

力学的自己増殖型ニューラルネットワークを用いた 状態空間構成における汎化学習

日大生産工 ○根本 洋輔 日大生産工 山内 ゆかり

1 まえがき

近年、斉藤らが提案した自己増殖型ニューラルネットワーク (Self-Organizing Incremental Neural Network: SOINN) が注目されている。従来手法ではSOINNを用いて行動則の学習が行われていたが学習の過程においてまだ改善の余地が見受けられた。

本研究では力学的自己増殖型ニューラルネットワークを用いた状態空間構成における汎化学習を提案し、その有効性を検証する。

2 SOINNによる行動則の獲得

SOINNによる行動則の獲得とは、SOINNを状態推定に用いて、迷路でのエージェントの行動を強化学習によって学習させ行動則を獲得させるというものである。以下にSOINNのアルゴリズムを示す。

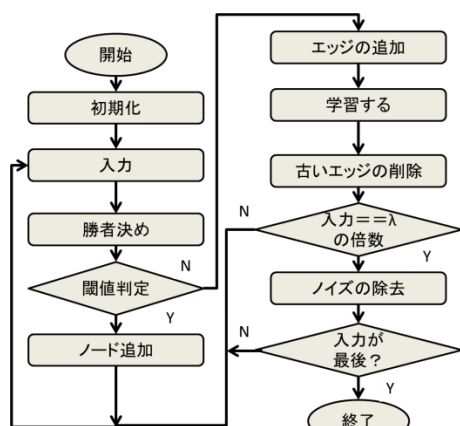


図1 SOINNのアルゴリズム

2-1 Profit Sharing

強化学習を応用する際に必要である、状態観測に不確実性や不完全性を伴う場合における頑健性と素早い学習が可能であるものがPSである。本研究ではPSを用いる。

PSは報酬を得たときにエピソード単位で使ったルールを一括で更新する。

$$\omega(s_i, a_i) \leftarrow \omega(s_i, a_i) + r * f_i \quad (1)$$

$\omega(s_i, a_i)$ はエピソード系列において i 番目の状態 s_i で選択した行動 a_i に対するルールの重みである。

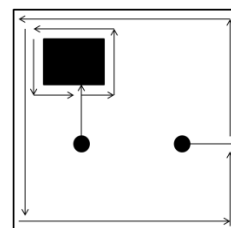
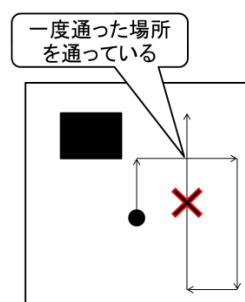
f_i は強化関数で一般的には次式で表される。

$$f_n = \frac{1}{S} f_{n-1} \quad (n = 1, 2, \dots, W - 1) \quad (2)$$

W はエピソードの最大長、 $1/S$ は報酬割引率を表している。

3 提案手法

本研究はエージェントの移動の仕方を変え、学習の促進を図るというものである。その提案手法の概要を図2に示す。



一度通った場所へは行かない 一番近い壁に向かっていき、その壁に沿って移動していく

図2 提案手法の概要図

提案手法は

- 同一座標には移動しない
- 一番近い壁に向かって移動し、壁に沿って移動していく

この提案手法を検証し、実用性を確かめる。

4 実験環境

エージェントは

- ・上下左右の4方向を状態価値に基づきルール選択で確率的に移動する
- ・障害物や壁の中には進めない

The Generalization Learning in State Space Construction
with Dynamics-based Self-organizing Incremental Neural Network

Yousuke NEMOTO and Yukari YAMAUCHI

- ・ゴールに着いたら報酬を受け取る

以下に行動則獲得のアルゴリズムを示す。

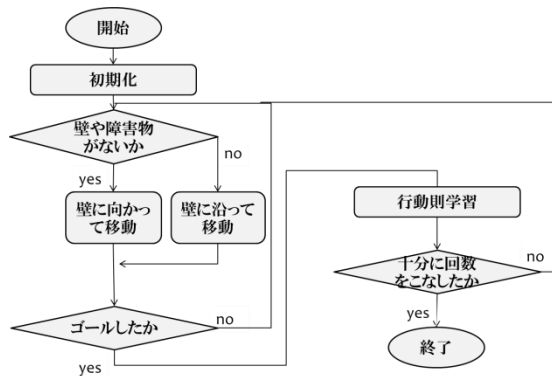


図3 行動則獲得のアルゴリズム

行動選択し実行するまでの過程を1ステップ、スタートからゴールに到達するまでの過程を1エピソードとした。30000ステップを超えてもゴールに到達しない場合は、その試行は失敗と見なして強制的に終了させ次のエピソードに進む。実験は論文[2]を参考に、40×40の正方形に障害物とエージェントを配置した、3種類の迷路でシミュレーションを行うこととした。以下に3種類の迷路の図を示す。

