

# 不平等ゲームにおいて進化した動的なシグナリングによる 行動の調整

## Coordination by Evolved Dynamic Signaling in a Game with Unequal Equilibrium Payoffs

高野雅典<sup>1,2\*</sup> 有田隆也<sup>1</sup>  
Masanori Takano<sup>1,2</sup> and Takaya Arita<sup>1</sup>

<sup>1</sup> 名古屋大学

<sup>1</sup> Graduate School of Information Science, Nagoya University

<sup>2</sup> 日本学術振興会特別研究員

<sup>2</sup> JSPS Research Fellow

**Abstract:** Communication is necessary for coordination among agents. We construct an agent based model where agents transmit continuous signals each other, and conduct evolutionary simulations in a game with unequal equilibrium payoffs. The dynamics of each other's signaling are analyzed.

### 1 はじめに

あるエージェントが自分の行動を他のエージェントの行動に依存させていて、他のエージェントにとってもそれが同様である場合、相手がどうするかを知ることなしにうまくいく自分の行動を決定することは難しい。このような状況において、相互に行動を調整するための手段として、コミュニケーションは有効と考えられる [8]。例えば、2体のエージェントが道ですれ違うような状況において、エージェントが互い衝突してしまわないようにするために、相互に自分のとるであろう行動（進行方向、避ける程度など）について情報を示すこと、相手を観察することを相互に行うことによって、それらの行動の組を衝突しないような行動の組へ調整できると考えられる [5]。この調整をしている際、互いのとるであろう行動の組が一方または両者にとって都合のよくない場合、それを調整するために交渉をする必要があり、情報の提示・観察を繰り返し行う、つまり動的に行う必要があると考えられる。

進化や学習などの適応によるエージェント集団のコミュニケーション成立のためには、情報を示すために発信されるシグナルと発信者の意図・状態の対応関係を獲得することが一つの大きな問題である [7, 8]。このような明示的な意味を持たないシグナルを用いて、動的な行動の調整に関する研究として例えば [6, 4, 2, 1] がある。それらでは関わるエージェントにとって調整

表 1: 対称化した「男女の争い」ゲーム

	A	B
A	0, 0	1, $\alpha$
B	$\alpha, 1$	0, 0

された行動の組は全員に平等であるという状況を扱っている。しかし、交渉という状況を考えると、例えば男女の争いゲームのような両者の最良の行動の組は異なるが、両者の行動の選択を調整しなければ両者にとって不利益という状況のように、関わるエージェントにとって調整された行動の組に不平等さがある場合も考えられる。このとき、エージェントの利害は完全には一致していないので、シグナルと発信者の意図・状態の対応関係の獲得はより難しくなると予想される。

本稿では不平等な調整の状況の最も単純な場合として男女の争いゲーム（表 1）を考える。表 1 ではエージェントの戦略と利得の組が対称になるように示している。  $0.0 < \alpha < 1.0$  の場合、このゲームでは両者が異なった戦略を選択しなければ共に利得を得られないが、その場合でも戦略 B を選択すると相手より利得が低くなる。したがって、自分が戦略 A を選択し、相手が B を選択する場合が最適である。この不平等さは  $\alpha$  が大きいほど弱くなり、1 になるとなくなるので、 $\alpha$  はゲームの平等度を表すといえる。以降ではゲームの結果が両エージェントが利得が得られる戦略の組（AB または BA）になることをゲームが成功するという。

\*連絡先：名古屋大学大学院情報科学研究科  
名古屋市中区  
takano@alife.cs.is.nagoya-u.ac.jp

まず、コミュニケーションをしない場合について述べる．男女の争いゲームにおける進化的に安定な戦略 (ESS) [3] は戦略  $s$  を選択する確率  $P(s)$  が  $(P(A), P(B)) = (1/(\alpha+1), \alpha/(\alpha+1))$  の戦略である [10]．その ESS 同士の対戦の平均利得は  $\alpha/(1+\alpha)$  であり、ゲームの成功率は  $2\alpha/(1+\alpha)^2$  である．ゲームの前にコミュニケーションをしない場合では  $\alpha = 1$  (不平等でない場合) でもゲームの成功率は 0.5 である．これはゲームに参加している 2 者が共に相手がどちらの戦略を選択するかわからないためであり、事前にコミュニケーションを行うことでゲームの成功率は高くなると考えられる [9]．ただし、前述したように  $\alpha < 1$  のときゲームが成功する戦略の組は不平等な利得をエージェントに与える．そのとき、ゲームに参加するエージェントにとって都合のいい結果 (自分が  $A$ , 相手が  $B$ ) は互いに異なり、交渉によって自分に都合のいい結果を導くことが重要になる．

