

進化計算を用いたロボットの行動学習に関する一考察

曾我紗知子[†] 小林一郎[‡]

[†]お茶の水女子大学理学部情報科学科

[‡]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

実世界で作業を行うロボットは、新しい環境に直面した時でも自律的に合理的な行動をすることが求められる。近年、そのようなロボットの行動を制御するコントローラを学習するのに、進化計算を利用する進化ロボティクスが盛んに研究されている。進化ロボティクスとは、遺伝的アルゴリズム、遺伝的プログラミングなどの進化計算や強化学習、ニューラルネットワークなどを用いて環境に対して柔軟な行動を選択できる、ロボットコントローラを設計する技術である。

本研究では、ニューラルネットワークと遺伝的アルゴリズムを用いシミュレータ上の様々な環境に応じて適切な行動を行うようになる、ロボットの行動規則の学習について考察を行う。

2 ロボットコントローラの構成

本研究では、図1に示すニューラルネットワークで構成されているコントローラを遺伝的アルゴリズムによって進化させる。

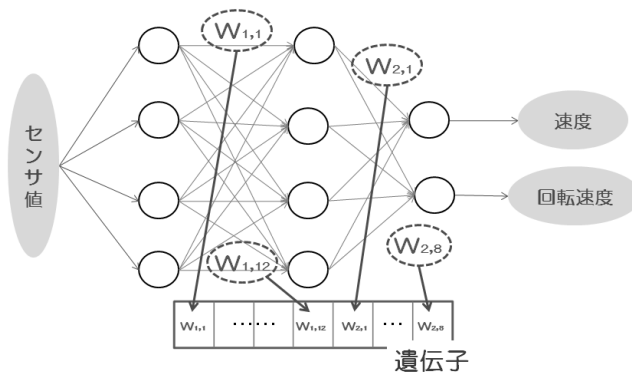


図1: ロボットコントローラ

コントローラは、単一方向にのみ信号が送られる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

A Study on Robot's Behaviour Learning based on Evolutionary Computation

[†]Sachiko SOGA(g0720527@is.ocha.ac.jp),

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

[†]Dept. of Information Sciences, Faculty of Science, Ochanomizu University, 2-1-1 Ohtsuka Bunkyo-ku Tokyo 112-8610

[‡]Advanced Sciences, Graduate School of Humanities and Sciences, Ochanomizu University, 2-1-1 Ohtsuka Bunkyo-ku Tokyo 112-8610

ンサ値を入力とし、速度と回転速度を出力とする。遺伝的アルゴリズム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。

3 ロボットコントローラの学習

図2にロボットコントローラの学習の概要を示す。

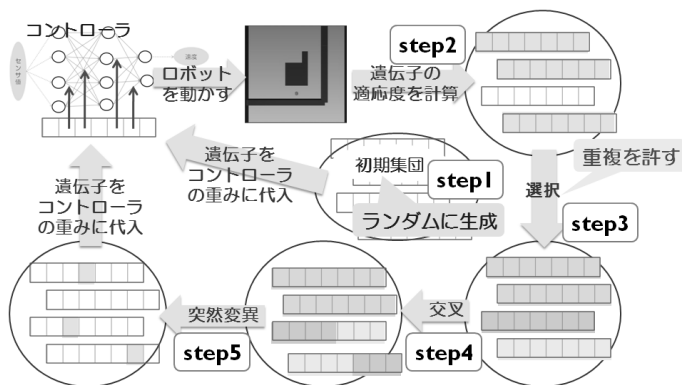


図2: ロボットコントローラの学習

以下に上記概要に沿って、コントローラの学習プロセスを示す。

step 1. 初期集団を生成

-1.0~1.0の実数値を結合荷重の数24個をランダムに発生させ、1つの遺伝子の中に代入する。この個体を50個生成する。

step 2. 適応度の計算

コントローラのノード間の結合荷重に1つの個体の値を代入し、シミュレータ上で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。これを50個全ての個体で行う。

step 3. 選択

ルーレット選択を行い、遺伝子集団から重複を許して次世代の個体候補を選び、新しい個体集団を生成する。新しい集団も元の個体集団の個数を保持する。また、ここでは最も適応度の高い1個体に対してエリート保存選択をしている。

