ブランチ制御を考慮したパラメータ可変一般化学習ネットワーク

正員 平 澤 宏太郎 (九州大学)

非会員 衛 藤 広 宣 (九州大学)

正員 胡 敬 炉 (九州大学)

正員 村 田 純 一 (九州大学)

非会員 熊 慶 宇 (九州大学)

Universal Learning Networks with Varying Parameters Considering Branch Control

Kotaro Hirasawa, Member, Hironobu Eto, Non-member, Jinglu Hu, Member, Junichi Murata, Member, Qingyu Xiong, Non-member (Kyushu University)

Universal Learning Network (ULN) which is a super set of supervised learning networks has been already proposed. Parameters of ULNs are trained in order to optimize a criterion function as conventional neural networks, and after training they are fixed as constant parameters. In this paper a new method to alter the parameters, therefore in a special case, to control the branch connection depending on the network flows is presented to enhance flexibility of the networks. In the proposed method, there exist two kinds of networks, the first one is a basic network which includes varying parameters and the other one is a network which calculates the optimal varying parameters, therefore decides the branch connection in a special case depending on the network flows of the basic network.

From simulations where parameters of a neural network are altered and branch connection in the neural network is determined by a fuzzy inference network, it is shown that the proposed network has higher representation abilities than the conventional networks.

キーワード : neural networks, fuzzy neural networks, learning, varying parameters, branch control, flexibility

1. はじめに

筆者等は教師付きニューラルネットワークを静的・動的構成、前向き・後向き学習、高次微分計算、マルチブランチ、任意の非線形関数、任意の時間遅れ等で一般化した一般化学習ネットワークをすでに提案してきた⁽¹⁾⁽²⁾。一般化学習ネットワークでは通常のニューラルネットワークと同様、非線形関数内のパラメータを学習により求め、適正な値に固定している⁽³⁾。本論文では、これ等パラメータを固定するのではなく、別のネットワークによって計算された状況に応じた適切な値に設定し、かつ、その特別な場合としてネットワークのブランチの接続・非接続を状況に応じて変更するブランチ制御を考慮したパラメータ可変一般化学習ネットワークを提案する。

ブランチ制御を考慮したパラメータ可変一般化学習ネットワークでは、基本のネットワークとパラメータ計算用のネットワークが対になって存在し、パラメータ計算用のネットワークが基本ネットワークのパラメータの値を信号の伝

播に応じて適応的に計算する事により、ネットワークの柔軟性を向上させることを目指している。

又、パラメータ計算の特別の場合として信号の伝播に応じて基本ネットワークのブランチの接続・非接続を制御する機能分散型構成が可能となり、このことによりさらなるネットワークの柔軟性の向上を目指している。

ネットワークの非線形写像の柔軟性を向上させるためのアーキテクチャーとして、従来高次ネットワーク⁽⁴⁾⁽⁵⁾、Group Method of Data Handling 法 (GMDH)⁽⁶⁾⁽⁷⁾、モジュラーネット⁽⁸⁾⁽⁹⁾および筆者等により Learning Petri Network (LPN)⁽¹⁰⁾等が提案されている。

高次ネットは主に画像データの認識を対象に入力変数の積を考慮した非線形変換を行い、画像の移動、回転、拡大、縮小に対して強いシステムを構築している。GMDH は入力変数の可能な組み合わせの中から最適な組み合わせを自己組織的に構成し、モジュラーネットでは簡単な構造を持つエキスパートネットと呼ばれる複数のモジュールの出力を可変重みで結合する事によりネットワークの柔軟性向上を

図っている。

又, LPNではトークン信号の流れの制御によりネットワークの機能局在を実現しこれにより表現能力の優れたネットワークを構築している。上記のシステムはいずれも, システムの目的に合った非線形アーキテクチャの実現に独自の工夫を凝らしている。

