

## 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

服部恭史 高丸尚教 星野達哉

Yasushi HATTORI, Hisanori TAKAMARU, Tatsuya HOSHINO

## 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が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することは困難である。そこで進化論的計算手法を用いて環境に適応した、より高次の行動規則を獲得可能なシステムの構築が研究されている。その中でも平澤らが提案した遺伝的ネットワークプログラミング<sup>(1)</sup>(以下 GNP) は、エージェントの行動規則を高次の抽象ネットワークを用いて有向グラフで表現し、適応度に応じて進化させることで従来のものより性能が向上したことが確かめられている。さらに、タスク実行中に得られる有益な情報もエージェントの行動に反映させるために、強化学習を組み込んだ強化学習型も提案されている。しかし、GNP には、性能面だけにおいても(1)初期に与える有向グラフの大きさなどの初期パラメータと(2)適応度を求める関数や、後に述べる強化学習型 GNP に用いる報酬関数によって性能が左右される問題点が残されている。

この問題に対して平澤らは、ノードの総数が動的に変化するノード数可変型 GNP<sup>(2)</sup>を提案している。しかし、適応可能な個体に対してもノード数が変化し、抽象ネットワークを元の状態に復元することが不可能などの安定性に問題があると我々は考える。そこで、本研究では、新たに中性のノードの概念を導入して有向グラフの構造を適応度の変化率に応じて改善するモデルを提案する。本論文では、まず平澤らが提案しているいくつかの GNP のモデルについて説明した後、これらを拡張した提案モデルの説明と結果を示す。

## 2. 平澤らのモデル

本節では、GNP について構成・概要と遺伝子型による有向グラフの記憶、進化をさせる遺伝子オペレータの3つに分けて説明をする。

GNP は、行動規則を表現する抽象ネットワークとしてエージェントの行動が記される複数のノードとノード間を繋ぐネットワークを用いて構成される。

次に、有向グラフを遺伝子構造として記憶する方法について説明する。GNP では、ノードが持つ情報とノードから遷移先の情報を一次元配列に記憶し、これをノードの数だけ用いる。ノードが持つ情報を、Node Gene として (n1) ノードの種類 (n2) 関数種の番号 (n3) 実行時間の3種類である。遷移情報は、Connection Gene として (c1) 遷移先の ID 番号 (c2) 遷移時間の2種類を1つにして、分岐の数だけ記憶される。

最後に遺伝子オペレータによる進化方法について説明する。

GNP では、ノードの遷移先を変更することで進化を図る。つまり、Connection Gene の1番目である遷移先の番号を変更することで行動規則を変化させる。遺伝子オペレータには、(A) 適応度が高いエリートの複製、(B) 2個体からランダムに選択された遺伝子を入れ替える交叉、(C) 1個体がランダムに選択したノードの遷移先をランダムに変更する突然変異の3種類が行われる。

上記の他にタスク実行中にも有益な情報を学習に反映させられる強化学習型モデルが提案されている。<sup>(3)</sup> このモデルでは、ノードの遷移時に全てのノードから最も重みのあるノードを選択し、適切な報酬を与えることで学習させる。これにより、進化による大域的な探索と学習による局所的な探索をそれぞれが担うことができ、通常の GNP と比べて適応速度が促進されることが発表されている。

## 3. 進化論的計算手法が抱える問題点

この系のモデルには、大きく分類して2つの問題点を抱えている。

1. GNP は、初期に与えられる関数種の数と、関数種の配分、遺伝操作の割合などの初期パラメータによって性能に差が生じる。
2. 強化学習用の報酬関数と GNP 用の適応度関数によって性能に差が生じる。

1つ目の問題への対策として、平澤らがノードの総数が変化するモデルとしてノード数可変型 GNP を提案している。しかし、このモデルでは、ノードの活性度、全ノード数に対する関数種の割合などのパラメータのもと、追加・削除オペレータを実行している。そのため、削除されたノードが保持していた遷移情報も削除され、さらに、対象のノードにアクセス可能なノードの遷移情報もランダムに変更させ、新しく追加されたノードの遷移情報も、ランダムに決定されるため、追加されたノードが有効となるためには、遺伝オペレータの突然変異に頼るしかなく、確率的である。これでは、ノード数の増減による影響が遅れて現れるため、そのノードが適切な処置であるのかは不確定となる。さらに、遷移情報が保持されないため、削除されたノードを再び挿入した場合、以前の状態に復元することができない。

2つ目の問題が意味するものは、複数タスクを満たさなければならない問題や、動的に目的が変化するようなより現実的な、シミュレーションなどに対応させ難い。そこで報酬ライブラリを各個体の適応度に応じて動的に構築する手法を検討しているが、ここでの議論は行わないものとする。

