

# 複数の隠れニューロンの生成が可能な成長型ニューラルネットの構築

大分大学 山崎 高志、柴田 克成

## The Construction of Growing Neural Network by Generating Multiple Hidden Neurons

Takashi Yamazaki and Katsunari Shibata, Oita University

Abstract: Growing neural network in which each growing neuron grows its axon towards a target neuron to acquire an appropriate structure for a given task in parallel with learning has been proposed<sup>1)</sup>. It could deal with only one hidden neuron set and it was shown that EXOR problem was learned by the neural network. In this paper, a technique is proposed to extend the growing neural network so as to deal with more than one hidden neuron set. It is confirmed that EXOR problem was learned on the condition of a limited number of the direct connections from the input neurons to the output neuron. The acquired structure was compared by the regular network and some difference found between them.

### 1 はじめに

ニューラルネットは、学習能力と汎化能力が高いが、タスクに適した構造決定が難しい。ニューラルネットの構造決定方法として、これまでに遺伝的アルゴリズム<sup>2)</sup>や情報量基準<sup>3)</sup>を用いる方法などが試されてきた。しかしこれらの方法では、学習中に構造獲得することが難しい上、ある程度構造を既定する必要があり、設計者の関与による自由度の低下が考えられる。

一方、生体では、NGF(Nerve Growth Factor)やBDNF(Brain-derived Neurotropic Factor)などの化学物質が神経回路形成に関与していると言われている<sup>4)5)</sup>。それらの化学物質は、標的ニューロンから分泌され、濃度勾配を形成すると軸索は、濃度勾配にしたがって成長し、標的ニューロンとの結合を形成する。筆者らは、タスクの学習と共に神経回路が形成されていくのではないかと考えている。そこで、BP(Error Back Propagation)法<sup>6)</sup>を拡張した"成長型ニューラルネット (Growing Neural Network:GNN)"を提案し、自律的な構造の獲得の実現を目指してきた<sup>1)7)</sup>。

これまでにGNNは、隠れニューロンを有する場合の成長・学習を可能にするための方法として2つの方法を提案している。一つは、入力とも出力とも結合のない隠れニューロンから誤差信号と無関係の一定量の物質を分泌し、入力ニューロンが結合した後に、出力ニューロンから分泌する誤差信号に依存する物質によって、隠れニューロンが結合を形成する方法である<sup>7)</sup>。もう一つは、入出力ニューロン間のダイレクト結合の形成、学習後に隠れニューロンを出力ニューロンから生成する方法である<sup>1)</sup>。生成された隠れニューロンは、予め出力ニューロンと結合を持ち、誤差信号を受け取り物質を分泌する。それらの方法によってEXOR問題を解決したことが示されている。後者は、学習によって入力-隠れニューロン間の結合

を獲得できるため、必要に応じた隠れニューロンを獲得でき、より自然な方法と筆者らは考えている。しかし、これまでは隠れニューロン数が複数となる場合が想定されていなかった。

本論文では、後者の方法を用いたGNNで複数の隠れニューロンを生成し、学習する方法を提案する。そして、EXOR問題で、入出力ニューロンのダイレクト結合数に制限を設けることで、2つ以上の隠れニューロンを必要とさせた上で、成長型ニューラルネットが正常に動作するか調べた。さらに、得られた構造を、逐次隠れニューロンを追加する通常のニューラルネットによって得られた結果と比較した。

### 2 成長型ニューラルネット

#### 2.1 構造の獲得手順

成長型ニューラルネット(GNN)の成長は、Fig.1に示すように段階を重ね、構造を複雑化する。各段階では、学習と成長が同時進行する。

最初は、隠れニューロンのないダイレクト結合の段階である。出力ニューロンを標的ニューロンとし、入力ニューロンの軸索を成長、結合させる。エラーが十分に減少しない場合は、一定時間後に出力ニューロンが隠れニューロンを1セット生成し、次の段階に移行する。そのとき、入力ニューロンは、出力ニューロンから遠ざけ、隠れニューロンを入出力ニューロン間に配置する。隠れニューロン

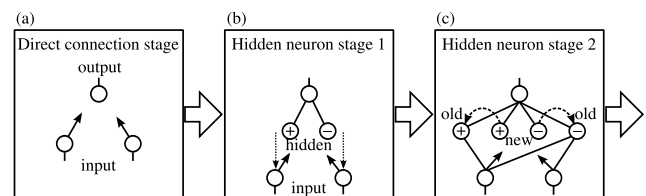


Fig. 1: Stage transition in the growth of a GNN

は、出力ニューロンと正の結合を有するものと負の結合を有する2つで1セットの用意ニューロンを用意し(理由は後述)、それぞれ2つの隠れニューロンは、生成した直後から出力ニューロンと一定の結合を持っている。そのため、隠れニューロンは、出力ニューロンよりBP法<sup>6)</sup>にしたがって計算された誤差信号を受け取る。そして、その誤差信号にしたがった量の物質を分泌する。入力ニューロンの軸索は、拡散された物質の濃度勾配と入力信号によって成長する。