本論文で述べるブランチ制御を考慮したパラメータ可変一般化学習ネットワークも同様ネットワークの表現能力の向上のためのアーキテクチャの工夫に関するものであるが, 具体的には, 入力の変化等の環境の認識をファジィ推論ネットワークで行い, その結果によりニューラルネットワークのパラメータの変更とその特別なケースとしてブランチの接続を可変にするアーキテクチャを採用している。従って, 前述の基本ネットワークがニューラルネットワークであり, パラメータ計算用ネットワークがファジィ推論ネットワークとなる。

ブランチ制御を考慮したパラメータ可変一般化学習ネットワークのねらいは, 通常のニューラルネットワークより多くの種類の非線形関数の写像が容易に実現出来るようにアーキテクチャの柔軟性を向上させる事であるが, そのためには, パラメータを可変にしブランチ制御を行うネットワークにファジィ推論ネットワークを使用して, ネットワークの構造にある種の見聞情報を導入しやすくしている点が重要である。換言すると, ブランチ制御を考慮したパラメータ可変一般化学習ネットワークでは通常のニューラルネットワークのように規模(学習パラメータの数)を大きくする事によって表現能力を向上させるのではなく, パラメータの計算とブランチを制御することによりネットワークの規模を適正に保ちつつネットワークの表現力の柔軟性の向上を図っている。

又, 提案するネットワークではメンバーシップ関数, シグモイド関数, 掛算等の各種の非線形関数を用いているため, 任意の非線形関数の取扱いが統一的に可能な一般化学習ネットワークの枠組でシステムを表現している。

学習については, 勾配法を活用する事も可能であるが, 本論文ではブランチの接続・非接続といった非連続システムをも取り扱っているためランダム探索法の一つである RasID 学習法⁽¹¹⁾を使用している。文献[11]では, ネットワークの学習において, RasID が BP 等の勾配法より同等またはそれ以上の高い学習効率を示す事が記載されている。

本論文の構成は, 次のとおりである。第2章で, 提案するブランチ制御を考慮したパラメータ可変一般化学習ネットワークの基本構成と学習方式について述べる。第3章はシミュレーション結果とその考察, 第4章はまとめである。

2. ブランチ制御を考慮したパラメータ可変一般化学習ネットワーク

〈2・1〉 構 成 ブランチ制御を考慮したパラメータ可変一般化学習ネットワークは, 基本ネットワークとパラメータ計算用ネットワークより構成され(図1を参照), そ

図1 ブランチ制御を考慮したパラメータ可変一般化学習ネットワークの構造

Fig.1. Structure of Universal Learning Networks with Varying Parameters Considering Branch Control

れ等が一般化学習ネットワーク⁽¹⁾⁽²⁾で表現されている。一般化学習ネットワークは, 連続微分可能な任意の非線形ノードが任意のサンプル遅れ時間を介して任意にマルチ接続された離散時間型のネットワークで, 静的および動的ネットワークを統一的に扱う, バックワード伝搬学習とフォワード伝搬学習を体系的に行う, 一次微分のみならず高次微分の計算が可能等の特徴を有している。今回はその中でも特に任意の非線形関数の取扱いが統一的に出来る点を考慮し一般化学習ネットワークの枠組を活用している。

文献(1)(2)によると, 一般化学習ネットワーク内の j ノードの t サンプル時刻の出力 $h_j(t)$ は(1)式で表される。

$$h_j(t) = f_j(\{h_i(t - \tau_{ij}(p)) | i \in J_F(j), p \in B(i, j)\}, \{r_n(t) | n \in N(j)\}, \{\lambda_m | m \in M(j)\}) \quad (1)$$

$$j \in J, \quad t \in T$$

- $h_j(t)$: j ノードの t 時刻の出力値
- $r_n(t)$: n 番目の外部入力変数の t 時刻の値
- λ_m : m 番目のパラメータの値
- f_j : j ノードの非線形関数
- $\tau_{ij}(p)$: i ノードから j ノードへの p 番目の
ブランチの遅れ時間
- J : ノード番号の集合
- $J_F(j)$: j ノードへ接続するノードの番号の集合
- $J_B(j)$: j ノードから接続するノードの番号の集合
- $N(j)$: j ノードへ接続する外部入力変数の
番号の集合
- $M(j)$: j ノードに含まれるパラメータ
の番号の集合
- $B(i, j)$: i ノードから j ノードへ接続する
ブランチの番号の集合
- T : サンプル時刻の集合

