リカレントニューラルネットワークの創発的学習手法†

喜多川 健*・本間 経康**・阿部 健***

An Emergent Learning Method for Recurrent Neural Networks

Ken KITAGAWA*, Noriyasu HONMA** and Kenichi ABE***

We propose a learning method for recurrent neural networks with dynamics. The core of this method is to keep a complexity of the network dynamics in the vicinity of the edge of chaos. To investigate the properties of the dynamics effectively and explicitly, novel stochastic parameters defined by combinations of the standard parameters such as individual connection strengths and thresholds are introduced, and then relations between complexities of the dynamics and the stochastic parameters are revealed. The standard parameters are changed by the core part based on the relations and also according to the global error measure. Some examples suggest that the method is practical one for temporal supervised learning tasks and therefore the dynamics of the edge of chaos are effectual for learning of the recurrent networks.

Key Words: recurrent neural networks, chaos, complex system, learning

1. はじめに

近年、従来のトップダウン的なシステム設計手法にかわる新たな手法を目指して、人工生命や創発システム⁸⁾などのボトムアップ的な手法が研究されている。創発システムで基本となるのは、システムを構成する要素間あるいは外部環境との間で、複雑で動的な相互作用（大域的な秩序形成とそれによる拘束）が存在することである。このような相互作用による複雑な振舞いは、複雑系（複雑適応系）と呼ばれるシステムの特徴的な振舞いとして報告され注目されてきた¹⁴⁾。

複雑系の一つであるフィードバック結合をもつリカレントニューラルネットワーク（**RNN**）は、構成ニューロン数が等しいフィードフォワード（階層型）ニューラルネットワーク（**FNN**）に比べシナプス結合の総数が多いため自由度が高く、その複雑な相互作用が生み出す複雑なダイナミクスを利用した動的システムの同定問題、制御問題などへの応用が期待されている。しかし、パラメータ自由度の増加は、ネットワークの表現能力の向上をもたらすが、冗長な自由度は汎

化および収束速度などの点で学習に悪影響を及ぼすことが知られている¹¹⁾。一方、従来の**RNN**に対する学習アルゴリズム^{12), 13), 16)}は、**FNN**の学習に対して有効なバックプロパゲーション（BP）アルゴリズムに基づいているため、高い自由度、とくに冗長な自由度が生み出すニューロン間の相互作用に起因する複雑なダイナミクスが十分に考慮されているとはいえず、計算に必要なメモリ量、計算コストおよび収束速度などの問題点が指摘されている²⁾。したがって、実際にはネットワーク構造を限定する（例えば、あるニューロン間にはシナプス結合をもたせない）など、なんらかの制限条件を用いる場合が多い。しかし、このような条件は、**RNN**の利点である高い自由度をあらかじめ制限することになり、また、どのような制限条件が良いかの一般的な基準もない¹¹⁾。

Unnikrishnan¹⁵⁾らは、Alopexと呼ばれる階層型とリカレント型の両方に適用できる学習アルゴリズムを提案している。このアルゴリズムは、ネットワークがダイナミクスをもつかもたないかに関係なく学習可能であるが、シナプス結合重みの変化量を統計的な誤差から計算するため、計算量および収束速度の点で問題が残っている。

著者らは先に、二値の複雑系のダイナミクスについて検討し、ホロンネットワークの自己組織化における複雑系の潜在的な能力について報告した³⁾。つまり、複雑系ではパラメータを調整することで秩序的振舞いから無秩序的振舞いまで種々のダイナミクスが観測され、秩序と無秩序との境界（edge of chaos）において、物性などの分野での相転移に相当するような現象が観られる。相転移の付近に存在するダイナミクスは多様性に富み、その重要性が指摘されている^{3), 6), 9)}。

† 第34回計測自動制御学会学術講演会で発表（1995・7）

* NTTデータ通信（株） 東京都千代田区一ツ橋 1-1-1 パレスサイドビル 8F

** 東北大学医療技術短期大学部 仙台市青葉区星陵町 2-1

*** 東北大学工学部 仙台市青葉区荒巻字青葉

* NTT DATA Corporation, Chiyoda-ku, Tokyo

** College of Medical Sciences, Tohoku University, Aoba-ku, Sendai

*** Faculty of Engineering, Tohoku University, Aoba-ku, Sendai

(Received September 3, 1996)

