遺伝的アルゴリズムを用いた

階層型ニュラル・ネットワークの

パターン学習構造の簡単化

森元映治[1]・平岡淳[2]・大崎榮喜[1]・中村誠[1]・和田憲造[3]
[1]水産大学校海洋機械工学科
[2]水産大学校研究科
[3]機械工学科

三層構造の階層型ニューラル・ネットワークにマトリクス上のパターンを学習させ、学習の過程でネットワークの構造を推移させることにより高速化をはかることを検討した。パラメーターの探索は遺伝的アルゴリズムによりおこなうため、各ユニットのシナプス荷重と閾値を遺伝子の染色体に対応づけた。451 の遺伝子座からなる 10 組の染色体を用い学習の効率化を高めるため突然変異を考慮して最適化をおこなった。学習に影響をおよぼす因子として初期値の乱数幅、突然変異条件、ネットワーク構造条件を変化させ、それらの効果を数値的に検討した。

Key Words : genetic algorithms, neural networks, architecture, optimization, simplification structure,

1.はじめに

ニューラル・ネットワークの教師あり学習では与えられたデータに対してシナプス荷重と閾値を変更し、対象問題のもつ変数関係を最もよく説明する方向へ探索をおこなう。このためネットワークの規模が大きくなるにしたがって探索パラメータ数が増加し、必然的に学習に要する時間が増加することとなる。特にパターン学習をおこなう場合 [1]の基本的方法として各画素に対応する値をネットワーク入力に対応させるとネットワーク構造が大きくなり、全体の計算量が増え学習を完了するまでの時間が増大する。そのため、本報では計算量軽減 [2)-4)]の一方法として学習中情報伝達に寄与する割合の低下したネットワーク結合を徐々に除き、より簡略な構造へ移行させることによって高速化を計ることを検討した。

2.アルゴリズム

(1)ニューラル・ネットワーク

マトリクス上に描かれたパターンを認識するために、パターンに数値を割り当て両者の対応関係を学習させる。マトリクス各格子点の値をニューラル・ネットワーク入力、パターン値を教師データとして与える。学習はダイナミクスを含まないため階層型を用いる。初期状態では入力層、中間層、出力層の各シナプスは隣接する層のシナプスと完全に結合している。学習過程では入力デ―タと対応する教師データを反復して与え、ネットワーク出力と教師データの誤差を減ずる方向へシナプス荷重と閾値を探索パラメータとして適宜変化させてゆく。これをおこなう代表的方法としては誤差逆伝播法 [5]があるが、ここでは遺伝的アルゴリズムを用いて探索する。

(2)遺伝的アルゴリズム

遺伝的アルゴリズムは遺伝子座に対象とする問題の情報を配列し生物の遺伝子生成を模擬した行程をとらせることにより評価関数を最小化(もしくは最大化)させる探索手法である。そのアルゴリズムを以下に示す [6]。

Step1:染色体の決定染色体の要素列に、対象とする問題の情報(数値、文字)を割り振る。

Step2:初期染色体集団の決定前 Step で決めた遺伝子型をもとに、複数個の要素列の異なる染色体の生成を行う。

Step3:各個体の適応度の評価あらかじめ定めた方法により、生成された染色体について適応度の計算を行う。

Step4:淘汰処理　得られた適応度をもとに、評価の低い染色体を、一定の基準にしたがって排除し消滅させる。

Step5:増殖処理　淘汰処理によって排除・消滅した個体に代えて、適応度の高い個体と同じ内容を持った染色体を淘汰処理した分だけ増殖させる。

Step6:交叉　　2 つの染色体間で遺伝子の組み替えを行い、先代の形質を受け継いだ個体を生成する。

Step7:突然変異　遺伝子の一部を強制的に特定の確率・方法で変化させ新しい個体の生成を行い、染色体集団としての多様性を持たせ、よりよい解をもつ個体の発生を促す。