従来の一般化学習ネットワークでは, 評価指標が最小となるように学習によって最適なパラメータ λ_m を決定して

いるが、ブランチ制御を考慮したパラメータ可変一般化学習ネットワークでは、基本ネットワークであるニューラルネットワークのパラメータ λ_m をパラメータ計算用ネットワークであるファジィ推論ネットワークが計算する。

従って、パラメータ λ_m は入力状況に応じて可変となる。換言すると、ブランチ制御を考慮したパラメータ可変一般化ネットワークでは λ_m を計算するためのファジィ推論ネットワークのパラメータ μ_q をニューラルネットワークの評価指標が最小となるように学習し、その結果、ファジィ推論ネットワークによって環境に応じた適切な λ_m が自動的に計算される事になる。

又、ファジィ推論ネットワークの後件部の平均値を表すパラメータを強制的に零に設定する事により、特定の入力に対してニューラルネットワークのパラメータ λ_m を零にする事が出来る。従って、ニューラルネットワークのブランチの接続・非接続の制御がネットワークの入力に依存して可能となる。

なお、ニューラルネットワークの各パラメータとファジィ推論ネットワークの各出力ノードは 1 対 1 に対応しており、ファジィ推論ネットワークの入力はニューラルネットワークの入力と同じとする。

〈2・2〉学 習 ファジィ推論ネットワークのパラメータの学習は勾配法ではなくランダム探索法の一種である RasID(Random search with Intensification and Diversification)⁽¹¹⁾で行う。

一般化学習ネットワークの勾配法で勿論ファジィ推論ネットワークのパラメータを学習することも可能であるが、ブランチ制御を考慮したパラメータ可変一般化学習ネットワークではニューラルネットワークとファジィ推論ネットワークが混在し、又、ニューラルネットワークのパラメータがファジィ推論ネットワークにより計算されるため掛算処理のノードが存在し、勾配法の計算がやや複雑になる。一方、各種のニューラルネットワークの学習に RasID が適切である例が数多く報告されており⁽¹¹⁾、従って、本論文ではファジィ推論ネットワークのパラメータの学習に RasID を適用している。

〈2・3〉ニューラルネットワークとファジィ推論ネットワーク 前述のようにブランチ制御を考慮したパラメータ可変一般化学習ネットワークでは基本ネットワークとしてニューラルネットワーク、パラメータ計算用ネットワークとしてファジィ推論ネットワークを使用した。

以下本論文で使用した多次元ガウス関数を利用したファジィ推論ネットワークおよびニューラルネットワークの記号について説明する。

今、 $\mathbf{x} = (\cdots, x_u, \cdots)$ 、 y を上記ネットワークの入力ベクトルおよび出力とする。

● ファジィ推論ネットワーク

ファジィ推論ネットワークでは下記の Q 個のファジィルールを与える。

if $\mathbf{x}$ is X_q then y is Y_q $q \in Q$

ここで：

X_q : メンバーシップ関数 $f_{xq}(\mathbf{x})$ を持つ
ファジィ集合

Y_q : メンバーシップ関数 $f_{yq}(y)$ を持つ
ファジィ集合

であり、 $f_{xq}(\mathbf{x})$ および $f_{yq}(y)$ は次の多次元および一次元のガウス関数である。

$$f_{xq}(\mathbf{x}) = \exp \left[-\frac{(\mathbf{x} - \boldsymbol{\mu}_{xq})^T \boldsymbol{\Sigma}_{xq}^{-1} (\mathbf{x} - \boldsymbol{\mu}_{xq})}{2} \right]$$