(Revised June 30, 1997)

Fig. 1 Fully connected recurrent neural networks

本稿では、二値系での結果を拡張して連続値の複雑系である **RNN** の edge of chaos でのダイナミクスの特徴が、表現能力を損なわない、冗長な自由度に対する制限条件となることを明らかにし、これを利用した創発的な学習手法を提案する。はじめに、ダイナミクスの複雑さを操作するパラメータを導入する。提案手法は、導入したパラメータによりダイナミクスの複雑さを edge of chaos に保ちつつ、簡単な BP 法と併用して学習を進める。非線形システムの出力推定問題に適用したシミュレーション結果より、連続値系の edge of chaos におけるダイナミクスの、制限条件としての効果と、提案手法が収束速度の改善に有効であることを示す。

2. リカレントニューラルネットワーク

システムの振舞いの複雑さは、システムを構成する各要素の振舞いの複雑さのみではなく、要素間の相互作用にも依存する³⁾。ニューラルネットワークにおける各ニューロン間の相互作用は、シナプス結合を介した情報通信により行われるとみなすことができる。したがって、この結合重みを改変することによって、さまざまな相互作用を実現することができると考えられる。

一方、静的な認識問題や推定問題などへの応用が進んでいるフィードバック結合をもたない従来の代表的な階層型ネットワーク、あるいは最適化問題などへ適用されている対称結合ネットワーク⁵⁾は、フィードバック結合をもっていないその結合構造に制限があるため、必然的に相互作用の構造にも制限を受けている。

本研究では様々なダイナミクスを表現できるモデルとして、各ニューロンの写像関数がシグモイド関数で、次式で時間発展してゆく **Fig. 1** に示すような結合トポロジー制限のないリカレントニューラルネットワークを考える。

$$x_i(t+1) = \frac{1}{1 + \exp(-\alpha_i s_i(t))}, \quad (1)$$

$$s_i(t) = \sum_j w_{ij} x_j(t) + w_{iu} u_i(t) + \theta_i, \quad (2)$$

$$y_i(t) = x_i(t), \quad (t = 1, 2, \dots). \quad (3)$$

ここで、 x_i , s_i , y_i はニューロン i の状態、入力、出力で、 N ,

Fig. 2 Return map of a random output of a neuron.

w_{ij} , α_i , u_i , w_{iu} , θ_i は、それぞれ要素数、要素 j との結合重み係数、写像関数の傾き、外部入力とその結合重み係数、および閾値である。

3. ダイナミクスのパラメータ

ランダムプリアンネットワーク⁶⁾およびセルオートマトン⁹⁾においては、ネットワークにおける要素間の結合線の総数、すなわち相互作用の強さと、ネットワーク内を実際に流れる信号の多様性に関するパラメータで系のダイナミクスを統計的に特徴づけることができる³⁾。この概念を **RNN** のような連続値系に拡張し、ダイナミクスを特徴づけるパラメータとして、以下で述べる $\overline{\sigma^2}$, l_θ , r_{in} の三つを導入する。

3.1 入力分布の分散 $\overline{\sigma^2}$

RNN においては、各要素に結合している結合重み w_{ij} によって入力 s_i がとる値の範囲が決められ、その範囲が広がればそれだけ出力される値の範囲も広くなり、それと同時にニューロンの飽和特性による非線形性が強くなる。この各ニューロンへの入力値の範囲を表す入力分布の分散値をダイナミクスを特徴づけるパラメータとして導入する。

はじめに、ネットワークを構成する要素の数と、各要素の出力時系列との関係を実験的に考察してみる。(1), (2) 式のパラメータ w_{ij} , θ_i などをランダムに設定した **RNN** (以下ランダムネットワークと呼ぶ) を多数準備すると、要素数が増えるにしたがって他のニューロンからの入力信号の影響が支配的になることより、ほとんどのランダムネットワークの要素ニューロンの出力信号は **Fig. 2** のように一様ランダムに近くなってゆく。したがって、ここでは任意の要素の出力時系列が $[0, 1]$ の一様ランダムであると仮定して、ニューロン i の入力 s_i の分布 G_i (分散 σ_i^2 , 平均値 m_i) を次式のように求めた¹⁾。

$$G_i : \begin{cases} \sigma_i^2 = \frac{1}{12} \sum_j^N w_{ij}^2, \\ m_i = \frac{1}{2} \sum_j^N w_{ij} + \theta_i. \end{cases} \quad (4)$$

(4) 式で表される各要素の入力 S_i の分散 σ_i^2 のネットワー

Fig. 3 Average mutual information $\overline{MI}$ and Lyapunov exponent λ over σ^2 space.