本研究では、このような不平等な調整的状況における動的なコミュニケーションの進化的獲得とそのコミュニケーションによる調整のダイナミクスに焦点を当てる．行動の前に相互にシグナルを送受信し、その後同時に行動をするというモデルを設計し、男女の争いゲームを用い進化シミュレーションを行った．そして、獲得された調整行動はどのようにしてゲームが成功するような行動の組を形成していくか、そのダイナミクスについて解析した．

## 2 モデル

一回の対戦でエージェントは 1 対 1 で 1 回だけゲームを行う．その対戦では、まずシグナルの送受信を繰り返し行った後、戦略を決定しゲームをする．ゲームの前にシグナルの送受信を行う期間を以下では交渉期間という．集団内でゲームの総当たり対戦を行い、その対戦の平均得点をそのエージェントの適応度とした．そしてそれに基づき進化させた．

エージェントは状態として、直接戦略を表す  $a \in \mathcal{R}$  と戦略とは直接関係無いが相手に観測されるシグナル  $\phi \in \mathcal{R}$ 、内部状態  $I \in [-0.5, 0.5]$  の 3 つを持つ． $a$  は交渉期間の後、その値がそのエージェントの戦略  $s$  となる．本稿では 2x2 ゲームを用いるので、 $a \leq 0$  の場合は  $s = A$ 、 $a > 0$  の場合には  $s = B$  を選択するとした (表 2)．エージェントの制御にはニューラルネットワークを用いた．

ニューラルネットワークの入力は相手のシグナル  $\phi_{other}$  と自分の内部状態  $I_{self}$ 、出力は  $o_a, o_\phi, o_I$  である (出力の範囲はすべて  $[-0.5, 0.5]$ )． $I_{self}$  は前の時刻のニューラルネットワークの出力のうちの一つ  $o_I$  である．エージェントの入出力の関係を図 1 に示す．これら出力に

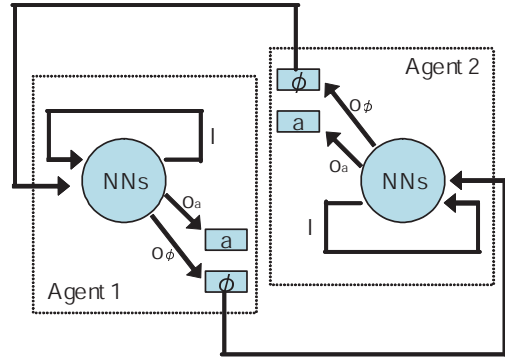


図 1: エージェントの構成

表 2: 利得行列

	A	B
A	$a_{11}, b_{11}$	$a_{12}, b_{21}$
B	$a_{21}, b_{12}$	$a_{22}, b_{22}$

よりエージェントは自分の状態  $a, \phi, I$  を次のように変更する．

$$a(t+1) = a(t) + o_a d \quad (1)$$

$$\phi(t+1) = \phi(t) + o_\phi d \quad (2)$$

$$I(t+1) = o_I \quad (3)$$

$d$  は定数． $t$  の範囲 (交渉期間) は  $[0, T]$  とする．これを  $t = 0$  から  $T$  まで (交渉期間の間) 繰り返し、 $a(T)$  の値を前述のように戦略  $s$  として評価しゲームを行う．