$$f_{yq}(y) = \exp \left[-\frac{(y - \mu_{yq})^2}{2\sigma_{yq}^2} \right]$$

$$\boldsymbol{\mu}_{xq} = (\cdots, \mu_{xqu}, \cdots)^T$$

$$\boldsymbol{\Sigma}_{xq}^{-1} = [\eta_{xquv}]$$

$\boldsymbol{\mu}_{xq}$: ファジィルール前件部の平均を表す
パラメータベクトル

$\boldsymbol{\Sigma}_{xq}^{-1}$: ファジィルール前件部のばらつきを表す
パラメータ行列

μ_{yq} : ファジィルール後件部の平均を表す
パラメータ

σ_{yq}^2 : ファジィルール後件部のばらつきを表す
パラメータ

Q : ファジィルールの番号の集合

従って、非ファジィ化方法として重心法を考えるとファジィ推論ネットワークの出力 y は (2) 式のようになる。

$$y = \frac{\sum_{q \in Q} [f_{xq}(\mathbf{x}) \sigma_{yq} \mu_{yq}]}{\sum_{q \in Q} [f_{xq}(\mathbf{x}) \sigma_{yq}]} \quad \cdots \cdots \cdots (2)$$

● ニューラルネットワーク

$$y = \sum_{m \in M} f_m(\mathbf{x}) \lambda_m + \theta \quad \cdots \cdots \cdots (3)$$

$$f_m(\mathbf{x}) = \frac{1 - e^{-\varphi_m(\mathbf{w}_m^T \mathbf{x} + \theta_m)}}{1 + e^{-\varphi_m(\mathbf{w}_m^T \mathbf{x} + \theta_m)}}$$

$$\mathbf{w}_m = (\cdots, w_{mu}, \cdots)^T$$

$f_m(\mathbf{x})$: シグモイド関数

$\mathbf{w}_m, \varphi_m, \theta_m$: シグモイド関数のパラメータ

λ_m, θ : 出力を合成するパラメータ

M : ニューラルネットワークの
中間層のノード番号の集合

図2 ブランチ制御を考慮したパラメータ可変一般化学習ネットワークの具体例
Fig.2. An Example of Universal Learning Networks with Varying Parameters Considering Branch Control

(2)(3) 式を比較すると、ファジィ推論ネットワークとニューラルネットワークの相異はノード関数が局所的であるかどうかの違いであり、又、ファジィルールの数 Q 、ニューラルネットワークの中間層のノードの数 M がネットワークの複雑さを表す指標となっている ((2)(3) 式の定義によると Q 、 M は本来集合であるが、便宜的に $|Q|$ 、 $|M|$ を Q 、 M で表現する場合がある。)

なお、ファジィ推論ネットワークでは前件部のパラメータ μ_{xq} 、 Σ_{xq}^{-1} と後件部のひろがりを表すパラメータ σ_{yq}^2 は先見情報により固定し、後件部の平均を表すパラメータ μ_{yq} のみを学習パラメータとした。

具体的には、(2) 式で計算された y が (3) 式の λ_m となり、学習の対象となるパラメータは、ニューラルネットワークの w_m 、 θ_m および θ とファジィ推論ネットワークの μ_{yq} である。

〈2・4〉 ブランチ制御のアーキテクチャ ブランチ制御を考慮したパラメータ可変一般化学習ネットワークのブランチ制御の機構を本節で述べる。

ファジィ推論ネットワークの後件部のメンバーシップ関数の平均値 μ_{yq} の 1 部を零に固定する。固定とは学習をしないという意味である。

今ニューラルネットワークの特定のパラメータ λ_m に対

図3 前件部のメンバーシップ関数
Fig.3. Membership Functions of IF Part

応するファジィ推論ネットワークの出力を考え、この出力の後件部のメンバーシップ関数の平均値を μ_{yq} 、 $q \in Q$ とする。この Q 個の平均値のうち特定の q_0 番目の平均値 μ_{yq_0} を零に設定すると、前件部のメンバーシップ関数 $f_{xq}(x)$ がローカル性を持つガウス分布であるため、

$$X = \{x | f_{xq_0}(x) = \text{大 and } f_{xq}(x) = \text{小 } q \neq q_0\} \quad (4)$$

