

進化計算を用いた複数ロボットによる行動学習に関する取り組み

曾我 紗知子[†] 小林 一郎[†][†]お茶の水女子大学大学院人間文化創成科学研究科理学専攻情報科学コース

1 はじめに

近年、ロボットの行動学習に進化計算を用いる手法が研究されている進化計算の1つに遺伝的アルゴリズム (Genetic Algorithm : GA) があるが、この手法での学習は、学習結果の正解が適応度関数によって一意に表現することができる場合に、獲得したい解を得ることが可能である。しかし、実世界での行動と数式で表現される適応度関数を一対で表すことができない場合、つまり、学習結果の正解が複数存在する場合に、自律的な行動の学習を行うことが可能であるかは不明である。本研究では、複数ロボットの自律的な編隊行動の学習について、ロボット3台で三角形の編隊を組むことを課題とし、シミュレータ Simbad[1] を利用して、ロボットコントローラの学習を行い、考察する。

2 単体ロボットの学習

2.1 コントローラの構成

ロボットのコントローラは、図1に示すニューラルネットワークで構成され、GAによって進化させる。コントローラは、フィードフォワード型の入力層4、中間層4、出力層2のニューラルネットワークで構成する。ロボットが取得するセンサ値を入力とし、速度と回転速度を出力とする。GAを利用する際に必要となる遺伝子には、入力層、中間層、出力層全てのノード間の24個の結合荷重の値を遺伝子座に入れる。

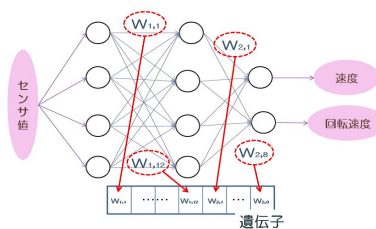


図1: ロボットコントローラ

2.2 学習設定

結合荷重を $[-1.0, 1.0]$ の実数値とし、実際にロボットを動作させ、設定した適応度関数に従って適応度を計算する。ルーレット選択とエリート保存選択を採用し、次世代の個体を選ぶ。交叉率 0.6 で、一点交叉を

An Approach to Learning Multiple Robots' Behaviors using Evolutionary Computation

[†] Sachiko SOGA(soga.sachiko@is.ocha.ac.jp)

[†] Ichiro KOBAYASHI(koba@is.ocha.ac.jp)

Advanced Sciences, Graduate School of Humanities and Sciences, Ochanomizu University (†)

2-1-1 Ohtsuka Bunkyo-ku, Tokyo 112-8610, Japan

を行い、突然変異率 0.1 で、遺伝子座の値を $[-1.0, 1.0]$ のランダムな値に変更する。これらを設定した世代数分繰り返す。

2.3 静止物体からの距離に基づく学習

静止物体を2つ置いた環境(図4)の下で、ロボットがどちらの静止物体からも等しい距離を保つ学習を行う。この時、ロボットがどちらの静止物体とも、静止物体同士の距離(今回は2m)と同じ距離を保つことが良いこととする。

2.3.1 実験設定

ロボットの初期位置は図4の通りである。

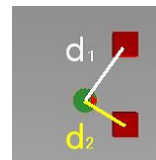


図2: 訓練環境

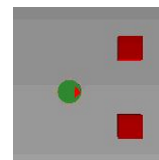


図3: 学習結果

本実験における適応度関数は以下の様に設定する。

$$f(d_1, d_2) = \sum_{t=0}^{2500} \begin{cases} -(\frac{d_1}{d_2} - 1)^2 + 1 & (1.9 < d_1, d_2 < 2.1) \\ 0 & (otherwise) \end{cases} \quad (1)$$

$f(d_1, d_2)$: 適応度関数

d_1 : ロボットから1つ目の静止物体までの距離

d_2 : ロボットから2つ目の静止物体までの距離

式(1)は、ロボットが2つの静止物体と2m前後の距離になった時に加算され、2つの静止物体との距離が等しくなる時に一番高い適応度が加算されるように設定されている。また、壁や静止物体に衝突した際には適応度を0にし、ステップ数が2500になるまで計算を続ける。静止物体2つと等距離を保つことと、衝突回避を学習するために、このように設定した。

2.3.2 実験結果

500世代まで学習を継続させたところ、一番適応度が高いコントローラは157世代目で得ることができた。このコントローラを利用したロボットは、静止物体2つから離れる動作をするように学習ができていた。静止物体をセンシングできる範囲でロボットの初期位置を変えた時にも、一定の距離を保つことに成功した。

