

2 記憶長戦略 2×2 ゲームにおける対戦及び 戦略適応のローカリティが 創発する互惠相に付与する影響

脇山宗也[†] 谷本潤^{††}

戦略更新時に作用するローカリティとゲーム対戦時に作用するローカリティが、社会の帰結にどのような影響を与えるのかを、数値実験によって明らかにした。エージェントは 5bit の 2 記憶長で表現される戦略に基づいて協調(C)、裏切り(D)のどちらかの行動を取り、遺伝的アルゴリズムにより戦略更新を行いながら、Prisoner's Dilemma, Chicken, Hero, Leader, Stag Hunt, Trivial の何れかの構造を持つゲームを行う。2 つのローカリティを独立に変化させて数値計算を行った結果、戦略更新時のローカリティは ST 互惠の構築に、対戦時の局所性は R 互惠の構築に作用し、2 つのローカリティは独立に作用するものではないことが示された。

How localities of interaction and adaptation affect on reciprocity phase in various 2 x 2 games played by agents equipped with 2-memory length strategies?

Motoya Wakiyama[†] and Jun Tanimoto^{††}

2x2 games covering Prisoner's Dilemma, Chicken, Hero, Leader, Stag Hunt and Trivial game classes, where agents have strategy expressed by 5 bit 2-memory length, are numerically investigated with a motivation of how localities for game interaction and strategy adaptation influence on ST-reciprocity and R-reciprocity (Tanimoto & Sagara; BioSystems 90 (3), 2007). The more enhanced R-reciprocity is observed with the stronger locality for game interaction when the relatively stronger locality for strategy adaptation is assumed. Whereas, the more enhanced ST-reciprocity is emerged with the stronger locality for strategy adaptation when the relatively weaker locality for game interaction is levied. In the numerical experiment, those two localities were dealt with independently.

1. 緒言

2×2 ジレンマゲームが network 上で行われる、所謂、ローカリティ（局所性）が存在する状況下では、互いに協調（Cooperation, C）を出す R 互惠が創発することが知られている。ローカリティには、2×2 ゲームの対戦時と、協調（C）と裏切り（Defection, D）のどちらを出すかを決定する戦略を適応（学習）時の 2 段階で R 互惠を促す効果がある。既往研究[1][2]では、対戦を行う範囲と戦略を適応させる際に参照対象とする範囲を変えることで創発相が異なることを明らかにしており、両範囲が一致する場合、最も R 互惠が効果的になるとされている。これらの研究ではローカリティを network 上で定義しており、対戦や戦略の母集団をサブクラスに分割する意味でのローカリティは想定されていない。また、ネットワークゲームと云う前提からプレイヤーの戦略は 1bit（C か D か）で表現されているが、戦略表記を多記憶長としたときのローカリティの特性については検討されていない。戦略が 2 記憶長となると、R 互惠だけでなく、Chicken 型ジレンマの Leader や Hero で効率的に高利得を上げる ST 互惠[3]が創発することが解っているが、ST 互惠に及ぼすローカリティの影響は興味あるところだろう。

本論では、一連の数値実験により、2 記憶長戦略をもつエージェントが様々なクラスの 2×2 ゲームを繰り返す際、対戦及び戦略適応時のローカリティを独立に変化させることで、創発相にどのような影響が顕れるのかを明らかにする。

2. モデルの概要

エージェントが互いに 2×2 ゲームを繰り返しながら、ゲームによる獲得利得がより大きな戦略が適応的に進化するエピソードをモデル化する。総エージェント数 A_g 人の系において、ローカリティを導入することで R 互惠や ST 互惠（互いに協調と裏切りを交互に出すことで互いに協調を出し合う R 互惠よりも高い利得を得る、より高度な互惠関係[3]）が創発するか否かを検討する。

2.1 ゲーム構造の表記

エージェントは 2×2 ゲームを行う。協調（C）もしくは裏切り（D）を離散手として出す 2×2 ゲームの構造は、自他の組手（D,D）、（C,D）、（D,C）、（C,C）に対する利得をそれぞれ P（Punishment）、S（Saint）、T（Temptation）、R（Reward）とするなら、

[†] 構造計画研究所・修士（工学）

Kozo Keikaku Engineering Inc., M.Eng.

^{††} 九州大学大学院総合理工学研究院・教授・工博

Prof., Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Dr.Eng.

マトリクス $M = \begin{bmatrix} R & S \\ T & P \end{bmatrix} = [R, S, T, P]$ で表現される。ゲーム構造 M が表現する PD, Chicken

などあらゆる 2×2 ゲーム空間は、ギャンブル性のジレンマ強度を表す D_g , リスク回避性のジレンマ強度を表す D_r を用いて、式(1)で表される。なお、式(1)において $p_0 \geq 0$ である[4].

$$\begin{aligned} P &= -p_0 && \dots (1-1) \\ S &= -p_0 - Dr && \dots (1-2) \\ T &= p_0 + Dg && \dots (1-3) \\ R &= p_0 && \dots (1-4) \end{aligned}$$

2.2 エージェントの戦略

各エージェントは戦略を有する。戦略は状況に応じて C か D かの手を規定する行動ルールであり、本論では前回の自他の組手 (P, R, S, T) を入力に、C か D かの行動を出力する状態遷移機械 (Finite State Machine, FSM) で表現する。FSM は前回の自他手組の 4 通りに加えて、さらに一連の 2×2 ゲームの初手に、C か D のどちらを出すかを規定する $4+1=5\text{bit}$ の情報で表現される (図 1)。TFT は C|DDCC, PAVLOF は C|CDDC で表される。全エージェントは各々 1 つの戦略を持ち、戦略の更新は 1 エージェントを 1 つの遺伝子と見なして、遺伝的アルゴリズム (Genetic Algorithm, GA) による。遺伝子の適合度は、1 世代中に獲得した総利得をゲーム回数で除した値とする。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, 1 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}) とする。

