

マルチエージェント環境におけるコミュニケーションの段階的創発

天正新二郎[†] 前川 聡^{††} 吉本潤一郎^{†,†††} 柴田 智広[†] 石井 信[†]

[†] 奈良先端科学技術大学院大学 情報科学研究科 〒630-0192 奈良県生駒市高山町 8916-5

^{††} 情報通信研究機構 〒619-0289 京都府相楽郡精華町光台 3-5

^{†††} 沖縄大学院大学 先行的研究事業 〒904-2234 沖縄県具志川市州崎 12-22

E-mail: †{shinji-t,tom,ishii}@is.naist.jp, ††maekawa@nict.go.jp, †††jun-y@irp.oist.jp

あらまし 近年、コミュニケーション創発のメカニズムを計算機シミュレーションによって解明しようという構成論的アプローチの研究が盛んである。コミュニケーション創発における計算論的問題の1つは、メッセージの発信側と受信側の双方が、同じ時期にコミュニケーション成立のための機能を獲得していなければならないことである。そこで本研究では、マルチエージェント競合環境における進化・学習エージェントのコミュニケーション創発メカニズムを提案する。このメカニズムにおいて重要な点は、エージェントが即応的に対処すべきである環境的要因が環境中に存在することである。環境的要因に対処するために獲得された機能は、エージェントによって間接的に利用され、コミュニケーションを段階的に創発する。環境中に環境的要因としての危険がある場合には、それが無い場合と比べて、コミュニケーションが創発しやすく、環境への適応度も高くなることをシミュレーションによって示す。

キーワード コミュニケーション, 競合環境, 段階的創発, 強化学習, 進化

Gradual emergence of communication in a multi-agent environment

Shinjiro TENSHO[†], Satoshi MAEKAWA^{††}, Junichiro YOSHIMOTO^{†,†††}, Tomohiro SHIBATA[†],
and Shin ISHII[†]

[†] Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0192, JAPAN

^{††} National Institute of Information and Communications Technology
3-5 Hikaridai, Seika, Soraku, Kyoto 619-0289, JAPAN

^{†††} Initial Research Project, Okinawa Institute of Science and Technology
12-22 Suzaki, Gushikawa, Okinawa 904-2234, JAPAN

E-mail: †{shinji-t,tom,ishii}@is.naist.jp, ††maekawa@nict.go.jp, †††jun-y@irp.oist.jp

Abstract Synthetic approaches have been popular to uncover mechanisms for the emergence of communication. One of the computational problems for the emergence of communication is that both a message sender and a receiver need to finish acquiring behavioral functions which enable communication to happen. This report presents a mechanism that deals with this computational problem in a competitive multi-agent system. The key point of the mechanism is there exists some environmental and fatal factors which agents have to cope with immediately. Once acquiring a function to cope with the environmental factors, agents divert it to another issue and then communication emerges. Simulation results supported our insights : (1) a fatal factor in environment helps agents to emerge communication with others; and, (2) communication brings higher fitness to agents.

Key words communication, competitive environment, gradual emergence, reinforcement learning, evolution

1. はじめに

計算機シミュレーションを用いた構成論的アプローチから、生物のコミュニケーション創発を解明しようという研究が近年

進められてきている。MacLennan は、計算機上の人工的な環境で動物行動学的現象を再現する Synthetic Ethology という方法論を応用し、シンボル選択とシンボルの意味づけに着目した生物のコミュニケーションメカニズムやその役割の解明を試

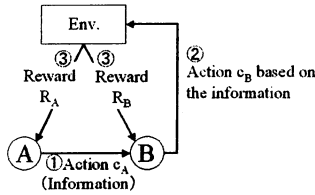


図1 コミュニケーションの概要図

みている [1], [2]. ここで、シンボルとは情報の送受信における抽象化されたプロトコルを指す。

本研究では、MacLennan によって提案された生物のコミュニケーションモデルを基に、マルチエージェント競合環境におけるコミュニケーション創発の仕組みについて議論する。

2. コミュニケーション創発

2.1 コミュニケーションとは

本研究では、コミュニケーションを以下のように定義する。

図1に示されるように、環境 Env. 中に直面するポジション A とポジション B があり、ある2体のエージェントがこれらのポジションに存在している状況を考える。ポジション A (ポジション B) にいるエージェントをエージェント A (エージェント B) と呼ぶものとし、この2体のエージェントは以下の流れで行動を実行するものとする。