本論文では、エラーが十分に減少しない状態が一定時間経過した場合に、既存の隠れニューロンを遠ざけてできるスペースに新たに隠れニューロンを生成し、隠れニューロン2セットの段階に移行する。本論文における各ニューロンの位置については、Fig.3、5、および6に示す。エラーが十分減少するまで、上の手順を繰り返す。

## 2.2 成長と学習

成長型ニューラルネットの成長は、BP法<sup>6)</sup>による学習の拡張にあたる。Fig.2に成長と学習の流れを示す。はじめに、すでに形成された結合によって出力は計算され、誤差信号は、通常のBP法<sup>6)</sup>によって計算する。

各標的ニューロンは、誤差信号を受け取り、それにしたがって物質を分泌し、各ニューロンは、形成された濃度勾配にしたがって、各成長ニューロンは、軸索を成長させる。誤差信号は、負となり得るが、負の物質濃度は、考えづらいため、正負別々の物質を想定した。正負の濃度  $u$  は、次の拡散方程式により求められる。

$$\frac{\partial u_{x,y}^i}{\partial t} = \text{div}_{x,y}^i + D\nabla^2 u_{x,y}^i \quad (i: p \text{ or } n) \quad (1)$$

$u_{x,y}^i$  は位置  $(x, y)$  の濃度、 $i$  は正 ( $p$ ) または負 ( $n$ ) の物質を表わす添字とする。 $\text{div}$  は誤差に依存する物質の分泌密度、 $D$  は拡散係数とする。

物質の拡散時間を考慮し、軸索の成長には、成長ニューロンの出力  $o$  そのものではなく、その1次遅れ  $S$  を

$$\tau_S \frac{dS_i}{dt} = -S_i + o_i \quad (2)$$

として計算して用いた。ここで、 $o$  は各成長ニューロンの出力、 $\tau_S$  は出力の時定数、 $i$  は入力ニューロンの添字 ( $i = 0, \dots, \text{NODE}$ ,  $\text{NODE}$ : 入力ニューロンの数) とする。

軸索の伸長は、この出力の1次遅れ  $S$  と軸索先端の濃度勾配から計算する。また、軸索伸長の安定のため、さ

- |   |
|---|
| <ol style="list-style-type: none"> <li>1. Forward calculation</li> <li>2. Propagation of error signal</li> <li>3. Diffusion of substance according to error signal</li> <li>4. Axon growth</li> <li>5. The formation of connection is judged</li> <li>6. Modification of connection weight<br/>(if neuron has connections)</li> </ol> |
|---|

Fig. 2: The flow of growing and learning. This sequence is repeated in each stage

らに1次遅れを導入し、軸索伸長速度  $A$  と軸索先端位置  $a$  を

$$\tau_a \frac{dA_i}{dt} = -A_i + \nabla(u^p - u^n) \cdot \text{flag}_i \cdot S_i \quad (3)$$

$$\frac{da_i}{dt} = \xi \cdot A_i \quad (4)$$

と求めた。ここで、 $\tau_a$  は軸索伸長の時定数、 $\xi$  は軸索伸長係数、 $\text{flag}_i$  は正のニューロンなら1とし、負のニューロンなら-1とする。これによって、各成長ニューロンは正負を考慮し、正の成長ニューロンは、正の誤差信号を分泌する標的ニューロンに軸索を伸長させる。逆に負の成長ニューロンは、負の誤差信号を分泌する標的ニューロンに軸索を伸長させる。結合の判定は、軸索先端が、標的ニューロン周辺の微小エリアに到達したときとした。

結合部はBP法によって学習させることから

$$\frac{dw_{ji}}{dt} = \eta_c \cdot \delta_j \cdot o_i \quad (5)$$

で計算した。ここで、 $w_{ji}$  は結合元の  $i$  番目のニューロンから結合先の  $j$  番目のニューロンへの重み、 $\eta$  は学習係数とする。ただし、生理学的知見である興奮ニューロンや抑制ニューロンを想定し、それぞれ重みが0以上しかなりえない正ニューロンと重みが0以下しかなりえない負ニューロンを生成することにした。

