  $R$  互恵はジレンマが大きくなると (図で  $r$  大) 直ぐに崩壊し解析解と同様裏切り合いの Nash 均衡に陥ってしまうのに対して,  $\theta=-20\text{deg}$  付近の Hero では大ジレンマに対しても同様に  $S \cdot T$  互恵がある割

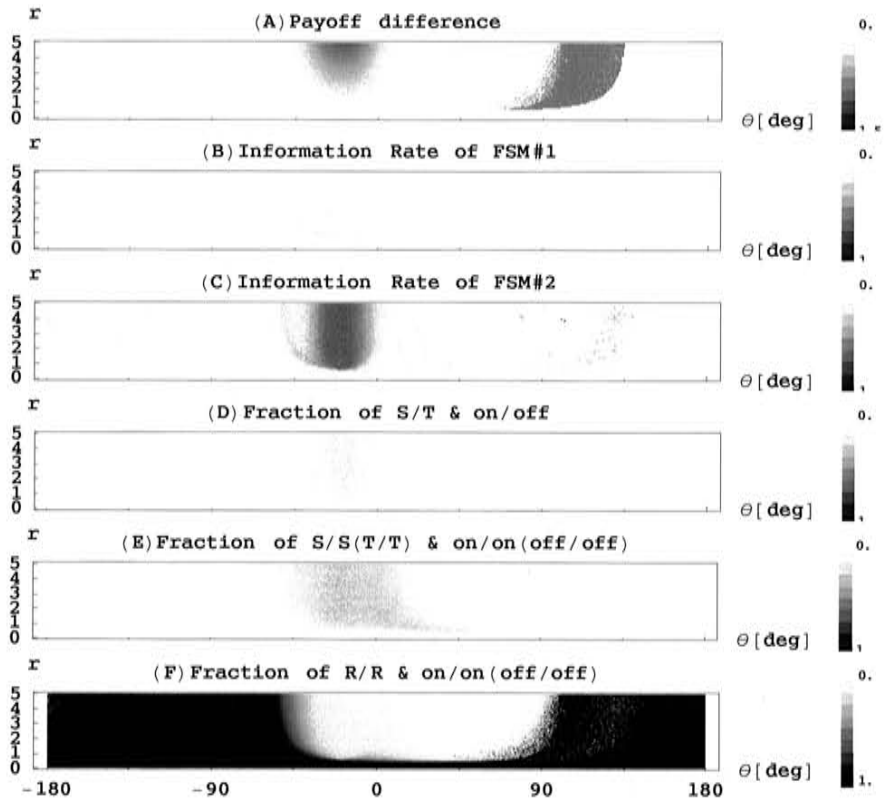


図 5 2x2 ゲーム全域における結果

Fig.5 Holistic pictures covering a 2x2 game world for (A)payoff difference between the present model and analytic solution by Replicator Dynamics, (B)information rate of FSM#1, (C)information rate of FSM#2, (D)fraction of  $S/T$  & on/off, (E)fraction of  $S/S(T/T)$  & on/on(off/off) and (F)fraction of  $R/R$  & on/on(off/off).

合で維持されている (図 5(D))。以上のことから、本モデルで再現されたコミュニケーションに基づく互恵の創発は、CAD 型 S・T 互恵を前提にする Hero と Leader で大きな意味を持っていると推量される。パレート最適からくる Chicken 型のジレンマが存在するという条件 ( $T > R$ ) に加えて、 $S+T > 2R$  が成立することが、これらのゲームの特徴である。推量を巡らすと、従来、進化ゲームでプロトタイプ的に取り上げられてきた PD より、寧ろ 2 人のプレーヤ間で異なる役回りを時間方向に均等に分担することで効率的にジレンマを解消出来る Hero や Leader のような状況が、シグナリングとその認知機構を進化させ、原初のコミュニケーションを創発させたのかも知れない。