(1) エージェント A は、エージェント B に何らかの情報を送信するための行動 $c_A \in A$ を実行する。ここで、 A はエージェントが実行可能な行動の有限集合を表している。

(2) エージェント B は、エージェント A からの情報を受けて行動 $c_B \in A$ をある条件付き確率分布 $P(c_B|c_A)$ にしたがって実行する。

(3) この結果として、エージェント A に対して収益 $R_A \in \mathcal{R}$ がある条件付き確率分布 $P(R_A|c_A, c_B)$ にしたがって与えられ、エージェント B に対して収益 $R_B \in \mathcal{R}$ がある条件付き確率分布 $P(R_B|c_A, c_B)$ にしたがって与えられる。

今、自己の外部への相互作用を全く与えない、あるいはその度合いが最も小さな基準行動 c_{Stay} が与えられているものとする。この時、エージェント間のコミュニケーションが成立しているとは、以下の2つの条件が成り立つこととして定義される。

$$\begin{cases} \sum_{c_B} P(c_B|c_A)E[R_A|c_A, c_B] \\ \geq \sum_{c_B} P(c_B|c_{Stay})E[R_A|c_{Stay}, c_B] \\ E[R_B|c_A, c_B] \geq E[R_B|c_A, c_{Stay}] \end{cases} \quad (1)$$

なお、基準行動 c_{Stay} は問題によって定める必要がある。本研究のシミュレーションにおいては後で述べる。また、ここでの行動 $c_A(c_B)$ をエージェント A (エージェント B) のコミュニケーション行為と呼ぶ。

2.2 コミュニケーションの段階的獲得

実際の生物のように学習や進化などの適応能力をもつエージェント間でのコミュニケーション創発問題を扱う場合、以下

の困難性が存在することが知られている [1], [2].

(1) 情報の送信側と受信側が同時期に適切な行動を獲得していないとコミュニケーションは発現しない。特にコミュニケーション行為にコストがかかる場合はより困難である。

(2) 2エージェント間でのコミュニケーション方法を全エージェント間での統一的なコミュニケーション方法とするのが難しい。

本研究では、これらの困難性が以下の条件を満たす環境的要因によって解消されているという仮説を立て、それについて議論を進める。

コミュニケーションの段階的獲得

環境中にコミュニケーションの目的とは直接的には無関係な要素 D が存在するものとする。このとき、

(1) 全てのエージェントが、要素 D に対処するための行動を獲得することで情報受信側のシンボルの意味が統一的に固定される。

(2) 情報送信側は、(1) で統一された受信側の反応を利用し、目的 P を実現するために必要なシンボルを送信する機能を獲得する。

(3) (1) と (2) により、要素 D への対処行動を利用した、目的 P のためのコミュニケーションが段階的に獲得される。

このような段階的創発は、生物の進化過程でも見られるものである。例えば、鳥は飛翔の際にうずくまり、尾を上げ、翼を広げるという典型的な動作をする。しかしながら、進化過程を経て、この飛翔動作は敵意をもつ相手への威嚇を表現する目的としても利用されている [3].

3. モデル

本研究では、Synthetic Ethology および計算論的な観点からこの段階的創発を検証する。そのために、我々は以下で示されるマルチエージェント競合環境モデルを構築し、計算機シミュレーションを行った。

3.1 環境の概要

同じ能力をもつエージェント群、資源、危険によって構成されるマルチエージェント競合環境を想定する。各エージェントは、環境中に存在する資源の獲得、危険の回避、他エージェントとの資源競合といった状況を経験し、報酬の蓄積である適応度 (fitness) を高めるように学習を行う。ある時刻を経ると、各エージェントが期間内に獲得した適応度を指標として、次世代に生き残るか、死滅するかが決まる。エージェントは世代交替を繰返し、進化によって環境に適応する。各エージェントの目的は、ある期間内において自己の適応度を最大化することであるが、エージェント集団の目的は、世代を経て適応しやすいエージェントを残すことである。

このような競合環境は、鳥などの種間の縄張り争いのモデルとみることができる [3]. 縄張りを争う種では、お互いの干渉を少なくし、遺伝的適応度の損失を最小にするような進化が種全