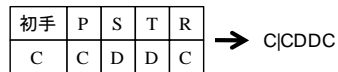


図 1 FSM の表記例

2.1 ゲーム対戦相手, 戦略更新時の参照範囲の制限

エージェントがゲームを行う対戦する相手, および GA を適用する Gene Pool に制限を加えることで、系に 2×2 ゲーム対戦時に作用するローカリティ (以下, 対戦ローカリティ) と戦略更新時に作用するローカリティ (以下, 戦略更新ローカリティ) を導入する。

エージェントはシミュレーション開始時に F 個の対戦フィールドの内、いずれか 1 つにランダムで割り振られる。この時、対戦フィールド内の人数はいずれも同一とする。すなわち、1 対戦フィールドのエージェント数は A_g/F である。エージェントは L 回の繰り返しゲームを同一対戦フィールド内の全エージェントと同数回行い、既述の

ように適合度は 1 世代中に獲得した総利得の 1 ゲーム当たり平均値とする。

同様に、 G 個の Gene Pool 内の人数が等しくなるように、全エージェントをエピソード開始時にランダムで割り振る。すなわち、1 対戦フィールドのエージェント数は A_g/G である。1 世代終了時に Gene Pool 単位で GA が適用され、戦略は更新される。

Gene Pool への割り振り, および対戦フィールドへの割り振りは、無相関に行われる場合と、相関がある場合の 2 つの方法を想定する。無相関の場合、Gene Pool と対戦フィールドへの割り振りは完全に独立して行う。相関を持たせる場合は Gene Pool を作成後、ある Gene Pool に属する全エージェントは特定の他 Gene Pool に存在するエージェントと総当たりで対戦を行う。そのため、相関がある場合、 G と F の値は等しい (図 2)。

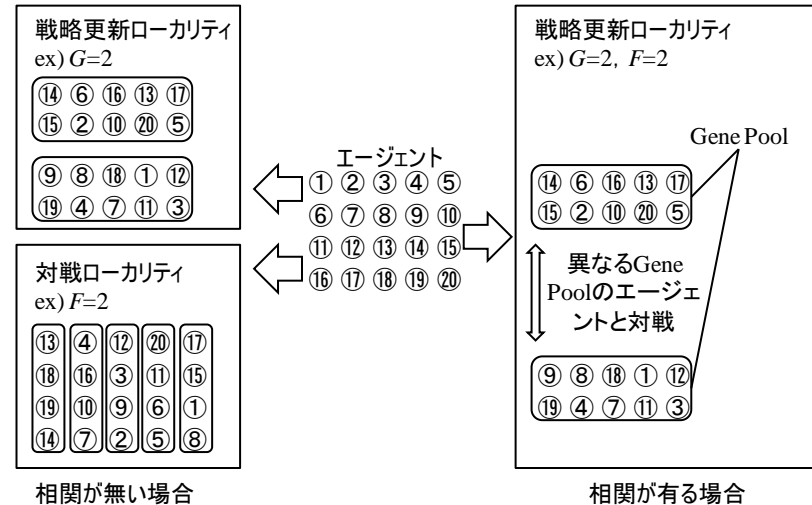


図 2 ローカリティの作成方法の例

3. 数値実験

パラメータの設定値を表 1 に示す。 F および G は $\{1, 2, 5, 10\}$ と 4 段階で変化させ、ローカリティの影響を検討した。シミュレーションは予備実験により系が疑似均衡に達することを確認した上で 1000 世代を繰り返し上限とし、951 世代から 1000 世代までの 50 世代の平均を 1 エピソードの結果とする。解析はエピソード 5 回のアンサンブル平均とする。

本稿では $D_g \in [-2, 5]$, $D_r \in [-5, 2]$ の間で共に 0.1 刻みで変化させながらシミュレーションを行った. この範囲に含まれるゲームクラスは, 図 3 に示すように, 主として Chicken 型のジレンマであり, 加えて Prisoner's Dilemma (PD), Stag Hunt (SH), Trivial である. 特に Chicken 型ジレンマゲームのクラスは, $S+T > 2R$ の条件を満たす, すなわち, ST 互恵のインセンティブがある Leader ($T > S$) と Hero ($S > T$), その条件を満たさない pure-Chicken に 3 分される.

$A_g = 10000$	$L = 100$	$T_B = 0.01$
$p_0 = 0$	$p_{cross} = 0.5$	$p_{mut} = 0.01$

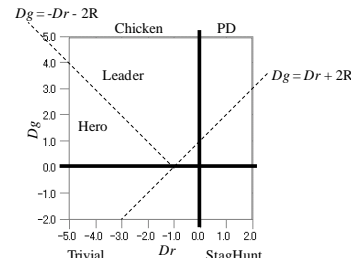


図 3 ゲーム領域図

4. シミュレーション結果及び考察

4.1 ローカリティが存在しない場合 ($G=1, F=1$)

Well-mixed な状況下でのゲームであり, 戦略が C か D かの 1bit 表記であれば Replicator Dynamics による解析解と一致する.

図 4 に均衡時の 1 エージェント 1 ゲーム当たりの平均利得 (社会平均利得), P の生起率, R の生起率, ST の生起率を表す. これらをもとに図 5 に 3 つの創発相にクラス分けした相図を示す.

領域 A は, ジレンマのない Trivial ゲーム, Chicken 型ジレンマがないという意味で微弱なジレンマしかない SH[4] およびジレンマの弱い Chicken, PD の一部で, 平均利得がほぼ 0.5 ($=R$) となっている. これは, 5bit と云う記憶長戦略の効果により, ジレンマが弱いゲームでは, 公平な Pareto である R 互恵が達成されていることを意味する [3].