ク全体での平均を次式で定義する.

$$\overline{\sigma^2} \doteq \frac{1}{N} \sum_{i=1}^N \sigma_i^2. \quad (5)$$

パラメータ $\overline{\sigma^2}$ は, すべての相互作用の強さ w_{ij} の 2 乗平均に比例し, マクロにみたときのシステムの相互作用の大きさ, 非線形性の強さを表すパラメータである.

Fig. 3 は要素数 N が 100 の RNN において, 平均値 m_i が 0 となるよう θ_i を設定したときの, $\overline{\sigma^2}$ と, $y_i(t)$, $y_i(t-1)$ の相互情報量 MI_i の平均値 $\overline{MI}$, リヤプノフ指数 λ ¹⁷⁾ との関係を示したものである. MI_i は $y_i(t)$ を 16 値の離散値過程³⁾ とし, シヤノンのエントロピー H を用いて (6) 式により算出した.

$$MI_i(y_i(t); y_i(t-1)) = H(y_i(t)) + H(y_i(t-1)) - H(y_i(t), y_i(t-1)). \quad (6)$$

また, 相互情報量の平均値 $\overline{MI}$ は次式で計算した.

$$\overline{MI} \doteq \frac{1}{N} \sum_{i=1}^N MI_i. \quad (7)$$

Fig. 3 より, $\overline{\sigma^2}$ の値が大きくなると $\lambda > 0$ であり, 時間発展がカオス的で信号が有限の時間に初期の記憶を失うことより, $\overline{MI}$ は小さな値になっている. また, $\overline{\sigma^2}$ が小さくても $\overline{MI}$ は小さな値をとるが, これは系のエントロピーが小さな値であるためであり, 振舞いとしては秩序立ったものになり, $\lambda < 0$ となっている. 一方, 両者の中間の $\lambda \approx 0$ 付近では $\overline{MI}$ が高い特徴的な部分が存在し, 本稿ではこの部分をカオスの辺縁 (edge of chaos) と呼ぶ. このように $\overline{MI}$ を高く保つことは, 自身の過去の出力値との相関が強いという意味でダイナミクスが限定されるとともに, ネットワークのパラメータ値も限定され, 高い自由度に対する一つの制限条件とみなせる.

3.2 相互作用の異質性 r_{in}

$\overline{\sigma^2}$ は, 重み係数の二乗平均, すなわちその符号ではなく, 絶対値にのみ依存したパラメータである. ランダムネットワークのように, 興奮性結合 ($w_{ij} > 0$) と抑制性結合 ($w_{ij} < 0$) がほぼ均等に存在するような大規模なネットワークでは信号の流れがその符号も含めてランダムになり, 相互作用のマクロな絶対値の大きさを表す $\overline{\sigma^2}$ がダイナミクスを操作するの

Fig. 4 Lyapunov exponent λ over variance $\overline{\sigma^2}$ space for some r_{in} values.

に有効に働く. しかし, たとえばすべての結合が興奮性であるようなネットワークでは以下に示すように Fig. 3 のようにならない場合もある. このことは, 相互作用の絶対値の大きさだけでなく, その符号もネットワークの振舞いの複雑さに影響することを意味している.

そこで, ネットワーク中の興奮性結合と抑制性結合の割合を示すパラメータ r_{in} を次式で定義する.

$$r_{in} \doteq \frac{\sum_{i,j} w_{ij}^-}{\sum_{i,j} w_{ij}^+ + \sum_{i,j} w_{ij}^-}. \quad (8)$$

ただし, w_{ij}^+ は興奮性の結合, w_{ij}^- は抑制性の結合の絶対値をそれぞれ表す. つまり r_{in} は全結合重みのなかで抑制性結合が占める重み付きの割合で定義される.

