

隠れニューロンの分離を伴う成長型ニューラルネットワーク

栗野 竜輔 柴田 克成

大分大学工学部電気電子工学科

〒870-1192 大分市大字旦野原 700 番地

E-mail: kurino0129@yahoo.co.jp, shibata@cc.oita-u.ac.jp

あらまし 筆者らは、タスクに応じた合目的な構造の自律的獲得を目指し、従来のニューラルネットにニューロンの成長の概念を導入した“成長型ニューラルネット”を提案してきた。生体のニューロンの成長をヒントに、教師あり学習である誤差逆伝播学習法において、伝播してきた誤差信号を物質として拡散し、その拡散によって生じた濃度勾配にしたがい、ニューロンを成長させる。誤差信号を拡散するので、問題に応じた適切な構造の獲得が期待される。本稿では、この成長型ニューラルネットを隠れニューロンを含むニューラルネットに拡張するために、隠れニューロン分離法を提案する。入出力間の結合形成後、出力ニューロンから隠れニューロンを分離し、分離した隠れニューロンは伝播してきた誤差を拡散する。シミュレーションにおいて、非線形分離問題の“EXOR”を学習させ、適切な構造の形成に至った。

キーワード 成長型ニューラルネット, 隠れニューロン分離法, 拡散, NGF, バックプロパゲーション

Growing neural network with generation of hidden neurons

Ryusuke KURINO Katsunari SHIBATA

Faculty of Engineering, Oita University, 700 Dannoharu, Oita, 870-1192 Japan

E-mail: kurino0129@yahoo.co.jp, shibata@cc.oita-u.ac.jp

Abstract The authors have proposed “growing neural network” in order to obtain a purposive structure for a given problem autonomously. Getting a hint from the growth of biological neurons, the propagated error signal that is generated based on back propagation (BP) learning is diffused as a substance. A growing neuron extends its axon according to the concentration gradient that is formed by the diffusion. Since the diffused substance depends on the error signal, it is expected that an appropriate structure for a given problem can be obtained. In this paper, “Hidden Neuron Generation Method” is proposed to expand the growing neural network to a neural network with hidden neurons. After making the direct connections between input and output, the output neuron generates a hidden neuron. Then, the hidden neurons diffuse a substance based on the propagated error signal. In simulations, it was conformed that an appropriate structure could be formed though the learning of non-linear separation problem of “EXOR”.

Keyword Growing neural network, Hidden Neuron Generation Method, Diffusion, NGF, Back Propagation

1. まえがき

われわれ人間の脳内には、数十億のニューロンが存在し、それらニューロン同士が複雑に結合を行っている。それに対し、生体の脳をモデルとした人工ニューラルネットでは、関数近似に用いる際、3層構造を使用するが多い。その理由として、合目的に適切な構造を決定することが難しいこと、さらに、中間層

ニューロン数を増やせば、どんな連続関数でも正確に近似できる[1][2]ことなどが挙げられる。

一方、人間の脳内では、認識、記憶、制御などあらゆる機能が、脳内における複雑なニューラルネット構造によって実現されており、それらさまざまな機能は、視覚野などのセンサに近いところから運動野までの間において、いろいろなレベルの抽象状態表現として学習に基づいて実現されているのではないかと考える。

また、センサからモータまでの一連の過程をニューラルネットで構成し、強化学習などの学習を通して獲得された時、隠れニューロンにおいて有益な抽象状態表現を形成できることが示めされている[3]。これらの抽象状態表現が、人間と知能ロボットの間に未だ存在する大きなギャップを埋める一つの鍵になるのではないかと筆者らは考える。よって、リカレント構造を含めた複雑な構造を自律的に獲得することが、知能ロボットの”脳”において今後ますます求められるに違いない。