領域 B, B' は, 効率的に S と T を取る ST 互恵により, 公平な Pareto である R を取るよりも高い利得を獲得している領域である. Tanimoto & Sagara [3] によれば, ST 互恵には 2 タイプがあると云う. すなわち, 戦略ペアで L 回繰り返すゲームの中で互いに S と T を交互に取り合う効率的な ST 互恵 (交互 ST 互恵) と, ある戦略ペアの対戦では一方的に S (或いは T) を取り他戦略との対戦時に R (或いは P) を取って一人負け

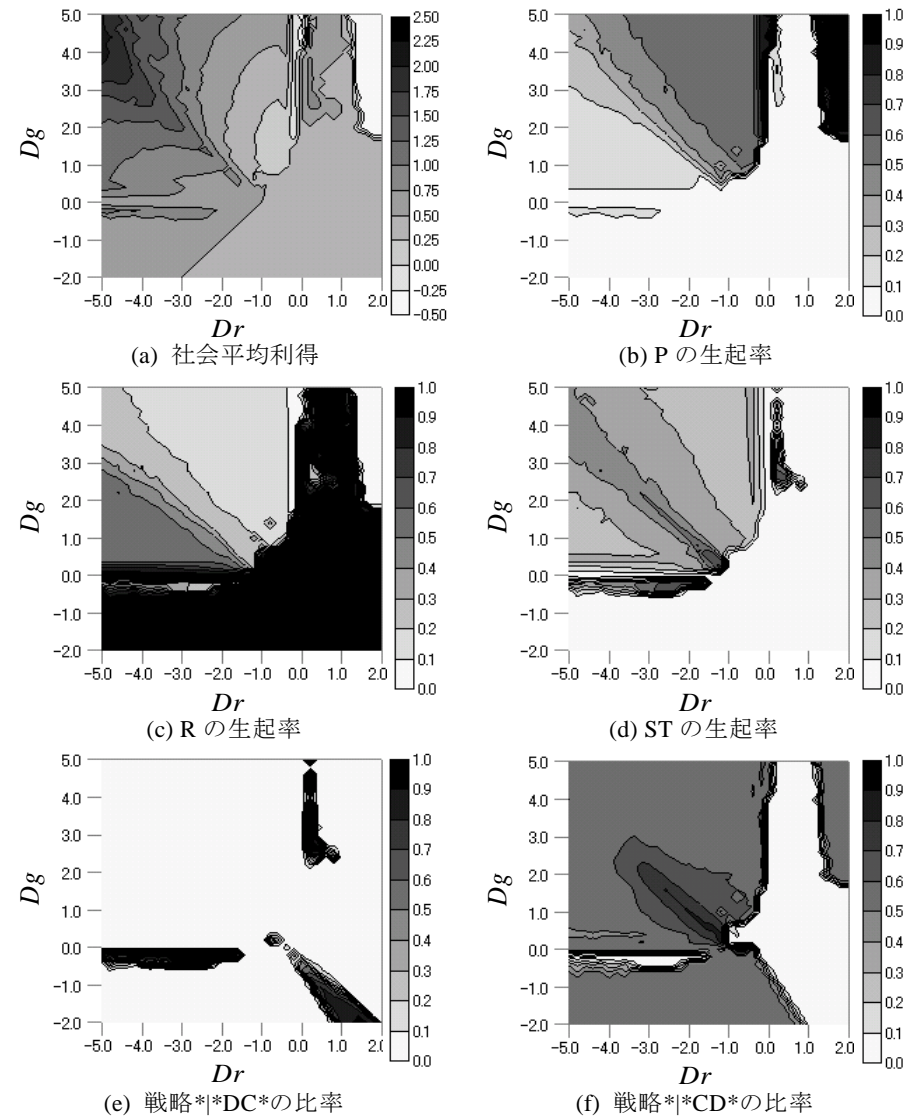


図 4 $G=1, F=1$ でのシミュレーション結果

を取り返す（一人勝ちを吐き出す）ことで帳尻を合わせる ST 互恵（確率 ST 互恵）である。このうち Trivial にある領域 B' では、均衡時の戦略分布を確認したところ、*DC*（*は wild card で C か D の何れかであることを意味する。以下、交互 ST 互恵戦略）の戦略が 9 割以上を占めており、効率的な交互 ST 互恵を創発していた。対して領域 B では、均衡時の戦略分布は*DC*（以下、確率 ST 互恵戦略）が 5 割を超え、高利得を上げ得ると言う意味では交互 ST 互恵には劣る確率 ST 互恵が創発していることが確認された。確率 ST 互恵戦略の存在比が 0.5 と稍小さいのは、ある確率 ST 互恵戦略から T を取り続けることで高利得を獲得し、フリーライダーとして存在する戦略が併存するためである（但し、他戦略との対戦では低利得しか上げられず、適合度は確率 ST 互恵戦略と同等となる）。なお、ST 互恵のインセンティブが有る領域は、 $D_g = D_r + 2R$ より左側にあるエリア ($S+T>2R$)、すなわち R 互恵より獲得利得が大きくなる領域のみである。

Chicken 型ジレンマ D_g の大きいエリアに顕れる領域 C は、PD クラス内と PD-Chicken 境界部の 2 カ所に創発する。前者では、ジレンマが大きくなって、5bit 記憶長戦略でも R 互恵は達成し得ず Nash 均衡にトラップされてしまうためである。PD と Chicken の境界部では、ST 互恵を創発させる領域 B（領域 C の左側）と R 互恵する領域 A（領域 C の右側）との狭間で、確率 ST 互恵から R 互恵に転相する境界で P の生起頻度が大きくなることで利得が 0.5 を下回ることにより生じる。

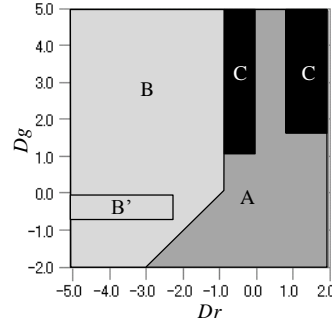


図 5 創発相パターン

4.2 相関の無いローカリティが存在する場合

図 6~11 は、それぞれが戦略更新ローカリティ、対戦ローカリティを変えた場合の均衡時の社会平均利得、P の生起率、R の生起率、ST の生起率、戦略*DC*（交互 ST 互恵戦略）の比率、戦略*CD*（確率 ST 互恵戦略）の比率を表す。ローカリティの強さはグラフ下に示した G と F の値で表され、値が大きいほどグループの分割数が多くなり、ローカリティが強くなる。従って、図中の下側のパネルほど戦略更新ローカリティが強くなり、右側のパネルほど対戦ローカリティが強いことを意味する。