## 2.3 離散化

物質の拡散、軸索の伸長は、平面上のみと仮定し、物質の濃度勾配の概算は、計算領域を離散化し、濃度  $u$  は格子点上のみ計算した。正負の物質の濃度  $u$  は、Eq.(1) を離散化した Eq.(6) により算出した。

$$\Delta u_{t+1,x,y}^i = \left\{ \text{div}_{x,y}^i + \frac{D}{\Delta x^2} \left( -4u_{t,x,y}^i + u_{t,x+1,y}^i + u_{t,x-1,y}^i + u_{t,x,y+1}^i + u_{t,x,y-1}^i \right) \right\} \Delta t \quad (6)$$

分泌量  $\text{div}$  は、標的ニューロンのある格子点では

for the substance  $u^p$

$$\text{div}^p = \rho \cdot \delta \quad \text{if } \delta > 0 \quad (7)$$

$$\text{div}^p = 0 \quad \text{otherwise}$$

for the substance  $u^n$

$$\text{div}^n = -\rho \cdot \delta \quad \text{if } \delta < 0 \quad (8)$$

$$\text{div}^n = 0 \quad \text{otherwise.}$$

とする。ただし、標的ニューロンは格子点以外には存在しないとし、 $\rho$  は湧き出し係数とする。 $\Delta x$  は格子間隔とする。標的ニューロンのない格子点では、 $\text{div} = 0$  とする。誤差の正負を  $i = p, n$  で表している。濃度勾配  $\nabla u^p$  および  $\nabla u^n$  は次式により近似的に求めた。

$$\nabla u_{x,y}^i \approx \frac{1}{2\Delta x} \left( \begin{array}{c} u_{x+1,y}^i - u_{x-1,y}^i \\ u_{x,y+1}^i - u_{x,y-1}^i \end{array} \right) \quad (i = p \text{ or } n) \quad (9)$$

ただし、軸索先端の位置は、離散化しないものとする。軸索先端のある最寄りの格子点の濃度勾配を参照し、軸索伸長の計算は Eq.(3) により計算した。

### 3 シミュレーション

#### 3.1 シミュレーション条件

本研究では、前の研究<sup>7)</sup>と同じタスクである EXOR 問題を扱ったが、複数の隠れニューロンが生成したときの動作に注目するため、入出力のダイレクト結合数を制限し、複数の隠れニューロンを必要とさせる条件にした。入力パターンは、各パターンをランダムに変化させ、1s 与え続けた。学習終了条件は、誤差が  $10^{-5}$  以下の状態が 1,000 パターン連続したときとした。ここで誤差は、入力の変化した直後の 2 乗誤差のことを指す。

GNN は、ダイレクト結合が 3 つとなるに十分な時間である 2,000s のとき、隠れニューロンを 1 セットの段階に移行する。隠れニューロン数が 2 セット以降は、隠れニューロンが 1 セット当たりの期間を 10,000s とし、終了条件を満たしていなければ、出力ニューロンから隠れニューロンを生成する。このとき、既存の隠れニューロンからの物質の分泌は、簡単化のためなくした。

比較のため用いた通常の NN の学習では、1 つの入出力間のダイレクト結合を持たせた状態で、通常のニューロンによって、学習を行った。ただし、通常の NN では、隠れニューロンを GNN と同じ期間ごとに追加した。

また、入力パターンに乱数を用いているため、乱数系列を変えて 10 回ずつシミュレーションを行った。

Tab.1 に本実験で用いた各定数を示す。各定数は、前の研究との比較のため同じにした。

項目	値
学習係数 $\eta$	0.2
刻み時間 $\Delta t$	0.001[s]
格子間隔 $\Delta x$	0.01[mm]
拡散係数 $D$	0.1 [mm <sup>2</sup> s <sup>-1</sup> ]
湧き出し係数 $\rho$	10 [s <sup>-1</sup> ]
軸索伸長係数 $\xi$	0.1 [mm <sup>2</sup> s <sup>-1</sup> ]
入力への時定数 $\tau_S$	0.1 [s]
出力への時定数 $\tau_a$	$2 \times 10^{-5}$ [s]

#### 3.2 結果

はじめに、成長と学習の結果を示す。ダイレクト結合の段階の軸索の成長軌跡とその Y 座標の時間的変化を Fig.3 および 4 に示す。Fig.5 は、隠れニューロンが 1 セットの段階の軸索の軌跡を示す。Fig.6 は、隠れニューロンが 2 セットの段階の場合である。ニューロンのラベルは、例えば 'h1+' は、最初に生成された隠れニューロンで、さらに出力ニューロンと正の重みの結合を形成する隠れニューロンである。'a-' は、'a' の入力信号を出力するニューロンであり、標的ニューロンの負の誤差信号に対し成長し、負の重みの結合を形成することを意味する。