筆者らは、通常のニューラルネットにニューロンの成長の概念を導入した成長型ニューラルネット (NN) を提案してきた。そして、生体の脳内において、NGF(Nerve Growth Factor:神経成長因子)等の化学物質が、ニューロンの成長を促進する役割を担う[4]という知見を基に、誤差逆伝播学習 (BP learning) 法において、逆伝播してきた誤差信号に基づいて物質を拡散し、形成された濃度勾配にしたがい、軸索を伸長させる。誤差信号を拡散することで、与えられた問題に応じた適切な構造の獲得が期待される。そして、まず、簡単な論理関数である“AND”と“OR”を成長型 NN によって学習させ、2層構造の形成を確認した[5]。

しかし、非線形を必要とするその他の問題を解決するためには、隠れニューロンが必要である。前述の方法を用いて隠れニューロンを含む構造を形成しようとしても、(1)隠れニューロンは入力信号の情報を持たず、軸索を適切に成長させることができない、(2)隠れニューロンは誤差信号の情報を持たず、入力ニューロンは隠れニューロンに軸索を成長させることができないという2つの問題点がある。したがって、隠れニューロンを含む構造を形成するためには、初めに、入力または出力ニューロンのどちらかと、何らかの方法で結合を形成することが必要である。

そこで、まず、入力-隠れニューロン間の結合を形成するという考えに基づいた伸長促進物質拡散法を提案し[6]、非線形分離問題である”EXOR”の学習に適用した結果、適切な構造の形成を確認した。しかし、多層構造形成時には、ほとんどの構造を誤差信号によらずに決めることになるため、柔軟な構造決定が難しい。

本稿では、誤差信号のみで隠れニューロンを含む構造の形成が可能な隠れニューロン分離法を提案する。そして、シミュレーションにおいて、“EXOR”を学習させ、適切な構造が形成されるかどうか検証を行う。

2. 隠れニューロン分離法に基づく成長型 NN

われわれ人間の脳内には、複雑な多層構造はもちろん、反射のように、脊髄を経由するだけの単純な経路も存在する。筆者らは、生体における脳内のネットワーク構造は、2層、3層、4層と層を重ねて形成され

ていくと推測し、形成されたネットワークにおいて、層数が少ないところでは、反射のような単純な処理を、多層部分では、思考などの複雑な処理を行っているのではないかと考える。そこで、本稿では、まず入出力間の2層構造を形成し、その後、隠れニューロンを分離生成して、入力-隠れ-出力間の3層以上の構造を順次形成することができるアルゴリズムとして、隠れニューロン分離法を提案する。

2.1 アルゴリズム

図1に隠れニューロン分離法に基づく隠れニューロンを含む成長型 NN の概念を示し、図2にそのアルゴリズムの流れを示す。初めに、出力と誤差信号を計算する。もし、出力ニューロンが十分な結合を持っていないければ、出力ニューロンは誤差信号を物質として拡散し、入力ニューロンはその拡散によって形成された濃度勾配にしたがい軸索を伸長させ、出力ニューロンとの結合を求める。そして、入出力間の結合が形成されると、その重み値をBP法で0から学習する。

次に、出力ニューロンは、入出力間の結合だけでは誤差が残る場合、隠れニューロンを分離する。その時、隠れ-出力ニューロン間の初期重み値を0とする。そして、分離した隠れニューロンは、隠れ-出力ニューロン間の重み値を学習しつつ、伝播してきた誤差を拡散し、形成された濃度勾配にしたがい、入力ニューロンは軸索を伸長させる。入力ニューロンが隠れニューロンと結合を形成した場合も、重み値0から通常のBP法にしたがい、学習を始める。

また、ここでは、基本的な成長機能の確認に重点を置くため、以下のような設定とした。(1)隠れニューロンの分離時期は、人間が決定した。(2)隠れニューロンの分離後、出力ニューロンは拡散を止めた。(3)隠れニューロンの分離時に、入力-出力ニューロン間の距離を大きくし、その間に隠れニューロンを配置した。そして、入力ニューロンは、既に出力ニューロンと結合している部分は、そのまま残して学習を続けるとともに、新たに軸索を伸長させた。

