

## 自己組織化回路素子 (SONE) への教師あり学習の付与

金 天海\* 菅野 重樹\*

早稲田大学 理工学術院

機械工学専攻 菅野重樹研究室

東京都新宿区大久保 3-4-1

{tennkai,sugano}@sugano.mech.waseda.ac.jp

尾形 哲也\*\*

京都大学 情報学研究科

知能情報学専攻 奥之研究室

京都市左京区吉田本町

ogata@i.kyoto-u.ac.jp

**Abstract**—自律型ロボット制御に用いる学習器への機能要求である、効果的な出力の自律的模索、単純な外部パラメータ、小さな計算コスト等を実現できる手法として、著者らは自己組織化回路素子 SONE を提案している。SONE はネットワーク素子間の強化信号伝播をベースとして回路構造を自己組織的に獲得するアルゴリズムである。従来 SONE による学習は強化学習に限定されており、基本特性の解析や強化学習以外の学習への応用が困難であった。そこで本稿では SONE による教師あり学習法を提案する。提案手法により SONE の挙動の解析を行い、追加学習、時系列学習等への有効性が確認できた。提案手法により扱えるタスクの幅も広がった。

## I. はじめに

近年、ニューラルネットワーク (NN) 等のネットワークを用いた自律型ロボットの制御技術は大きく発展している。自律型ロボットの学習制御装置に求められる要求機能は様々であり、その一部として以下のような機能が挙げられる。

- 1) 効果的な出力の自律的模索
- 2) 単純な外部変数
- 3) 小さな計算コスト

これらの機能要求を実現する手法としては、Direct-vision-based reinforcement learning (DVB-RL) や Neuro genetic algorithm (NGL) が比較的有効である。しかしながら DVB-RL はタスク・環境に応じた調整を必要とする、NGL は多くの計算コストがかかる等の問題がある。そこで筆者らは上記の要求機能の全てを満足する枠組みとして、自己組織化回路素子 (SONE) を提案した [1], [2]。

SONE に基づいたネットワーク上の各素子は、ローカルルールによって自律的にネットワーク構造の変更を行う独立したエージェントとして構成される。各素子に強化信号伝播機能を持たせることにより、全ての素子について評価値の算出が可能であり、その評価値に基づいた回路素子の生成・淘汰によって回路構造の最適化を行える。SONE に基づいて開発した自己組織化論理回路 (SOLC) は 1 ステップにつきノード数 (V) + リンク数 (E)  $O(V+E)$  のオーダーで計算可能である。さらにはオンラインに強化学習を行うことができる。よって、SONE は上記の要求機能の全てを満足する枠組みであるといえる。

しかしながら SONE による学習は強化学習に限定されており、特性の解析が困難であった。また応用できるタ

スクも限定されていた。そこで本稿では新たに SONE を用いた教師あり学習法を提案する。本稿で提案した手法により、SONE によって汎化、追加学習、時系列学習ができることがわかった。応用できるタスクも広がった。

## II. SONE による教師あり学習

本稿では、SONE に基づいて開発した SOLC を用いる。SOLC は And ノード、Or ノード、反転リンク、非反転リンクによって自律的に論理回路を構成するシステムである。SOLC のネットワークは行動フェイズと学習フェイズを備えており、行動フェイズでは入力に従って出力を計算する。学習フェイズでは人や環境といった外部環境より強化信号を受け取り、その報酬を各素子の評価に反映する。さらには、より多くの報酬 (正の強化信号) を得られるよう、回路構造の変更を行う。

従来の SOLC では、3-bit 演算の学習やロボットシミュレーションによる衝突回避が扱われてきた。これらの実験では、外部より得られた強化信号は出力層に対して均等に付与されている。しかしながら教師あり学習を行うためには、出力層のそれぞれの素子が異なる目標値に対して学習を行う必要があり、従来の方法では困難である。そこで本稿では従来の学習フェイズ (強化学習フェイズ) に加えて、教師あり学習フェイズを提案する。

従来研究によれば、SOLC の各素子は有効度 50% を上回るリンク (報酬期待値が罰期待値を上回るリンク) を生存させることにより、自らの受け取る報酬量を増加させることができる。さらに各素子は有効度 50% を下回るリンクの削除機能、そして新たなリンクの探索機能を有する。よって SOLC の各素子は報酬に対して貪欲な性質を有しているといえる。そこで教師あり学習フェイズでは SOLC の出力と目標値が同じである場合には出力ノードへ報酬 (1) を異なる場合には罰 (-1) を付与して学習を行う。

この実装法ではネットワーク内部の報酬系が二つのフェイズで共有できるため、一つのネットワークに二つのフェイズを共存させることができる。

## III. 実験

本稿では汎化、追加学習、時系列学習の特性を調べるために軌道の学習に関する実験を行った。

実験に用いた軌道は二次元平面上の円軌道 (C-track) と八の字軌道 (I-track) であり、C-track と I-track を交互に周回する C& I-track に関する実験も行った。ノイズの有無により全 6 通りの実験を 10 回ずつ行い、収束点を CP として、その結果を Fig.I と Fig.II にまとめた。

Enhancement of Self-Organizing Network Elements for Supervised Learning

\* Chyon Hae Kim and Shigeki Sugano are with the Department of Mechanical Engineering, Waseda University, 3-4-1 Shinjyuku-ku Okubo, Tokyo, Japan {tennkai,sugano}@sugano.mech.waseda.ac.jp

\*\*Ogata Tetsuya is with the Graduate School of Informatics, Kyoto University, Yoshida-honmachi sakyo-ku, Kyoto, Japan ogata@i.kyoto-u.ac.jp