## 3 シミュレーション

### 3.1 設定

ニューラルネットワークの重みを実数値の遺伝的アルゴリズム [11] を用いて進化させた．エージェントの初期状態 ( $t = 0$ ) では  $a(0), \phi(0), I(0)$  を 0.0 とし、シグナルを送受信する回数  $T$  は 200、 $d$  は 0.05 とした．ニューラルネットワークは 3 層であり、入力ニューロンは 2 つ ( $\phi_{other}, I_{self}$ )、中間ニューロンは 4 つ、出力ニューロンは 3 つ ( $o_a, o_\phi, o_I$ ) とした．各層にバイアスニューロンがあるので重みの数は 31 となる．遺伝的操作として、突然変異は変異率 0.05 で平均 0、標準偏差 0.5 の正規分布に従って与え、交叉は交叉率 0.8 の二点交叉、選択はルーレット選択を用いた．遺伝子長はニューラルネットワークの重みの数になる．以上の方法で個体数 100 で 8000 世代実行した．利得行列は表 1 の  $\alpha$  ( $[0.0, 1.0]$ ) を 0.1 刻みで変え進化を行った．

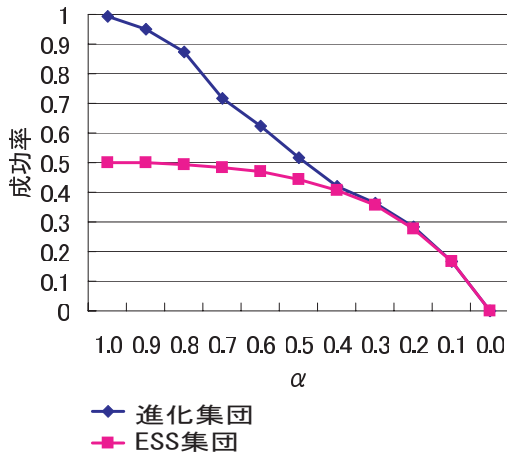


図 2: ゲームが成功した割合

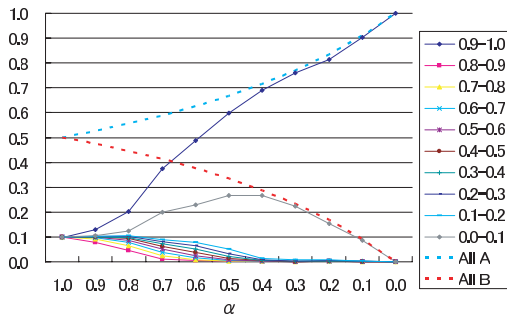


図 3: 戦略 A の選択割合別のエージェントの頻度

### 3.2 シグナリングの効果の基本的傾向

進化の結果得られた集団において、どのような相互作用をしているかについて示す。各  $\alpha$  について進化を 16 試行を行った。以降の節で示している値は 16 試行の平均である。

まず、進化の結果のおおまかな傾向として、各  $\alpha$  についてゲームの成功率と戦略に関する集団の分布について示す。以下で示している値は世代 7000 から 8000 の平均値である。

図 2 に進化によって得られた集団（以下進化集団）と ESS で構成される集団（以下 ESS 集団）の  $\alpha$  ごとのゲームが成功した割合を示す。進化集団は  $\alpha = 1$  でほぼ割合 1 でゲームが成功し、 $\alpha$  が小さいほどゲームが成功した割合は小さくなった。ここで、ESS 集団の成功率と比較すると、 $\alpha \geq 0.5$  において進化集団は ESS 集団を上回り、 $\alpha < 0.5$  では  $\alpha$  が小さくなるほど進化集団は ESS 集団に漸近していることがわかる。

図 3 に、各  $\alpha$  において集団で総当たり対戦をした場合に戦略 A を選択した割合をエージェントごとに求め、その頻度を求めたものを示す。加えて、純粋戦略の平