Diffusion of error signal

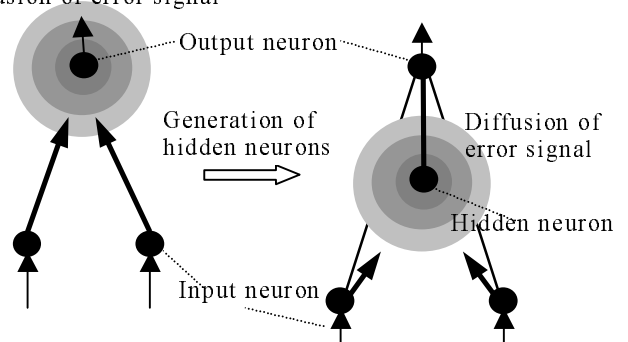


図1 隠れニューロン分離法に基づいた成長型 NN
Fig.1 Growing NN based on the Hidden Neurons Generation Method.

- 1.各ニューロンの位置の決定
- 2.各ニューロン間の結合の有無を決定
- 3.入力と教師の設定
- 4.出力と誤差の計算
- 5.誤差信号の伝播
- 6.誤差信号の拡散
- 7.軸索伸長計算
- 8.各ニューロン間の結合の有無を確認
- 9.重み値の更新（結合したニューロンに対して）
- 10.拡散ニューロンは、学習が進まないと判断した時、隠れニューロンを分離
- 11.ステップ3に戻る

図2 隠れニューロン分離法に基づいた成長型 NN の流れ
Fig.2 Flow of the Growing NN based on the Hidden Neurons Generation Method.

(1) 拡散計算

成長型 NN では、BP 法において計算された誤差信号を物質として拡散する。誤差信号は正または負になりうるが、現実的な問題として、負の拡散は考えにくいので、ここでは、以前の論文[5]と同様に、正と負の誤差信号に対する物質を用意し、独立して拡散する。また出力と隠れニューロンにおける誤差信号 δ の計算は、

$$\delta_j = -\frac{\partial E}{\partial net_j} = (d_j - o_j) \cdot f'(net_j) \quad (\text{Output neuron}) \quad (1)$$

$$\delta_j = f'(net_j) \sum_i w_{ij} \cdot \delta_i \quad (\text{Hidden neuron}) \quad (2)$$

とする。ここで、 d_j は教師信号、 o_j は出力、 net_j はニューロン j の内部状態、 $f'(net_j)$ は出力関数の微分値、 E は誤差関数、 j は隠れまたは出力ニューロンのユニット番号、 i は結合元ニューロンのユニット番号とした。そして、計算した誤差信号を湧き出しとして、正負2種類の物質の拡散を計算する。それぞれの物質を u^p (p : 正の誤差信号)、 u^n (n : 負の誤差信号) と記述し、それぞれの拡散は、拡散方程式にしたがい、

$$\frac{\partial u_{x,y}^k}{\partial t} = \text{div} + D \nabla^2 u_{x,y}^k \quad (3)$$

として計算する。ここで、 $u_{x,y}^k$ は座標 (x,y) での濃度、 $k=p$ または n 、 div は物質の湧き出し、 D は拡散係数とする。出力または隠れニューロンが存在する場所では、 $\delta > 0$ の時、 u^p に対して $\text{div} = \rho \delta$ 、 u^n に対して $\text{div} = 0$ 、 $\delta < 0$ の時、 u^n に対して $\text{div} = -\rho \delta$ 、 u^p に対して $\text{div} = 0$ とし、拡散物質の量は誤差信号 δ に比例させた。また、それ以外の場所において、

$$\text{全ての物質に対して } \text{div} = 0$$

とする。ここで、 ρ は湧き出し係数とする。前述のように、簡単のため、隠れニューロン分離後は、出力ニューロンの拡散は止める。

(2) 軸索伸長計算

入力ニューロンの出力と軸索先端における濃度勾配にしたがい、軸索は伸長する。ここでは、生体のニューロンにおいて、シナプスで発生する神経伝達物質によって、興奮または抑制のどちらの結合を形成するニューロンかが決められていると言われている[7]ことを考慮し、正負2種類のニューロンを用意した。正のニューロンは正の結合のみを、負のニューロンは負の結合のみを形成する。また、軸索伸長計算では、計算時間を離散化したことにより、濃度勾配が急激に変化するところで、軸索の伸長が止まるという問題を解決するために、濃度勾配とニューロンの出力との積に一次遅れを導入し、ニューロンの成長 A と定義した。また、拡散遅れの調整のため、ニューロンの出力にも一次遅れを導入して、ニューロンの状態 S と定義した。そして、軸索伸長の計算は、