#### 4.3. Norm の共進化モデルとの関連

本モデルのような 2 層の知識構造を有する共進化モデルとして、Chalub らの Norm のモデル<sup>16)</sup>がある。彼らのモデルではエージェントは幾つかの島に分居し、Norm を共有している (進化プロセスで島間の Norm コピーやエージェントの移住が起こるようになっている)。島内でランダム対戦するゲームにおけるプレーヤ手 (C か D か) は対戦相手の Image Score を引数とする FSM で規定されている (行動戦略)。各島の Norm は、自他の Image Score が与えられたとき自手として C (D) 出すことが自らの Image Score を good にするのか bad にするのかを決めている。行動戦略は個体レベルで、Norm は島レベルで適応が起こる。この共進化により、例えば、“裏切り者に C 出すことが bad な行為 (逆に裏切り者に D 出すことは寧ろ good な行為)” (2 次フリーライダー対策) とか、“裏切り者が裏切り者に D 出すことで彼の Image Score が good に復活する” (悪党の改悛プロトコル) と云った Norm が進化し、これにより頑強な協働社会が創発する。

彼らの行動戦略が本モデルにおける FSM #1, Norm は FSM #2 に相当する。ただし、Chalub らのモデルでは島内でランダム対戦するエージェント間では Norm は共有されているのに対して、本モデルでは FSM #2 は FSM #1 同様、個体ごとに異なる (その意味ではタフなゲームである)。また、彼らのモデルでは、Norm は島内のある人の Image Score を good にするか、bad にするかを第三者の視点に立って規定している (島内で共有しているから第三者的視点になる) のに対して、本モデルではライト点滅は自らの FSM #2 により規定される。これはライト点滅を Image Score の good か bad かに準えて考えると、自ら有する規範に応じて Image Score の good/bad を切り替えている自己情報発信に相当する。喩えて言えば、悪党なのに自分は good だと宣言することが可能な構造になっており、実際、本モデルで見たようにこの点を衝いた mimic が侵入することで、PD では R 互恵が頻繁に崩壊、再生する振動状況が起きた。何が Image Score 上 good なのか bad なのかを規定する Norm の視点、すなわち“自分の Image Score を自ら決めて自己宣伝” (本論のような構造) するの、 “対戦相手の Image Score を自分が評価する” のかは、進化の帰結を決める上で大きく影響する。Ohtsuki & Iwasa<sup>17)</sup> は間接互恵における reputation dynamics を考

慮するモデルを構成し、この点に関する示唆を与えている。

#### 5. 結論

コミュニケーションに基づく互恵創発のメカニズムを明らかにするために、ゲーム手 (C か D か) とライト点滅の各行動出力を、夫々、自他のライト点滅、前回ゲーム帰結 (P, R, S, T) を情報入力力で規定する状態遷移機械で与える 2 層共進化モデルを構築し、以下の結果を得た。

- (1) 学習など一切の協調サポート機構のない Replicator Dynamics による解析解と較べると、様々なゲーム構造でジレンマを緩解し互恵を創発させる効果が観察された。
- (2) PD における R 互恵では、頻繁な mimic の侵襲に晒され、安定した協調には至らず振動的な振る舞いを見せるのに対して、Hero と Leader ではある一定の割合で比較的安定した CAD 型 S・T 互恵が創発する。
- (3) R 互恵におけるライト点滅は時間方向に static な Tag 的 “対象” として作用しているに過ぎないのに対して、CAD 型 S・T 互恵ではコミュニケーションとして機能しており、伝送される情報量も大きい。
- (4) 前項のことから、原始的コミュニケーションの進化には、異なる役回りを時間方向に分担することで R 互恵より高利得が上げられる Hero や Leder のようなジレンマ状況が深く関わっていた可能性が示唆される。
- (5) 提示した 2 層の共進化モデルと Norm の進化を扱った Chalub や Ohtsuki & Iwasa の先行研究との関連について論及した。

