

4N-2 学習環境の変化に適応可能な GA による

ニューラルネットワークの構造決定法

○大枝真一 市村匠 寺内睦博 高濱徹行 磯道義典

広島市立大学大学院 情報科学研究科

1 はじめに

BP(Back Propagation)学習, GA(Genetic Algorithms)は広く知られているが, これらの研究のほとんどは静的な実験環境上で評価されているに過ぎず, 動的性質を持つ環境上での議論は乏しいのが現状である. 佐々木らは, この点に着眼し, 動的環境におけるエージェントの進化と学習における考察を行っている^[1]. しかしながら, そこではニューラルネットワーク(Neural Network: NN)の特有の問題である学習パラメタ値の設定, ネットワーク構造などには触れていない. 本論文では, 動的環境でたとえ学習中に教師信号に変化が及んだ場合でも, 初期状態から学習を行うのではなく, ネットワーク構造や結合荷重ベクトル, 学習パラメタ値などの継承を考慮した適応的な学習を可能とする方法を提案する.

2 学習環境の変化に適応可能な GA によるニューラルネットワークの構造決定法

最適化手法の一つである GA を用い, NN の構造を決定する方法が提案されている^[2]. GA により, 隠れニューロン数 n , 結合荷重ベクトル \mathbf{W} , 学習係数 η , シグモイド関数の傾き ε , 慣性係数 α を探索する.

2.1 遺伝子型

図 1 は, 表 1 に示す学習パラメタ値と各層間の結合荷重 w を実数値として持つ染色体を表している.

2.2 遺伝的操作

GA の探索は選択, 交叉, 突然変異の 3 つの操作からなる. 本論文では, 選択はエリート戦略とルーレット選択を用いた. エリート戦略は, 集団中で適合度の高い上位 10% の個体を次世代に残す. 交叉に

は一様交叉を用いた. [2] の突然変異の操作では局所突然変異と大域突然変異の 2 種類を定義しており, それぞれ表 2, 表 3 で与えられる.

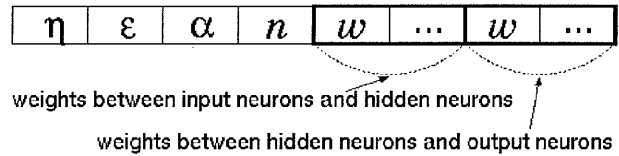


図 1. 遺伝子型

表 1. 各学習パラメタ

η	$0 < \eta \leq 1.0$	$\eta \in \mathbf{R}$
ε	$0 < \varepsilon \leq 1.0$	$\varepsilon \in \mathbf{R}$
α	$0 < \alpha \leq 1.0$	$\alpha \in \mathbf{R}$
n	$2 \leq n \leq 20$	$n \in \mathbf{N}$

表 2. 局所突然変異

$\{\eta, \varepsilon, \alpha\} = \{\eta, \varepsilon, \alpha\} + r_1$	$\{-0.1 \leq r_1 \leq 0.1\}$
$n = n + r_2$	$\{-5 \leq r_2 \leq 5\}$
$w = w + r_3$	$\{-1.0 \leq r_3 \leq 1.0\}$

表 3. 大域突然変異

$\{\eta, \varepsilon, \alpha\} = r_4$	$\{0.0 \leq r_4 \leq 1.0\}$
$n = r_5$	$\{2 \leq r_5 \leq 20\}$
$w = r_6$	$\{-5.0 \leq r_6 \leq 5.0\}$

2.3 遺伝機構

[1] では, ラマルク型とダーウィン型と呼ばれる 2 種類の遺伝機構を提案している. ラマルク型では, 学習により修正された結合荷重を次世代へそのまま遺伝するのに対し, ダーウィン型では, 修正された結合加重を次世代へ伝えるのではなく, 自分の親から受け継いだ遺伝子を次世代の子を作る際の遺伝子として GA の遺伝的操作が行われる. また, 動的環境においてはダーウィン型が適していると報告されている.

2.4 情報量基準 AIC に基づく適合度関数

ここでの GA の探索は, 適したネットワーク構造を探索することに他ならない. したがって, ネットワーク構造の優劣に応じた適合度を割り当てる必要がある. そこで, NN を評価するために, 情報量基準

AICを用いる^[3].