$$\tau_s \frac{dS_i}{dt} = -S_i + o_i \quad (4)$$

$$\tau_a \frac{dA}{dt} = -A_i + \nabla(u^p - u^n) \cdot \text{flag}_i \cdot S_i \quad (5)$$

$$\frac{da_i}{dt} = \xi \cdot A_i \quad (6)$$

とする。ここで、 \mathbf{a} は軸索先端における位置ベクトル、 ξ は軸索伸長係数、 flag は正のニューロンの時 1、負のニューロンの時 -1 とした。 τ は時定数である。

(3) 重み値の更新

ニューロン間の結合が形成された後、および、隠れニューロンを分離した時には、いずれもシナプスは、重み値 0 から学習を始める。重み値の更新は、

$$\frac{dw_{ji}}{dt} = -\eta \frac{\partial E}{\partial w_{ji}} = \eta \cdot \delta_j \cdot o_i \quad (7)$$

として、通常の BP 法と同様の計算を行う。ここで、 w_{ji} はニューロン i からニューロン j の重み値とする。

3. シミュレーション

隠れニューロン分離法を伴った成長型 NN のアルゴリズムを用いて、非線形分離問題である EXOR を学習させ、パターンの提示順を固定した場合と、ランダムにした場合でシミュレーションを行った。

3.1 設定

シミュレーションは、 $0.6(\text{mm}) \times 0.6(\text{mm})$ の領域で行い、格子間隔 $\Delta x = 0.01(\text{mm})$ として、 61×61 の格子点で拡散物質の濃度分布を離散化して差分法に基づいて近似計算した。また、時間 $\Delta t = 0.001[\text{sec}]$ で離散化して計

算した。図3にシミュレーション環境と各ニューロンの位置を示す。図中の白抜き矢印は、隠れニューロン分離による入力ニューロンの移動を示す。正と負の入力ニューロン"+a"と"-a"、"+b"と"-b"は同位置に配置した。また、500(sec)経過後に隠れニューロンを分離し、出力ニューロンの下(30,25)に配置した。図4はパターン提示順をランダムにした時の入力と教師の時間変化の一例を示す。図のように、1(sec)ごとに学習パターンを切り替える。パラメータは、拡散係数 $D = 0.01$ (mm^2/sec)、湧き出し係数 $\rho = 10$ ($1/\text{sec}$)、軸索伸長係数 $\xi = 0.1$ (mm^2/sec)、時定数 $\tau_a = 10$ (sec)、 $\tau_s = 0.1$ (sec)、学習係数 $\eta = 0.2$ ($1/\text{sec}$)と設定した。

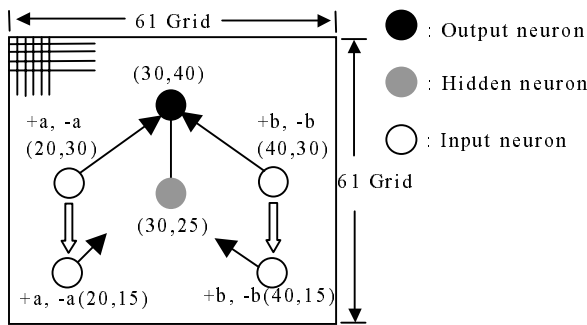


図3 シミュレーション環境と各ニューロンの位置
Fig.3 Simulation environment and position of each neuron.

