

NGL法に基づいたリカレントニューラルネットワークの学習

4 T-3

浜道 典孝 山口 文彦 中西 正和

慶應義塾大学大学院 理工学研究科 計算機科学専攻

1. はじめに

階層構造ニューラルネットワークの学習においては、誤差逆伝播法 [1] の有用性が確認されている一方、解が誤差超曲面上の極小値に陥ってしまうことがある点、また、ネットワークの構造決定を設計者の主観に頼っている点などが問題点として指摘されており、それらの改善を目的として、生物の進化（選択淘汰・突然変異）の過程に基づいた遺伝的アルゴリズムを導入する研究が行われている。

本研究では、リカレントニューラルネットワークの学習において、遺伝的アルゴリズムを用いたアルゴリズムの一つである Neurogenetic Learning (NGL) 法を適用することを提案する。

2. 遺伝的アルゴリズムの適用

遺伝的アルゴリズムは多点探索による最適値探索を行う。この性質を利用した、階層構造ニューラルネットワークの重みの学習や構造決定が試みられている [2, 3, 4, 5]。

2.1 重み調節への適用

遺伝的アルゴリズムが大域的な探索を行うことから、ニューラルネットワークの重み調節の際に遺伝的アルゴリズムを用いることにより、局所解に陥るのを防ぐことが期待できる。

Montanaらは重みとバイアスの実数値を染色体中の遺伝子に対応させた (図1) [2]。

重みの学習は以下の繰り返しにより行われる。

1. 染色体が表す重みをネットワークにマッピングする。

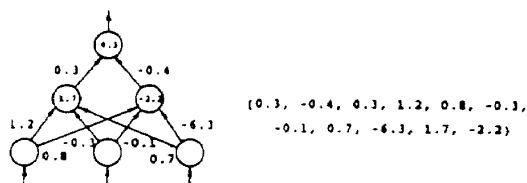


図1: 染色体へのコーディング

2. ネットワークに1エポック分のパターンを提示する。
3. ネットワークの適応度を求め、遺伝的操作を行う。

2.2 構造決定への適用

Millerらはネットワークの構造を結合行列を用いて表現し、結合行列の要素を染色体に直接対応させた (図2) [3]。

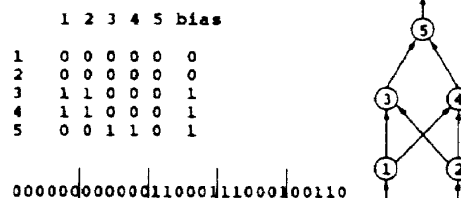


図2: 結合行列による表現

2.3 Neurogenetic Learning 法

北野は遺伝的アルゴリズムの適用において、重みの学習と構造決定を融合させた手法として、Neurogenetic Learning (NGL) 法を提案した [4]。

NGL法では染色体にネットワークの構造を直接表現せず、ネットワークの生成ルールを表現している (文法エンコード法)。染色体は生成ルールを表現した断片 (図3) の集合である。

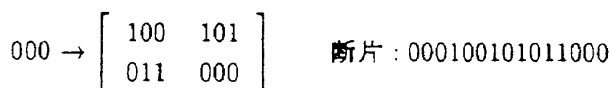


図3: 生成ルールの例とその断片

Training Recurrent Neural Networks using Neurogenetic Learning
 Noritaka HAMAMICHI Fumihiko YAMAGUCHI Masakazu NAKANISHI
 Department of Computer Science, Graduate School of Science and Technology, Keio University 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223-8522, Japan

学習は以下の繰り返しにより行われる。

1. 生成ルールを用いて、染色体から接続行列を得る。
2. 接続行列と CAM 行列から構造と重みの決定を行う。
3. 誤差逆伝播法などにより重みの調節を行う。
4. 適応度を求め、遺伝的操作を行う。

接続行列 (図4) の対角成分はユニットのタイプを表現し、それ以外は結合を表現する。また、CAM (cell adhesion molecule) 行列 (図5) はそれぞれのユニットのタイプからユニット間の重みを決定するために用いられる。