Step8:繰り返し　Step3 に戻り、次世代として同様の手順を繰り返す。設定された目標値、または世代数を満たすまで以上の処理を繰り返す。

(3) パターン学習

学習パターンの画素数およびパターン数は今回使用した処理系の制約から 4×4 の正方マトリクス状の 16 画素からなる 4 パターンとした。したがって各画素をニューラノレ・ネットワークの 16 の入力シナプスに対応させた。4 種のパターンに対応させた数値を教師データとし、ネットワーク出力と比較する。中間層のユニット数を 25 個とし、1 出力を持つ 16×25×1 の階層型ネットワークを構成する。遺伝的アルゴリズムではシナプス荷重、閾値の総数 451 データを染色体の遺伝子座に対応させる。交叉方法としては単数点交叉を用い、学習の効率化のために突然変異を与えた。

学習では 10 本の染色体を用いた。適応度は教師データと出力データの差をとり 4 パターン分の合計から算出する。シナプス荷重と閾値の更新に遺伝的アルゴリズムを用いる。すなわち、ニューラル・ネットワークの 1 組の探索パラメータ値を遺伝情報とし 1 本の染色体を構成し探索を実行する。学習過程において適応度がある設定値に達したとき情報伝達に寄与する割合の低下したシナプスを切断しネットワーク構造の簡単化をはかる。

学習に対して影響を与える因子として、遺伝的アルゴリズムに関しては初期値の乱数幅 I_r、突然変異を起こす染色体の数 C_h、突然変異を起こす遺伝子の個数 G_e、突然変異において変異を与える乱数幅 R_a を、またニューラル・ネットワークについてはシグモイド関数の傾斜を決定する係数 I_c、ネットワーク構造を変更する時に消去するシナプス荷重の上限値 E_r、を考慮した。

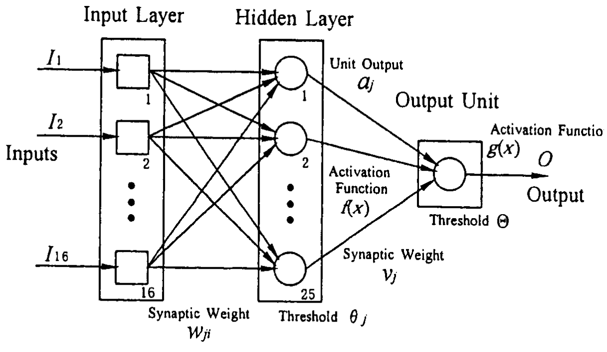

Fig.1　Neural Network Structure

(4) 計算手順

計算のためのネットワーク構造とプログラムのフローチャートを各々Fig.1、Fig.2 に示す。

①初期値設定

ニューラル・ネットワークの探索パラメータ w_{ji}、v_j、θ_j、Θの初期値を与える。w_{ji}、v_j は、おのおの入力の値と乱数によって任意の幅を持つ値に設定する。

②外部入力とパラメータ入力

学習条件を設定するパラメータ I_r、C_h、G_e、R_a、I_c、E_r、及び 4×4 の教師出力と入力パターンをデータファイルにより与える。

③出力計算

現時点の染色体での出力を以下のように計算し、適応度を算出する。

q 番目の入力パターンに対する p 番目の染色体について考え、中間層および出力ユニットにおけるシグモイド関数を次式で定義する。

$$f(x) = \frac{1}{1+\exp(-x)} \quad (1)$$

$$g(x) = \frac{3}{1+\exp(-I_c \cdot x)} \quad (2)$$

中間層のユニット出力;

$$a_j^{pq} = f\left(\sum_i I_i w_{ji}^{pq} - \theta_j^{pq}\right) \quad (3)$$

ネットワーク出力;

$$o^{pq} = g\left(\sum_j a_j^{pq} v_j^{pq} - \Theta^{pq}\right) \quad (4)$$

以上より適応度を次式で求める。

$$T_p = \sum_q \left| o^{pq} - d^q \right| \quad (5)$$

④順位入れ替え

ニューラル・ネットワークからのデータは 3 次元のディメンションを使用しているので遺伝的アルゴ

リズムで使用する染色体の 1 次元データとして w_{ji}、v_j、θ_j、Θの順に並べ替える。

⑤淘汰と増殖

10 本の染色体に適応度の高い順から番号を割り振る。淘汰、増殖処理では、エリート保存選択法を単純化したものを用いた。エリーーート保存選択法では、適応度の最も低いものからいくつかの染色体を適応度の最も高い物から順に置き換える。今回は、10 番目と 9 番目を淘汰し、1 番目と 2 番目を増殖した。