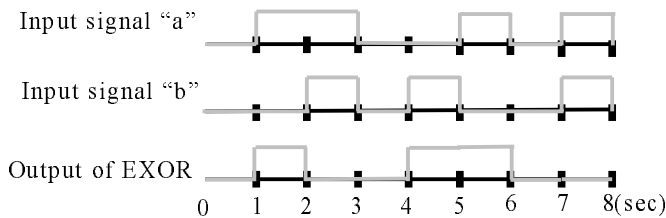


図4 学習パターンの時間変化の例(EXOR)
Fig.4 Timing chart of a sample training pattern (EXOR)

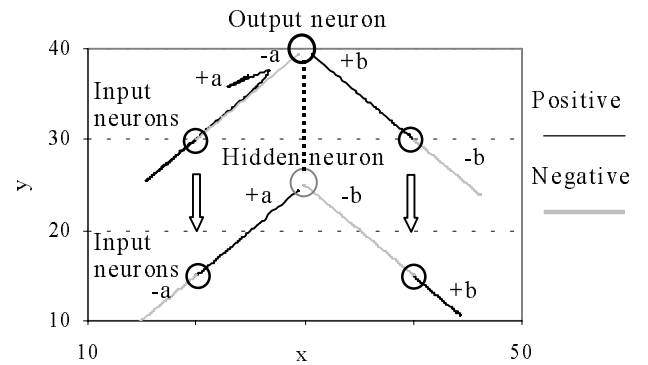
3.2 パターン提示順を固定した場合

図5にニューロンの分離と成長の様子を示す。ここでは、分離した隠れニューロンは正のニューロンとした。図5(a)は各入力ニューロンの軸索先端の軌跡を示す。y座標30から出ている軌跡が隠れニューロン分離前であり、y座標15から出ている軌跡は隠れニューロンの分離に伴い、入力ニューロンを移動させた後の軌跡である。図(b)は各入力ニューロンの軸索先端におけるy座標の時間変化を示す。また、表1は結合形成前における学習パターンと発生した誤差について示し、図6は誤差の変化、図7は入出力間の重み値と入力-隠れ間の重み値と隠れ-出力間の重み値の変化を示す。図中の矢印は、各入力ニューロンが結合した時のタイ

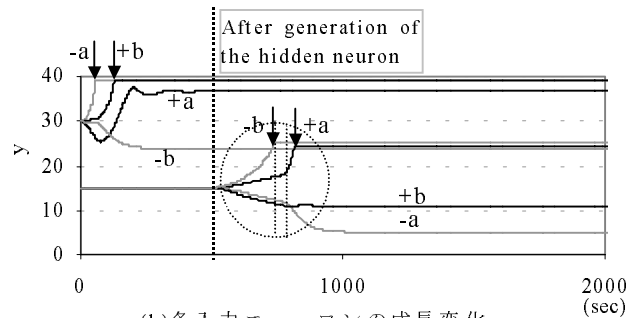
ミングを示す。また、結合が全くない場合でも、出力ニューロンのバイアスが学習を行うため、パターンが切り換わっても誤差が増大してもすぐに誤差が減少する。そのため、図6では判りにくいですが、誤差の変化は実際にはパルス状になっている。

図5(a),(b)をみると、隠れニューロンからの誤差信号の拡散にしたがい、各入力ニューロンが軸索を伸ばさせる。初めに、入力"-a"のニューロンが80(sec)で出力ニューロンと結合を形成した。Table1を見ると、入力'a'が1の時、誤差は0か負となる。パターン(3)で誤差が0になるのは、パターン(2)でバイアスが学習し、出力がほぼ0.9となっているからである。つまり、本来入力-出力間に相関がないEXORでも、提示順を固定したことによって、入力と誤差の間に相関が生じたことになる。これによって、早い時期に結合が形成できたと考えられる。その後、入力"+b"と"+a"のニューロンが、出力ニューロンに向かって軸索を成長させ、入力"-b"は逆の向きに軸索を成長させた。その後、入力"+b"が144(sec)で出力ニューロンと結合を形成した。

そして、入力ニューロンが出力ニューロンと結合後、500(sec)経過後、隠れニューロンが分離され、隠れニューロンは伝播してきた誤差信号に基づき、誤差信号の拡散を行った。その拡散によって形成された濃度勾配にしたがい、各入力ニューロンは軸索を伸ばさせた。



(a)各入力ニューロンの軸索の軌跡
(a)The loti of each input neuron's axon.