step 4. 交叉

交叉率0.6で一点交叉を行う。

step 5. 突然変異

突然変異率0.1で突然変異を行う。

step 6. step 2. へ戻る

step2～step5を設定された世代数になるまで繰り返す．設定された世代数になったら終了する．

4 ロボットコントローラの学習環境

本研究では，Web上でオープンソースとなっている，進化ロボティクスの研究・教育用3D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う．Simbadは，ロボットの他，壁や箱などの障害物を設置することができ，様々な環境の下でロボットコントローラの学習を行うことができる．ロボットには，センサ（視覚センサ（カラー単眼カメラ），ソナー，光センサ，衝突検出用のバンパー）が取り付け可能となっており，制御対象としては，速度，回転速度，回転角度，左右の車輪となる．今回のシミュレーションで利用するのは，センサ（ソナー）と速度，回転速度である．

5 コントローラ学習実験

本研究では，ロボットが環境の中で壁に衝突しないで動くことを目標とした．

5.1 学習仕様

今回は，図3に示す環境でロボットコントローラの学習を行った．

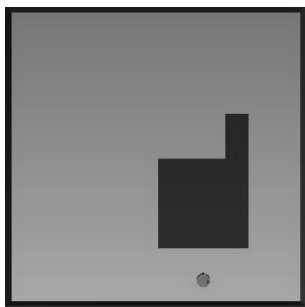


図 3: 学習環境

適応度関数の設定は以下の通りである．

$$f = \sum \{t + (\pi - |r|) + s + k\} + 10 \times d$$

f : 適応度

t : 速度

$|r|$: 回転速度の絶対値

s : センサ値の平均

k : センシング 1step 間の移動距離

d : 総移動距離

適応度関数は，ロボットが大きく回転せず，高速に前進し，次のセンサ情報が入力されるまでの移動距離と，総移動距離が長くなる場合に，適応度が高い値をとるように設定される．長い距離を移動した場合に適応度が高くなるよう，経験的に移動距離に10の値を掛けると良いことがわかった．また，壁に衝突した際には適応度を0にする．

今回は，コントローラを学習させる世代数を100世代と500世代にして実験を行った．コントローラを進化させる前には動き始めてからすぐに壁に衝突していたのに対して，進化させた後には単純な環境の中で壁に衝突せずに動くことができるようになった．

5.2 100 世代

100世代進化させたコントローラは，学習させた環境（図3参照）では壁に衝突しないで動作できるようになったが，得られたコントローラを使って，学習させた環境から少し変化させた環境で動作させると，壁への衝突が観察された（図4参照）．このことから，学習した環境よりも，複雑な環境において行動するには，まだ十分な学習がなされていないことがわかる．

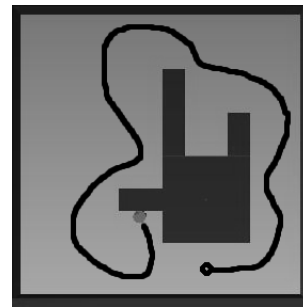


図 4: ロボットの動作軌跡

5.3 500 世代

同様に，図3の環境で500世代まで学習を続けると，図4に示す環境でも壁に衝突せずに動作することが観察された．学習した環境には存在しなかった壁に近づくとき，センサからの情報を得て回避し，袋小路に入り込むと，そこから抜け出す動作を見せた．これにより，世代数を増やして学習したことにより，100世代のときと比べ，複雑な環境でも動作可能なコントローラが獲得されたと考える．

6 おわりに

本研究では，1つの環境でロボットコントローラを進化計算により学習させて，その環境に適したコントローラを得た．今後，様々な異なる環境において学習を行い，その環境それぞれに適したコントローラを構築し，新しい環境にロボットが置かれた際，学習によって獲得されたコントローラ群の中からその環境に合ったものをロボット自身が選択して，動作できるようにしていきたいと考える．

参考文献

- [1] <http://simbad.sourceforge.net/>
- [2] Louis Hugues, Nicolas Bredeche, Simbad: an Autonomous Robot Simulation Package for Education and Research, Simulation of Adaptive Behavior, Rome: Italy(2006)
- [3] S.Nifi and D. Floreano, Evolutionary Robotics, MIT Press, 2000.