	1	2	3	4	5	6
1	010	001	011	100	011	000
2	001	110	001	000	101	010
3	100	010	001	000	010	010
4	000	001	001	110	000	001
5	001	100	000	000	101	001
6	001	000	001	000	001	010

図4: 接続行列

	000	001	010	011	100	101
000	3.800	3.500	-2.320	3.620	1.480	-2.020
001	1.720	-3.060	-1.840	-3.580	1.320	1.020
010	-0.680	-1.620	-2.960	-3.420	2.440	-1.700
011	3.960	-3.700	3.280	-1.980	-1.880	2.300
100	-3.560	0.300	2.800	1.700	2.440	-0.420
101	3.320	-0.820	-1.520	0.260	0.680	1.020

図5: ランダムに生成された CAM 行列

NGL 法には以下のような特徴がある。

- 文法エンコード法により規則的な結合パターンが生成されやすく、下位構造をコピーしたり移動したりすることができる。
- 重みと構造を一緒に扱えるため、良いネットワークが現れたとき、その構造と重みを一つの組として扱うことができる。

北野は、NGL 法が直接エンコード法に対して収束速度、拡張性ともに優れていることを実験により示している。

3. 本研究の手法

リカレントニューラルネットワークは、ユニット間に相互結合を持つことにより時系列情報を学習させることが可能なネットワークの一つである。

本研究では、リカレントニューラルネットワークの学習に NGL 法を元にした文法エンコード法を導入す

ることを試みる。

NGL 法をリカレントニューラルネットワークに適用するにあたって、以下の二点について変更を加える。

- 接続行列の解釈
- 重みの調節に用いるアルゴリズム

接続行列

階層型ニューラルネットワークに適用する場合は、接続行列の上三角成分しか用いていなかったが、下三角成分も結合の表現とみなして利用する。また、対角成分についても各ユニット自身への結合の表現とみなす。

重みの調節

誤差逆伝播法は、そのままの形ではリカレント型のニューラルネットワークに対しては適用できない。そこで、本手法では BPTT (Back Propagation Through Time) 法 [6] や BPQ (Back propagation based on Partial Quasi-Newton) 法 [7] などを重みの調節に用いることにする。

4. おわりに

現在、提案手法について実装、及び実験を行っている。結果については発表時に報告を行う。

参考文献

- [1] Rumelhart, D. E., Hinton, G. E. and Williams, R. J.: Learning Internal Representation by Error Propagation, in *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, Vol. 1, pp. 318-362, MIT Press (1986).
- [2] Montana, D. J. and Davis, L.: Training Feedforward Neural Networks Using Genetic Algorithms, in *Proc. of the 11th International Conference on Artificial Intelligence*, Vol. 1, pp. 762-767, Morgan Kaufmann (1989).
- [3] Miller, G. F., Todd, P. M. and Hegde, S. U.: Designing Neural Networks Using Genetic Algorithms, in *Proc. of the 3rd International Conference on Genetic Algorithms*, pp. 379-384, Morgan Kaufmann (1989).
- [4] Kitano, H.: Neurogenetic Learning: An Integrated Method of Designing and Training Neural Networks Using Genetic Algorithms, *Physica D*, Vol. 75, pp. 225-238 (1994).
- [5] 高橋裕樹, 中嶋正之: 遺伝的アルゴリズムを用いた最適な階層型神経回路網の構造決定に関する一考察, *電子情報通信学会論文誌 (D-II)*, Vol. J79-D-II, No. 11, pp. 1920-1928 (1996).
- [6] Williams, R. J. and Zipser, D.: A learning algorithm for continually running fully recurrent neural networks, *Neural Computation*, Vol. 9, No. 1, pp. 123-141 (1997).
- [7] 斎藤和巳, 中野良平: 2次学習アルゴリズム BPQ によるリカレントネットワーク学習とガウス混合分布推定, *電子情報通信学会論文誌 (D-II)*, Vol. J81-D-II, No. 3, pp. 538-546 (1998).