また、図4の位置で静止物体をロボットに変更し、ロボット3台全てに獲得したコントローラを使用して

動作させると、それぞれが等距離を保とうと動き、三角形の編隊を実現させることができた(図3)。さらに、初期位置を横一列などに変更した環境で動作させても、三角形の形になった。

2.4 静止物体からの比に基づく学習

次に、環境の大きさに応じて、編隊の大きさを自由に変更できるように、ロボットと2物体との距離を一定距離ではなく、ロボットからそれぞれの物体までの等距離となるように学習させることにする。

図4と同じ環境において、それぞれの物体までの距離が1:1になった場合に、適応度が加算されるように変更する(式(2))。

$$f(r_1, r_2) = \sum_{t=0}^{2500} \begin{cases} -(\frac{r_1}{r_2} - 1)^2 + 1 & (0.9 < \frac{r_1}{r_2} < 1.1) \\ 0 & (otherwise) \end{cases} \quad (2)$$

$f(r_1, r_2)$: 適応度関数

r_1 : ロボットから1つ目の静止物体までの距離

r_2 : ロボットから2つ目の静止物体までの距離

結果として、想定された行動の学習が行われず、ロボットは、動き始めてから終始2物体から離れ続け、最後には壁に衝突した(図5)。どの長さの距離でも、2物体からそれぞれ等距離で離れていれば適応度が高くなるため、ロボットは静止物体から離れ続ける動作を学習してしまったと考えられる。

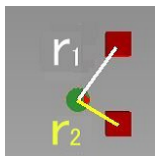


図4: 比での訓練環境

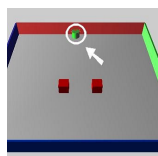


図5: 比での学習結果

3 複数ロボットの学習

2.4節の実験では学習ができなかったことを受け、静止物体をロボットに変更し、それぞれのロボットがお互いに距離を保つ学習を行う。

3.1 コントローラの構成

今回の学習では、入力層8, 中間層8, 出力層2のニューラルネットワークでコントローラを構成する。入力はセンサ値4つと、他の2台のロボットとのx座標の差, y座標の差で、出力は移動速度と回転速度である。学習で計算するのは、それぞれのノード間の結合荷重80個である。

3.2 学習設定

ロボットを複数台にし、それぞれが学習を行うため、学習には並列分散遺伝的アルゴリズム[2]を使用する。

今回のロボット数は3台とし、それぞれのロボットが20個体ずつ、全部で60個体の遺伝子で学習を行う。選択, 交叉, 突然変異は2.2節と同じ設定とし、選択後に他のロボットに移住する。最も適応度の高かった1個体を毎世代移住させる。移住のペアはランダムに決定することとした。適応度関数は式(2)を利用する。式(2)では静止物体からの距離としていた部分を、今回は他の2台のロボットとのそれぞれの距離とする。

3.3 実験結果

185世代まで学習を行うと、三角形の編隊を形作るようになった(図6)。

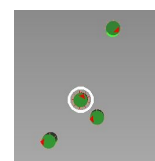
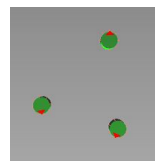


図6: 複数台の学習結果 図7: 外乱を与えた場合

次に、比で学習しているため、編隊の形が一度崩れた後にも、また編隊を違う大きさで組むことができるかどうかの確認をするため、もう1台のランダムに動き回るロボットを用意して、編隊を組もうとするロボットに外乱を与えた。先ほどとは異なる大きさの三角形を作るような行動を期待していたが、ロボット3台は再び三角形を形作ることはできなかった。

4 おわりに

本研究では、ロボットのコントローラを進化計算により、自律的に三角形の編隊を組むことが可能であるかを検討した。ロボット1台の学習に対し、得られたコントローラを複数台のロボットに使用し、三角形を形成させた。しかし、静止物体との学習では、相手も移動物体だった場合、お互いに少しずつ動いて距離を保つ、という行動がうまく学習できない。そこで、複数台のロボットで学習を行うことにしたが、学習時のコントローラはそれぞれ違う物が使われており、お互いにそれらの動きに合わせて学習を行っていたため、全てのロボットに獲得したコントローラを使用して動作させると、期待した学習結果の動作は得ることができなかった。

今後は、学習時のロボット3台のコントローラを同じ物にし、全体の行動の評価を行うことにより、編隊行動の実現を目指す。

参考文献

- [1] <http://simbad.sourceforge.net/>
- [2] Reiko Tanese, Distributed Genetic Algorithms, Proc. 3rd International Conf. Genetic Algorithms, pp434 ~ 439, 1989.