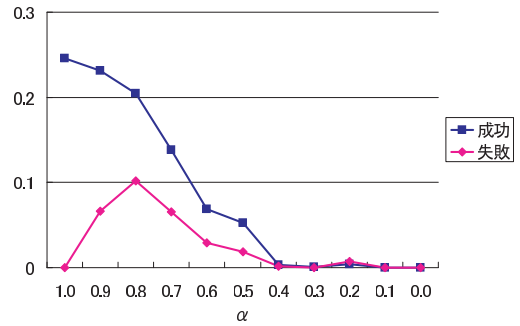


図 4:  $t = T$  における  $\phi$  と  $s$  の相互情報量

衡点の集団（以下純粋戦略集団）の A と B の割合も示す。純粋戦略の平衡点は A の割合  $= 1/(\alpha + 1)$ 、B の割合  $= \alpha/(\alpha + 1)$  である。進化集団は  $\alpha = 1$  で頻度は一様になった。 $0.5 \geq \alpha \geq 1.0$  では、 $\alpha$  が小さいほどほとんど一方の戦略しか選択しないエージェントの頻度（0.0 - 0.1, 0.9 - 1.0）が多くなった。さらに  $\alpha$  が小さくなり、 $0.0 \geq \alpha < 0.5$  ではほとんど A を選択するエージェントの頻度（0.9 - 1.0）が多くなった。ここで、純粋戦略集団の各戦略の割合と比較すると、 $\alpha < 0.5$  では  $\alpha$  が小さくなるほど 0.0 - 0.1, 0.9 - 1.0 は純粋戦略集団の各戦略の割合に漸近した。

以上より、 $\alpha$  が大きいと創発したコミュニケーションによる調整が極めてうまくいくが、0.5 より小さくなると急激にコミュニケーションが無効化し、交渉しない 2 つの純粋戦略の平衡点の状態と同様の集団に漸近していると考えられる。

### 3.3 シグナルに関する相互情報量の分析

前節で示されたようなコミュニケーションによる調整が、どのように行われているか、相互情報量を用いてシグナルの持つ意味について調べる。

まず、交渉期間後 ( $t = T$ ) の  $\phi(T)$  と選択した戦略  $s$  の相互情報量について調べた。対象とした集団は各進化の試行によって得られた最終世代である。相互情報量は各個体の全対戦の  $\phi(T)$  と  $s$  に関して求め、それを全個体で平均した。 $\phi(T)$  は -6 から 6 の範囲で幅 0.12 に区切って離散化し求めた。図 4 に各  $\alpha$  について相互情報量を示す。 $\phi(T)$  による情報とゲームの成功率の関係を調べるために、同図ではゲームが成功した場合 (AB または BA) と失敗した場合 (AA または BB) で分けて示している。成功した場合には  $\alpha$  が小さくなると相互情報量が減少した。特に、 $\alpha < 0.5$  では成功・失敗にかかわらず相互情報量はほぼ 0 になった。これは、 $\alpha$  が小さくなると純粋戦略のようなエージェントが集団内に多くなるためであると思われる。 $\alpha \geq 0.5$  では、ゲームが成功した場合は失敗した場合に比べ相互

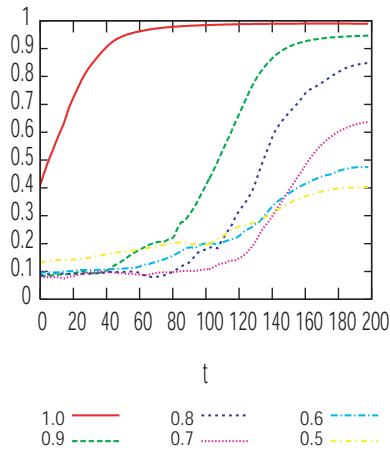


図 5: 相互情報量の時間変化

情報量が高くなった。以上より、 $\alpha \geq 0.5$  ではゲームが成功する場合、最終的なシグナルの値  $\phi(T)$  にそのエージェントの選択した  $s$  に関する情報が示されているといえる。

つぎに、エージェントの  $\phi(t)$  と選択した戦略  $s$  の相互情報量が交渉期間の間どのように変化し、前述のように  $\phi(T)$  が戦略  $s$  の情報を示すようになるかを調べた。図 5 に相互情報量の変化を示す。相互情報量がある点  $t$  で増加しているのがわかる。ここから、エージェントには相手に自分の戦略の情報を示し始める  $t$  に関する閾値  $t'$  があると考えられる。そして、 $t'$  は  $\alpha$  が小さいほど遅くなった。