Fig.3 および 4 で示されているダイレクト結合の段階において、入力'b-'はおよそ 800s に出力ニューロンと結合

し、その後入力'a+'が結合を形成している。入力'b+'は標的ニューロンから反対向きに成長したが、最終的には結合に至っている。2,000s の時点では入力'a-'は結合を持っていない。しかし、もし次の段階に進ませず、そのまま時間が経過すると'a-'は、必ず出力ニューロンと結合に至った。さらに Fig.8 より、最終的に'b-'と'a+'は、重みがおよそ 0.0 になっている。

Fig.5 より、隠れニューロンが 1 セットの段階において、ニューロン'h1-'の周辺で入力'b+'の不安定な挙動がみられるが、最終的に入力'a-'と入力'b+'は、ニューロン'h1-'と結合に至っている。

Fig.6 より、隠れニューロンが 2 セットの段階において、入力'a-'はいったん'h1+'と'h2+'間に位置するが、最終的にニューロン'h2+'と結合した。ニューロン'h2-'は、3 つの入力ニューロンと結合した。

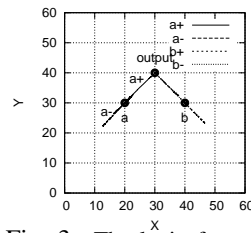


Fig. 3: The loci of axon tips at the direct connection stage

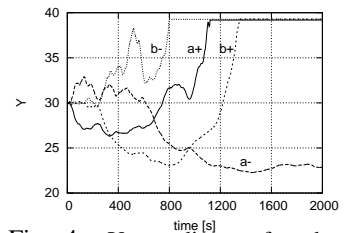


Fig. 4: Y coordinate of each axon tip at the direct connection stage

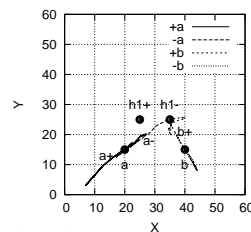


Fig. 5: The loci of axon tips at the 1st hidden neuron stage

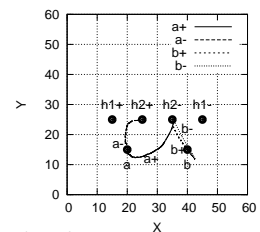


Fig. 6: The loci of axon tips at the 2nd hidden neuron stage

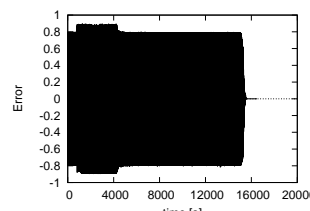


Fig. 7: Temporal change of the error

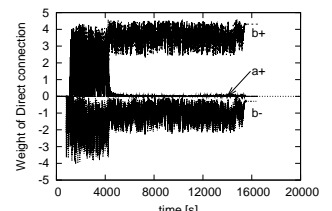


Fig. 8: Temporal change of the weight of direct connection

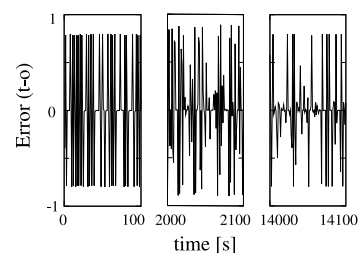


Fig. 9: Temporal change of the Error. Time is magnified.

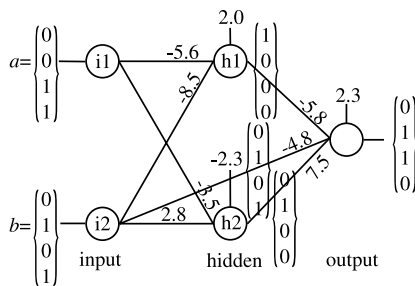


Fig. 10: Acquired structure in the case of one direct connection from input in the regular neural network after learning

Fig.7 および 8 は、誤差およびダイレクト結合の重みの時間的変化である。Fig.7 より、誤差は成長ニューロンの結合・学習にかかわらず、減少していないようにみえるが、詳細を表わすため、時間軸を伸ばして表示した Fig.9 を見ると、0、2,000、14,000s と、大きな誤差が出現する頻度が減少していることがわかる。