(b)各入力ニューロンの成長変化
(b) Change of the growth of each input neuron.

図5 各入力ニューロンの成長
Fig.5 The growth of each input neuron.

まず、入力“-b”のニューロンが 740(sec)で隠れニューロンと結合を形成した。その後、入力“+a”のニューロンが軸索を隠れニューロンに向けて成長する速度を大きくした。一方、入力“-a”と“+b”のニューロンは隠れニューロンに対して、逆の向きに成長した。その後、入力“+a”のニューロンは、815(sec)で出力ニューロンと結合を形成した。そして、誤差が最終的にほぼ 0 となり、入力“-a”と“+b”のニューロンの成長は止まった。

また、図 5(a)を見ると、隠れニューロン分離前の入力“+a”の軸索先端の軌跡が、入出力方向以外に曲がっている。これは、解析した結果から、拡散ニューロンの近くで、かつ、軸索の戻りが大きい時に、離散化誤差の影響が出るためとわかった。また、図 5(b)中の円で囲まれている部分を見ると、入力“-b”が結合後、すぐに重み値の学習が進み、誤差のパターンが変化したにもかかわらず、少し時間をおいてから各ニューロンが成長の仕方を変更している。この遅れは、入力“-b”の結合が重み値 0 から学習を始めており、結合直後は重み値が小さく、誤差があまり変化しないことによる。

表 1 各入力ニューロンに対する教師信号と誤差
Table 1 The training signal and the error for each input pattern.

Pattern	Input 'a'	Input 'b'	Training	Error
(1)	0	0	0.1	0
(2)	0	1	0.9	Positive
(3)	1	0	0.9	0
(4)	1	1	0.1	Negative

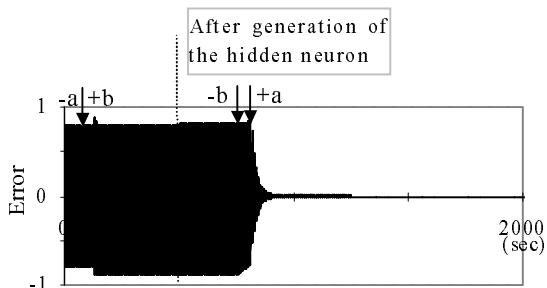


図 6 誤差の変化
Fig. 6 Change of the error.

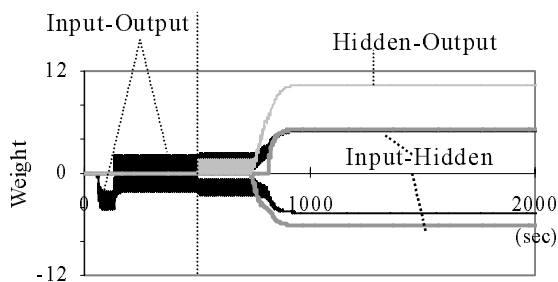


図 7 重み値の変化
Fig. 7 Change of the weights

3.3 パターン提示順をランダムにした場合

パターンの提示順を固定した場合には、入力と誤差の間で、相関が生じることで適切な構造が形成されたと考えることができる。一方、入力をランダムにした場合には、平均的にみると軸索の伸長と縮退に差が出ず、成長がより難しくなると予想される。

ここでは、正と負 2 つの隠れニューロンが同時に分離されるとし、分離するタイミングは 10000(sec)後とした。また、分離後のそれぞれの位置を正(25,25)、負(35,25)として配置した。それ以外の設定は、(3.1)の設定と同じである。

図 8 にニューロンの分離と成長の様子を示す。図 8(a)は各入力ニューロンの軸索先端の軌跡を示し、図 (b)は各入力ニューロンの軸索先端における y 座標の成長変化を示す。図中の矢印は、各入力ニューロンが結合した時のタイミングを示す。また、図 9 に誤差の変化、図 10 に入力-出力間の重み値、図 11 に入力-隠れ-出力間におけるそれぞれの重み値の変化を示す。