となる X ではファジィ推論ネットワークの出力は非常に小さな値を持つ事になり、従ってパラメータ $\lambda_m \cong 0$ となって、ニューラルネットワークの m 番目のブランチが接断される事になる。

このように、1 部のファジィルールの後件部の平均値パラメータを零に固定する事により、状況に応じてニューラルネットワークのブランチの接続・非接続の制御が可能な機能分散型のネットワークを構成する事が出来る。

3. シミュレーションと考察

ブランチ制御を考慮したパラメータ可変一般化学習ネットワークについて 2 次元入力関数近似の表現能力の柔軟性の検証を行った。シミュレーションでは、ネットワーク全体の学習パラメータの数を同じにして、つまりネットワークの規模を同じにして、通常のニューラルネットワークとブランチ制御を考慮したパラメータ可変一般化学習ネットワークの表現能力を比較をした。今回のシミュレーションではニューラルネットワークのパラメータ w_m 、 θ_m 及び θ と、ファジィ推論ネットワークの後件部のメンバーシップ関数の平均パラメータ μ_{yq} を学習させた。従って、学習パラメータの総数は、 $M(3+Q)+1$ となる (図2参照)。ファジィルールの前件部のパラメータについては、平均値パラメータ μ_{xq} が入力空間に一樣に分布するように、また、ばらつきパラメータは各ファジィルールの前件部のメンバーシップ関数が適切な重なりを持つように設定した (図3参照)。

教師関数 1

教師関数 2

図4 教師関数
Fig.4. Teaching Functions

学習方法は全てのシミュレーションにおいて、ランダム

探索法の一つである, RasID(Random search with Intensification and Diversification) 法⁽¹¹⁾ を用いた。

関数近似に用いた教師関数 1, 教師関数 2 を図 4 に示す。ネットワークに与えた教師信号は, x 軸に 17 点, y 軸に 17 点の組合せである 17×17 の 289 点を与えた。

〈3・1〉 ブランチ制御無しのパラメータ可変一般化学習ネットワークの検証 この節では, ブランチ制御無しのパラメータ可変一般化学習ネットワークの検証を行う。

表 1 教師関数 1 の M, Q の組合せの違いによる比較

Table 1. Comparison of MSE of Teaching Function 1 with Various Combinations of M and Q

M	Q	MSE[10^{-4}]
4	12	14.05
5	9	6.23
6	7	18.11
10	3	39.42
12	2	42.48
15	1	77.03

(M, Q)=(15, 1)

(M, Q)=(5, 9)

図 5 教師関数 1 の学習結果

Fig. 5. Realized Functions of Teaching Function 1

〈3・1・1〉 シミュレーション結果 (教師関数 1) 教師関数 1(図 4) について, 学習パラメータの総数が 61 となる組合せ, $(M, Q)=(4, 12), (5, 9), (6, 7), (10, 3), (12, 2), (15, 1)$ の 6 通りについて, 学習パラメータの初期値を範囲 $[-1:1]$ の一様乱数系列により 5 通り変えてシミュレーションを行った。M, Q はそれぞれ, ニューラルネットワークの中間層のノードの数, ファジィ推論ネットワークのルールの数である。組合せ $(15, 1)$ が通常の中間層が 1 層のニューラルネットワークに相当する。表 1 に学習後の 5 個の乱数系列に対する平均 2 乗誤差 MSE(Mean Square Error)

を示す。

また, 通常のニューラルネットワークである $(M, Q)=(15, 1)$ と表現能力が一番高い $(M, Q)=(5, 9)$ について, 学習後の出力結果を図 5 に示す。なお, 図 5 は 5 個の乱数系列の中で MSE が最小の値を示す図形である。