r_{in} はネットワーク内を流れる信号の多様性に関するパラメータであり, $\overline{\sigma^2}$ を補助する役割を果たす. 一般に $\overline{\sigma^2}$ の増加に伴い, ダイナミクスも複雑になり λ も増加するが, この傾向は $r_{in} = 0.5$ のとき最も顕著になる (Fig. 4). 相互作用の符号に関しては, 連続時間の二要素系で詳しく解析され, 相互作用の重み係数が異符号 ($w_{ij}w_{ji} < 0$) のとき系の振舞いが複雑 (この場合リミットサイクル) になると報告されている¹⁰⁾. この二要素系での結果は, ランダムネットワークにおいては, 興奮性結合の平均値と分散は, 抑制性結合の絶対値の平均値と分散にそれぞれ等しくなることを考慮すると, $r_{in} = 0.5$ のとき, すなわちネットワーク内に異符号の相互作用が存在する確率が高くなるほど $\overline{\sigma^2}$ によるダイナミクス操作が効率的に作用することと一致する.

3.3 写像関数のバイアス l_0

以上二つのパラメータ $\overline{\sigma^2}$, r_{in} は重み係数 w_{ij} によってネットワークのダイナミクスを変化させるためのものである. 一方, 閾値も恒等入力 (常に 1 を出力するニューロン) との結合重みと解釈できるので, これを用いたダイナミクスの操作が可能である. 各ニューロンの入力分布である (4) 式の平均値 m_i を 0 になるよう閾値を調整することにより, その出力分布は 0.5 を境に対称になることが期待され, ニューロ

ンの写像関数を異質 (heterogeneous)⁹⁾ なものにすることができる。

したがって、すべてのニューロンの入力分布の平均値 m_i が 0 のとき、ネットワーク内を実際に伝わる信号は最も多様になると考えられる。このときの各要素の閾値の値 θ_{ic} は (4) 式より次式のように与えられる。

$$\theta_{ic} = -\frac{1}{2} \sum_j^N w_{ij}. \quad (9)$$

第 i 要素が、ニューロン i の閾値であるベクトル θ と、第 i 要素が θ_{ic} であるようなベクトル θ_c を次式で定義する。

$$\theta = [\theta_1, \theta_2, \dots, \theta_N]^T, \quad (10)$$

$$\theta_c = [\theta_{1c}, \theta_{2c}, \dots, \theta_{Nc}]^T. \quad (11)$$

ネットワークのダイナミクスは、 θ が θ_c と離れ、各ニューロンが等質性 (homogeneous), すなわち 0.5 よりも大きいもの、あるいは小さいもののどちらか一方が出力されやすくなることによって、秩序的になってゆくと考えられる。そこで、この異質性の指標となるパラメータ l_θ を θ と θ_c とのユークリッド距離で次式のように定義する。

$$l_\theta \equiv \|\theta - \theta_c\| = \sqrt{\sum_{i=1}^N (\theta_i - \theta_{ic})^2}. \quad (12)$$

この l_θ を大きくすると秩序立ったネットワークになり、逆に l_θ を小さくすると、ネットワークの振舞いは複雑になることが期待される。

Fig. 5 は w_{ij} を固定して (θ_{ic} を固定して) θ_i を変化させることにより、 l_θ によるネットワークダイナミクスの変化を測ったものである。 σ^2 のときと同様に edge of chaos での特徴的な部分が観られる。

4. 創発的学習手法

ここでは、前節で示した提案パラメータとネットワークのダイナミクスの関係を利用した学習手法を提案する。

Fig. 6 は、提案パラメータ $\sigma^2, r_{in}, l_\theta$ を適当に変化させてさまざまなネットワークを構成したときのダイナミクスの様子を、 λ を横軸に、 $\overline{MI}$ を縦軸にプロットしたものである。 $\overline{MI}$ を高く保つことは、高い自由度に対する制限条件とみなせることはすでに述べたが、Fig. 6 から $\lambda \simeq 0$ とすることも相対

Fig. 5 Average mutual information $\overline{MI}$ and Lyapunov exponent λ over l_θ space.

Fig. 6 Average mutual information $\overline{MI}$ versus Lyapunov exponent λ .

