

# ファジクラシファイアシステムを用いたロボットの行動学習に関する一考察

金子貴美<sup>†</sup> 曾我紗知子<sup>‡</sup> 小林一郎<sup>‡</sup>

<sup>†</sup>お茶の水女子大学理学部情報科学科

<sup>‡</sup>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

本研究では、ロボットがある環境において獲得した行動知識を他の環境においても再利用可能にすることを目的とする。ロボットが獲得した知識を他の環境に転用するにあたって、獲得された知識が人に可読であり、かつ様々な状況に対して部分的に適用可能であることが望ましいため、学習メカニズムにファジクラシファイアシステムを用いて、if-then 規則で表わされるロボットの行動知識を獲得する。そして、他の類似した環境下にその知識を適用することで、様々な環境で既に得た知識が再利用可能かどうかを検討する。

## 2 学習環境

環境に合わせて行動を学習させる手段としては、ニューラルネットワークと遺伝的アルゴリズム(以下GA)を組み合わせるなど様々な方法が考えられるが、人にも可読性が高く、プロダクション規則でコントローラが構成されるクラシファイアシステムで学習させる。また、プロダクション規則をファジ推論規則とするファジクラシファイアシステム(以下FCS)を適用することで、環境に柔軟に反応でき、かつ、少ない規則数でコントローラを構成することを目指す。

## 3 実験環境

### 3.1 実験仕様

実験環境として Simbad ロボットシミュレータ [1] を用い、ファジクラシファイアシステムを構築した。

実験内容として、まず、ランダムに生成したルール 60 個のルール群(ルール A とする)を用意する。次に、図 1 の訓練環境で 1 周するようになるまで学習させたルール群(ルール B とする)を獲得する。得られた 2 つのルール群を図 2 に示す学習用環境で、それぞれ 1 周するようになるまで学習させ、学習効率を比較する。

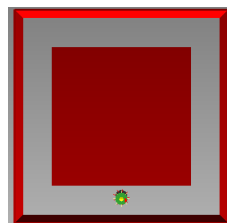


図 1: 訓練環境

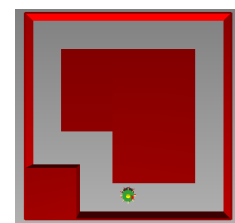


図 2: 学習用環境

### 3.2 FCS による学習

実験に用いた FCS を図 3 に示す。

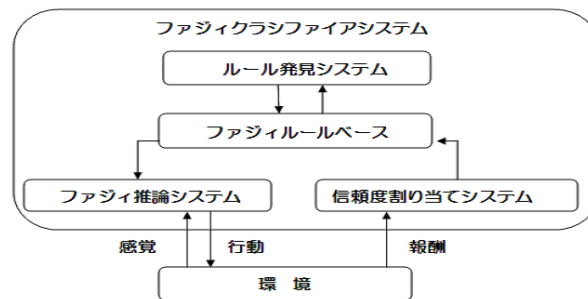


図 3: ファジクラシファイアシステムの構成

以下に、本研究における FCS の実装を各部の内容について説明する。

#### • ルールベース

ルールは、図 4 に示される、5 入力 2 出力 2 評価値の 9 遺伝子座で表現する。

メンバーシップ関数の値					角度	速度	信頼度	適合度
S	N	D	S	S	0.0	0.5	1.0	1.5

入力(前件部)
出力(後件部)

図 4: 遺伝子座の設定

5 入力はロボットに備え付けられた 5 つのセンサーから得られる値となり、センサーの入力値は 3 種類のメンバーシップ関数: Safe(S), Neutral(N), Danger(D) のいずれかで表現される(図 5 参照)。2 出力はそれぞれロボットの回転角と進行速度となり、2 評価値はルールの信頼度および、ファジ推論

A Study on Learning a Robot Controller by Fuzzy Classifier System

<sup>†</sup>Kimi KANEKO(your-email@is.ocha.ac.jp),

<sup>‡</sup>Sachiko SOGA(soga.sachiko@is.ocha.ac.jp),

<sup>‡</sup>Ichiro KOBAYASHI(koba@is.ocha.ac.jp)

<sup>†</sup>Dept. of Information Sciences, Faculty of Science, Ochanomizu University, 2-1-1 Ohtsuka Bunkyo-ku Tokyo 112-8610

<sup>‡</sup>Advanced Sciences, Graduate School of Humanities and Sciences, Ochanomizu University, 2-1-1 Ohtsuka Bunkyo-ku Tokyo 112-8610

