

# 様々な環境におけるロボットの行動学習に関する基礎的検討

曽我紗知子 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

進化計算を用いてロボットコントローラを学習させる際、ロボットは与えられた環境において満足に行動できるように学習することができるが、さらに複雑な環境にそのコントローラを適用した際には再度の学習を要する。本研究では、これを避けるため、複雑な環境を構成する部分環境に基づき構築された訓練環境を設定し、その環境で複雑な環境を代替させる。この際、ロボットを効率良く学習させるために、どのような訓練環境を設定したらよいかについて調査する。

## 2 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、視覚センサ(カラー単眼カメラ)、ソナー、光センサ、衝突検出用のパンパーが取り付け可能となっており、制御対象としては、左右の車輪の速度、回転速度、回転角度である。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。

## 3 ロボットの行動学習

### 3.1 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、遺伝的アルゴリズムによって進化させる。

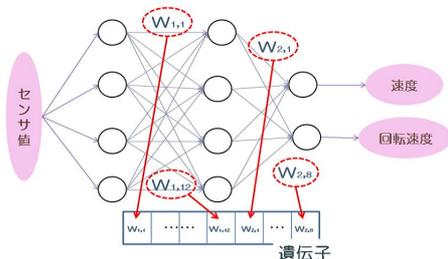


図 1: ロボットコントローラ

### A Study on Learning Robot Behaviors under Various Circumstances

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

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。

### 3.2 学習過程

結合荷重は  $[-1.0, 1.0]$  の実数値であり、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。次にルーレット選択で次世代の個体を選ぶが、更にエリート保存選択を行っている。その後、交叉率 0.6 で、一点交叉を行う。最後に突然変異率 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。

## 4 行動学習実験

### 4.1 学習仕様

文献 [2] における、壁への衝突回避問題で定義された適応度関数を参考にして、本実験における適応度関数を以下の様に設定する。

$$f = \sum \{t + (\pi - |r|) + s + k\} + 10 \times d \quad (1)$$

$f$ : 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 1 ステップ間の移動距離

$d$ : 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。コントローラは、センシングの 1 ステップ毎に入力値を代入し、出力値を計算している。それによって、適応度関数も 1 ステップ毎に累積している。設定したステップ数が終了するか、ロボットの速度が 0, または 1 ステップ間の移動距離が 0 になった時点で、適応度関数の累積計算を終了し、最後に総移動距離を足し合わせている。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 の値を掛けている。また、壁に衝突した際には適応度を 0 にし、終了条件が満たされるまで計算を続ける。終了条件の 1 つであるステップ数は 2500 と設定した。

#### 4.2 静的な環境の変更

訓練環境 (図 2) とその環境を基に、直進を必要とする要素と道幅が狭い要素を加えて訓練環境より複雑にした環境 (図 3) の 2 つを用意する。

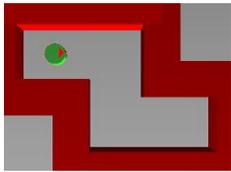


図 2: 訓練環境

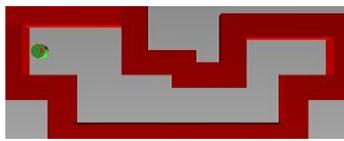


図 3: 複雑な環境

訓練環境で 20 世代まで学習を行うと、壁を避けて環境内を巡回するコントローラを獲得できた。また、図 3 の環境でも学習を行うと、適応度関数が一番高くなるまでの世代数が 120 世代となった。次に、訓練環境で得られたコントローラを使って、ロボットを図 3 の環境で動かしてみると、こちらの環境でもロボットは壁に衝突することなく、滑らかに動き続けた。このことから、「壁への衝突を避けて移動する」学習をするためには、その行動が実現されるのに十分な環境が与えられていれば良いことがわかる。つまり、学習目的とする行動が単純な場合は、ロボットは複雑な環境で学習を行うよりも、その複雑な環境を想定して単純な環境を構築し学習を行う方が効率的であると考えられる。

#### 4.3 動的な環境変化

ここでは、環境が逐一変化し続けている、動的な環境での学習について検討する。今回は学習を行うロボットの他に、複数台のロボットをランダムに動作させ続けることにより、動的な環境とした。また、移動物体との衝突を回避することに重点をおき、式 (1) から総移動距離を除いた以下の適応度関数を設定した。

$$f = \sum \{t + (\pi - |r|) + s + k\} \quad (2)$$

まず初めに、訓練環境を (図 4) とし、シミュレーションを行う。6 台で学習を行った結果、30 世代目で、壁に衝突しないことに加えて、ロボット同士も衝突しないコントローラを得ることができた。このコントローラでは、ロボットはある程度環境内を巡回し、衝突しそうなになると避ける動きを見せている。5 台では、40 世代目で得ることができ、6 台よりも学習にかかる世代数が多くなるのがわかる。また、ロボットは環境内を巡回せず、回転をしており、他のロボットと衝突しそうな場合に避ける動きをしている。2 台～4 台の場

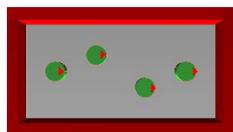
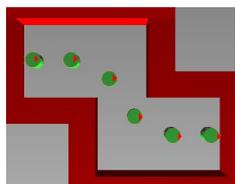


図 4: 移動物体回避の訓練 図 5: 移動物体回避の環境

合では、ロボットは衝突しない半径で回転をする、という学習結果となり、衝突しそうな時に避ける、という行動は学習できなかった。これは、同じ環境においてロボットの台数が増えることでロボット同士が衝突する機会が増えたことにより、学習する機会が増えたため適切に動作するコントローラを得るまでの世代数が少なくなったと言える。この結果から、ロボットが少ない台数の時には、図 4 の訓練環境が広く、学習の機会が十分に与えられなかったため、学習が不十分になってしまい、適切に動作するコントローラが得られなかったということがわかった。

次に、図 4 の環境よりも約半分の大きさの図 5 の環境で学習を行った。ロボットが 5 台の時には、23 世代で適応度が一番高いコントローラを得ることができたが、学習時に他のロボットに衝突しすぎてしまうため、その場で回転し続ける学習が行われ、良い学習結果は得られなかった。4 台では、37 世代が適応度が一番高いコントローラであり、ある半径で回転して、他のロボットと衝突しそうな場合には避ける、という学習結果となった。3 台では、49 世代が適応度が一番高いコントローラであり、衝突を避ける動きも少し見せるが、最終的には壁やロボット同士で衝突していた。2 台では、うまく学習が行えなかった。このことは、与えられた訓練環境においては、ロボットを 4 台用いてコントローラを学習させることが良いということを示しており、それ以上あるいは以下の台数では効率的な学習が行えないことを示している。

図 4 の環境で 6 台のロボットで学習を行ったコントローラをそのまま利用し、図 5 の環境で 4 台にして動作確認をしたところ、正しく動作していた。動的な環境でも、うまく学習が行えなかった場合、動作させたい環境を基にし、訓練環境を構築して学習を行うことで、動作可能にすることができるということがわかった。

#### 5 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得て、訓練環境から他の環境へと移し、考察を行った。訓練環境が学習する間に変化をしない、かつ、目的が単純な場合は、複雑な環境を想定した単純な環境で学習を行うと良いことがわかった。また、訓練環境が学習する間に変化をする場合は、十分な学習の機会が必要となり、環境と動く物体の数の関係を工夫する必要がある。今後は、目的を複雑なものにし、環境が変更されても動作可能なコントローラを獲得することを考えている。

#### 参考文献

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