TABLE I  
CONCLUSION OF LEARNING (NOISE-FREE)

Data	Average steps	Average error (bit)	Average error (real number)	Number of nodes at CP	Number of links at CP
C-track	88.3	0.3	$6.25 \times 10^{-7}$	70.8	244.0
I-track	1064.9	0.3	$4.93 \times 10^{-5}$	82.2	306.3
C& I-track	923.1	1.0	$2.78 \times 10^{-4}$	149.6	536.7

TABLE II  
CONCLUSION OF LEARNING (5 PERCENT NOISE)

Data	Average steps	Average error (bit)	Average error (real number)	Number of nodes at CP	Number of links at CP
C-track	196.3	0	0	91.5	318.2
I-track	1068.7	0	0	153.9	537.5
C& I-track	2779.8	1	$2.08 \times 10^{-6}$	448.1	1725.3

さらに追加学習に関する実験では、C-track を完全に学習させた後 I-track を学習させ、I-track が完全に学習できた時点で再び C-track を学習させる実験を SOLC、リカレントニューラルネットワーク (RNN) の両方で行った。RNN は 2-20-2 のトポロジーを持つ三層型であり、コンテキスト層は 5、学習は BPTT で行っており、学習率は 0.01 である。SOLC に実数表現を学習させるため、そのインタフェースには 16bit の A/D, D/A 変換を用いることとした。ネットワークへの入力には現在位置であり、ネットワークの出力は次に移動するべき位置とした。ネットワークが出力を行うと次の位置へ移動でき、その位置座標 (ノイズが乗った場合にはノイズの乗った位置座標) がネットワークへ入力される。

SOLC にとって C-track は静的な問題であり、それぞれの入力に対して 1 対 1 に対応する出力を学習する問題である。Fig.2 にこの実験により得られたネットワークを示す。左右に並ぶ点はそれぞれ入出力用端子であり、それ以外の中間素子は全て SONE の自己組織化により得られる。しかしながら、I-track は時系列を含んだ問題であり、八の字軌道の中心点では入出力を 1 対 1 に定めることはできない。この軌道を学習するためには、SOLC 内部に再帰結合によるメモリ構造が形成され、移動の履歴に応じた学習が必要である。

#### IV. 考察

ノイズの乗った軌道学習では、ネットワークが受け取る入力信号は毎回異なり、単に軌道を覚えるだけでは解へ達することはできない。従って、学習軌道を汎化することで未知の軌道へと対応する必要がある。Fig.II において C-track と I-track は誤差が 0 になるまで収束しており、SOLC に汎化能力が有ることを示している。

先に述べた追加学習実験の学習履歴が Fig.1 である。Fig.1 のオーバーシュートしている部分が C-track, I-track の切り替え点である。実験からエラーが収束することがわかり、SOLC は新しいタスクの学習中に古いタスクの記憶を保持しており、次の学習に利用できることがわかった。また、RNN との比較から SOLC のタスク保持能力の高さと、収束の早さが伺える。I-track の中心点を正しく学習するためには、複数ステップの軌道の履歴を何らかの形で保持することが必要である。しかしながら本実験では、SOLC の各素子が入出力を保持できる期間は 1 ステップであるため、この学習を行うためには再帰結合が正しく形成され、ネットワークのループによって記憶を保持する必要が

Fig. 1. Error history

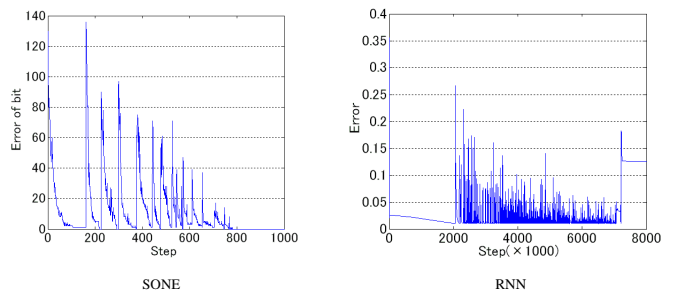
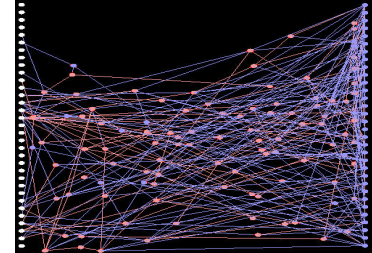


Fig. 2. Self-organized network (C-track)



ある。Fig.II によれば SOLC は I-track に関しても誤差を 0 にまで収束させることができる。よって、SOLC は再帰結合を伴ったメモリ構造を獲得することができた。

#### V. おわりに

本稿では、SONE における教師あり学習法の提案を行った。提案した手法によって汎化、追加学習、時系列学習への有効性を確かめた。従来の特性であるトポロジーの自己組織化、小さな計算コスト、オンライン学習も同時に実現できており、SONE の扱えるタスクの幅が大きく広がった。今後はさらに長時間の時系列学習や、より複雑なタスクへの適用を行っていきたい。

#### REFERENCES

- [1] Chyon Hae Kim, Tetsuya Ogata, and Shigeiki Sugano: "Self-Organizing Algorithm for Logic Circuit based on Local Rules", Transaction of the Society of Instrument and Control Engineers Vol. 42 No.4, 2006.
- [2] Chyon Hae Kim, Tetsuya Ogata, and Shigeiki Sugano: "Efficient Organization of Network Topology based on Reinforcement Signals" (Accepted), IEEE/RSJ Proceeding of the International Conference on Intelligent Robotics and Systems, 2006.