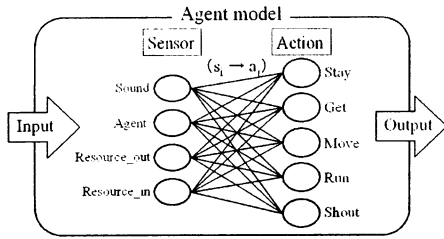


図2 エージェントモデル

体を保存する上で理想的となるからである。

3.2 エージェントモデル

各エージェントは、以下の機能を有するものとする。

3.2.1 現在の状況とセンサ知覚

エージェントは、各時刻において以下の5種類の状況に遭遇する。

- (1) 何もない状況
- (2) 環境中の資源と直面している状況
- (3) 他エージェントと直面している状況
- (4) 他エージェントが占有する資源と直面している状況
- (5) 大きな音のする危険と直面している状況

エージェントはある状況に遭遇すると、センサ入力を通じて以下の4種類の情報を同時に知覚することができる。

- Sound : 環境中に大きな音になっているかどうか
- Agent : 他エージェントと直面しているかどうか
- Resource_out (以下 R_o とする) : エージェントが資源と直面しているかどうか
- Resource_in (以下 R_i とする) : エージェント自らが既に資源を占有しているかどうか

これらのセンサ入力は、 $S = (s_1, s_2, s_3, s_4)$, $s_i \in \{0, 1\}$ ($i = 1, \dots, 4$) という2値ベクトルとして記述される。ここで、 $s_i = 1$ ($s_i = 0$) は i 番目の情報が知覚されている (知覚されていない) ということを示す。

3.2.2 行動選択

エージェントの選択可能な行動の集合は $A \equiv \{a_1, \dots, a_5\}$ であり、それぞれ以下のように定義される。

- $a_1 \equiv$ Stay : 現在の状況に居続ける。本研究における基準行動とする。
- $a_2 \equiv$ Get : 環境中の資源を獲得する。あるいは相手エージェントに攻撃する。
- $a_3 \equiv$ Move : 他の状況に遷移するが、危険や他エージェントからの退避はできない。占有資源は放棄される。
- $a_4 \equiv$ Run : 他の状況に遷移し、危険や他エージェントから退避できる。コストを支払う上に占有資源は放棄される。
- $a_5 \equiv$ Shout : コストを支払い、大きな音を出す。

各エージェントは、センサ入力 S から、進化・学習によって適応的に変化する結合荷重を通じて各行動の選択確率を計算する。そして、この選択確率に基づいてひとつの行動を選択する。以下に、行動選択のアルゴリズムを示す。

- (1) 環境からの情報を知覚し、それをセンサ入力 S とする。

- (2) 各行動の効用値を以下で計算する。

$$u_j = \sum_{i=1}^4 w_{ij} s_i \quad (2)$$

ただし、 w_{ij} は i 番目のセンサから j 番目の行動への結合荷重であり、 u_j は現在の入力における j 番目の行動の効用値である。

- (3) 選択可能な行動集合 A から行動 a_j を以下の確率 π_j で選択する。

$$\pi_j = \frac{e^{u_j}}{\sum_{b=1}^5 e^{u_b}} \quad (3)$$

3.2.3 適応度

各エージェントは、選択された行動に応じて、環境から資源を獲得できたり、あるいはコストを支払ったりする。この獲得報酬の増減を適応度 (fitness) として表現する。時刻 t における適応度を f_t 、その時に獲得される強化信号を r_t 、占有資源を q_t とすると、適応度 f_t は以下のように定義される。

$$f_t := f_{t-1} + r_t + q_t \quad (4)$$

各時刻 t における強化信号 r_t は以下で定義される。

$$r_t := R - C - D \quad (5)$$

ここで、 R, C, D はある定数であり、それぞれ以下のような意味を持つ。

- R : 獲得された資源量を表す。資源 R を獲得した場合は強化信号に加算される。
- C : 行動にかかるコストを表すものであり、行動 Run と行動 Shout にはそれぞれ c_r と c_s というコストがかかる。
- D : 戦闘によるダメージ量を表すものであり、他エージェントとの戦闘が起こる場合および危険を回避しない場合には、それぞれ、ダメージ d と d_d を被ることになる。