的に $\overline{MI}$ を高くする一つの方法であることが確認できる。

また、 $\lambda \simeq 0$ で $\overline{MI}$ が広範囲に分布しており、離散値系の場合^{3), 9)}と同様に、連続値の複雑系である RNN においてもカオスの辺縁でのダイナミクスが特徴的であることがわかる。ネットワークにダイナミクスを学習させるためには、目標のダイナミクスと何らかの意味 (指標) で等価なダイナミクスをネットワーク内部に表現する必要がある。差分方程式 (微分方程式) は、過去の値と未来の値の関係を規定していることを考慮して、指標として相互情報量を用いることにすると、 $\lambda \simeq 0$ の付近にネットワークのダイナミクスを保つという制限条件は、単純なダイナミクスから、複雑なダイナミクスまでの広範囲なダイナミクスを効率的に表現することができ、学習に有効である⁷⁾。

1 節でも触れたように、創発とは、自律的に振舞う個体 (要素) 間および環境との間の局所的な相互作用が大域的な秩序を発現し、他方、そのように生じた秩序が個体の振舞いを拘束するという双方向の動的過程により、新しい機能、形質、行動が獲得されることである⁸⁾。学習過程における大域的秩序とは、誤差を減少させるパラメータ改変であり、ダイナミクスの複雑さを edge of chaos に保つこと (制限条件) は、大域的秩序の拘束の下で、個体 (ニューロン) の振舞い (パラメータ改変) を決める自律的基準であるといえる。

従来の、学習目標との双方向の相互作用がない、静的な制限条件に対し、提案する創発的学習手法 (Emergent Learning Algorithm: ELA) では、この大域的秩序 (誤差減少) と制限条件でもある自律的基準 ($\lambda \simeq 0$) とによるパラメータの改変によって生み出されるネットワークダイナミクスの変化が創発的な動的過程を取り、広範囲のダイナミクスを効率的に表現できるパラメータ値付近での学習を可能とする。具体的なパラメータ改変量は、(13), (14) 式で計算する。

$$\delta w_{ij} \doteq \begin{cases} \delta^e w_{ij} + c_1 \lambda^3 |\delta^e w_{ij}|, & (i r_{in} \leq 0.5), \\ \delta^e w_{ij} - c_1 \lambda^3 |\delta^e w_{ij}|, & (i r_{in} > 0.5), \end{cases} \quad (13)$$

$$\delta \theta_i \doteq \begin{cases} \delta^e \theta_i + c_2 \lambda^3 |\delta^e \theta_i|, & (\theta_{ic} \leq \theta_i), \\ \delta^e \theta_i - c_2 \lambda^3 |\delta^e \theta_i|, & (\theta_{ic} > \theta_i). \end{cases} \quad (14)$$

ここで c_1, c_2 は定数で、 $\delta^e w_{ij}, \delta^e \theta_i$ は通常の BP 法により計算された誤差情報からの修正量である。また、 $i r_{in}$ はニューロン i に結合している抑制性結合の重み付き割合を表し、(15)

式で求める.

$$i_{rin} = \frac{\sum_j^N w_{ij}^-}{\sum_j^N w_{ij}^+ + \sum_j^N w_{ij}^-}. \quad (15)$$

(13), (14) 式で第二項がダイナミクスを $\lambda \simeq 0$ に保つための修正項である. この項は, 3節で述べたパラメータと λ との関係をもとにしている. すなわち, $\lambda < 0$ のときは, σ^2 を増加し l_θ を減少させる方向に, $\lambda > 0$ のときは, その逆方向にパラメータを変更する項である. ここで, (13), (14) 式の第二項を λ ではなく λ^3 に比例させているのは, $\lambda \simeq 0$ を edge of chaos とみなして, 第二項による改変量を λ に比例させた場合よりも小さな値に抑えるためである. パラメータの更新は (16), (17) 式で行った.

$$w_{ij}(\tau+1) = \beta_i(\tau)(w_{ij}(\tau) + \delta w_{ij}(\tau)), \quad (16)$$

$$\theta_i(\tau+1) = \theta_i(\tau) + \delta \theta_i(\tau), \quad (17)$$

$$\beta_i(\tau) = \sqrt{\frac{\sigma_i^2(\tau) + \delta^a \sigma_i^2(\tau)}{\sigma_i^2(\tau)}}. \quad (18)$$

ただし

$$\delta^a \sigma_i^2(\tau) = C \exp\left[-\frac{\{i_{rin}(\tau) - 0.5\}^2}{2c_4}\right] \lambda(\tau)^3 \delta \sigma^2, \quad (19)$$

$$C = -\frac{c_3}{\sqrt{2\pi c_4}}. \quad (20)$$

ここで c_3, c_4 は定数, $\tau (= 1, 2, \dots)$ は学習ステップ, $\delta \sigma^2$ は微小量の定数である. β_i は, 提案パラメータ σ^2 の値を調節するための係数で, (19) 式の正規分布関数 (i_{rin} の関数) を用いて (18) 式で計算される.