表 1 を見て分る通り, 通常のニューラルネットワークである, $(M, Q)=(15, 1)$ からファジィルールの数が増加していくにつれて, 表現能力が向上していることがわかる。MSE が最小となる $(M, Q)=(5, 9)$ の場合は, 表現能力が通常のニューラルネットワークの約 12 倍になっている。また, $(M, Q)=(4, 12)$ の場合は $(5, 9)$ の場合より悪くなっている。つまり, ニューラルネットワークの中間層のノード数 M とファジィルールの数 Q の組合せには最適な組合せが存在する。

〈3・1・2〉 シミュレーション結果 (教師関数 2)

教師関数 2 (図 4) について, 前回と同様学習パラメータの総数が 61 となる組合せ, $(M, Q)=(4, 12), (5, 9), (6, 7), (10, 3), (12, 2), (15, 1)$ について, 学習パラメータの初期値を範囲 $[-1:1]$ の一様乱数系列により 5 通り変えてシミュレーションを行った。表 2 に学習後の 5 個の乱数系列に対する平均 2 乗誤差 MSE(Mean Square Error) を示す。

また, 通常のニューラルネットワークである $(M, Q)=(15, 1)$ と表現能力が一番高い組合せ $(M, Q)=(6, 7)$ について, 学習後の出力結果を図 6 に示す。なお, 前述のシミュレーションと同様, 図 6 は 5 個の乱数系列の中で MSE が最小の値を示す図形である。

表 2 を見ると, 組合せ $(M, Q)=(6, 7), (10, 3), (12, 2)$ において, ファジィルールの数が増加していくにつれて, 表現能力が向上していることがわかる。そして, $(M, Q)=(6, 7)$ のところで表現能力が一番高くなっている。つまり, 前述のシミュレーションと同様ニューラルネットワークの中間層の数 M とファジィルールの数 Q のいずれかが極端に小さくない組合せ $(M, Q)=(6, 7)$ で表現能力が一番高くなっている。

なお, 教師関数 1 の場合に通常のニューラルネットワークに対応する $(M, Q)=(15, 1)$ と $(M, Q)=(5, 9)$ との間で 10 倍強, 又, 教師関数 2 の場合には, $(M, Q)=(15, 1)$ と $(M, Q)=(6, 7)$ との間で 6 倍強の MSE の差が生じた理由は, 比較的複雑な教師関数 1 および 2 の学習にとっては, 規模が小さすぎる $M=15$ のニューラルネットワークを基準にパラメータ可変一般化学習ネットワークの評価を行ったためである。

換言すると, 通常のニューラルネットワークでは規模が小さく学習が進まぬ場合でも同規模のパラメータ可変構造のネットワークでは学習が可能な事を示し得る。

学習は勾配法ではなく RasID を用いているが, 文献 [11] に記載されているように RasID は勾配法以上の学習能力を備えている。従って, 学習法の違いによりパラメータ固定構造と可変構造の差が生じた訳ではない。

ルールにはほぼ均等に割り当てている。

なお、ブランチ制御を行うためのファジィ推論ネットワークの後件部の平均値パラメータをランダムに選定しても、通常のニューラルネットワークより良好な表現能力が得られる事が報告されているが⁽¹³⁾、どのパラメータを零にするのかを学習により見出すのは今後の課題である。

〈3・2・1〉 シミュレーション結果 (教師関数 1) 教師関数 1 について、 $(M, Q)=(4, 12), (5, 9), (6, 7), (10, 3), (12, 2)$ について、それぞれブランチ制御を考慮したシミュレーションを行った。具体的には、零に固定するファジィルール後件部の平均パラメータの数を増して (従って、パラメータ数一定の条件下ではファジィルール数 Q が増加する) ニューラルネットワークの中間層から出力層への接続率を変更したシミュレーションを行った。