#### 4. 提案モデル

本モデルでは、通常の処理ノード、判定ノード、初期ノードの3つのノードに対して、さらに、いずれの機能も果たさない中性ノードの概念を導入する。ノードが必要に応じて予め定義された関数種あるいは中性ノードに状態を変化させる。これにより全体としての構造の再構築を図る。通常ノードには、その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遷移する。ノードには、それぞれの潜在的活性度を表す potential 値を新たに加え、この potential 値がある上限値以上に高まると他の中性ノードを自身と同じ働きをするノードに変化させる。逆に下限値以下になると中性ノードに変化する。potential 値の更新は、1世代毎に行われ、各個体の適応度の変化の符号に応じて使用したノードのみが更新される。適応度が負の方向に連続して進むとノードの活性度が下げられ、中性ノードが増加し構造の再構築が始まる。逆に適応度が正の方向に進むと、ノードが中性ノードに変化することがないため、結果的に構造が維持される。このモデルでは、すぐに別の関数種に変化するのではなく、一旦中性ノードに変化するため、適応度によって元の状態に戻る可能性がある。

なお、本研究では、構造の安定と学習の促進を確かめるために、あえて進化フェーズを行わずに、複数のエージェントが関数種の配分だけを共有し、単エージェントがタスク実行中に遷移情報のみを強化学習させ実験を行った。

#### 5. 実験

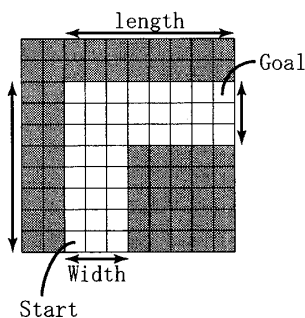


図1 L字クランク問題

単純なモデルで検証するために図??のようなL字クランク問題を考える。正方マップ上に長さ20、幅10のL字クランクを作成し、スタートとゴールを準備する。エージェントの総数を100体として全て同一のスタートからゴールに向かう。ただし、エージェントの向きは、世代毎にランダムに変更される。1世代あたりの行動ステップの上限を3000として20世代実行する。各世代毎に求める適応度は次式で求める。

$$fitness = \sum_{i=0}^n \text{Min} \left\{ \frac{A^*(P_g, P_{it})}{A^*(P_g, P_0)} \right\} \quad (1)$$

なおnは、エージェントの総数tは一世代中のあるステップ、A\*(始点, 終点) A\* は最短パスを示す

表.1 エージェントの関数種と配分

	関数種	初期配分
Ahead	前進	5
Back	後退	5
TurnLeft	左に向きを変える	5
TurnRight	右に向きを変える	5
CheckBarricade	前方の障害物を調べる	10
CheckUntrampled	前方が未知か調べる	10
Neutral	中性ノード	50

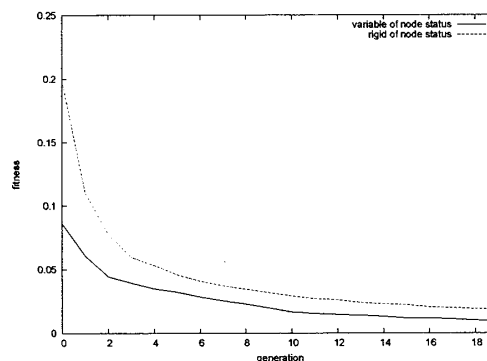


図2 L字クランク問題に対してノード状態を可変にした場合と固定にした場合の適応度の比較

#### 6. 結果

図2は、L字クランク問題に対して提案モデルであるノード状態が変化させた場合とノード状態が変化しない場合の適応度に対する20ステップ間の推移を表している。この図から初期に関数種の配分が適応度によって変化させたことで、固定である場合と比べて適応速度が早いことが確かめられる。このことから、初期に関数種の配分が決定することで学習に影響のあるノードが制御可能であることを示唆している。

#### 7. まとめ

本論文では、いずれの役割も果たさない中性ノードの概念を導入し、適応度に応じてノード状態を変化させ関数種の配分を変化させるモデルを提案した。関数種の配分が固定の状態と比較して学習が促進されることが確かめられた。しかし、初期に与えるノードの総数については可変ではないため、今後検討する必要がある。

#### 文献

- (1) 平澤宏太郎, 大久保雅文, 片桐広伸, 胡敬煥, 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, Vol.121, No.6 (20010601) pp. 1001-1009
- (2) 片桐広伸, 平澤宏太郎, 胡敬煥, 村田純一:「ノード数可変型 Genetic Network Programming」, 電気学会論文誌 C, Vol.123, No.1 (20030101) pp. 57-66
- (3) 間普真吾 1 平澤 宏太郎 2 古月 敬之 3:「強化学習を用いた遺伝的ネットワークプログラミングとそのエージェントの行動生成における性能評価」, 情報処理学会論文誌, Vol.46, No.12(20051215) pp. 3207-3217