5. シミュレーション結果

提案手法を未知システムの出力推定問題に適用した. シミュレーションでは, **RNN** と **FNN** をカスケードに接続したネットワーク⁴⁾を用い, **FNN** の出力を推定値 $\hat{y}$ とし, **FNN** は BP 法で学習した. **RNN** の学習は提案手法でおこない, その際 $\delta^e w_{ij}$, $\delta^e \theta_i$ は逆伝搬されてきた誤差情報を 1 時刻のみ遡る簡便な BP 法によって計算した. ここでの目的は, 未知システムの入力 $u(t)$ と出力 $y(t)$ の時系列をネットワークに与えて, 過去 D 時刻間の推定誤差の和を表す (21) 式の評価関数を小さくすることである.

$$e(\tau) = \frac{1}{D} \sum_{t=\tau-D+1}^{\tau} |y(t) - \hat{y}(t)|. \quad (21)$$

ただし簡単のため, 学習ステップ τ の間隔と時刻 t の間隔は等しいとする. この際, 汎化という点から学習を評価するため, τ における学習では, 未知システムの $t = \tau - D + 1 \sim \tau$ までの入出力時系列を与えて, 各学習ステップでの学習パターンが異なるようにした. また, 各パラメータ値は自律的基準を優先するように $c_1 = c_2 = 1000$, 入力分布の形より $c_3 = 80$, $c_4 = 0.001$, $\delta \sigma^2 = 0.0005$ とし, $D = 100$ とした.

Fig. 7 Error e and Lyapunov exponent λ as functions of learning step τ for temporal XOR problem with delay. The case of learning by Emergent Learning Algorithm (ELA) are shown as dotted lines and solid lines show the case of Back-Propagation Algorithm (BP).

5.1 遅延のある XOR 問題

(22) 式で表される遅延のある XOR の出力推定を行なった.

$$y(t+1) = u(t-2) \oplus u(t-3),$$

$$u(t) \in \{0, 1\}. \quad (22)$$

Fig. 7 は, **RNN** の初期ダイナミクスを秩序相に設定して学習を行なったときの $e(\tau)$ と $\lambda(\tau)$ の変化を, ダイナミクス操作を行わない通常の BP 法と比較して示したものである. 図より, カオスの辺縁 $\lambda(\tau) \simeq 0$ で学習が効率良く行われているのがわかる. 一方, 通常の BP 法を用いた場合には, ダイナミクスが秩序相に固定されてしまい学習がうまく行われていないのがわかる. この結果は学習過程において $\lambda(\tau) \simeq 0$ に保つことが有効であることを示唆している.

5.2 状態量フィードバックのある XOR 問題

(23) 式で表される状態量フィードバックのある XOR の出力推定を行なった.

$$x(t+1) = u(t) \oplus x(t), \quad (u(t) \in \{0, 1\}),$$

$$y(t) = x(t). \quad (23)$$

Fig. 8 は **RNN** の初期ダイナミクスを無秩序相に設定して学習したときの $e(\tau)$, $\lambda(\tau)$ である. $\tau \simeq 700$ 付近で, うまく汎化できていないため誤差が大きくなっているが, カオスの辺縁と秩序相をさまようことで, ネットワークパラメータを急激に変化させながら, 再び $\lambda(\tau) \simeq 0$ に回復し, その後はより誤差が少なくなるように学習している.

6. ま と め

RNN のダイナミクスの複雑さを調整するパラメータを導

Fig. 8 Simulation result of temporal XOR problem with state feedback.

