

準静的変化過程における多階層マルチエージェントモデルの構築

星野 達哉[†] 服部 恭史[†] 高丸 尚教[†]

[†] 中部大学大学院工学研究科情報工学専攻

1 はじめに

エージェントが与えられた環境に適応する行動規則を獲得するためには、設計者自身に問題と手法の両方の知識が要求される。しかし、すべての問題に対して適切な対応を指示することは困難である。そこで近年進化論的計算手法などの学習メカニズムを用いて環境に適応し、より高次の行動規則を獲得可能なシステムの研究が行われている。

我々は、平澤らが提案した遺伝的ネットワークプログラミング [1] (以下 GNP) に中性ノードの概念を導入し、従来法に比べ適応度が 37% 向上させることに成功した [2]。

本研究では、昨年我々が提案した中性ノード概念を持つマルチエージェントに対して、小空間での初期学習のみでより大きな探索空間に適応可能なモデルを提案する。具体的には、コントロール集合を用いた自己相似学習空間を与え経路探索問題に対してロバストな特性を持つようにする。

2 自己相似学習空間について

本研究では自己相似学習空間としてフラクタルの 1 種であるコントロール集合を採用する。これをエージェントの学習環境に適応する際にはコントロール集合を 2 次元とし、狭い学習空間内においてもさまざまなパターンについての学習を行うこと出来るようにしている。

3 モデルの概要

本研究で用いるモデルでは、通常の GNP で用意されている処理ノード、判定ノード、初期ノードの 3 つのノードに加えて中性ノードと言う概念を導入している。中性ノードを除くノードには割り当てられた関数を実行するための実行時間が決められているが、中性ノードは実行時間を要せずに次のノードへと遷移させ

る。また、中性ノードには関数が割り当てられておらず何も実行はしない。なお、中性ノードには通常の関数に変化する機能が用意されているが、本論文では使用しないので説明は割愛する。

3.1 学習と進化について

本モデルでは、強化学習の 1 つである Profit Sharing を用いて各ノード間のネットワークの学習を行う。本モデルでは、各エージェントが保持している関数種の数を遺伝子として、これを交叉と突然変異を用いて進化を行う。

4 評価実験

まず、1. 初期学習環境として 2 次元コントロール集合を作成し、全エージェントは同一のスタートからゴールを目指すことを考える。ゴール地点は 2 つ用意されており、同一の環境にてスタート地点は固定のまま、ゴールをずらしてから学習させる。学習環境サイズの影響を調べるために、一辺 9,27,81,243 の 4 種類作成し、行動ステップをそれぞれ 50,200,1000,3000 としてシミュレーションする。

適応度が 0.7 以上を達成、または 1000 世代実行するまで行うものとする。適応度はエージェントが行動するごとに式 1 で算出している。

$$\text{適応度} = 1.0 - \left\{ \frac{A^*(P_{\text{present}}, P_{\text{goal}})}{A^*(P_{\text{start}}, P_{\text{goal}})} \right\} \quad (1)$$

ここで、 P_{start} をマップ上のスタート場所、 P_{goal} をマップ上のゴール場所、 P_{present} をエージェントの現在地、 $A^*(s, g)$: s から g までの最短経路を算出する関数とおく

2. コントロール集合における初期学習終了後、図 1 のような L 字クランク問題で評価を行う。正方マップ上に長さ 150、幅 30 の L 字クランクを作成し、スタートとゴールを準備する。L 字クランクには 5% から 20% まで 5% 刻みの割合で障害物を作成する。エージェントの総数は 100 体として、全エージェントは同一のスタートからゴールを目指す。ただし、エージェントの方向は各世代ごとにランダムに設定される。1 世代あたり

Hierarchical multi-agent model in quasi-static learning
Tatsuya HOSHINO[†], Yasushi HATTORI[†] and Hisanori TAKAMARU[†]

[†] Graduate School of Engineering, Chubu University
Kasugai-shi, 487-8501, Japan
tatsuya@ss.cs.chubu.ac.jp

の行動ステップの上限を 2000 とし、適応度が 0.7 以上を達成・または 1000 世代実行するまで行う。適応度はエージェントが行動するごとに式 1 で算出している。

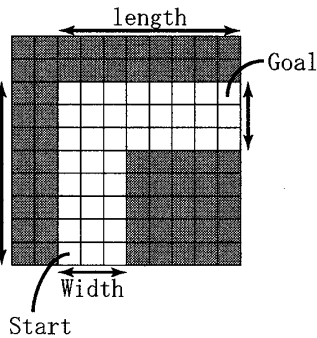


図 1: L 字クランク問題

表 1: 関数の種類と配分

関数名	機能	初期配分
Neutral	次のノードに遷移	60
Ahead	前進する	5
Back	後退する	5
TurnRight	右に向きを変える	5
TurnLeft	左に向きを変える	5
CheckBarricade	前方が障害物か	10
CheckUntrampled	前方が未知か	0

5 結果

表 2 に、適応度 0.7 を超えるまでに必要となる世代数について、学習環境と障害物との関係をまとめる。なお、×は適応度 0.7 未満になったことを示す。

この表から初期に学習を行う環境の大きさにかかわらず、未学習の時に比べて早く環境に適応することが出来ることを示している。このことから、初期学習を行うことの有用性が示された。

図 2 では 243 × 243 で学習を行ったエージェントの関数種の変化をグラフとし、その 1 部を抜粋したものである。この図によると、処理ノードである「Back」と判断ノードである「CheckBarricade」「CheckUntrampled」は世代が進むことで淘汰されている。これは経路探索問題においては判断ノードは不必要、または適切な判断を行えていないことを示唆するものである。

表 2: 実行結果

障害物率	学習環境				
	未学習	9*9	27*27	81*81	243*243
5%	29	1	1	2	1
10%	90	18	29	48	1
15%	306	151	261	271	151
20%	×	320	×	×	×

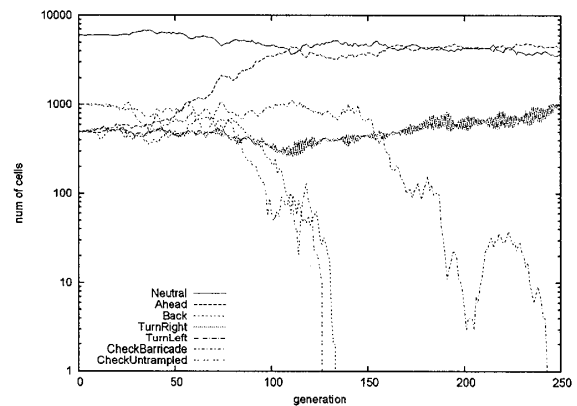


図 2: 関数種の変化

6 まとめ

本研究では、提案したモデルを用いてフラクタル環境で初期学習を行い評価を行った。初期学習は中性ノードを用いてノード種の適切な配分に寄与し、準静的環境への適応を高めるなどその有用性を確かめられた。しかし、初期学習空間と実際の探索空間の大きさなどの関係性については未知であるため、今後研究する必要がある。

参考文献

- [1] 平澤 宏太郎, 大久保 雅文, 片桐 広伸, 胡 敬炉, 村田 純一: "蟻の行動進化における Genetic Network Programming と Genetic Programming の性能比較", 電気学会論文誌 C, Vol121, No6 (20010601) pp.1001-1009
- [2] 服部 恭史, 高丸 尚教, 星野 達哉: "中性ノードを導入した遺伝的ネットワークプログラミングの拡張", 情報処理学会第 71 回全国大会 (2010)