表 3 を見ると、ニューラルネットワークの中間層のノードの数 M が 4, 5 と小さいとき、すなわちこのときはファジィ推論ネットワークのルール数 Q が大きいときであるが、そのときは、ブランチ制御の効果が見られないが、 M が大きく、すなわち、 Q が小さくなるとブランチ制御の効果があることが分かる。これは、ニューラルネットワークの中間層のノードの数 M が小さいと、それ以上ブランチの数を減少させる機能分散型構成に無理があると考えられるからである。

なお、表 3 の MSE は乱数系列を変えた 5 回のシミュレーションの最小値である。

〈3・2・2〉 シミュレーション結果 (教師関数 2) 教師関数 2 について、教師関数 1 と同様にシミュレーションを行った。

表 4 を見ると、ニューラルネットワークの中間層のノードの数 M が 10 以上になるとブランチ制御の効果が顕著になる事が分かる。これは教師関数 1 の場合と同様の傾向である。

表 3 および表 4 より、ブランチ制御の効果を確実に期待出来るのは、換言すると、ニューラルネットワークの中間層から出力層のブランチの接続率を小さくすればするほど MSE の減少が期待出来るのはニューラルネットワークの中間層のノード数 M が 10 以上の場合である事が分かる。

しかし、例えば表 3 の教師関数 1 の場合 $(M, Q)=(6, 9)$ で、従って、接続率 78% でシミュレーション中最も MSE の小さい結果が得られている。

このことは、全体的にブランチ制御の効果のあるニューラルネットワークの規模の周辺で、最も小さい MSE の期待出来る (M, Q) の組み合わせが存在する訳ではない事を示している。しかし、いずれにせよ

表 3 の $(M, Q)=(6, 9)$ 、接続率 78%、 $MSE=6.12 \times 10^{-4}$

表 4 の $(M, Q)=(5, 17)$ 、接続率 53%、 $MSE=17.95 \times 10^{-4}$ のケースにおいてブランチ制御を行うパラメータ可変一般化学習ネットワークが最適な MSE を得ている。

4. ま と め

本論文では、ネットワークの表現能力の向上を目指した、ブランチ制御を考慮したパラメータ可変一般化学習ネットワークを提案した。

ブランチ制御を考慮したパラメータ可変一般化学習ネットワークでは、ネットワークの入力あるいは信号の伝播状況に応じて適応的に学習ネットワークのパラメータを変更し、その特別な場合としてネットワークのブランチの接続・非接続を制御する点に特徴がある。

シミュレーションの結果より、ブランチ制御を考慮したパラメータ可変一般化学習ネットワークは、基本ネットワークとパラメータ計算用ネットワークの規模のバランスが適切などに表現能力が高くなり、さらに、基本ネットワークの規模が大きいときには、ブランチ制御を考慮すると表現能力の点でさらなる効果が得られる事が明らかとなった。今回のシミュレーションでは、基本ネットワークとしてニューラルネットワークを、パラメータ計算用ネットワークとしてファジィ推論ネットワークを使用し、特定のパラメータのみを可変パラメータとしたが、今後、基本ネットワークとパラメータ計算用ネットワークにラジアル基底関数ネットワーク等も含めた各種のネットワークを適用し、かつ適切な可変パラメータをどのように選択するか等についても検討していく予定である。

(平成 12 年 03 月 28 日受付、同 12 年 08 月 07 日再受付)