また、自己の占有資源 q_t は以下のように定義される。

- エージェントの占有する資源は、時間の経過とともに1ずつ消費される。すなわち $q_t := q_{t-1} - 1$ である。ただし、 $q_{t-1} = 0$ のとき $q_t := 0$ となる。
- 資源を獲得した場合には、 $q_t := R$ となる。
- 行動として Move あるいは Run が選択された場合、 $q_t := 0$ となる。

ここで、 R, C, D は競合的状况を設定するために $C \ll R \leq D$ という大小関係を持つ。このことから、エージェントはできるだけ D を避け、 R を多く獲得することが目的となる。

3.2.4 学習と進化

エージェントは、学習と進化という2つの時定数の異なる適応を行う。エージェントは、生存する世代内では強化信号 r_t に基づく強化学習を行う。具体的には、各時刻 t においてセンサ-行動間結合荷重 w_{ij} が以下の手続きで更新される。

- (1) TD-error δ を計算する。

$$\delta = r_t + \gamma V(S_t) - V(S_{t-1}) \quad (6)$$

ただし、 $\gamma \in (0, 1)$ は割引率であり、 S_t は時刻 t におけるセンサ入力である。 $V(S_t)$ は状態 S_t に関する状態価値関数である。

(2) 状態価値関数 $V(S_t)$ を更新する.

$$V(S_{t-1}) := V(S_{t-1}) + \alpha \delta \quad (7)$$

ただし, $\alpha \in (0, 1)$ は学習率である.

(3) 結合荷重 w_{ij} を更新する.

$$w_{ij} := w_{ij} + \beta \delta s_i \quad (8)$$

ただし, $\beta \in (0, 1)$ はステップサイズパラメータである. この学習は, actor-critic 法 [4] と呼ばれている.

進化の過程における適応は, actor-critic 法における学習可能性を変化させるためのものである [5], [6].

全てのエージェントがある固定された時刻を経過した後, エージェントは次の世代へ移る. このとき, 獲得した適応度 f_t の大きいエージェントは繁殖し, 適応度の小さいエージェントは死滅する. 世代間において, エージェントの総数は一定であり, 繁殖数と死滅数は同数である. エージェントは世代が変わる際には, 以下の手続きで適応度と状態価値関数が初期化され, センサー行動間結合の初期値が遺伝される.

(1) 適応度 f を以下で初期化する.

$$f := 0 \quad (9)$$

(2) 状態価値関数 $V(S)$ を以下で初期化する.

$$V(S) := 0, \quad \text{for all } S \quad (10)$$

(3) 初期結合荷重 w_{ij}^0 を以下で遺伝させる.

$$w_{ij}^0(T) := w_{ij}^0(T-1) + \epsilon \quad (11)$$

ただし, $w_{ij}^0(T)$ は T 世代目における結合荷重の初期値であり, ϵ はある白色ガウスノイズである.

(4) 結合荷重を以下で初期化する.

$$w_{ij} := w_{ij}^0(T) \quad (12)$$

3.3 競合的状況

本研究で想定した環境では, 資源を占有しているエージェント (防衛エージェント) とその資源を知覚しているエージェント (侵入エージェント) が対峙している状況で競合が生じる. この時, 侵入エージェントのセンサは “占有されていない資源” であるか, “他エージェントの占有する資源” であるかの区別がつかないため, 知覚のエイリアシングが生じる. したがって, 侵入エージェントが相手が占有している資源を獲得しようとする時, 防衛エージェントとの無益な戦闘が生じる.

一方, 防衛エージェントは侵入エージェントの存在を知覚可能である. 本来ならば, エージェント同士の戦闘は行動 Run によって回避すべきであるが, 自己の占有資源を放棄することは自己の適応度増加の点からして不利益になる. つまり防衛エージェントは, 可能な限り資源を放棄することなく戦闘を回避したい. そのためには, 侵入エージェントに自己の存在を知らせる何らかの方法をとる必要がある. よって, 行動 Shout を選択することが, 資源の種類を区別するための付加的情報を伝達し, 無益な戦闘を回避させるひとつの方法となる.