図 6 に交渉期間中に相互情報量が 0 から増加したときの変化を示す。ここでは典型的な例として  $\alpha = 0.9, 0.8$  に関してのみ示している。ただし、同図は戦略  $A$  を選択した割合 (0.0-0.2, 0.2-0.4, 0.4-0.6, 0.6-0.8, 0.8-1.0) ごとに平均して示している。  $A$  を選択する割合の高いエージェントほど閾値  $t'$  が大きくなり、集団内に  $t'$  に関して多様性があることがわかる。 $\alpha$  が小さいと  $A$  を選択する割合の分布は偏る (図 3) ためこの多様性は減少する。このことから、自分の戦略に関する情報を相手に示すことは、自分が  $B$  を選択することを相手に示すことであると考えられる。

$\alpha = 1.0$  では、ゲームが成功した場合のエージェントの利得は等しく、選択された戦略が異なればよいため、閾値  $t'$  が集団中で多様性を持っていればよいと考えられる。しかし、 $\alpha < 1.0$  では、ゲームが成功した場合のエージェントの利得は異なるため、閾値  $t'$  による区別に  $\alpha$  による利得差が関わってくる。そのため、 $t'$  は大きくなる方向へ進化したと考えられる。しかし、 $t'$  が大きいエージェント同士では、戦略の分担が間に合わずゲームは失敗しやすい。そのため、閾値  $t'$  のそれほど大きくないエージェントも集団中に一定割合で存在

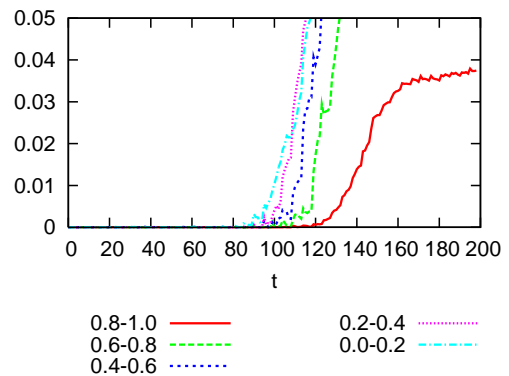
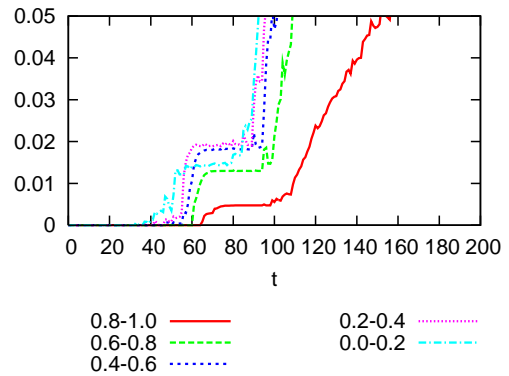


図 6: 戦略  $A$  を選択した割合別の相互情報量の時間変化 ( $\alpha = 0.9$  (上),  $\alpha = 0.8$  (下))

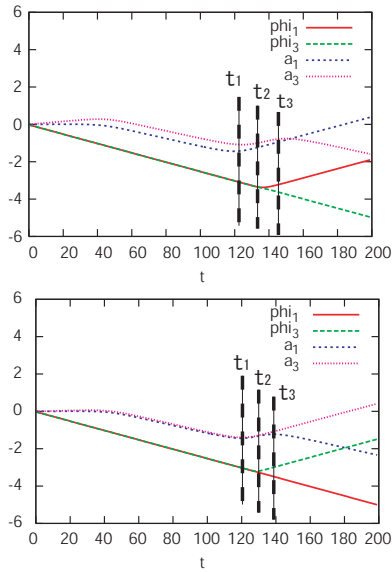


図 7: ゲームが成功した場合の 2 体のエージェントの  $\phi(t)$  と  $a(t)$  の典型的な軌道 ( $\alpha = 0.9$ )

して進化したと考えられる。その結果、 $\alpha$  が小さくなるほど、純粋戦略が増加し、相手次第という側面はなくなるため、 $\phi(t)$  の意味はなくなったと思われる。