⑥交叉

交叉では、単数点交叉法を使用し、1 番目と 10 番目、2 番目と 9 番目という順で選択し、2 本の染色体を一つの組として交叉を行う。交叉位置は 2～450 の値を 5 回の乱数により決定した。交叉位置を境に染色体の要素である遺伝子を相互に組み替えた。

⑦突然変異

今回の計算では C_h、G_e、R_a を任意に与えることができる。C_h により低い方から与えられた値だけ染色体に突然変異を発生させた。G_e、により 1 本の染色体中の遺伝子にいくつの突然変異を起こさせるかが設定できる。突然変異を起こす場所は乱数により選び出す。R_a は乱数により選ばれた値を元の値に加えて新しい遺伝子を生成する。

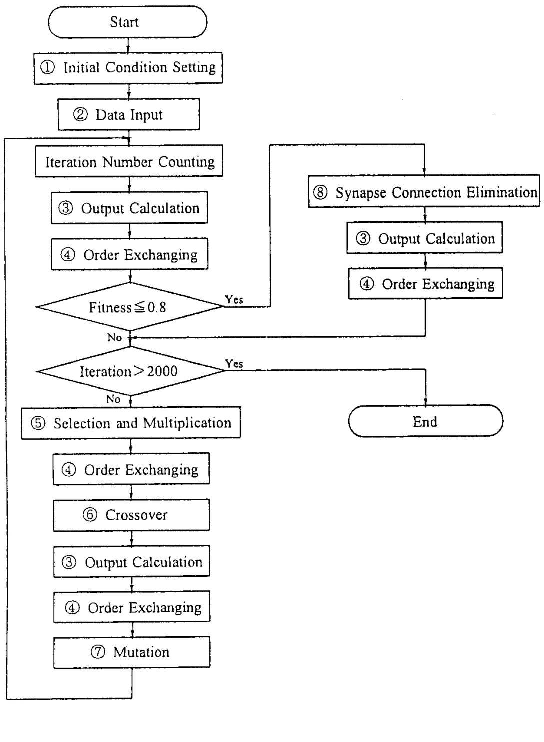

Fig.2 Flowchart

⑧シナプス荷重の消去

染色体の適応度が設定した値以下になったとき、染色体を構成する遺伝子情報中のシナプス荷重のうち消去パラメータ E_r、以下であるものを選び出し、結合を切断する。

3. 計算結果

ネットワークと遺伝的アルゴリズムの条件を変えてシミュレーションを行った。最終的なネットワーク出力、構造状態を Table.1 にまとめて示した。遺伝的アルゴリズムにおける条件としては 4 種のパラメータが学習に与える効果を検討した。

I_iの効果: ： 初期乱数幅 I_iを 1、1.5、2、4 に設定した場合の学習状況を比較検討した(Fig.3、Fig.4、Fig.5、Fig.6)。I_iが小さい場合には学習速度が遅く

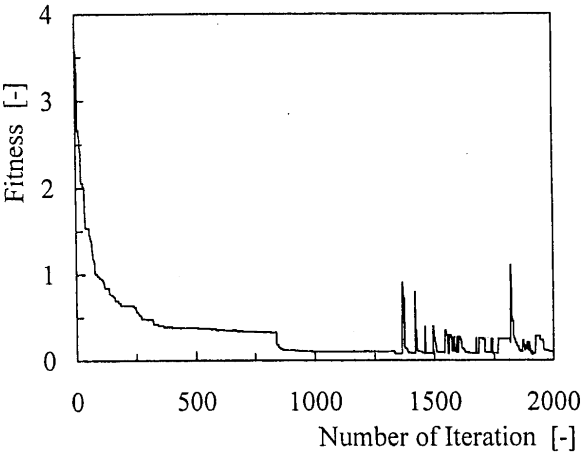

Fig.3 Leaning process with network structure modification (Result 1)

(I_i =1, C_h =2, G_e =5. R_a =5, I_c =0.5, E_r=0.8)