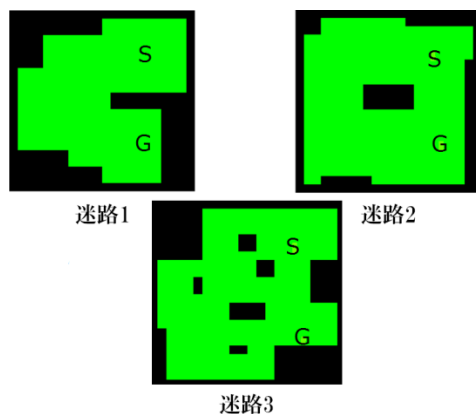


図4 迷路1,2,3

図4の迷路のSはスタートで初期のエージェントの位置もスタートにある。Gはゴールである。

5 実験結果および検討

図5にProfit Sharingのみの結果、図6にProfit SharingとSOMによる状態空間構成の結果を示す。

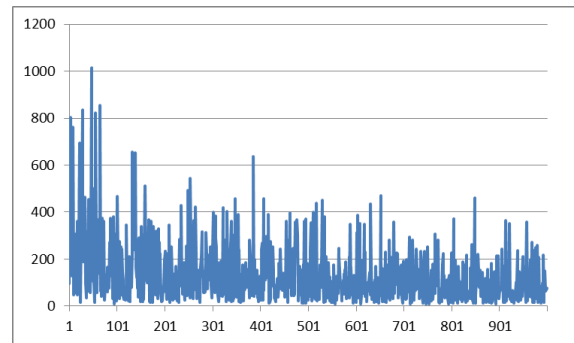


図5 PSの各エピソードでの行動回数

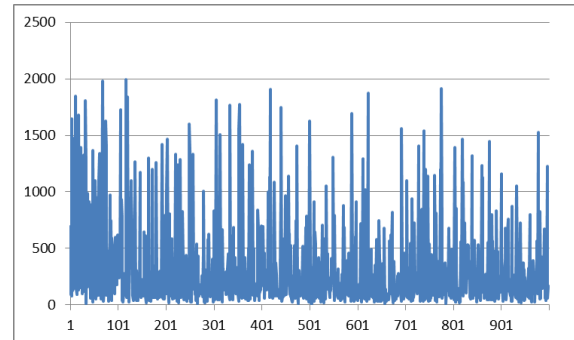


図6 PS+SOMの各エピソードでの行動回数

PSの学習は段々と収束していったが、PS+SOMの学習では収束には至らなかった。これはSOMが予めノードを用意して、そこから学習するため、行動則が上書きされるからだと考えられる。しかしSOINNは予めノードを用意しておかなくてもよい。よってSOMではなくSOINNを使うことによって解決できると考えられる。

6 まとめ

本研究では力学的自己増殖型ニューラルネットワークを用いた状態空間構成における汎化学習を提案し、迷路での行動則獲得実験を行った。PSでは学習が収束し、ゴールまでの行動則がスムーズに獲得できた。しかしPS+SOMでは学習が収束せず、行動則の獲得がスムーズにしていなかった。

「参考文献」

- 1) 斉藤 史哲、長谷川 修：「行動則の獲得状況を考慮した自己組織化マップによる状態空間構成法」，日本知能情報ファジィ学会誌：20(3), 369-378, 2008-06-15 日本知能情報ファジィ学会
- 2) 斉藤 史哲、長谷川 修：「力学的自己増殖型ニューラルネットによる知覚・行為系列に基づく包括アーキテクチャ上の状態空間構成」，日本知能情報ファジィ学会誌：22(2), 266-278, 2010-04-15 日本知能情報ファジィ学会