

# 5 J-7 センサ情報と記憶により行動選択を行うクラシファイアシステムによる自律移動ロボットの学習

渡辺勝 山田新一 藤川英司 志田晃一郎  
武蔵工業大学

## 1 はじめに

ゴールの位置を知らない自律型ロボット (Khepera<sup>TM</sup>) が、スタート位置からゴール (光源) にたどり着くまでの行動を学習する。

ロボット工学の分野では強化学習がよく用いられる。本研究では、強化学習法の一つである遺伝的アルゴリズム (GA) を用いたクラシファイアシステムを採用する。GA に基づく機械学習法にはルールセットを集団として扱うミシガンアプローチとルールセットを個体として扱うピッツアアプローチがあるが、システム全体に安定性があり、オンライン学習に適したミシガンアプローチを用いる。[1]

従来の移動ロボットの学習では、センサ入力に対する行動を学習するだけのものが多いが、過去の情報 (行動の記憶) を利用することにより、単なる条件反射でない行動の学習を目指す。

また、学習の効率を考え、従来のクラシファイアシステムとは違い、初期ルールを作らずに、ロボットが直面した状況において逐次ルールを作成する方法をとる。

## 2 クラシファイアシステム

クラシファイアは一種のプロダクションルールであり、次のようなものである。

```
<classifier> ::= <condition> : <action>
IF <condition> then <action>
```

もし、条件を満足するクラシファイアが複数ある場合、報酬によって割り当てられた各クラシファイアの強度によりこの競合を解消する。そして、あ

る程度学習が進み、各クラシファイアの強度に十分な変更がなされた時点で、強度を適合度として遺伝的アルゴリズムを適用する。これにより、無駄なクラシファイアは淘汰、突然変異、有効なクラシファイアを交叉することにより新たなクラシファイアの生成が行われる。以上の仕組みにより、クラシファイアシステムでは戦略や手法といったある問題に対する解法をルール群という形で得ることができる。

## 3 自律移動ロボット

Khepera(図1, 2) は、人工生命、人工知能研究で広く使われている小型移動ロボットである。周囲には8つの赤外線接近センサと光センサを兼ねたセンサを装備している。ロボットには2つの車輪があり、DCモータによってコントロールされる。

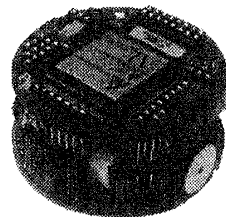


図1: Khepera 概観

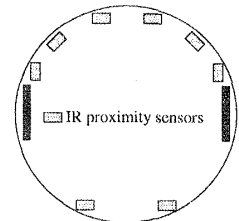


図2: センサの位置

クラシファイアシステムを適用するにあたって、<condition> と <action> を以下のように定義する。赤外線接近センサの値は近いと遠いで2値化し、2つのセンサのORをとって前後左右の4つで状態を表す。光センサは最も光の強い方向 (前後左右) を状態とする。ただし、光センサの返す値がほとんど変わらない場合は、光の強い方向が分からない状態とする。行動は、前進、その場で左右に 30° 回転、180° 回転の4つとする。

## 4 自律移動ロボットの学習

前節で、定義した <action> は単純なものだけであり、センサから得られる情報だけで行動を起こしたのでは複雑な行動をすることは難しい。そこで、過去の記憶を利用して行動選択を行う。つまり、クラシファイアを次の様に定義する。

$\langle \text{classifier} \rangle ::= \langle \text{condition} \rangle : \langle \text{memory} \rangle : \langle \text{action} \rangle$   
 IF <condition> and <memory> then <action>

また、従来の方法では初期クラシファイアをランダムに生成するため、ロボットと環境の関係から絶対に起こりえないクラシファイアまで作成してしまい無駄が多いと考えられる。そこで、ロボットの入力情報によるメッセージリストにマッチングするクラシファイアが存在しなければ、そこで初めて行動をランダムに決め、ルールを生成を行う。

次に、ロボットが障害物に衝突、または、いっそうに前進しようとしなない場合、その時に実行されたクラシファイアを淘汰し、スタートからそれまでの間に実行され淘汰されなかった全てのクラシファイアに使用されたことに対する報酬1を与え、スタートに戻す。ゴールに到達した場合は、スタートからゴールまでの全てのクラシファイアに報酬10を与える。そして、学習が十分に進んだ場合(スタートしてから3000step), 以下に示す遺伝的アルゴリズムを適用する。

1. 親をルーレット選択で選び、交叉を行う
2. 確率『 $1/(10+\text{強度})$ 』で、突然変異を起こす
3. 強度が0のクラシファイアを淘汰する

図3に本移動ロボットにおける学習の手順を示す。

## 5 シミュレーションと評価

本研究ではWWW上で公開されている Khepera シミュレータ [3] を使用した。図4は、138世代に現れたロボットの行動である。簡単なL字型の環境ではあるが、自律型ロボットが、スタート位置からゴールにたどり着くまでの行動を学習することができた。今後は、さらに複雑な環境での学習を試みたい。また、記憶を用いたために探索空間が増大してしまい、学習が困難になるという問題が発生したため、遺伝的アルゴリズムの適用法などを工夫する必要がある。なお、本研究はITS研究の一部である。

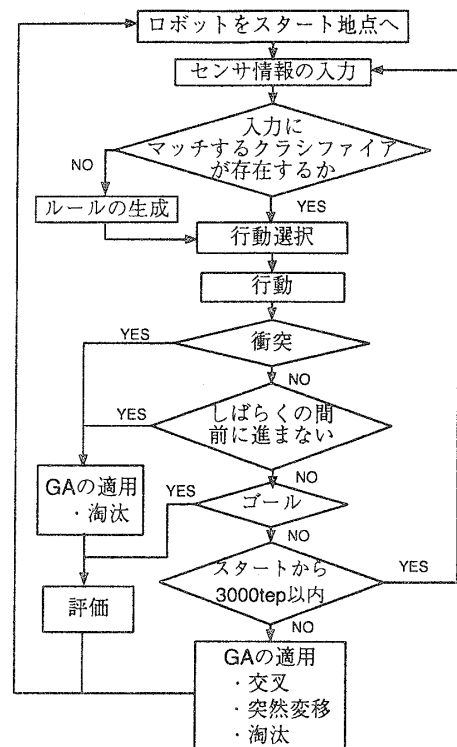


図3: 移動ロボットの学習手順

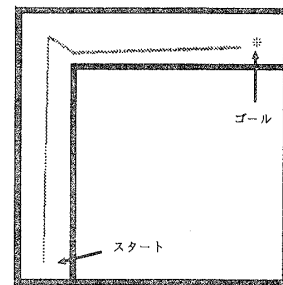


図4: シミュレーション結果

## 参考文献

- [1] 坂和正敏, 田中雅博, 遺伝的アルゴリズム 朝倉書店 (1995)
- [2] 小沼克治, 星野 力, 分類システムによるロボットの行動の進化学習 人工知能研究会資料 p.9-14 (1995)
- [3] O.Michel. Khepera Simulator v.2 User Manual. Universit of Nice-Sophia Antipolis [http://in1suns1.epfl.ch/lami/team/michel/khep-sim/] (1996)