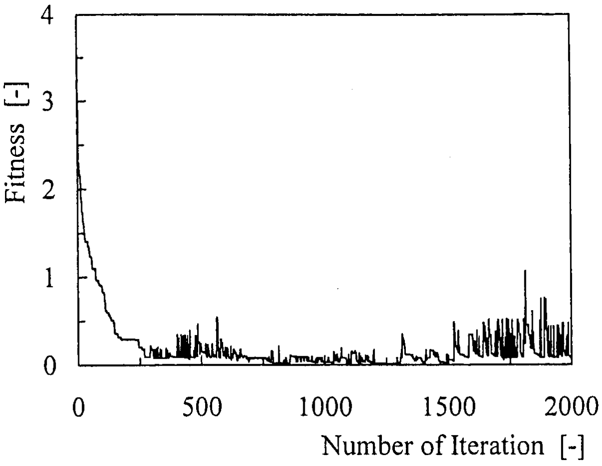

Fig.4 Learning process with network structure modification (Result 2)

(I_i =1.5. C_h =2. G_e =5, R_a =5, I_c =0.5, E_r=0.8)

構造の変更開始も遅いが最終的には大きな構造変更が行われている。しかし出力の誤差が大きく適切な構造に至っていない。I_i が大きい場合には学習が早く進み、構造変更も早期の段階から行われるが、遺伝子構造が安定せず従って適応度の安定が悪い。構造変更は最終的には少ないが誤差は大きい。乱数幅が両者の中間の状態では適度に構造変更が行われ、出力誤差も小さく良好な学習が達成されたと思われる。

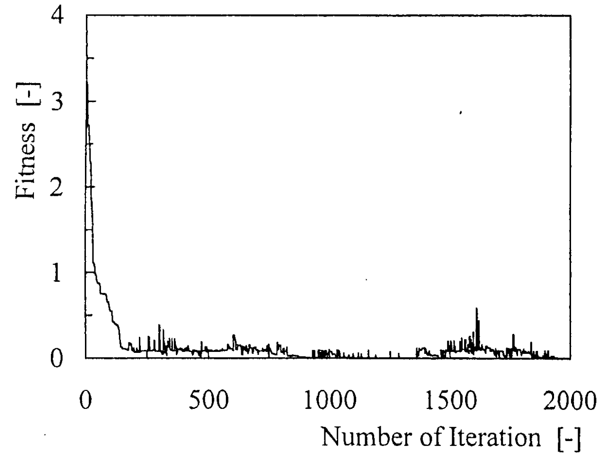

Fig.5 Learning process with network structure modification (Result 3)
(I_i =2, C_h =2, G_e =5, R_a =5, I_c =0.5, E_r=0.8)

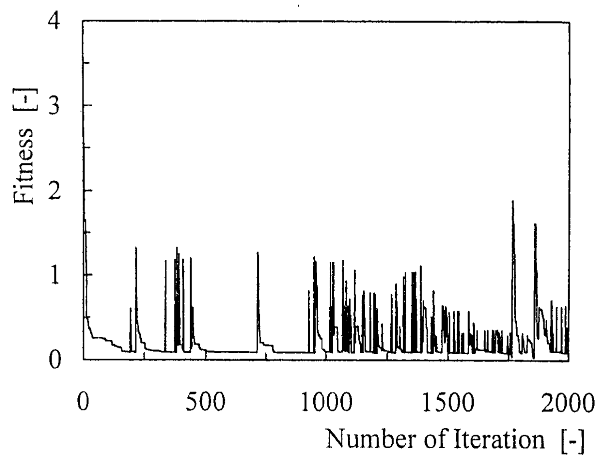

Fig.6 Learning process with network structure modification (Result 4)
(I_i =4, C_h =2, G_e =5, R_a =5, I_c=0.5, E_r=0.8)

C_h の効果:突然変異をおこす染色体の数を 1、2、5 として比較検討した(Fig.7、Fig.4、Fig.8)。C_h が大きくなるにつれて学習速度は上昇する傾向を示した。構造の変更も C_h が大きいほど学習の初期から行われる。変更の程度は一様ではないが出力誤差には大きな違いは見られなかった。