戦略更新ローカリティの強さによって、前述した領域 A、B、C の様相は異なる。対戦ローカリティが比較的弱い場合（図中左側、a, b, e, f, i, j, m, n）では、戦略更新ローカリティが強いほど（図の下パネルほど）、領域 B、B' において獲得利得が大きくなり、ST 互恵がより強く顕れる。また、領域 C の間にも ST 互恵によって高利得を獲得する領域が顕れ、領域 B が拡大しており、中でも対戦ローカリティが無い場合 ($F=1$, 図中最左列パネル) でより明らかである。ローカリティが変化した場合の戦

略分布の違いを図 9, 10 から確認すると、戦略更新ローカリティが強くなるほど、Chicken の領域で確率 ST 互恵戦略の比率が増加しており、PD の領域では交互 ST 互恵戦略の比率が増加していることが分かる。また、Chicken の領域で ST の生起率だけで

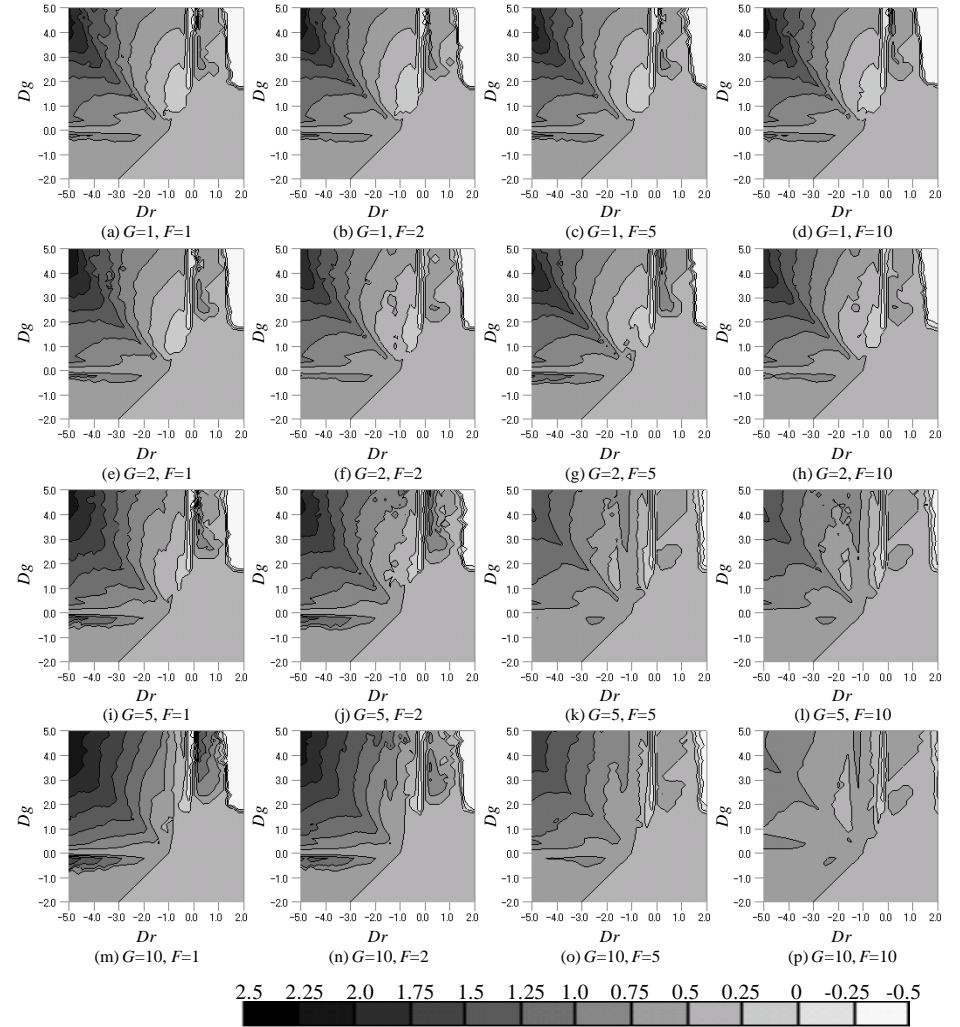


図 6 各ローカリティにおける社会平均利得

なく、 R の生起確率も増加している。以上のことから、戦略更新時のローカリティは、互惠関係を崩壊に導く裏切り戦略が社会全体に一気に拡散することを防ぐことで、互惠関係の容易にさせていると推測される。そのため、強い戦略更新ローカリティが導

入されることで ST 互惠関係を築いているエージェント群に対して裏切りを出し続けて利益のみを食うフリーライダーは駆逐され、より広範な領域で ST 互惠が構築されるようになったものと考えられる。

