

進化型ニューラルネットワークによるモジュラーロボットの行動制御

平野 祐一†

長尾 智晴‡

† 横浜国立大学工学部

‡ 横浜国立大学大学院環境情報研究院

1 はじめに

複数の準同一なロボットモジュールから構成されるモジュラーロボットが構造の組み替えによる柔軟性や生産性の面などから期待され、近年盛んに研究されている [1][2]。しかし、人手でその制御プログラムを構築することは困難であり、それを自動的に行うシステムが求められている。本報告では、Neural Network(NN)と遺伝的アルゴリズムを用いて学習を行う進化型 NN によって、モジュラーロボットの行動制御プログラムの自動獲得を行う手法を提案する。

2 提案手法

今回の報告で行う実験は2次元のシミュレーション環境で行われる物体の運搬問題である。使用するモジュラーロボットの基本構成モジュールは円形で移動することができ、他のモジュール1つとの結合と分離ができる。このモジュラーロボットの制御プログラムにおいては柔軟な構造の変化や構造ごとの行動が求められるが、単一の NN では実現が難しい。そこで、提案手法ではセンサ入力や行動出力などのモジュールの機能を限定した複数の環境下で個々に学習を行い、それらの学習結果である NN を保存する。その後、目標とする環境下でそれらの NN のどれを用いるかを判断する NN を学習する。

2.1 使用するニューラルネットワーク

今回使用する NN である Real-valued Flexibly Connected Neural Network (RFCN) [3] は従来研究において連続値空間におけるエージェント制御問題に適用され、その有効性が示されている。RFCN は入出力に実数値を用いることができることに加え、各ニューロンも単純なしきい値関数だけでなく任意の関数を設定することが可能である。

2.2 学習方法

提案手法における学習は2段階に分けられる。第1段階では、モジュールの機能を限定した状態で、モジュール数や運搬物の異なる複数の環境下での学習を従来の RFCN を用いて行う。機能の限定の具体例としては、モ

ジュラーロボットが一定の長さで結合しているときの運搬や、分離して結合することができない状態にあるときの運搬などが挙げられる。これらの学習結果として得られた RFCN を以降では機能 RFCN と呼ぶ。

第2段階では、目標とする環境において、第1段階で得た機能 RFCN 群のうちどれを用いるかを動的に決める行動決定用 RFCN を学習する。行動決定用 RFCN の出力は対応する機能 RFCN の使用優先度であり、これが最も高いものが使用され、その機能 RFCN に基づいた行動を行う。

2.3 行動決定用 RFCN について

機能 RFCN はある行動規則を記述したものであり、行動決定用 RFCN の優先度が頻繁に切り替わってしまうのは好ましくない。そこで、行動決定用 RFCN では、出力を蓄積関数に固定することで出力値の急激な変化を抑える。この蓄積関数を式 (1) に示す。さらに、目標環境における望ましい初期状態を補正用メモリとして用意し、同時に学習を行った。補正用メモリを決定する式を式 (2) に記述する。式 (1)、式 (2) において I は入力、 O は出力、 Mem はメモリの値を示す。 G はニューロンが持つ補正値を、 N はノードに対する入力数を、 T は最大ステップ数を表し、 n と s はそれぞれ出力番号とステップ数を表す。また、行動決定用 RFCN の例を図 1 に示す。

$$O_{n,s} = \sum_{k=1}^N I_k \times G + O_{n,s-1} + Mem_n \quad (1)$$

$$Mem_n = \frac{\sum_{s=0}^T O_{n,s}/(s+1)}{T} \quad (2)$$

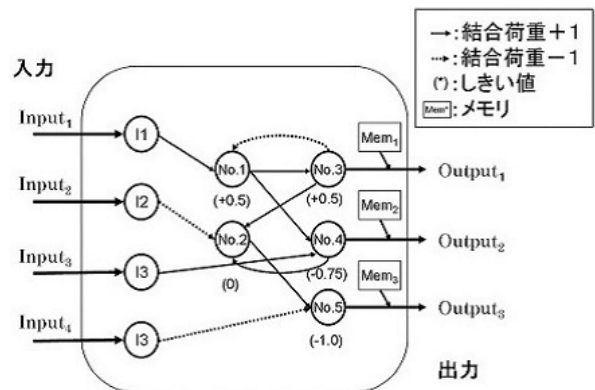


図 1: 行動決定用 RFCN の例

Motion Control for Modular Robots Using Evolutionary Neural Network

Yuichi Hirano† Tomoharu Nagao‡

†School of Engineering, Yokohama National University

‡Graduate School of Environment and Information Sciences, Yokohama National University