しかしながら, このような競合回避のコミュニケーションが創発されることは自明ではない. その理由のひとつとして, エージェントの発声行為はコストを要することがあげられる. 競合的状況において, 防衛エージェントが警告の意味で声をあげたとしても, 侵入エージェントはそれを警告と認識できずに戦闘になってしまうかもしれない. 防衛エージェントは, 声をあげたにも関わらず戦闘になってしまうということは, その状況において最もリスクの大きい行動を選択したことになり, この方策を獲得することは難しくなる. また, マルチエージェント環境では, あるエージェント間では, 防衛エージェントが行動 Shout を選択することによって侵入エージェントが行動 Run を選択してくれるかもしれないが, 他の侵入エージェントは行動 Get を選択してくるかもしれない. このような場合は, コミュニケーションが成立する保証はなく, しかもエージェント数が増えるにつれて困難な問題となる.

本研究においては, 以上のような困難性を含む環境においても, 環境的要因としての危険の存在によってコミュニケーションが創発されるかどうかについてシミュレーションによって調査した.

4. シミュレーション

4.1 シミュレーション設定

計算機シミュレーションでは, 危険の有無がコミュニケーションの創発に影響を与えるかどうか注目して, 我々の仮説を検証した. 具体的には, 環境中に “危険あり” の場合と “危険なし” の場合についての比較実験を行い, 以下 2 点について検証した.

(1) 競合的状況における全エージェントの行動選択傾向

(2) 全エージェントの世代内獲得適応度の総和

表 1 は, シミュレーション時のパラメータ設定を示したものである.

4.2 各状態における全エージェントの行動選択傾向

まず各状態における全エージェントの行動選択傾向を調べるにあたり, 競合的状況における状態 Agent+R_i (防衛側の知覚) と状態 Sound+R_o (防衛側が声を発したときの侵入側の知覚) 以外の状態においては, 危険の有無に関わらず行動選択傾向に大きな違いは表れず, 獲得適応度の差を見る際に本質ではないと考えた. よって, ここでは状態 Agent+R_i と状態 Sound+R_o に限定して行動選択傾向を調べることにした. 以下で示される行動選択傾向を表現するグラフにおいて, 横軸はシミュレーションにおける世代を, 縦軸は実行した行動の選択割合 (ある行動の回数/全行動選択回数) をそれぞれ表している.

まず, “危険あり” 環境において競合的状況が生じた時の行動選択傾向を調べた. 図 3 上は状態 Agent+R_i での行動の選択割合を, 図 3 下は状態 Sound+R_o での行動の選択割合を, それぞれ示したものである. 適応過程を通して, 状態 Agent+R_i では行動 Shout が最も多く選択されるようになることが読みとれる. これは, 相手の状況, つまり状態 Sound+R_o において相手が行動 Run をとる傾向が高くなっているということが同

表 1 シミュレーションにおけるパラメータ設定

エージェント数	100
遷移確率 $P_i (i = 1, \dots, 4)$ “危険あり” (なし, 危険, agent, 資源)	(0.55, 0.05, 0.2, 0.2)
“危険なし” (なし, 危険, agent, 資源)	(0.6, 0.0, 0.2, 0.2)
世代間時刻ステップ数 t	5000
最大世代ステップ数 T	500
状態知覚数 S	4
行動数 A	5
行動 Run のコスト c_r	1
行動 Shout のコスト c_v	1
危険からのダメージ d_d	20
戦闘によるダメージ d	5
資源獲得による報酬 R	5
占有資源 q	毎ステップ 1 ずつ消費
学習率 α	0.9
割引率 γ	0.9
ステップサイズパラメータ β	0.5
1 世代交替における繁殖/死滅数	5 / 5
ガウシアンノイズ ϵ	平均 0, 分散 0.01

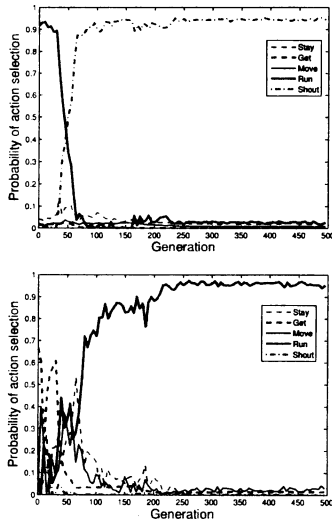


図 3 競合状況における行動選択 (危険あり)