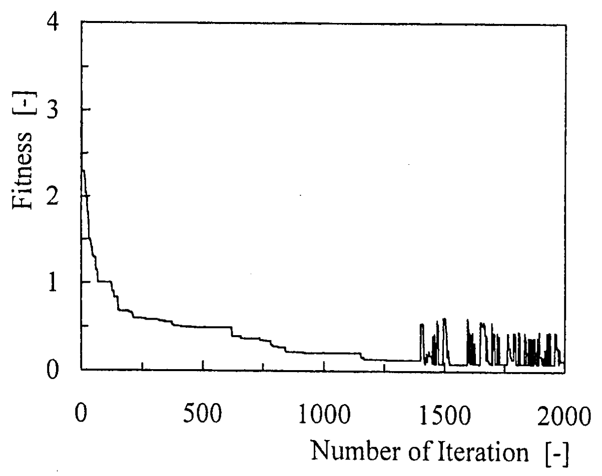

Fig.7 Learning process with network structure modification (Result 5)
(I_i = 1.5, C_h=1, G_e =5, R_a =5, I_c =0.5, E_r=0.8)

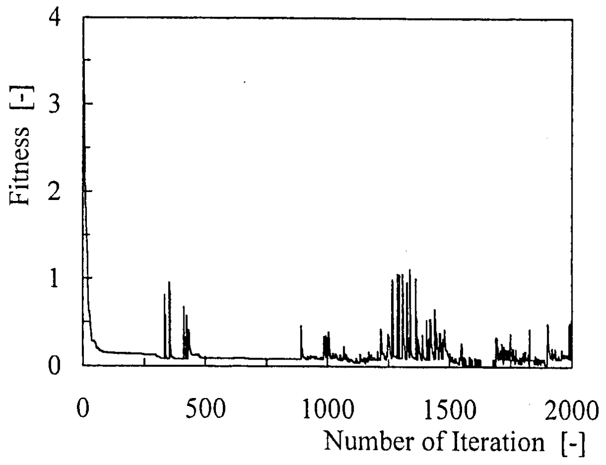

Fig.8 Learning process with network structure modification (Result 6)
(I_i =1 .5, C_h =5,G_e =5, R_a =5, I_c =0.5, E_r =0.8)

G_e の効果:突然変異をおこす遺伝子の個数を 5、10、30 として検討した(Fig.4、Fig.9、Fig.10)。G_e が大きい場合消去されたシナプス数はかなり多い。しかし出力誤差は他の結果と比較しても今回の計算範囲で特に良好とはいえず、改善が適切に行われたとは言い難い。これは突然変異をおこす遺伝子が増えると適切な適応度値を持つ染色体が出現しがたく、世代交代を重ねても改善されないままある状態に飽和してしまうためと考えられる。

R_a の効果:突然変異のため変更を加える乱数の幅を 3 から 10 と変えると最適な値が存在し、それ以外では学習速度が低下した。(Fig.11、Fig.4、Fig.12)R_a が小さい場合は特に遅く構造変更はうまく行われない。また大きすぎる場合には適当な値をとる確率が低下し、学習がうまく行われなかった。

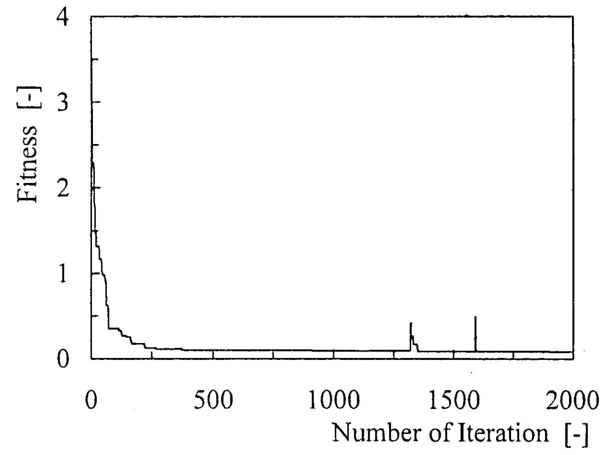

Fig.9 Learning process with network structure modification (Result 9)
(I_i =1.5, C_h =2, G_e =10, R_a =5, I_c =0.5, E_r=0.8)

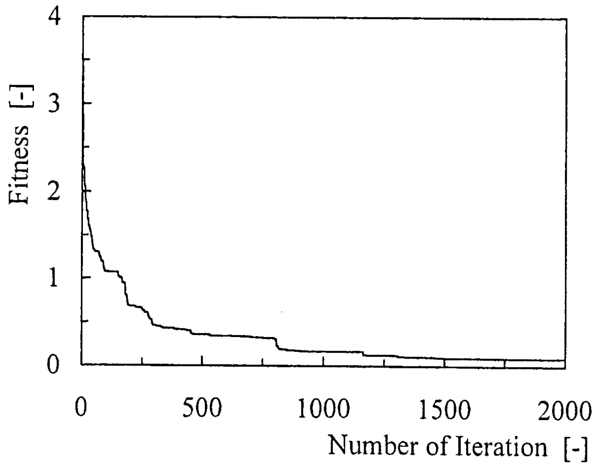

Fig.11 Learning process with network structure modification (Result 11)
(I_i =1.5, C_h =2, G_e=5, R_a =3, I_c =0.5, E_r=0.8)