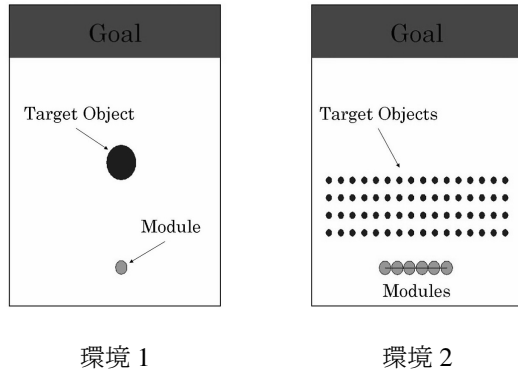


図 2: 第 1 段階学習環境

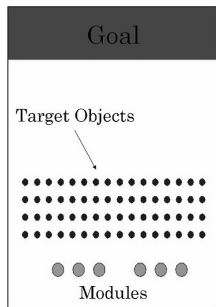


図 3: 第 2 段階学習環境

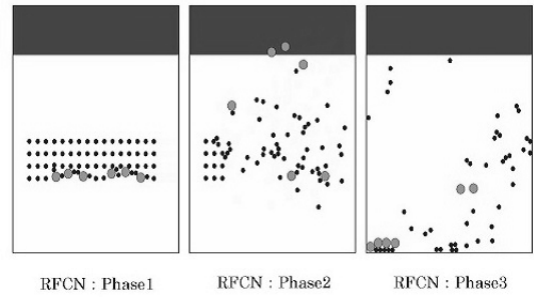
3 実験設定

複数のオブジェクトを目的地までロボットが運搬する問題に対して提案手法を用いて学習を行う。第 1 段階の実験環境は図 2 に示すような環境 1 と環境 2 の 2 つを用意し、最終目標となる第 2 段階には図 3 に示す環境を用意した。使用するモジュールは連結するためのアームを 1 つ、モジュールの連結を許すコネクタを 1 つと、全方位に向けての距離センサを等間隔に 12 個、対象オブジェクト毎に種類を分けると計 36 個もつ。モジュール同士は連結と切り離しが可能であり、単体であっても 360 度自由な方向への移動が可能である。第 1 段階では一部の入出力が制限されているため、結合状態では分離が、分離状態では結合ができない。第 2 段階では切り替えが行われる際に使用する機能 RFCN に応じて分離または結合が行われる。

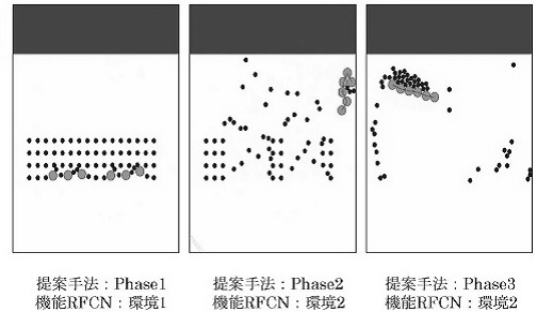
なお、実験における評価関数は次の式 (3) のとおりである。オブジェクトの運搬達成数を n 、運搬しきれなかった l 番目のオブジェクトの縦方向の運搬距離を y_l 、運搬を達成したオブジェクト 1 つに対する報酬を R 、環境中の全オブジェクト数を MO とする。

$$Fitness = \sum_{k=0}^n R + \sum_{l=0}^{MO-n} y_l \quad (3)$$

なお、本実験は 2 次元の物理シミュレーション環境下において行われている。物理演算エンジンは Box2D[4] を用いた。



平均評価値：15056.6



平均評価値：22685.5

図 4: 獲得した行動

4 実験結果

第 2 段階における学習結果の最良個体と単一 RFCN で学習した場合の最良個体の行動例と 10 試行の平均評価値を図 4 に示す。図 4 から、行動決定用 RFCN が学習環境において適切な機能 RFCN を選択できていることが確認できた。単独の RFCN を用いた場合よりも学習を 2 段階に分けて行ったほうが評価値や行動からモジュールの特性を活かしていることが確認できた。

5 まとめ

モジュラーロボットの行動制御に関して、進化型 NN である RFCN を用いて 2 次元の物理シミュレーション環境で運搬問題に対する学習を行った。RFCN を機能と行動決定用に分けて学習した結果、単一の RFCN で学習するよりも効果的であることが確認できた。今後はより複雑な環境下において行動決定用 RFCN がどこまで適切に機能 RFCN の切り替えを行うことができるのか検証していく予定である。

参考文献

- [1] Satoshi Murata, Eiichi Yoshida, Akiya Kamimura, Haruhisa Kurokawa, Kohji Tomita and Shigeru Kokaji: "M-TRAN: Self-Reconfigurable Modular Robotic System," IEEE/ASME Transactions on Mechatronics, Vol. 7, No. 4, pp. 431-441, 2002.
- [2] M. Yim, Y. Zhang and D. Guff: "Modular robots," IEEE Spectrum, vol.39, no.2, pp.30-34, 2002
- [3] 白川真一, 長尾智晴: "RFCN による連続値空間上での自律エージェントの行動制御", IEEJ Trans.EIS, Vol.127, No.5, pp.762-769, 2007.
- [4] <http://box2d.org/>