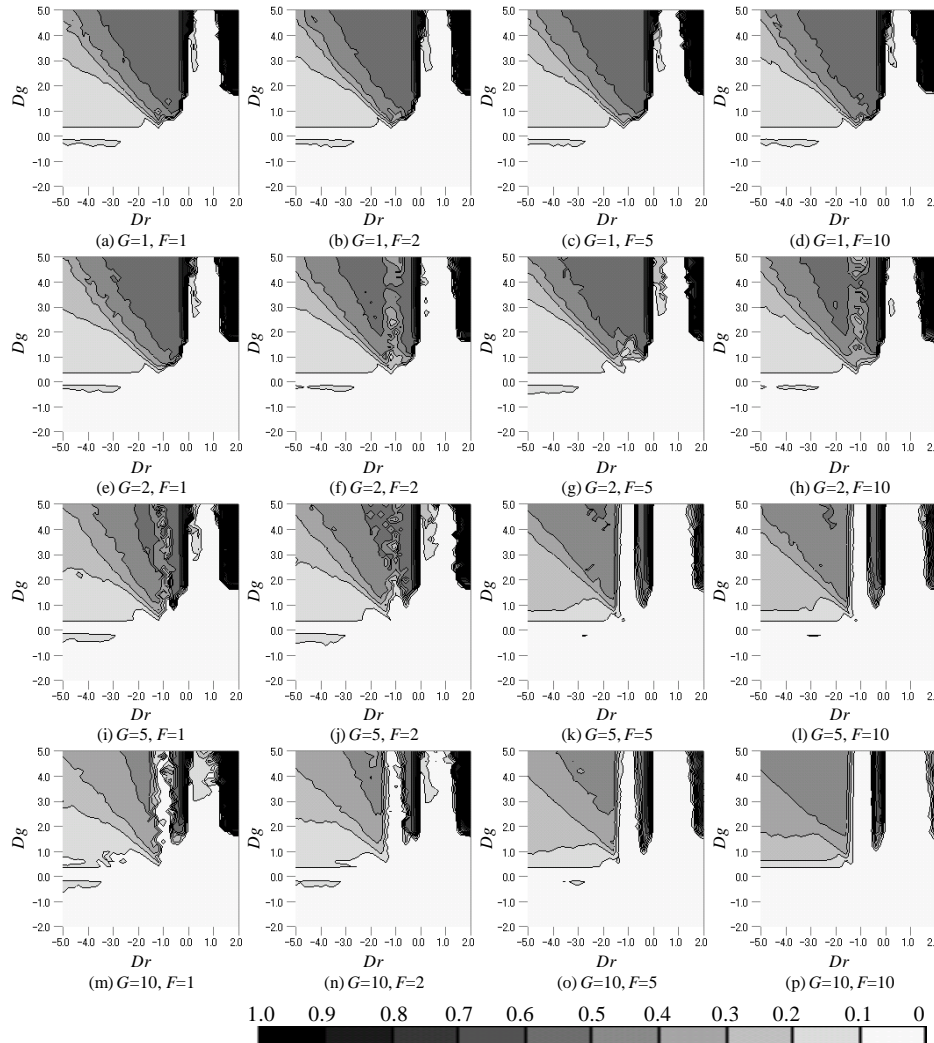


図 7 各ローカリティにおける P の生起率

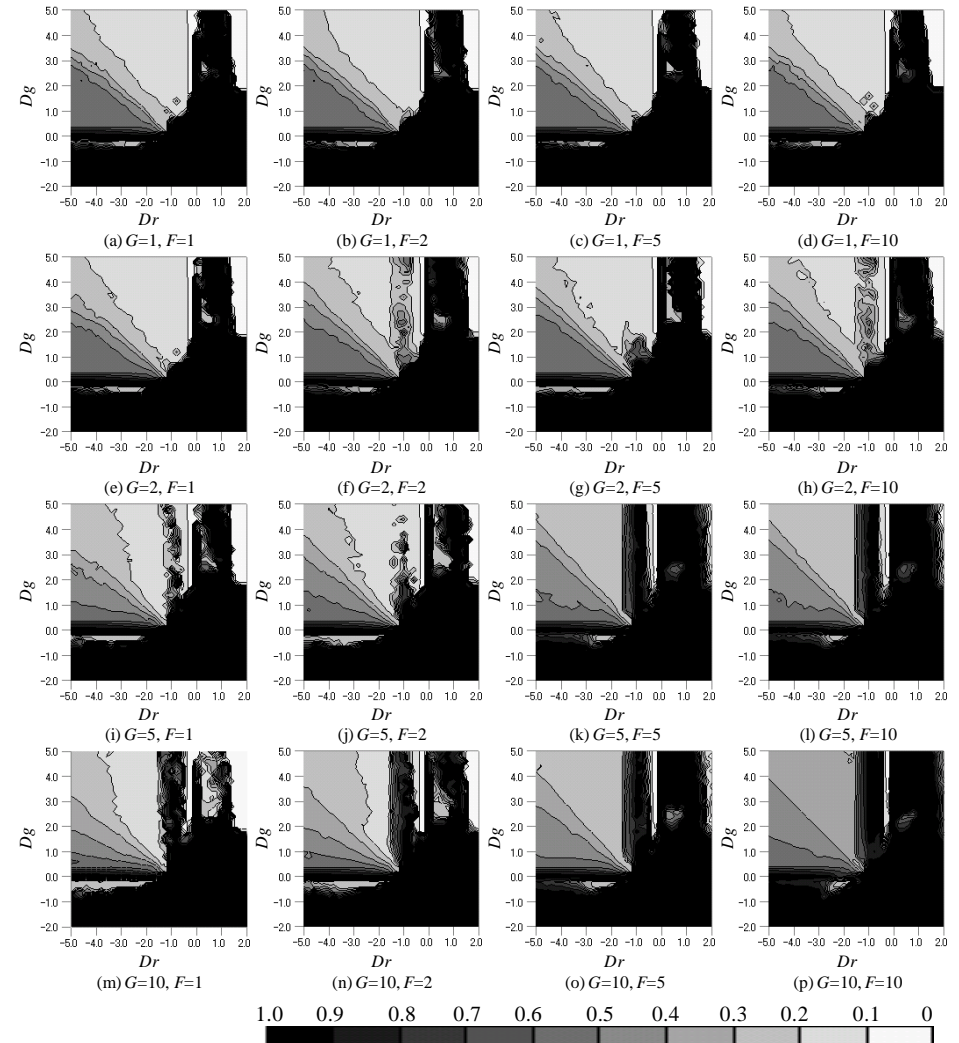


図 8 各ローカリティにおける R の生起率

一方、対戦ローカリティが強い場合（図中右側，c, d, g, h, k, l, o, p），戦略更新ローカリティが強まるほど領域 B の獲得利得，及び ST の生起率が低下し，ST 互恵が形成されにくくなっている．その代わりに PD 領域において顕れていた領域 C は，