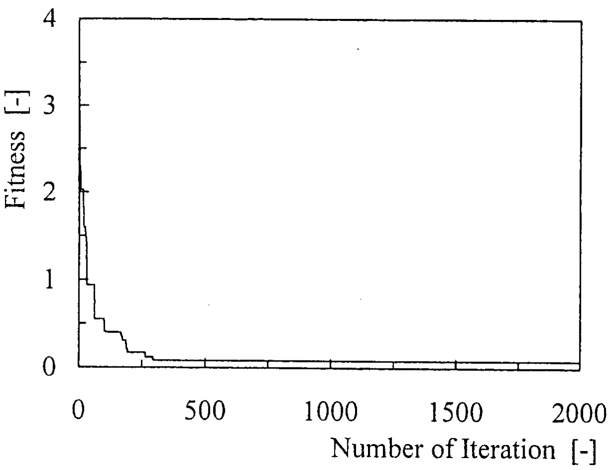

Fig.10 Learning process with network structure modification (Result 10)
(I_i = 1.5, C_h =2, G_e =30, R_a =5, I_c =0.5, E_r=0.8)

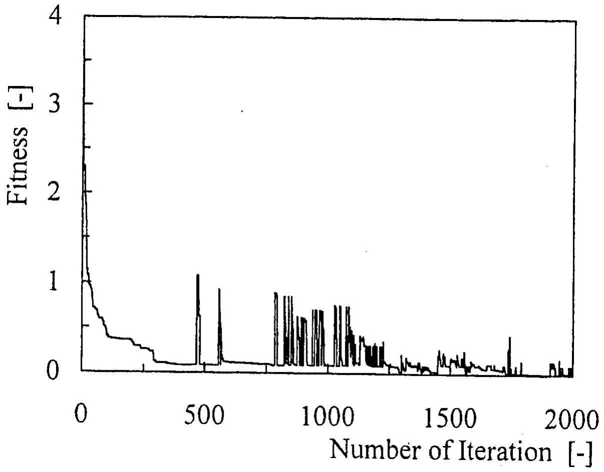

Fig.12 Learning process with network structure modification (Result 12)
(I_i = 1.5, C_h =2, G_e =5, R_a =10, I_c =0.5, E_r=0.8)

Table 1 Network results after learning with GA

Result No.	Network outputs [-]				Mean Error [-]	Fitness [-]	Elimitation Rate [-]
	0	1	2	3			
1	1.40E-01	-1.96E-01	5.97E-02	6.69E-04	1.88E+00	9.98E-02	73.2
2	6.67E-03	-3.36E-01	-1.59E-01	6.32E-05	1.96E+00	8.42E-02	33.0
3	4.29E-04	-1.34E-03	3.40E-03	1.23E-04	1.87E+00	5.10E-03	35.9
4	2.77E-02	-3.89E-01	-2.20E-01	2.03E-08	1.99E+00	8.50E-02	16.6
5	4.73E-02	-2.83E-01	-9.32E-03	3.24E-02	1.90E+00	1.12E-01	60.1
6	1.33E-03	-2.84E-01	2.32E-01	3.86E-06	1.86E+00	3.92E-01	35.0
7	9.43E-04	1.72E-02	4.10E-02	1.62E-03	1.86E+00	8.16E-02	28.8
8	3.00E-05	-1.69E-03	-1.25E-04	6.56E-06	1.87E+00	1.85E-03	39.0
9	4.27E-02	2.33E-01	1.76E-01	6.82E-02	1.77E+00	8.27E-02	6.2
10	1.10E-01	-7.22E-01	3.63E-01	1.29E-02	1.91E+00	8.25E-02	0
11	1.18E-01	-2.11E-01	2.16E-01	4.97E-02	1.83E+00	9.46E-02	0
12	1.00E-05	-2.78E-03	1.15E-02	9.33E-08	1.87E+00	8.92E-02	34.4
13	5.45E-02	6.68E-01	2.43E-01	3.56E-05	1.75E+00	1.07E-01	53.4
14	6.73E-02	3.14E-02	-1.44E-01	1.55E-06	1.91E+00	1.91E-01	52.5

