遺伝的アルゴリズムによる 階層型ニューラルネットワークの構造選択

芝 直 樹*・小 谷 学*・赤 澤 堅 造*

Designing Multi-layered Neural Networks Using Genetic Algorithm

Naoki SHIBA*, Manabu KOTANI* and Kenzo AKAZAWA*

Neural networks have been often used for pattern recognition since the error back-propagation algorithm (BP) has been proposed. One of difficulties to use the network is how to design the structure of the network. For this difficulty, various structural learning algorithms have been proposed. It, however, is considered that these learning algorithms based on BP are only local search. Another way is genetic algorithm (GA) which is good at search of large space since it uses many populations. It, furthermore, is not necessary for the fitness function to be continuous and differentiable. The object of this paper is to design the structure of the layered neural network using GA. We propose new coding method and fitness function. The proposed method is applied to various Boolean problems and acoustic diagnosis of compressor. The results shows that the proposed method is effective for designing the structure of the network.

Key Words: genetic algorithm, neural network, acoustic diagnosis

1. はじめに

誤差逆伝搬法 (BP 法)¹⁾を学習方法とする階層型ニューラルネットワークは、パターン分類などによく用いられている。この階層型ニューラルネットワークを音響診断のようなノイズを含む複雑で大規模な問題に適用する際には、ネットワークの適切な構造をどのように選択するのが重要な問題となる。ネットワーク構造が小さすぎると学習を十分に行うことができない。また、ネットワーク構造が大きすぎると一般的に汎化能力が劣化するとされている。しかし、問題に応じたネットワーク構造を予め適切に設定することは困難であるので、試行錯誤的に決めているのが現状である。

この問題に対して、評価関数にネットワーク構造の複雑さなどを付加する構造学習方法^{2)~7)}や遺伝的アルゴリズム (GA)^{8), 9)}を用いる方法が提案されている。

GA を用いて構造選択を行う方法は、コーディング方法により分類される。まず、ネットワーク構造と遺伝子型を 1:1 に写像する直接エンコード法とネットワーク構造の構築ルールを指定する文法エンコード法^{10), 11)}に大別される。さらに、直接エンコード法は強指定法^{12)~14)}と弱指定法^{15), 16)}に分類される。強指定法はネットワーク内の結合の有無を全て指定する方法であるので、大規模なネットワークに対しては遺

伝子が巨大になるため、探索が困難になるという問題がある。一方、弱指定法はネットワークの概要、すなわち、中間ユニットの数や結合の数などを指定する方法である。したがって、大きなネットワークをコーディングするには適しているが、ネットワークの細部を表現することが困難となる。また、文法エンコード法は構築ルールによってネットワーク構造を探索するために、計算量は軽減すると考えられる。しかし、遺伝子型と表現型が直接対応していないので、遺伝的操作が表現型にどのような影響を与えているのかが不明確になると考えられる。

このようなことから、本論文ではネットワークの細部を表現することが可能で、遺伝子型と表現型が直接対応している強指定法による階層型ネットワークの構造選択を目的としている。具体的には、巨大化した遺伝子でも探索を可能にするような、新たなコーディング方法と適応度関数を提案する。そして、種々の論理問題や複雑で大規模な問題としてコンプレッサの音響診断に提案方法を適用し、ネットワークの汎化能力も含めて評価する。

2. 遺伝的アルゴリズムによる構造選択方法

ここでは、遺伝的アルゴリズムによる階層型ニューラルネットワークの構造選択方法について述べる。

2.1 構造選択方法

本論文において、適用する構造選択方法の概要を **Fig. 1** に示す。このように、GA を用いてネットワークの構造を進化しながら、ネットワーク内の結合の大きさを BP 法により学

* 神戸大学工学部 神戸市灘区六甲台町 1-1

* Faculty of Engineering, Kobe University, Nada-ku, Kobe
(Received February 26, 1997)
(Revised April 9, 1998)

Fig. 1 Proposed algorithm

習する。

まず、階層型ニューラルネットワークの構造を示す遺伝子型の初期集団を設定する。つぎに、設定された遺伝子型にもとづいて個体を生成し、そのネットワーク内の結合の大きさをランダムに与える。このような初期集団に対し、各々の個体をBP法により一定回数学習させ、適応度を算出する。そして、適応度に応じて遺伝的操作を行う。このとき、学習された結合の大きさは、新たに生じた個体に引き継がれることとする。このような個体群をさらにBP法により学習させ、GAにより構造を探索することを最終世代まで繰り返す。