によって得られたルールの適合度の累計となる。

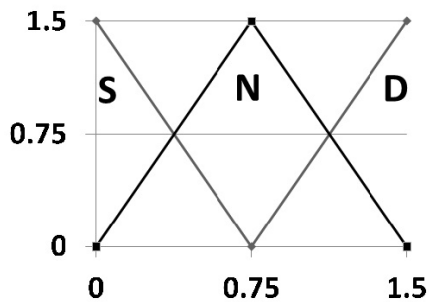


図 5: メンバシップ関数の設定

● ファジィ推論

センサーの入力を取りこみ、ルールベースの各ルールごとにメンバシップ関数を用いてファジィ推論を行い、値を出力する。推論方法は、重み 1 のファジィシングルトン型推論法を用いる。

● 信頼度割り当てシステム

目的の地点、および中継点まで到達したら正の報酬を、到達する前に壁に衝突したら負の報酬を与えるとする。また、ルールの適合度に応じて与える報酬の量を決めるとする。

$$\text{与える報酬} = \alpha \sum_x \text{ルール適合度}$$

ここで  $\alpha$  は、正の報酬の場合は各地点で設定した値、負の報酬の場合は最後に報酬を与えた地点からのユークリッド距離である。 $x$  は、最後に報酬を与えてから  $x$  番目のサンプリングタイムである。

● 遺伝的アルゴリズム

目的の地点に辿り着いた場合、もしくは衝突した場合、ルール群に報酬を付与した後に GA を適用し、ルール群を学習させる。

GA においては、信頼度の高さ上位 2/3 のルールは交叉対象とはせず保存しておき、下位 1/3 のルールに対して一点交叉・突然変異を施すとした。また、突然変異は、最初、前件部で行った後、前・後件部全体で再度行い、前件部が重複するルールが生成された場合は、新しく生成されたルールを削除するとした。

4 実験結果と考察

実験結果を表 1 に示す。

‘世代数’ は 1 周出来るようになるまでの学習にかかった世代の数、‘ルール数’ は学習後のルール群の個体数である。

表 1: 実験結果

事例	使用ルール	環境	世代数	ルール数
1	ルール A	訓練環境 (図 1)	30	51
2	ルール A	練習用環境 (図 2)	251	40
3	ルール B	練習用環境 (図 2)	111	40

事例 1 において、学習前の段階では、図 6 のようにすぐ壁に衝突した。また、事例 2 でも、学習前は図 6 同様にすぐ壁に衝突した。一方、事例 3 の学習前段階では、初めから右上にある図 7 の経路で進むことが出来たが、訓練環境と異なる部分になった途端衝突した。

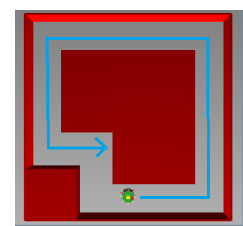
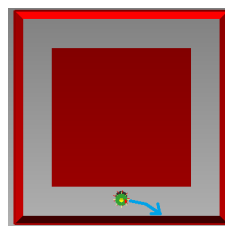


図 6: ルール A 学習前

図 7: ルール B 学習前

実験結果では、ランダムに構築したルールから学習させるより、類似した環境での既得知識を転用して学習させたほうが学習効率が悪くなっていることがわかる。また、図 7 のように、ルール B は学習用環境に移した時点で以前学習した環境と類似している部分では頑健に動作したことが分かる。これらのことから、類似した環境においては、既得知識を転用して学習させたほうが効率が良くなると考えられる。

5 まとめと今後の課題

本研究では FCS を用いて、ランダムに構築したルール群と学習用環境と類似する環境で学習させたルール群の学習用環境における学習効率を比較した。実験結果から、類似した環境においては、既得の知識を再利用して学習した方が学習効率が悪くなることがわかった。

今後は、他の環境・他の初期ルール等でも同様に実験し、既得知識の転用に関してより深い考察や検証を行い、学習した知識を部分的に転用する可能性について調べていきたい。

参考文献

[1] <http://simbad.sourceforge.net/>  
 [2] 高野敏明, 高瀬治彦, 川中普晴, 鶴岡信治: 強化学習における異目的タスク間での知識の転移に関する一考察, 第 27 回ファジィシステムシンポジウム, 9 月, 2011.  
 [3] 井上寛康, 高玉圭樹, 下原勝憲, 片井修: 知識を再利用するクラシファイアシステムの複数の環境への適応, 第 3 回 MYCOM, 2002.