次にニューラル・ネットワークの影響因子について検討した。

I_c の効果:シグモイド関数の傾斜を大きくとるとユニット飽和値に達する頻度が高まるため適応度が最適値に至らないうちに構造変化がおこり、結果として学習がうまく行われない傾向を示した(Fig.4、Fig.11、Fig.12)。

E_r の効果:構造を簡略化するため学習中に情報伝達に寄与する割合の低下した結合を遮断するための閾値を設定した。設定値以内に適応度が陥った場合シナプス荷重を強制的に 0 と置いてその後の学習をおこなった(Fig.13、Fig.4、Fig.14)。この結果、計算した値の中で E_r=1.2 の場合が出力誤差が最小となり、これは今回の全シミュレーション中で最も低い値であった。

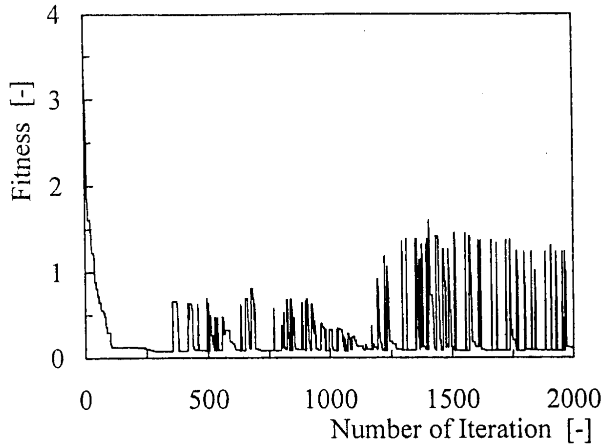

Fig.13 Learning process with network structure modification (Result 13)
(I_i = 1.5, C_h =2, G_e =5, R_a =5, I_c=1, E_r=0.8)

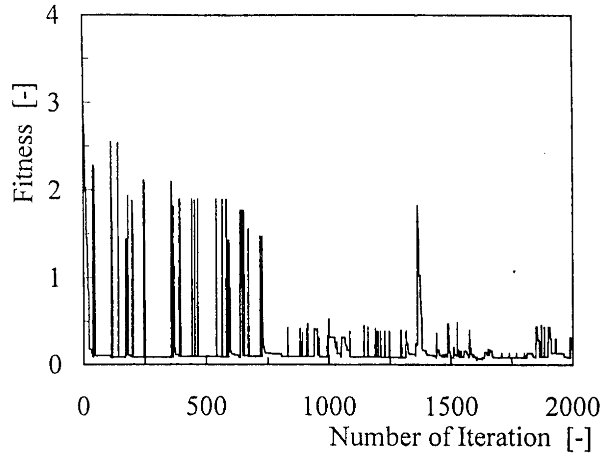

Fig.14 Learning process with network structure modification (Result 14)
(I_i=1.5, C_h=2, G_e=5, R_a=5, I_c=2, E_r=0.8)

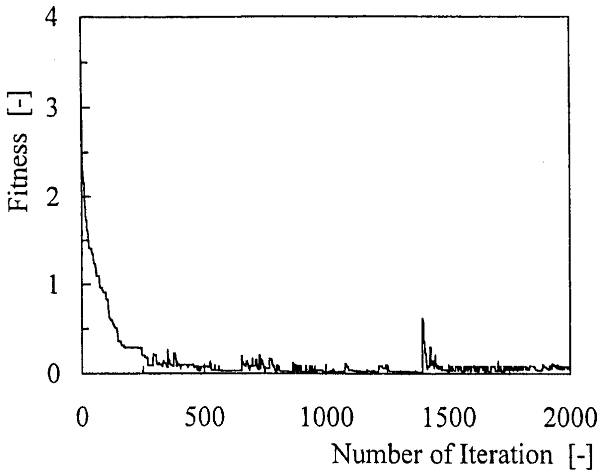

Fig.15 Learning process with network structure modification (Result 7)
(I_i =1.5, C_h =2, G_e =5, R_a =5, I_c=0.5, E_r=0.6)