図 8 を見ると、初めに、入力“-b”の軸索が伸長し、798(sec)で出力ニューロンと結合した。これは、パターン提示順のランダムな要素によって、偶然、入力“-b”が伸長するパターンが連続したと考えられる。乱数系列を変えると、別の入力に先に結合し、また、この結合によって、入力と誤差の間に相関関係が形成され、入力“+a”と“+b”のニューロンが出力ニューロンに向けて軸索を成長させた。そして、入力“+a”と“+b”のニューロンが、1120(sec)と 1354(sec)で、それぞれ出力ニューロンと結合を形成した。その結合後、入力“-a”のニューロンが、8678(sec)で出力ニューロンと結合を形成した。

図 10 に示すように、それぞれの入力ニューロンが出力ニューロンと結合した後、重み値 0 から学習を始めた。そして、10000(sec)経過後、出力ニューロンから隠れニューロンが分離され、その隠れニューロンは、伝播してきた誤差に基づいて誤差信号を拡散した。その拡散にしたがい、入力“+a”と“+b”のニューロンが負の隠れニューロンに向けて軸索を成長させた。一方、入力“-a”と“-b”のニューロンは逆の向きに軸索を成長させた。そして、入力“+b”のニューロンが、13882(sec)で負の隠れニューロンと結合を形成した。その後、入力“+a”は逆の向きに軸索を成長させ、また、入力“-a”は、負の隠れニューロンに向けて軸索を成長させた。その後、入力“-a”が負の隠れニューロンと 17077(sec)で結合を形成した。その結合形成後、各ニューロンの重み値が収束し、また、誤差がほぼ 0 になり、入力“+a”と“-b”のニューロンの成長が止まった。

また、図 8 を見ると、入力“+a”のニューロンの軸索先端における軌跡が大きく曲がっている。これは、正

と負の隠れニューロンからの拡散の影響が混在するため、両方の影響を受けた結果と考えられる。さらに、図 11 を見ると、両隠れニューロンの分離後、出力と負の隠れニューロン間の重み値が正の場合よりも大きくなっている。このため、正負の誤差信号の拡散が混在する中で、負の拡散が強まり、負の隠れニューロンとの結合のみが形成されたと考えることができる。

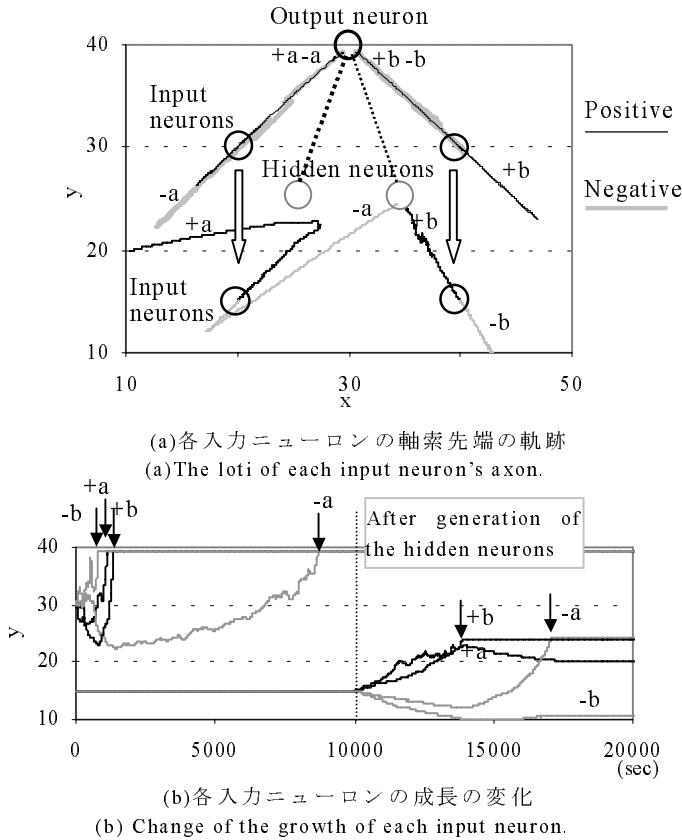


図 8 各入力ニューロンの成長
Fig.8 The growth of each input neuron.