#### 注記

<sup>11)</sup> 確率事象であるセンサー入力信号  $y_j$  のもとで  $x_k$  なる行動が生起する情報エントロピー  $H_{x_k|y_j}[\text{bit}]$  は以下で定義される。

$$H_{x_k|y_j} = -\sum_{j=1}^n \sum_{k=1}^m p(x_k, y_j) \cdot \log_2 p(x_k | y_j) \dots (2)$$

ここで、 $p(x_k, y_j)$  は複合事象生起確率、 $p(x_k | y_j)$  は  $y_j$  のもとで  $x_k$  が生起する条件付き確率を表す。また、情報量  $I_{x_k|y_j}[\text{bit}]$  とは、情報入力なしに計測されたエントロピー  $H_{x_k}$  と情報を与えられた状況下で行動に歪みが生じているエントロピー  $H_{x_k|y_j}$  との差、

$$I_{x_k|y_j} = H_{x_k} - H_{x_k|y_j} \dots (3)$$

で定義される。

#### 参考文献

- 1) Blackmore, S.: The Meme Machine, Oxford University Press, (1999).
- 2) Arita, T., Koyama, Y.: Evolution of linguistic diversity in a simple communication system, Artificial Life 4 (1), 109-124, (1998).
- 3) Arita, T., Taylor, C.E.: A simple model for evolution of communication. Evolutionary Programming V, MIT Press, 405-409, (1996).
- 4) Dorigo, M., Maniezzo, V., Colorni, A.: The ant system: Optimization by a colony of cooperating agents, IEEE Transactions on Systems, Man and Cybernetics - Part B 26 (1), 29-41, (1996).
- 5) Kawamura, H., Yamamoto, M., Mitamura, T., Suzuki, K., Ohuchi, A.: Cooperation search based on pheromone communication for vehicle routing problems, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences E81-A-6, 1089-1096, (1998).
- 6) Grim, P., Kokalis, T., Alai-Tal, A., Kibbi, M., St Denis, P.: Making meaning happen. Journal of Experimental & Theoretical Artificial Intelligence 16 (4), 209-243, (2004).
- 7) Lenaers, T., Jansen, B., Tuyls, K., Vylder, B.D.: The evolutionary language game: An orthogonal approach, Journal of Theoretical Biology 235, 566-582, (2005).
- 8) van Baalen, M., Janssen, V.: Common language or tower of Babel? On the evolutionary dynamics of signals and their meanings, Proc. of the Royal Society B 270, 69-76, (2003).
- 9) Axelrod, R.: The Evolution of Cooperation, New York: Basic book (1984).
- 10) Nowak, M.A., Bonhoeffer, S., May, R.M.: Spatial games and the maintenance of cooperation, Proc. National Academy Science USA Vol.91, 4877-4881 (1994).
- 11) Riebo, R., Colten, M.D., Axelrod, R.: Evolution of cooperation without reciprocity, Nature 414, 441-443, (2001).
- 12) Nowak, M.A., Sigmund, K.: Evolution of indirect reciprocity by image scoring, Nature 393, 573-577, (1998).
- 13) 安藤尚, 内田美穂, 網谷賢治: 進化学習エージェントによる協調行動とコミュニケーションの開発。整理モデル化と問題解決シンポジウム, 163-170, (2006).
- 14) Tanimoto, J., Sagara, H.: Relationship between dilemma occurrence and the existence of a weakly dominant strategy in a two-player symmetric game, BioSystems, . . . (2006).
- 15) Crowley, P.: Dangerous games and the emergence of social structure: evolving memory-based strategies for the generalized hawk-dove game, Behavioral Ecology 12 (6), 753-760, (2001).
- 16) Chalub, F.A.C.C., Santos, F.C., Pacheco, J.M.: The evolution of norms, Journal of Theoretical Biology 241, 233-240, (2006).
- 17) Ohtsuki, H., Iwasa, Y.: How should we define goodness?—reputation dynamics in indirect reciprocity, Journal of Theoretical Biology 231, 107-120, (2004).