入し、それにより秩序相、カオス相など種々の振舞いが表現できることを示した。これらのダイナミクス相のうち、相転移後のカオスの辺縁で RNN のダイナミクスの相互情報量 $\overline{MI}$ が広範囲に変化することを明らかにした。提案した学習法は、RNN をカオスの辺縁に保つという、表現能力を損なわない冗長な自由度に対する制限条件を併用することにより、収束速度および汎化能力も含めて効率的な学習を実現できることがシミュレーションにより示された。

RNN をバックプロパゲーション法で学習させる場合、構成要素数が増えるにしたがってメモリと計算時間がべき乗のオーダーで増えていく¹⁶⁾。一方、本研究でのネットワークの大域的性質を示すパラメータは、要素数が増えるとともに精度が増すという特徴をもち、併用する BP 法は簡単なものでも十分な精度が得られる可能性をもつ。したがって、提案手法は大規模なネットワークの構成、すなわちより高度な機能を実現するための設計手法へも適用可能である。

今回の手法は、提案パラメータとダイナミクスとの関係を定性的に考察することにより、カオスの辺縁での学習が有効である可能性を示した。数学的に厳密な解析および評価は今後の課題である。

参考文献

- 1) 阿倍 斎：応用数理統計学入門，倍風館（1985）
- 2) 銅谷 賢治：リカレントネットワークの学習アルゴリズム，計測と制御，**30-4**，296/301（1991）
- 3) 本間，佐藤，阿部，竹田：自律分散の適応手法によるホロンネットワークの進化について，計測自動制御学会論文集，**31-7**，908/915（1995）
- 4) N. Honma, K. Kitagawa and K. Abe: "Effect of Complexity on Learning Ability of Recurrent Neural Networks", Proc. of the 2nd AROB, 58/61 (1997)
- 5) J. J. Hopfield and D. W. Tank: "Neural" Computation of

Decisions in Optimization Problems, Biological Cybernetics **52**, 141/152 (1985)

- 6) S. A. Kauffman: Antichaos and Adaptation, Scientific American, August (1991)
- 7) 喜多川，本間，阿部：カオスの辺縁におけるリカレントニューラルネットワークの学習能力，システム情報関連合同シンポジウム前刷，103/106（1995）
- 8) 北村 新三：創発的機能形成のシステム理論，計測と制御，**34-7**，593/594（1995）
- 9) C. G. Langton: Computation at the Edge of Chaos: Phase Transitions and Emergent Computation, Physica **D 42** 12/37 (1990)
- 10) 中野 馨他：ニューロコンピュータの基礎，コロナ社（1990）
- 11) 西川，北村：ニューラルネットと計測制御，朝倉書店（1995）
- 12) B. A. Pearlmutter: Learning State Space Trajectories in Recurrent Neural Networks, Neural Computation, **1-2**，263/269 (1989)
- 13) F. J. Pineda: Generalization of Back-Propagation to Recurrent Neural Networks, Physical Review Letters, **59-19**，2229/2232 (1987)
- 14) R. Ruthen: Adapting to Complexity, Scientific American, **1**，110/117 (1993)
- 15) K. P. Unnikrishnan and K. P. Venugopal: Alopex: A Correlation-Based Learning Algorithm for Feedforward and Recurrent Neural Networks, Neural Computation, **6**，469/490 (1994)
- 16) R. J. Williams and D. Zipser: A Learning Algorithm for Continually Running Fully Recurrent Neural Networks, Neural Computation **1-2**，270/280 (1989)
- 17) A. Wolf, J. B. Swift, H. L. Swinney and J. A. Vastano: Determining Lyapunov Exponents from a time series, Physica, **16D**，285/315 (1989)

[著者紹介]

喜多川 健

1996 年，東北大学大学院工学研究科博士前期課程修了。同年 NTT データ通信株式会社入社。カオス理論による複雑適応系の解析および創発システムの研究に従事。

本間 経康（正会員）

1995 年，東北大学大学院工学研究科博士後期課程修了。同年同大学電気通信研究所研究員。96 年同大学医療技術短期大学部講師。最適制御，自律分散システム，人工生命などの研究に従事（工学博士）。

阿部 健一（正会員）

1969 年，東北大学大学院工学研究科博士課程修了。同年同大学工学部助手。74 年同助教授。84 年豊橋技術科学大学教授。93 年東北大学教授。ディジタル制御，適応・学習制御などの研究に従事。IEEE，電気学会，情報処理学会などの会員（工学博士）。