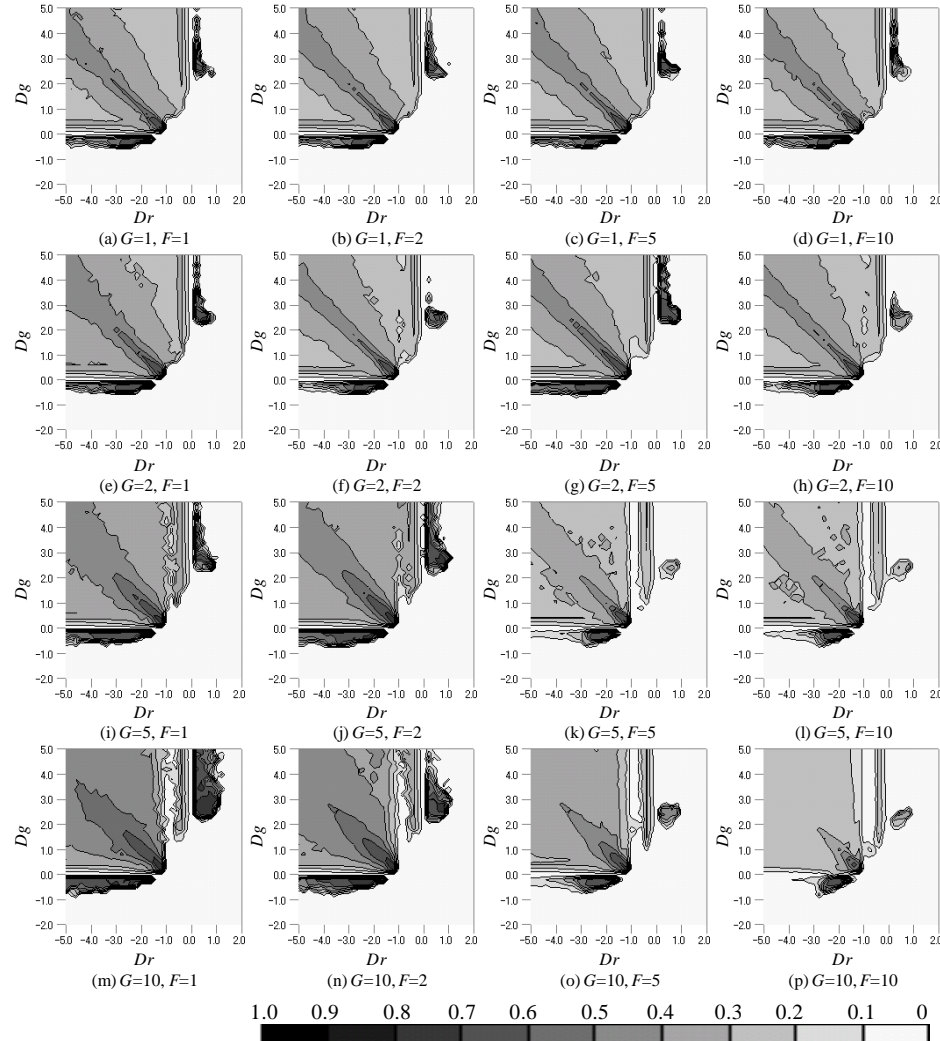


図 9 各ローカリティにおける ST の生起率

Nash 均衡から脱し，R 互恵を達成できるようになる．

即ち，対戦ローカリティが弱い状況では戦略更新ローカリティが強いほど ST 互恵は強化され，対戦ローカリティが強い状況では ST 互恵は創発しにくくなるが，R 互

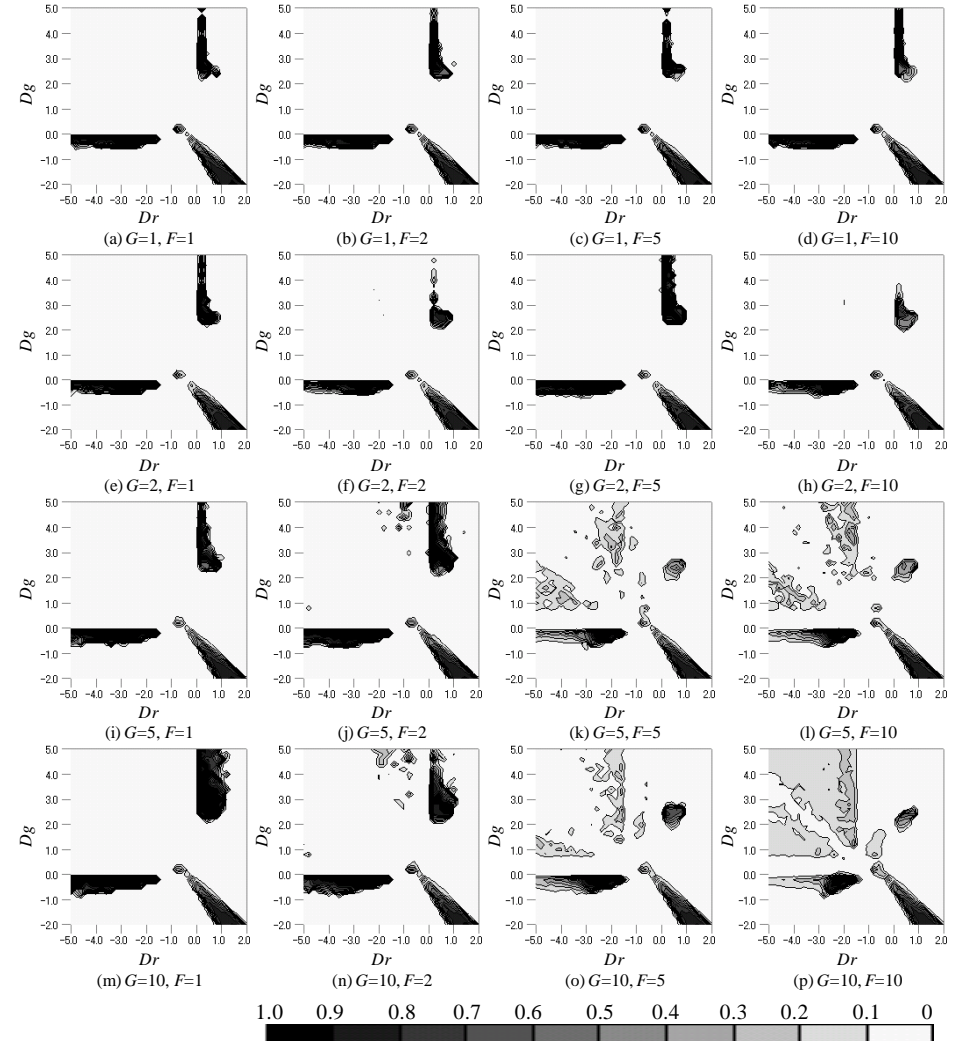


図 10 各ローカリティにおける戦略*DC*の比率

恵に達し得ない領域は圧縮される。

効率的な ST 互恵には Tanimoto & Sagara[3]が言うように交互 ST 互恵が望ましいが、それが不可能であれば次善の策としては確率的 ST 互恵が進化する。この相では S を