Fig.10 および 11 にそれぞれ、通常の NN および GNN の学習終了後の構造および重みを示す。Fig.11 の 4 個の隠れニューロンのうち、上 2 つは正の隠れニューロン、下 2 つは負の隠れニューロンである。かつこの値は、各入力パターンに対する出力を示す。

通常の NN の場合、各値の符号で正が多数なら出力が正となり、簡単に理解できる。10 回シミュレーションを行った結果、全てほぼ同じ重みとなり、2 個の隠れニューロン生成後に学習が終了する結果となった。

一方、GNN の場合は、出力の計算を各値の符号だけでは説明ができないことがわかった。GNN の場合、シミュレーションごとに構造と重みが変化するが、いずれの場合も 2 セットの隠れニューロンで学習できた。Fig.11 では、それぞれの隠れニューロンの出力を計算すると 'h1-' と 'h2+' がダイレクト結合、'h1+' がバイアスと同等の働きをそれぞれが行い、残りの隠れニューロンである 'h2-' が出力の ON/OFF を切り替えていることがわかった。

これより、GNN と通常の NN では、獲得できる構造に相違があることがわかった。しかし、GNN はタスクの学習可能な構造の自律的な獲得ができることを確認した。

通常の NN と GNN の違いを説明するには、次の 2 つを考慮する必要がある。一つは、通常の NN では、重みの符号が学習途中で反転することがあるが、GNN では、それができない。もう一つは、1 段階につき 1 つのニューロンは、1 つの結合しか形成することができない。1 つの段階につき正負 1 セットの隠れニューロンを生成するが、入力ニューロンの軸索がどちらかの隠れニューロンに結合を形成すると、もう片方へニューロンを成長させることができない。そのため、出力ニューロン数が複数の場合やリカレント構造など、高度な学習に必要な複雑な構造には対応できないと考えられるため、筆者らは改良が必要と考えている。

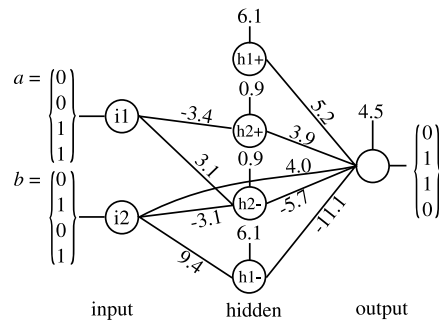


Fig. 11: Acquired structure in the case of growing neural network after learning

## 4 まとめ

成長型ニューラルネット (GNN) において、複数の隠れニューロンを生成し、学習する方法を提案した。その方法は、既存の隠れニューロンを移動させ、空いたスペースに新しい隠れニューロンを生成する。この方法を用い、ダイレクト結合数に制限を持たせた GNN で EXOR 問題を学習させたところ、2 セットの隠れニューロンの構造で学習ができた。獲得した構造では、GNN と通常の NN では、同等のニューロン数となったが、通常の NN の重みとは異なる結果となった。GNN において、各隠れニューロンは入力ニューロンとの結合が 1 つしか持たないために、このような成長をずっと考えている。今後、大規模で複雑なネットワーク構造の獲得を目指すには、成長ニューロンが複数の標的ニューロンとの結合できるようにするなど多くの問題点を解決する必要がある。

## 謝辞

この研究の一部は、日本学術振興会科学研究費補助金 #15300064 と #19300070 によって補助された。

## 参考文献

- 1) R.Kurino and K.Shibata, Growing neural network with generation of hidden units. *IEICE Tech.Report*, Vol. 103, No. 734, 2004. [109-114].
- 2) M.Kotani N.Shiba and K.Akazawa, designing multi-layered neural networks using genetic algorithm. *Trans. of SICE*, Vol. 34, No. 8, 1998. [1080-1087].
- 3) D.B.Fogel, An information criterion for optimal neural network. *IEEE*, Vol. 2, No. 5, 1991. [490-497].
- 4) R.W.Gundersen and J.N.Barrett, Neuronal chemotaxis: Chick dorsal-root axon turn toward high concentrations of nerve growth factor. *SCIENCE*, Vol. 206, , 1976. [1079-1080].
- 5) Brunso-Bechtold J.K. and Hamburger V, Retrograde transport of nerve growth factor in chicken embryo. *Proc. Natl. Acad. Sci. USA.*, No. 76, 1979. [1494-1496].
- 6) G.E.Hinton D.E.Rumelhart and R.J.Williams, Learning internal representations by error propagation. *The MIT Press, Cambridge, MA*, Vol. 1, , 1986. [318-362].
- 7) R.Kurino M.Sugisaka and K.Shibata, Growing neural network with hidden neurons. *Proc. of The 9th AROB*, 2004.