

ニューラルネットワークによる行動作成に基づいた 進化ロボット

2 T-4

有馬 徹 板倉 秀清 山口 智

千葉工業大学 情報工学科

1. はじめに

進化ロボットの研究は、人工生命等の分野で盛んに行われている。中でも、特定の環境の中で自律移動ロボットを進化させていく研究は、様々な興味深い問題を含んでいる。

一般に自律移動ロボットは、ロボット自身が持っているセンサやカメラを用いて自分の置かれている環境を認識し、目的に応じた行動を決定して、その行動を取るための命令を駆動系に与える。その環境から得られた様々なセンサパターンに対する適切な行動をロボットが学習していかななくてはならない。できるだけ直進する動きや障害物を回避する動きを同時に学習して行くことが求められてくる。

本研究では、行動作成ネットワークと行動レベルネットワークからなるロボットコントローラを作成し、行動レベルネットワークを遺伝的アルゴリズム(以下GA)により進化させることにより、自律移動ロボットの適切な行動を決定していく方法を提案する。

2. 進化ロボット

制御の対象とする進化ロボットとして小型の移動ロボットであるKheperaを利用した。Kheperaは図1のような直径55mmの円形をした、2つの車輪によって動くロボットである。このロボットは左右の車輪の回転速度をそれぞれに付いたモータで変えることによって、置かれた環境の中を自由に移動することができる。また前後に8個の赤外線センサを持ち、壁や障害物等を認識する。

本研究では、Kheperaを参考にシュミレータを作成し、シュミレーションによって進化アルゴリズムの評価を行なった。Kheperaのセンサは、センサと障害物の距離により0から1023の値を出力するが、今回はこれを0から7の値をとるように量子化して使用した。Kheperaは障害物に衝突しないよう、できるだけ直進するよう進化させる。

The method of behavior decision for evolutionary robots using the neural network Toru ARIMA, Satoshi YAMAGUCHI, Hidekiyo ITAKURA
Chiba Institute of Technology

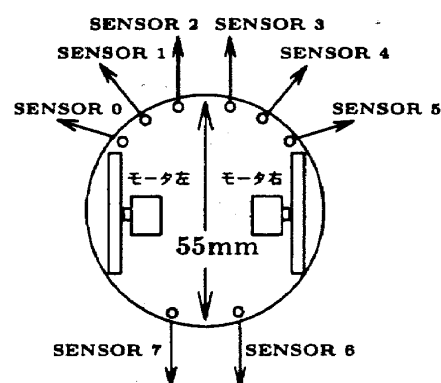


図1: Kheperaの概略図

3 進化アルゴリズム

この節では、行動作成ネットワークと行動レベルネットワークからなるロボットコントローラにより、自律移動ロボットが障害物を回避しながら前進する進化アルゴリズムを提案する。そのロボットコントローラのシステム構成図は図2のようにになっている。このシステムの入力はロボットからの8個のセンサの値である。出力はロボットの行動を制御する速度と、このコントローラが現在のロボットの置かれている状況に有効であるかどうかを決定する行動レベルの2つになっている。

このシステムで基本となっているのが、行動レベルネットワークと行動作成ネットワークである。行動レベルネットワークは、フィードフォワードネットワークにより構成されている。このネットワークからの出力は他のコントローラからの行動レベルと比較し、行動レベルの高いロボットコントローラが現在の自律移動ロボット制御を行なう。行動作成ネットワークは、現在のセンサ値に対する適切な行動をニューラルネットワークにより決定し、ロボットへと出力する。

本研究で使用する学習アルゴリズムは、部分観測マルコフ決定過程における強化学習を用いる。この学習アルゴリズムは、ロボットが周りの状況を完全には把握できない場合に適している。

強化信号決定は、現在のロボットのセンサ値と1ステップ前のセンサ値とを比較して決定した。1ステップ前のセンサ値より現在のセンサ値が増加している場合、強化信号は負の数とし、減少している場合を正の数とした。