2.4.1 AICによるネットワーク評価^[4]

最急降下法に基づくBP学習を最尤推定とみなすと、学習が収束したネットワークに対する対数尤度を最大対数尤度とみなすことができる。Fをパラメータ数としたとき、AICは式(1)で与えられる。

$$AIC = -2(\text{最大対数尤度}) + 2F \quad (1)$$

2.4.2 適合度関数

NNを式(1)の情報量基準AICに基づいて評価する。AICの値が小さいネットワークほど、適したネットワークといえ、GAで用いる適合度関数を式(2)と定義する。ここで、 λ_{\max} 及び λ_{\min} はAICの最大値と最小値である。また、uは集団の個体番号である。

$$Fitness_u = \frac{\lambda_{\max} - AIC_u}{\lambda_{\max} - \lambda_{\min}} \quad (2)$$

3 計算機実験

学習中に教師信号が変化する動的環境を表4と定義した。入力は{0,1}の6bitからなり、上位3bitで教師信号が決定され、下位3bitは教師信号の決定に作用しない。{*}は{0,1}のどちらでもよい。環境はAからEの順に20世代毎に変動する。

表4. 動的環境

Input Signal \ Teaching Signal	Teaching Signal							
	0	0	0	0	1	1	1	1
0	0	0	1	1	0	0	1	1
0	1	0	1	0	1	0	1	0
*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*
Environment A	0	0	0	0	1	1	1	1
Environment B	0	0	0	1	0	1	1	1
Environment C	0	0	1	0	0	1	1	1
Environment D	0	1	0	0	0	1	1	1
Environment E	1	1	1	1	0	0	0	0

実験では世代数を100世代、集団数を20個体、各個体の学習回数を500回とした。また、交叉率 $P_c = 0.5$ 、局所突然変異率 $P_{ml} = 0.025$ 、大域突然変異率 $P_{mg} = 0.025$ とした。 λ_{\max} 及び λ_{\min} は、前実験によりそれぞれ2000、-6000と決定した。遺伝機構には、ダーウィン型を用いた。

図2は、学習前と学習後における全個体の二乗和誤差の平均を示している。NNのみの場合、環境変動時に振動し、収束しないことがあった。しかし、提

案する手法では大きく振動することなく、全世代を通じて学習後の二乗和誤差の平均が低いことがわかる。表5は初期世代、終了世代での全個体の表現型の平均を示している。提案する手法により、終了世代での隠れニューロン数が平均2.25個の小さいネットワークを獲得した。

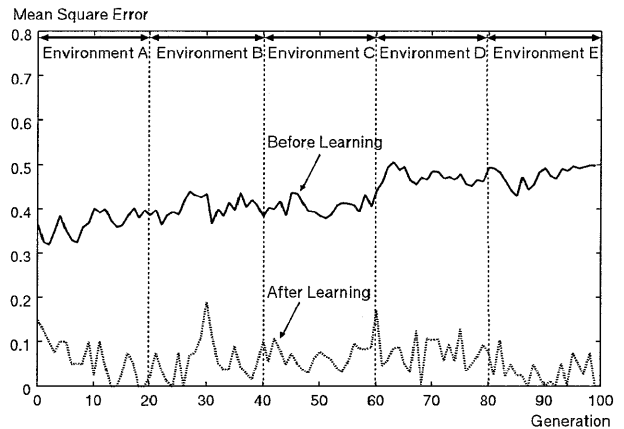


図2. 学習前後の二乗和誤差の平均

表5. 獲得した表現型の平均

	$\bar{\eta}$	$\bar{\varepsilon}$	$\bar{\alpha}$	\bar{n}	\overline{MSE}
初期世代	0.449881	0.470177	0.515680	10.2	0.150110
終了世代	0.035863	0.858270	0.992187	2.25	0.000030

4 まとめ

提案した手法が、環境に適したネットワーク構造を獲得するとともに、動的環境下の学習に有効であることが明らかになった。

参考文献

- [1] Sasaki, T. and Tokoro, M., "Adaptation toward Changing Environments: Why Darwinian in Nature?", Proc. of the 4th European Conference on Artificial Life (ECAL-97), pp. 145-153(1997)
- [2] Takahashi, H. and Nakajima, M., "Evolutional Design and Training Algorithm for Feedforward Neural Networks", IEICE Trans. information & systems, Vol.E82-D, No.10, pp.1384-1392(1999)
- [3] Akaike, H., "A new look at the statistical model identification", IEEE Trans. on Automatic Control, Vol.AC-19, No.6, pp.716-723(1974)
- [4] 栗田多喜夫, 本村陽一, "階層型ニューラルネットとその周辺", 応用統計学, Vol.22, No.3, pp.99-115(1993)
- [5] Oeda, S., Ichimura, T., Terauchi, M., Takahama, T. and Isomichi, Y., "Adaptive Evolutional Learning Method of Neural Networks using Genetic Algorithms under Dynamic Environments", Proc. of Knowledge-Based Intelligent Engineering Systems & Allied Technologies (KES'2000)(2000)(now printing)