時にわかる。この方策にしたがえば、防衛エージェントが行動 Shout をとり、かつ、侵入エージェントが Sound を知覚して行動 Run をとっているため、これらのエージェント間での無益な戦闘を回避できる。このとき、侵入エージェントの視点で考えると、“環境中の資源が大きな音を発する”ということを知覚することによって、“環境中の資源”と“他エージェントが占有している資源”との曖昧性が消されているといえる。2 エージェント間でこの対応づけに成功した時、コミュニケーションが創発したということができ、防衛エージェントの行動 Shout と侵入エージェントの行動 Run がそれぞれコミュニケーション

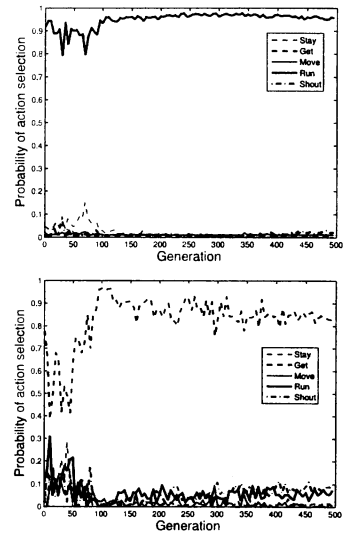


図 4 競合状況における行動選択 (危険なし)

ン行為に対応する。

本研究で設定した環境におけるコミュニケーションの段階的創発について議論する。本シミュレーションにおけるコミュニケーションの段階的創発とは、まず危険回避のための方策が全エージェントによって獲得された後に、その方策を利用してエージェント間の無益な戦闘を回避する方策が獲得される一連の過程といえる。

コミュニケーション行為を何らかの情報伝達手段として考えた場合、競合状況では、情報送信側では行動 Shout を選択する方策が、情報受信側では Sound を知覚して行動 Run を選択する方策が、ともに同時期に獲得されていなければならない。環境中の危険は、

- (1) 受信側の認識系を獲得
- (2) 送信側の出力系を獲得

というコミュニケーション機能の段階的な獲得を促すものといえる。なぜこのような段階的な獲得が実現されるかという理由は、エージェントの状態分割の方法にある。エージェントは、まず危険に対して状態 Sound を知覚して行動 Run を選択するように結合荷重を学習する。また別の状況では、資源 R.o を知覚して行動 Get を選択するように結合荷重を学習する。そして競合状況における侵入エージェントの知覚ははじめは状態 R.o であり行動 Get との間の結合荷重が強く学習されている。ここで防衛エージェントが行動 Shout を選択することによって侵入エージェントの知覚が状態 Sound+R.o となる。つまり、ここで行動 Run が選択されるようになる理由は、状態 Sound+R.o における状態 Sound と行動 Get との結合荷重が負となり、状態 Sound と行動 Run との結合荷重が大きく正に傾くからである、と推測される。その結果として、競合状況におけるコミュニケーションが容易に創発できたと考えられる。

次に、“危険なし”環境においての競合状況が生じた時の行動選択傾向を調べた。図 4 上は状態 Agent+R.i での行動の選

択割合を、図4下は状態 Sound+R.o での行動の選択割合を示したものである。この場合、状態 Agent+R.i における行動 Run を選択する傾向が高くなっている。これは、状態 Sound+R.o が知覚されたとしても行動 Get が選択する割合が終始高くなっており、状態 Sound から行動 Run への結合荷重が適切に学習されなかったからである、と推測される。状態 Agent+R.i では、行動 Shout や行動 Stay などを選択していても侵入エージェントが行動 Get を選択するために、占有資源を捨てて行動 Run を選択する行動の傾向が高くなっている。またここで行動 Run の行動傾向は、状態 Agent と行動 Run との間の結合荷重が大きな正の値をとるように学習されていることも理由のひとつと考えられる。

環境中に危険が存在する場合は、状態 Sound と行動 Run の間の結合荷重によって汎化が起こっているが、環境中に危険がない場合は、この状態 Sound と行動 Run の間の結合荷重が競合的状況での行動選択に与えるほど大きなものではない。よって状態 Sound+R.o においては状態 R.o と強い結合荷重をもつ行動 Get が出易くなる。以上のような原理が働き、“危険なし”の環境では、コミュニケーションが創発されなかったと考えられる。