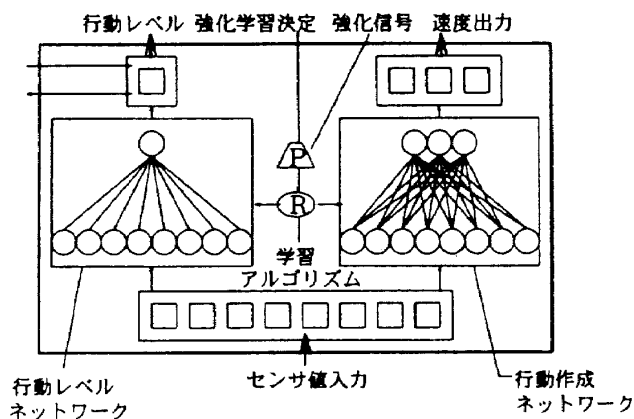


図 2: システム構成図

行動レベルネットワークの進化は、行動レベルネットワークの結合加重をGAを用いてトレーニングすることにより行なう。トレーニング方法はD.Floreano等によって提案された方法を用いる。遺伝子は、各ロボットコントローラの行動レベルネットワークの結合加重を図3のようにつなぎ合わせたものとなる。

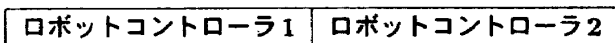


図 3: 遺伝子構成

4 シミュレーション

前節で提案した進化アルゴリズムに基づきシミュレーションを行なった。ここで用いたロボットコントローラは次の2つを用いる。

- **前進コントローラ**
行動作成ネットワークは、センサ値がどのような値でも簡単な直進行動のみを出力する。入力には8つのセンサ値で、出力はロボットをコントロールする速度
- **障害物回避コントローラ**
部分観測マルコフ決定過程における強化学習により、前進、後進、右回転、左回転、右回避、左回避等の行動を決定する。入力には8つのセンサ値で、出力はロボットをコントロールする速度

シミュレーションは、縦横600mmの正方形の箱の中に正方形の仕切によりKheperaの2倍の幅の通路を作り、そこを周回させる環境下で行なった。なお、Kheperaのセンサに対する蛍光灯等のノイズはないものとした。今回のシミュレーションでロボットを進化させる過程として、まず初期集団100個体生成し、それに基づいて行動レベルネットワークに結合加重を与えた。その結合加重を基に、行動レベルネットワークと行動生成ネットワーク共に強化学習を10ステップ行なった後、適応度を計算した。この動作を100個体分行ない、その後遺伝的操作をほどこし、それを50世代行なった。このときの評価関数は(1)のとおりである。

強化信号は、現在のセンサ値が1ステップ前のセンサ値より増加している場合は $R = -1$ とし、減少している場合は $R = 1$ とした。

$$\Phi = V/(S+1) \quad (1)$$

ここで、 $V = |(v_{left} + v_{right})|/(2 * v_{max})$, $0 \leq V \leq 1$ となるが、 v_{left} , v_{right} はロボットをコントロールするための左右の速度、 v_{max} は最大の速度、 S は $0 \leq S \leq 7$ で8個のセンサ値の中で最も高い値とする。

以上のような条件でシミュレーションを行なったが、適応度は0.01から0.24の間で推移した。詳細についてはプレゼンテーションでしめす。

5 おわりに

今回提案した進化アルゴリズムによるシミュレーションでは、前進、障害物回避の2つのロボットコントローラしかなかったが、今後の展開として移動物等を追いかけるようなロボットコントローラも作成していく方針である。

参考文献

- [1] 北野 宏明: “遺伝的アルゴリズム”, 産業図書, 1993.
- [2] 木村 元: “部分観測マルコフ決定過程下での強化学習”, 人工知能学会誌, vol. 11, no. 1, 1996.
- [3] D. Floreano and F. Mondada: “Evolution of homing navigation in a real mobile robot”, IEEE Trans. on Systems, Man and Cybern., vol. 26, pp. 396-407, 1996.
- [4] J. Urzelai and M. Dorigo and D. Floreano and M. Colombetti: “Incremental Robot Shaping”, International Symposium on Evolutionary Robotics pp. 193-203, 1998.