### 3.4 行動の調整方法の分析

以下では、どのように 2 体のエージェントが  $\phi, a$  を変化させ、戦略の選択を調整しているかを、具体的な場合を用いて説明する。図 7 に  $\alpha$  が 0.9 におけるゲームが成功したエージェントの組の  $\phi(t)$  と  $a(t)$  の典型的な軌道を示す。初期状態は  $(\phi(0), a(0)) = (0.0, 0.0)$  のため、原点から軌道は始まる。同図の  $\phi_{i_1}(t)$  と  $a_1(t)$  はエージェント 1,  $\phi_{i_2}(t)$  と  $a_2(t)$  はエージェント 2,  $\phi_{i_3}(t)$  と  $a_3(t)$  はエージェント 3 の状態をあらわす。同図では 2 つ軌道の  $a(t)$  の符号が  $t = T$  で異なることがゲームが成功したことをあらわす。

両図とも以下に示すことを行っていると考えられる。最初の段階でエージェントは  $a(t)$  は  $A$  を選択する方向へ動かす。しかし、両者が  $A$  を選択するとゲームが失敗になるため、図中の  $t_1$  において  $a(t)$  の速度を  $B$  方向へ変えはじめる。 $t_2$  において、一方のエージェントがその変更を  $\phi(t)$  の速度を逆方向へ変えることで示す。 $t_3$  において、もう一方のエージェントがそれに反応し、 $t_1$  で変えた  $a(t)$  の速度を  $A$  方向へ変える。ここで、 $t_2$  において  $\phi(t)$  の速度を変えたのは、上の図ではエージェント 1 であるが、下の図ではエージェント 3 である。つまり、エージェントの持つ閾値が異なることによって相手によって  $A$  を選択する場合、 $B$  を選択する場合が変わり、 $\phi(t)$  による調整がなされている。

## 4 まとめ

本論文では、動的シグナリングの中に意味を進化的に獲得する現象における不平等性の役割を検討するために、同時に行動する前にシグナルを一定期間相互に送るモデルを設計し、男女の争いゲームを用いて進化シミュレーションを行った。このゲームでは、両者が同じ戦略をすると両者の利得は 0 になるため、一方だけが戦略を変更する必要があるが、同時に変更してしまうことを避けるために、自分が戦略を変更したことを相手に伝えることで利得が増加する。実験の結果、自分の選択しようとしている戦略を変更したことを相手に伝える役割をシグナルが持つように進化したことが示された。利得の差が小さい、つまり相互作用が平等に近いならば、シグナルを送るタイミングの多様性のために、このコミュニケーションシステムは機能する。しかし、獲得する利得の差が開いて平等度が小さくなっていくと、タイミングを遅らせる選択圧が強まり、同時に、戦略の変更が間に合なくなるリスクが増加する。そして、コミュニケーションは崩壊し、相手によらない純粋戦略の共存という分極化したすみ分け社会に等しくなってしまう。

## 参考文献

- [1] A. Lorincz, V. Gyenes, M. Kiszlinger, and I. Szita. Mind model seems necessary for the emergence of communication. *arXiv:q-bio.NC/0611012*, Vol. 1, 2006.
- [2] Davide Marocco and Stefano Nolfi. Self-organization of communication in evolving robots. In *Proc. of ALIFE X (10th International Conference on the Simulation and Synthesis of Living Systems)*, 2006.
- [3] John Maynard Smith and George R. Price. The logic of animal conflict. *Nature*, Vol. 246, pp. 15–18, 1973.
- [4] John H. Miller and Scott Moser. Communication and coordination. *Complexity*, Vol. 9, No. 5, pp. 31–40, 2004.
- [5] Natalie Sebanz, Harold Bekkeringb, and Günther Knoblich. Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, Vol. 10, No. 2, pp. 70–76, 2006.
- [6] 柴田克成, 伊藤宏司. 利害衝突回避のための交渉コミュニケーションの学習-リカレントニューラルネットワークを用いたダイナミックコミュニケーションの

学習— 計測自動制御学会論文集, Vol. 35, No. 11, 1999.

- [7] Luc Steels. Evolving grounded communication for robots. *TRENDS in Cognitive Sciences*, Vol. 7, pp. 308–312, 2003.
- [8] Luc Steels. Experiments on the emergence of human communication. *TRENDS in Cognitive Science*, Vol. 10, pp. 347–349, 2006.
- [9] 岡田章. ゲーム理論. 有斐閣, 1997.
- [10] 石原英樹, 金井雅之. 進化的意思決定. 朝倉書店, 2002.
- [11] 有田隆也. 人工生命. 医学出版, 2002.