文 献

- (1) K. Hirasawa, J. Murata, J. Hu and C. Jin: Chaos Control on Universal Learning Networks, IEEE Trans. on Systems, Man, and Cybernetics-PART C, Vol.30, No.1, pp.95-104, January(2000)
- (2) K. Hirasawa, X. Wang, J. Murata, J. Hu and C. Jin: Universal Learning Networks and its Application to Chaos Control, Neural Networks, Vol. 13, No.2, pp.239-253, March(2000)
- (3) 西川, 北村: ニューラルネットと計測制御, 朝倉書店 (1995)
- (4) L. Spirkovska and M. B. Reid: Connectivity Strategies for Higher Order Neural Networks Applied to Pattern Recognitions, Proc. of IJCNN 90, Vol.1, pp.21-26 (1990)
- (5) L. Spirkovska and M. B. Reid: Coarse-Coded Higher Order Neural Networks for PSRI Object Recognition, IEEE Trans. Neural Networks, Vol.4, No.2, pp.276-283 (1993)
- (6) A. G. Ivakhnenko: The group method of data handling, A rival of the method of stochastic approximation, Soviet Automatic Control, Vol.13, pp.43-55 (1968)
- (7) A.G. Ivakhnenko: Polynomial theory of Complex system, IEEE Trans., Vol. SMC-1, pp. 364-378 (1971)
- (8) R.A. Jacobs and M.I. Jordan: Learning piecewise control strategy in a modular neural network architecture, IEEE Trans. on System, Man and Cybernetics, Vol. 23, pp. 337-345 (1993)
- (9) M.I. Jordan and R.A. Jacobs: Hierarchical mixture of experts and the EM algorithm, Neural Computation, vol. 6, pp.181-214 (1994)
- (10) K. Hirasawa, M. Ohbayashi, S. Sakai and J. Hu: Learning Petri Network and Its Application to Nonlinear System Control, IEEE Trans. on System, Man and Cybernetics, PART B, Vol. 28, No. 6, pp. 781-789 (1998)
- (11) J. Hu, K. Hirasawa, and J. Murata: RasID-Random Search

for Neural Network Training, Journal of Advanced Computational Intelligence, Vol.2, No.4, pp.134-139 (1998)

- (12) K. Hirasawa, J. Hu, J. Murata, and C. Jin: Universal Learning Networks with Varying Parameters, Proc. of IJCNN (1999).
- (13) K. Hirasawa, J. Hu, Q. Xiong, J. Murata, and Y. Shiraishi: Universal Learning Networks with Branch Control, Proc. of IJCNN, pp. 97-102(2000).

熊 慶 宇 (非会員) 1965 年 10 月 3 日生。1991 年中国重慶大学大学院修士課程修了。1999 年 4 月九州大学大学院入学。現在システム情報科学研究科電気電子システム工学専攻博士後期課程在学中。

平 澤 宏太郎 (正員) 1942 年 1 月 26 日生。1966 年 3 月九州大学大学院工学研究科電気工学専攻修士課程修了。同年 4 月 (株) 日立製作所入社日立研究所勤務。1991 年 8 月大みか工場主管技師長。1992 年 12 月九州大学システム情報科学研究科教授。現在に至る。電気学会, 計測自動制御学会, 情報処理学会, IEEE の各会員。工学博士。

衛 藤 広 宣 (非会員) 1974 年 1 月 22 日生。1998 年 3 月愛媛大学工学部電気電子工学科卒業。1998 年 4 月九州大学大学院入学。2000 年 3 月同大学大学院システム情報科学研究科電気電子システム工学専攻修士課程修了。同年 4 月日立プロセスコンピュータエンジニアリング (株) 入社。現在に至る。

胡 敬 炉 (正員) 1962 年 4 月 4 日生。1986 年中国中山大学大学院修士課程修了。同年 7 月同大学電子工学科助手。1988 年 11 月同講師。1993 年 10 月渡日。1997 年九州工業大学情報工学研究科博士後期課程修了。同年 4 月九州大学ベンチャー・システムラボラトリ非常勤研究員を経て, 同年 8 月同大学大学院システム情報科学研究科助手。現在に至る。主として学習ネットワーク理論, システム

同定とその応用に関する研究に従事。電気学会, 計測自動制御学会の各会員。情報工学博士。

村 田 純 一 (正員) 1959 年 1 月 19 日生。1986 年 3 月九州大学大学院工学研究科博士後期課程電気工学専攻修了。1986 年 4 月九州大学工学部助手。1988 年 4 月同大学工学部助教授。現在に至る。電気学会, 計測自動制御学会, システム制御情報学会, IEEE の各会員。主としてシステムのモデル化, 電力負荷予測に関する研究に従事。工学博士。