

## 5F-07 遺伝的アルゴリズムを用いたマルチエージェント問題に関する検討

西山 敦 梶川 嘉延 野村 康雄

関西大学工学部電子工学科

### 1. はじめに

近年、分散人工知能の分野においてマルチエージェント問題<sup>[1]</sup>の研究が盛んに行われている。このような研究におけるエージェントの自律的な学習方法として、進化的学習<sup>[2][3]</sup>が注目されている。この進化的学習とは、適応度を用いて最適解を探索するため、環境のマルコフ性に関係なく最適解を獲得することができると考えられる。しかし、このような研究のほとんどが木構造を用いたものであり、環境が複雑になると木構造が爆発し、探索が困難となる。そこで本稿では、配列を扱う遺伝的アルゴリズム(以下 GA)がマルチエージェント問題に適用可能であるかを検討する。

### 2. 遺伝的アルゴリズムの適用

GA をマルチエージェント問題に適用する場合、各状態に対して行動を割り当てることにより最適政策を探索する。しかし、1つの状態に対して1つの行動を割り当てた場合では、同じ状態でも最適行動が異なることがあるため、環境の変化に追従できないという問題が生じる。そこで、Fig.1に示すように各状態に対して2つの行動を割り当てることにより、その問題点を解決する。

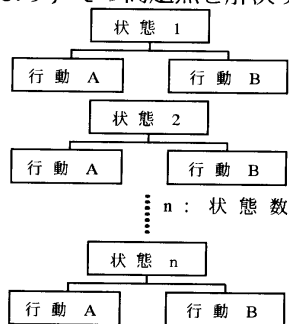


Fig.1 Gene of the agent.

A Study on Multi-Agent Problem with Genetic Algorithm.  
Atsushi Nishiyama, Yoshinobu Kajikawa,  
Yasuo Nomura  
Department of Electronics, Faculty of Engineering,  
Kansai University  
3-3-35, Yamate-cho, Suita-shi, Osaka 564-8680, Japan

### 3. 適応度の継承<sup>[4]</sup>

GA では、現世代の適応度だけを用いて個体の淘汰及び遺伝子操作を行うため、常に環境が変動するマルチエージェント問題では、前の世代で優良であった個体が次の世代では低く評価され淘汰されてしまう可能性がある。そこで本稿では、前の世代の適応度を継承させることにより適応度の変動を抑え、収束値の向上を試みる。式(1)に適応度の継承による更新式を示す。

$$fit_{new} = \alpha \times fit_{old} + (1 - \alpha) \times fitness \quad \dots\dots(1)$$

ここで、 $fit_{old}$  は1世代前の適応度、 $fitness$  は現世代の適応度(行動回数上限値-行動回数)とし、 $fit_{old}$  を一定の割合  $\alpha$  ( $0 \leq \alpha \leq 1$ ) で受け継がせた  $fit_{new}$  を現世代で実際に用いる適応度とする。

また、交叉によって作成された新しい個体の適応度は、受け継いだ遺伝子の割合に応じて式(2)で計算する。

$$fit_{old} = gl_a \times fit_{a,old} + gl_b \times fit_{b,old} \quad \dots\dots(2)$$

但し、 $fit_{old}$  を交叉後の適応度、 $gl_a$  を個体  $a$  の遺伝子を受け継ぐ割合、 $gl_b$  を個体  $b$  の遺伝子を受け継ぐ割合、 $fit_{a,old}$  を交叉前の個体  $a$  の適応度、 $fit_{b,old}$  を交叉前の個体  $b$  の適応度とする。

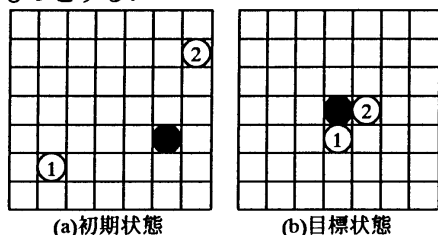
本稿では、式(1)と式(2)で更新された適応度を用いて個体を評価する。

### 4. 追跡問題

追跡問題とは、Fig.2に示すような  $n \times n$  の格子状トラス環境上で複数の追跡者エージェントが1体の獲物エージェントを捕獲することを目的とした協調問題解決型マルチエージェント問題のベンチマークである。