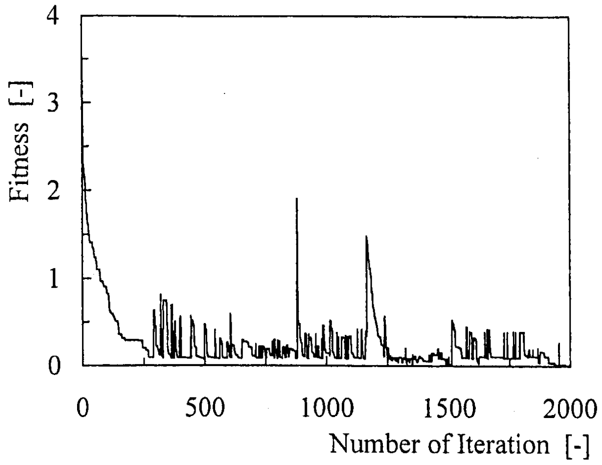

Fig.16 Learning process with network structure modification (Result 8)
(I_i =1.5, C_h =2, G_e =5, R_a=5, I_c =0.5, E_r=1.2)

4. おわりに

パターン認識学習のための階層型ニューラル・ネットワークのパラメータ探索を遺伝的アルゴリズムによりおこない、構造を簡略化する機能を持たせて学習の効率化をはかることを検討した。学習に関与する影響因子を考慮し数値的に感度解析した。シミュレーションは処理系の制約から小規模の範囲でおこなわざるを得なかったが、ニューラル・ネットワークおよび遺伝的アルゴリズムにおける各因子の効果については確認をすることが出来た。

なお本研究の計算の一部は蔵本賢一君(現:三光産業㈱)の労によることを付し謝意を表す。

記号

a_j　中間層 j の出力
C_h　突然変異を起こす染色体の数
d　教師データ
E_r　構造変更時に消去するシナプス荷重の上限値
f(x)　中間層ユニットのシグモイド関数
G_e　突然変異を起こす遺伝子の数
g(x)　出力ユニットのシグモイド関数
I_c　シグモイド関数の傾き係数
I_I　初期値の乱数幅
o　ネットワーク出力
R_a　突然変異のため変更を加える乱数
T　適応度
v_j　中間層ブから出力層へのシナプス荷重
w_{ji}　入力層∫から中間層ブへのシナプス荷重
Θ　出力層の閾値
θ　中間層プでの閾値
(添え字)
p　入力パターン番号
q　染色体番号

参考文献

1) C.G.Looney,Patten Recognition using Neural Networks,Oxford Univ.Press,1997.
2) 甘利俊一,向殿政男:ニューロとファジィ,培風館,1994.
3) 西川緯一,北村新三:ニューラルネットと計測制御,朝倉書店,1995.
4) 安居院猛,長尾智晴:ジェネティックアルゴリズム,昭晃堂,1995.
5) 石川眞澄:コネクショニストモデルの忘却を用いた一般構造学習,電気学会論文集 C112-6,pp.364-369,1992.
6) 渡辺嘉二郎,芹沢一雅,候麗雅:学習係数の適応調整によるニューラルネットワーク学習の高速化,計測自動,制御学会論文集 30-9,pp.1093-1099,1994.

(1998.11.13 受理)

SIMPLIFICATION OF PATTERN LEARNING STRUCTURE
FOR MULTI-LAYERED NEURAL NETWORKS
USING GENETIC ALGORITHM

Eiji MORIMOTO, Atsushi HIRAOKA Eiki OSAKI,
Makoto NAKAMURA, Kenzo WADA

Faster convergence and learning for neural networks have been studied by modifying the network connections to simpler structure during leaning process for two-dimensional pattern recognition. Matrices patterns consist of 16 pixels were learned by three layers neural networks. The synapse weights and thresholds were corresponded to the chromosome of gene to adopt genetic algorithm approach for parameter search in the networks. Parameter optimization was executed including mutation algorithm for efficient leaning by using 10 genes consisting of 451 loci. The effect of parameters such as width of random number for network initial condition, mutation indices, input-output condition index, structure index have been studied by numerical sensitive analysis.