

# 強化学習による Gauss-Sigmoid ニューラルネットの大域的情報表現の獲得

大分大学 前原 伸一 柴田 克成 杉坂 政典

## 1. はじめに

従来、主として行動のプランニングに用いられてきた強化学習は、ニューラルネットと組み合わせることにより、センサからモータまでの一連の処理を総合的に学習することが可能となる。しかし、この組み合わせは Boyan らにより、学習を不安定に導くと指摘された [1]。これに対し、RBF ネットなどの局所的な情報表現の使用が有効であることが示されているが、これらの方法は、大域的な情報を表現する手段を持っておらず、汎化能力が劣る。そこで、われわれは連続値入力の強化学習に Gauss-Sigmoid ニューラルネット (NN) [2] を用い、シグモイド型 NN よりもよい動作を獲得できることを確認した [4]。本稿では、Gauss-Sigmoid NN の大域的情報表現能力に注目し、RBF ネットと比較することでその有効性を検証する。

## 2. Hill-car タスク

本稿では、タスクとして hill-car 問題を用いる。このタスクは、推力が小さく、一度では頂上に到達できない車が、何度も左右に進み反動をつけて斜面の頂上に到達するというものである。そのため、ゴールの方向に最高推力を出し続ければゴールに到達できる領域とそうでない領域の境界部分においては理想的な状態評価関数と出力すべき推力は共に不連続となり、その近似には強い非線形性が要求される。

## 3. Gauss-Sigmoid ニューラルネット [2]

RBF ネットは優れた非線形関数近似能力を有するが、出力が局所的な情報の線形和で表されるため、大域的情報表現ができず、汎化能力も大きく劣る。一方、シグモイド型 NN では、大域的な表現が可能であるが強い非線形関数の近似には適していない。そこで、本稿では、RBF ユニットの出力をシグモイド型 NN の入力とする Gauss-Sigmoid NN を用いる。これにより、局所化された信号をシグモイド関数を出力関数とする中間層で統合し、大域的な表現の獲得が可能となる。

## 4. シミュレーション

学習は Actor-Critic アーキテクチャで行い、Actor と Critic を 1 つの Gauss-Sigmoid NN で構成した。そして、出力ユニットを 2 つ設け、1 つを動作、1 つを評価として扱った。Hill-car タスクにおいて、初期状態をくぼみ  $(x, v) = (0.0, 0.0)$  として、強化学習で 2000 回学習することによって、ほぼ最大推力で左右に進み、4 回の往復の後、ゴールに到達するようになった。その後、出力層に中間層との結合係数 0 のニューロンを新たに追加し、そのニューロンに対して教師あり学習を行った。教師信号は 図の小さい で示した状態に対し、学習によって獲得された評価値、推力値をそれぞれ用いた。2000 回の学習後と教師あり学習後の車の位置  $x$ 、車の速度  $v$  に対する出力の分布を図 1 に示す。ただし、図中の灰色の線は車の軌跡である。図から、Gauss-Sigmoid NN では、中間層内部に以前学

習した状態評価値の尾根と推力の情報を蓄えていることがわかる。一方、RBF ネットでは、学習が進まず、教師信号を与えた領域の近傍以外では出力は一定となった。このように、Gauss-Sigmoid NN では、中間層内部に以前学習した大域的な情報としての知識を蓄えることができる。したがって、この知識が活用できる別のタスクを学習する場合、学習を 1 からやり直す必要はなく効率的な学習が行えることが期待される。

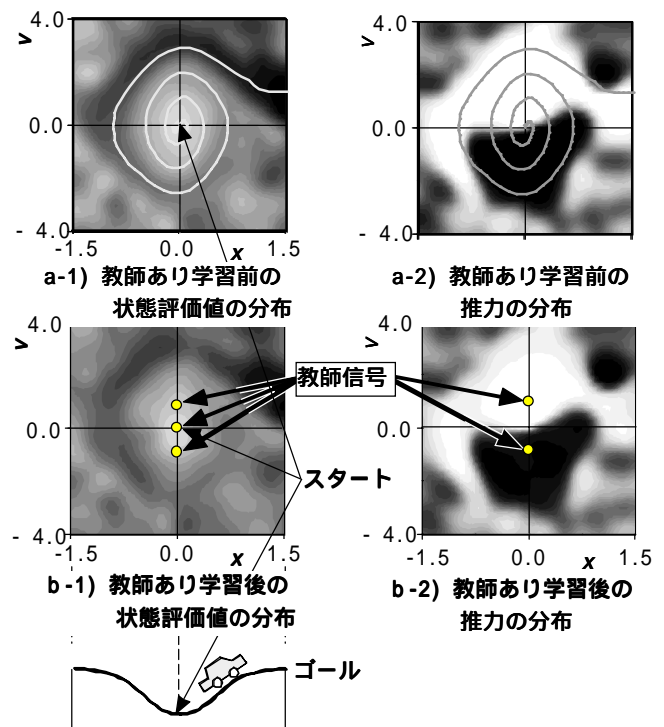


図 1. Gauss-Sigmoid NN の大域情報表現能力

## 5. 結論

本稿では、Gauss-Sigmoid NN が RBF ネットワークと比較して、優れた大域情報表現能力を備えている一例を示し、入力信号が大域的な連続値信号である場合に Gauss-Sigmoid NN を用いることで、中間層内部に必要な知識を蓄え、汎化が有効に働くことを確認した。しかし、タスクによってはその能力を十分引き出せない場合があるなど今後さらに研究を進めていく必要がある。

## 参考文献

- [1] Boyan, J.A. & Moore, A.W., "Generalization in Reinforcement Learning: safely approximating the value function", in Advances in Neural Information Processing Systems, Vol. 7, pp. 369-376, The MIT Press(1995)
- [2] 柴田克成, 前原伸一, 杉坂政典 "Gauss-Sigmoid ニューラルネットワーク" 第 13 回自律分散システム・シンポジウム資料, pp.133-138(2000)
- [3] 前原伸一, 杉坂政典, 柴田克成, "連続値入力強化学習における Gauss-Sigmoid ニューラルネットワークの有効性", 信学技報 100-688, pp.75-82 (2001)