4.3 全エージェントの世代内獲得適応度の総和

“危険あり”の場合と“危険なし”場合における全エージェントの世代内獲得適応度を図5に示す。横軸はシミュレーションにおける世代を、縦軸は1世代当たりの獲得適応度を、それぞれ表している。

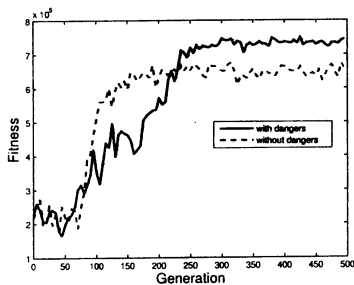


図5 獲得適応度の比較(危険の有無)

結果として、“危険あり”の場合の方が獲得適応度は高くなっている。なお前述のように、競合的状況以外の状態における行動選択傾向はコミュニケーションの有無に関わらずほぼ同等のものであり、獲得適応度にも大きな差は見受けられなかった。つまり、競合的状況における行動選択が獲得適応度の差を生じさせる主たる要因であると考えられる。

5. おわりに

本研究において、マルチエージェント競合環境における環境的要因である危険の存在が競合的状況におけるコミュニケーション創発を促すということを示唆し、シミュレーションによって確認した。コミュニケーションの創発には、情報の送信側と受信側の双方が同時期にコミュニケーションに関する機能を

獲得していなければならない、また、2エージェント間でのコミュニケーション方法を全エージェント間での統一的なコミュニケーション方法とするという困難性が存在する。これに対して、環境中にある環境的要因が存在することで、各エージェントはこれに対処する方策を獲得し、競合的状況におけるコミュニケーション創発を促進させるためにこの要因からの方策を利用した。

本研究の内容において、コミュニケーション行為は、情報の受信側・送信側それぞれひとつずつという単純な形でのシミュレーションであった。具体的には、エージェント間の競合的状況において、防衛エージェントは行動 Shout、侵入エージェントは行動 Run がそれぞれコミュニケーション行為であった。よって、エージェントのもつコミュニケーションシンボル数を増やしたより複雑なモデルにおいて環境的要因の存在がコミュニケーション創発に与える影響を調査することが今後の課題として考えられる。

また、マルチエージェント環境におけるコミュニケーションを考える際に、利他行動、特に血縁関係を含まない場合の互恵的利他行動[7]~[9]は、集団の生存確率の最大化に大きく寄与してくる概念である。マルチエージェント環境において、利他行動が創発されるにはどのような条件が必要であるか、また、利他行動をとる集団の形成過程はどのようなものかなどの調査は、本研究の延長にあるものとして考えることができる。

文 献

- [1] B. MacLennan : Evolution of Communication in a Population of Simple Machines, Technical Report CS-90-99, Computer Science Department, University of Tennessee, Knoxville, 1990
- [2] B. MacLennan : Synthetic Ethology -An Approach to the Study of Communication-, Artificial Life II, pp.631-658, 1991
- [3] エドワード・O. ウィルソン(著), 坂上 昭一, 宮井 俊一, 前川 幸恵, 北村 省一, 松本 忠夫, 粕谷 英一, 松沢 哲郎, 伊藤 嘉昭, 郷 采人, 巖佐 庸, 羽田 節子(翻訳) : 社会生物学, 新思案社, 1999
- [4] R. S. Sutton and A. G. Barto : Reinforcement Learning - An Introduction, MIT Press, Cambridge, MA, 1998
- [5] D. Ackley and M. Littman : Interaction Between Learning and Evolution, Artificial Life II, pp.487-509, 1991
- [6] 早川 充洋, 大森 清博, 前川 聡, 藤原 義久, 北村 新三 : 強化学習エージェントの進化と学習による環境への適応行動の発生, 計測自動制御学会 第13回自律分散システム・シンポジウム, 2001
- [7] リチャード・ドーキンス(著), 日高 敏隆 他訳 : 利己的な遺伝子, 紀伊国屋書店, 1991
- [8] 長谷川 寿一, 長谷川 眞理子 : 進化と人間行動, 東京大学出版, 2000
- [9] 中山 功一, 松井 博和, 野村 由司彦 : 動的隔離型 GA(DS-GA) の提案, 情報処理学会論文誌 数理モデル化と応用, Vol. 43, No. SIG 10(TOM 7), pp.95-109, 2002