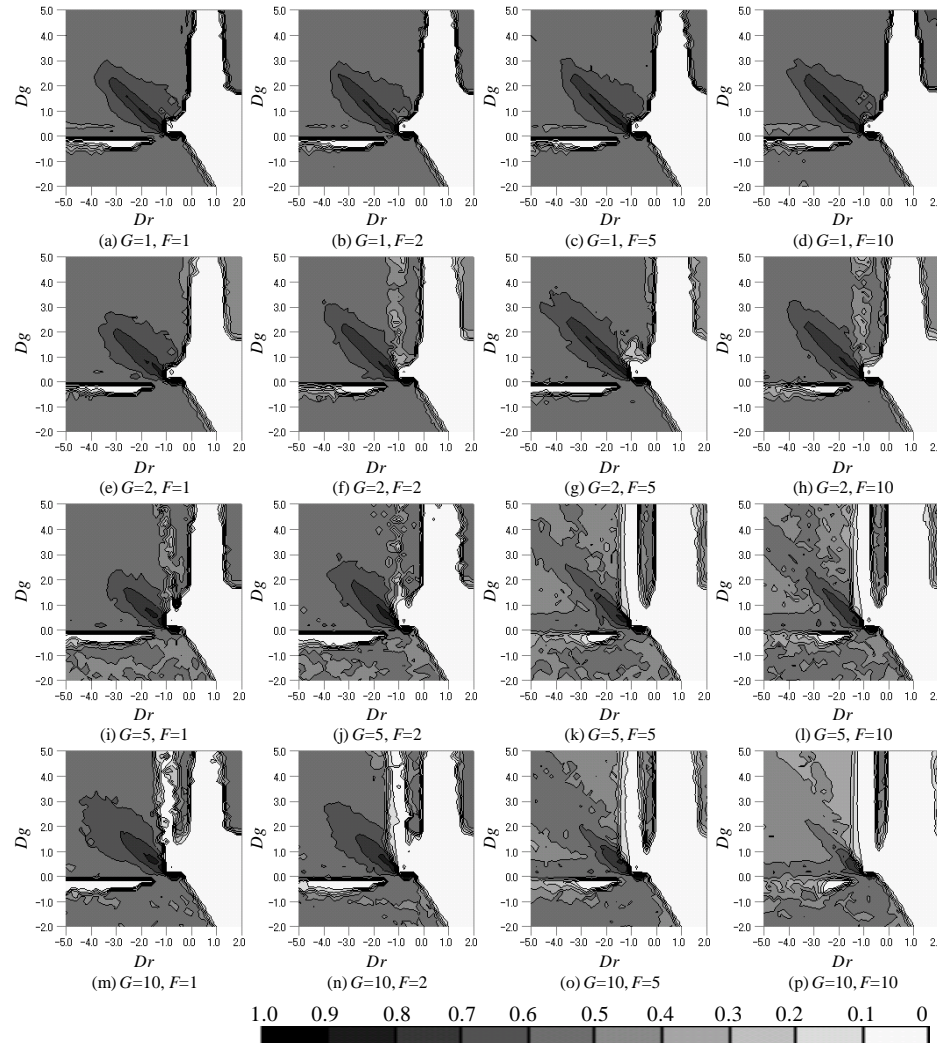


図 11 各ローカリティにおける戦略*CD*の比率

取り続ける戦略と T を取り続ける戦略が対戦し、別の対戦相手とゲームするときには前者は R を、後者は P を取って互いの平均利得の帳尻を合わせる。これはある意味で戦略による role-play が成立していることを意味する（これに対し交互 ST 互恵では 1 戦略内で貪り役と貪られ役を交互に交代する）。定性的には、このような戦略が適応的になるためには、戦略適応をグループに分けることは有効であろうと推量される。その際、ゲーム対戦がそのグループと無関連な別グループ内で行われることは、この role-play 戦略群の進化には却ってバイアスとして作用するだろう。従って、ゲーム対戦に局所性があるよりもグローバルに行われる ($F=1$) 方が ST 互恵の創発には好適な結果となったと思われる。図 9, 10 から、ST 互恵を行うインセンティブのある Leader と Hero のエリアについて均衡時の戦略分布を確認すると、戦略更新ローカリティと対戦ローカリティが共に強い場合（右下パネル）は、確率 ST 互恵戦略の比率が減少する代わりに交互 ST 互恵戦略が増加している。これは、無関連な強い 2 つのローカリティが導入されると戦略更新ローカリティによってフリーライダーとなる戦略が駆逐されるが、対戦ローカリティによって複雑になった対戦状況に対応することができないために確率 ST 互恵を行うことが困難になり、僅かでも高い利得を得ようと交互 ST 互恵戦略を採るエージェントが増加したためと考えられる。確率 ST 互恵戦略と交互 ST 互恵戦略が混在する状況でも、互いに ST を取り合うことは出来るが、この両戦略同士が対戦を行うと確率 ST 互恵戦略が S、交互 ST 互恵戦略が T を取った次のステップは共に R を、確率 ST 互恵戦略が T、交互 ST 互恵戦略が S を取った次のステップは共に P を取り、S と T を取り合う間に必ず R もしくは P を取るステップが入るために効率が低下してしまう。そのため、両ローカリティが強い状況下では ST 互恵がそれほど強く顕れなくなったのだろう。

以上を要約すると、R 互恵を enhance して領域 C を縮退させるには、戦略更新ローカリティが強いことを前提にして、対戦ローカリティを強くすることが有効である（結局、戦略更新ローカリティも対戦ローカリティも強くする）。ST 互恵を enhance して領域 B を拡張させるには、対戦ローカリティが弱いことを前提にして、戦略更新ローカリティを強くすることが有効である（結局、戦略更新ローカリティだけを強くする）。