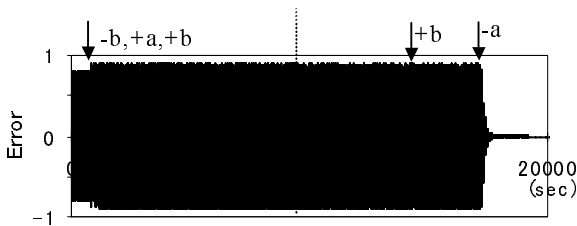


図 9 誤差の変化
Fig.9 Change of the error.

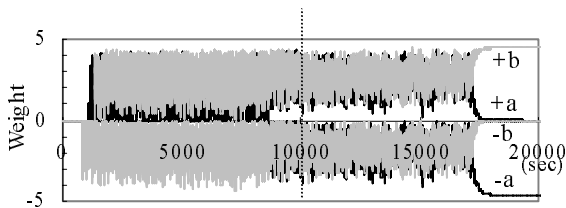


図 10 重み値の変化 (入力ー出力)
Fig.10 Change of the weights between the input neurons and the output neuron.

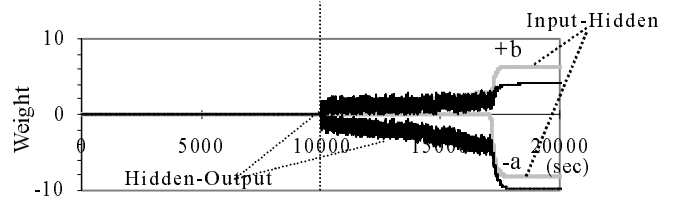


図 11 重みの変化 ((入力ー隠れ)、(隠れー出力))
Fig.11 Change of the weights
((Input- Hidden), (Hidden-Output))

あとがき

本稿では、隠れニューロンを含む適切な構造が形成可能である隠れニューロン分離法に基いた成長型 NN を提案した。シミュレーションにおいて、出力からの誤差信号の拡散により、入出力間の結合後、出力ニューロンは、隠れニューロンを分離し、その隠れニューロンは伝播してきた誤差を拡散した。そして、隠れニューロンを必要とする EXOR 問題を解くことができた。しかし、拡散誤差と伝播誤差の分配の問題や、学習に基づく隠れニューロンを分離する時期の決定などの問題がある。これらの問題点を解決することで、より複雑な構造の形成を求められる場合にできると考える。

謝辞

本研究は、日本学術振興会科学研究費補助金基盤研究(B)(14350227,15300064)の補助の下で行われました。ここに謝意を表します。

文 献

- [1] K.Hornik, Multiayer Feedforwad Networks are Universal Approximators, Nueral Networks, Vol.2, pp.359-366 (1989).
- [2] K.Funahasi, On the Approximate Realization of Continuous Mappings by Neural Networks, Neural Networks, Vol.2, pp.183-192 (1989)
- [3] 柴田克成、岡部洋一、伊藤宏司、ニューラルネットワークを用いた Direct-Vision-Based 強化学習ーセンサからモータまでー、計測自動制御学会論文 文集、Vol.37, No.2, pp.168-177 (2001)
- [4] 畠中寛、神経成長因子ものがたり、羊土社(1992)
- [5] R.Kurino, M.Sugisaka, K.Shibata, Growing neural network for acquisition of 2-layer structure, Proc.of IJCNN'03, pp.2512-2518 (2003)
- [6] R.Kurino, M.Sugisaka, K.Shibata, Growing neural network with hidden neurons, Proc. of the AROB'04, Vol. 1, pp. 144-147 (2004)
- [7] 後藤秀機、神経と化学伝達、東京大学出版会(1988)