Fig.2の環境において、(a)のように各エージェントをランダムに配置した状態を初期状態とし、(b)のように2体の追跡者エージェントが獲物エー

エージェントに隣接すれば獲物を捕獲したとし、追跡者エージェントに報酬を与える。また、全てのエージェントには Fig.3 で示す 5×5 の視覚範囲を設定し、獲物は視覚範囲内の全ての追跡者エージェントから遠ざかる方向へ移動する逃避的な行動を取るものとする。



①: エージェント 1 ②: エージェント 2  
●: 獲物

Fig.2 The hunting problem.

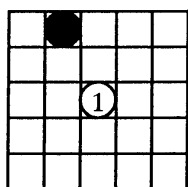


Fig.3 Observable range.

#### 4. シミュレーション結果

Table 1 の条件でシミュレーションを行った時の最優良個体の行動回数の推移を Fig.4 に示す。但し、①は 1つの状態に対して割り当てる行動数が 1つ、②は 1つの状態に対して割り当てる行動数が 2つとする。また、縦軸は対数軸とする。

Table 1 Condition of Simulation.

個体数	100
世代交代周期	50 試行で 1 世代
適応度の継承	$\alpha = 0.10$
突然変異確率	10%
行動数の上限値	10000
環境サイズ	7×7

Fig.4 より、①、②どちらも獲物捕獲パターンを獲得していることがわかる。しかし、Fig.4 は最優良個体の行動回数の推移であるため、次の世代で

もその個体が環境に適応しているとは限らない。そこで、学習結果が環境変化に追従しているかを確認するため、学習後の最優良個体を用いて獲物の捕獲を行った。その結果を Fig.5 に示す。

Fig.5 より、①は 1つの状態に対して 1つの行動しかできないため、環境の変化に追従できず、無限ループに陥っている。しかし②は、1つの状態において行動を 2つ割り当てているため、無限ループに陥らず環境の変化に追従できていることがわかる。以上のことより、本手法はマルチエージェント問題に適用可能であるといえる。

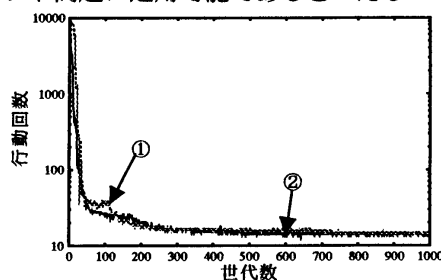


Fig.4 Convergence properties.

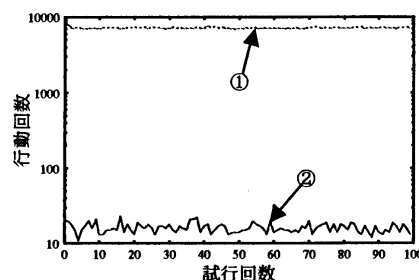


Fig.5 The results after learning.

#### 5. まとめと今後の方針

本稿では、GA がマルチエージェント問題に適用可能であるかの検討を行った。その結果、1つの状態に対して 2つの行動を割り当てることにより、無限ループに陥らず環境の変化に追従できることがわかった。今後は、追跡問題以外の問題に本手法が適用可能であるかを検討する。

#### 【参考文献】

- [1] 沼岡他：分散協調メディアシリーズ 11, 共立出版, 東京, 1998.
- [2] 佐藤他：信学技報, AI97-64, Jan.1998.
- [3] 鈴木他：情処学研報, 97-ICS-108, May 1997.
- [4] 村田他：人工知能学会誌, Vol.14, No.3, 1999.