4.3 相関が有るローカリティが存在する場合

図 11 に $G=10, F=10$ の相関があるローカリティのシミュレーションの均衡時の結果を示す。

社会平均利得を確認すると、ST 互恵を行うインセンティブが有る $D_g \geq D_r + 2R$ の広範な領域で ST 互恵が創発するようになり、領域 B が拡大している。ローカリティが無い ($F=1, G=1$) では Nash 均衡に陥っていた領域 C も ST 互恵が構築される領域 B となり、全体的に高利得が得られるようになっている。

これは戦略更新ローカリティと対戦ローカリティが完全に相関しているため、

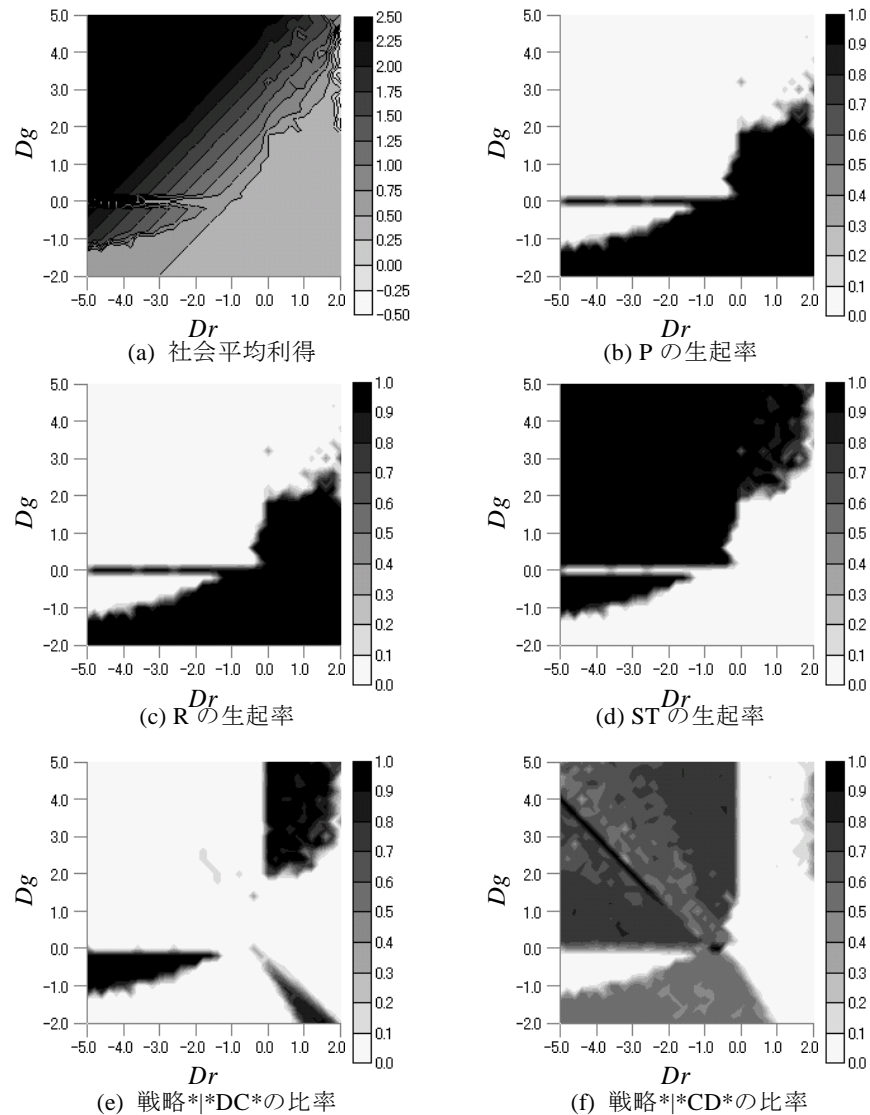


図 12 $G=10$, $F=10$ のシミュレーション結果

Chicken の領域では確率 ST 互恵を行う際に S と T のどちらを取るかを Gene Pool 単位で役割を切り替えることができるようになるため、確率 ST 互恵戦略である**CD*の比率が増加し、role-play が容易に達成され ST 互恵が容易に達成されるものと推測される。また、両ローカリティに相関がある場合、戦略更新の拡散範囲はローカリティが無相関な場合よりも限定的となるため、社会全体を Nash 均衡に陥れる裏切り戦略の拡散が抑えられ、互恵関係が構築されるようになる。

5. 結論

対戦、及び戦略更新時のローカリティを独立に変化させて 2 記憶長戦略付きの 2×2 ゲームを行い、以下の結果を得た。

- ローカリティが無い状況下では、Chicken 及び Trivial のゲーム構造を持つ領域で ST 互恵が創発し、PD のジレンマが強い領域では Nash 均衡にトラップされ、裏切り合いになる。その他のゲーム領域では R 互恵が構築される。この結果は Tanimoto & Sagara[3]の結果と一致する。
- 対戦ローカリティが弱い状況下で戦略更新ローカリティが強くなると ST 互恵が創発する領域が拡張する。一方、対戦ローカリティが強い状況下では戦略更新ローカリティが強くなるとともに ST 互恵が行われるゲーム領域は縮小するが、PD 内の Nash 均衡に陥っていた領域では R 互恵が構築される。
- 対戦ローカリティと適応更新ローカリティが無相関で共に強い場合、ST 互恵が構築されにくくなるが、両ローカリティに相関がある場合は ST 互恵が広範な領域で構築されるようになる。

参考文献

- 1) Hisashi Ohtsuki, Jorge M. Pacheco, Martin A. Nowak: Evolutionary graph theory: Breaking the symmetry between interaction and replacement, *Journal of Theoretical Biology* 246, pp.681-694 (2007).
- 2) Xiaojie Chen, Long Wang: Cooperation enhanced by moderate tolerance ranges in myopically selective interactions, *Physical Review E* 80 (2009).
- 3) Jun Tanimoto, Hiroki Sagara: A study on emergence of Coordinated Alternating Reciprocity in a 2×2 game with 2-memory length strategy, *BioSystems* 90 (3), pp.728-737 (2007).
- 4) Jun Tanimoto, Hiroki Sagara: Relationship between dilemma occurrence and the existence of a weakly dominant strategy in a two-player symmetric game, *BioSystems* Vol.90, No.1, pp.105-114 (2007).