なお、用いる階層型ネットワークは3層構造で、入力ユニットの入出力関数は恒等関数、中間・出力ユニットの入出力関数は $f(x) = 1/(1 + \exp[-x])$ のシグモイド関数である。

2.2 コーディング方法

一般的に強指定法ではネットワーク内の結合の有無をコーディングしている。本論文ではビルディング・ブロックが中間ユニットとそれに関する結合をもとに形成され则认为、中間ユニットの有無も含めてコーディングを行う。すなわち、中間ユニットの有無を示す遺伝子の直後に、その中間ユニットに関係する結合の有無を示す遺伝子をおくことにする。また、遺伝子はバイナリ値により表し、“0”は存在しないことを示し、“1”は存在することを示す。コーディング例をFig. 2に示す。

Fig. 2(a)の遺伝子型において、“○”で囲まれた遺伝子は中間ユニットの有無を示し、太い下線の遺伝子は入力層のユニットからその中間ユニットへの結合の有無を、細い下線の遺伝子はその中間ユニットから出力層のユニットへの結合の有無を示している。下線のない遺伝子は入力層のユニットから出力層のユニットへの結合の有無を示している。そして、

①101①011①11101

(a) Genotype

(b) Phenotype

Fig. 2 Representation of the network's structure in genotype and phenotype

中間ユニットを示す遺伝子が“0”のときは、そのユニットに関わる結合を示す遺伝子が“1”であっても、表現型では結合はないものとする。

このような遺伝子の表現型をFig. 2(b)に示す。この図において、実線は結合があり、破線は結合がないことを表す。さらに、細線は遺伝子において中間ユニットが存在しないことになっているので、表現型では存在しない結合である。すなわち、この例では中間ユニット数2個、結合数6個となる。なお、中間ユニットを示す“○”内の数字は中間ユニットの有無に関する遺伝子の情報を示している。

また、BP法による学習は表現型において存在するとされている結合に対してのみ行われる。

2.3 遺伝的操作

本節では、各遺伝的操作に関して述べる。

2.3.1 選択

選択方法として適応度比例戦略とエリート保存戦略を用いる。エリート保存戦略は、高い適応度を示す上位20%の個体を交叉・突然変異に影響されことなく次世代に残す。また、次世代を形成する残り80%は、適応度比例戦略により選択する。

さらに、適応度比例戦略を適用する際に、適応度に対して線形スケーリングを行う。本論文では、次式のように適応度を f をスケーリング変換された適応度 f' に変換する。

$$f' = f + b \quad (1)$$

最低の適応度を示す個体が確率 p で選択されるように b の値を定める。

$$b = \frac{p \sum_{j=1}^n f_j - f_{\min}}{np - 1} \quad (2)$$

ここで、 $f_{\min}$ は最低の適応度、 n は集団中の個体数である。

2.3.2 交叉

交叉方法として、1点交叉を用いる。ただし、交叉位置は前述のビルディング・ブロックを考慮して中間ユニットの有無を示す遺伝子の直前に限定する。Fig. 3に交叉の例を示す。

Fig. 3 Crossover operator

この図には遺伝子型とそのときの表現型も併せて示している。遺伝子型や表現型における太い縦線は、交差点を表している。このように、親 (P1, P2) の中間ユニットの有無を示す遺伝子の直前でのみ交叉を起こしているので、表現型では中間ユニットとそれに関連する結合の交換ということになる。

また、中間ユニットおよび結合の交換において、それらの結合の大きさも交換する。

2.4 突然変異

突然変異における個体の変化を Fig. 4 に示す。この図では、遺伝子型上の黒丸が突然変異の位置を表している。

まず、Fig. 4(a) は突然変異位置が中間ユニットの有無を示す遺伝子でおこるときの例である。中間ユニットの有無を示す遺伝子が“0”から“1”に変化するとき、その中間ユニットが出現し、それに関わる結合も遺伝子にもとづいて現れる。そのとき、結合の大きさは元の値を保持しているものとする。反対に、中間ユニットの有無を示す遺伝子が“1”から“0”に変化するとき、その中間ユニットが消滅するために、それに関わる結合の有無を示す遺伝子は変化しないが表現型における結合はなくなる。

つぎに、Fig. 4(b) は突然変異位置が結合の有無を示す遺伝子でおこるときの例である。結合を示す遺伝子が“0”から“1”に変化するとき、結合が新たに出現するので、そのときの結合の大きさはランダムに与える。また、結合の有無を示す遺伝子が“1”から“0”に変化するとき、その結合は消滅する。

2.5 適応度

ネットワークの能力を評価するために、教師信号と出力信号の差の2乗和を用いる。また、ネットワーク構造が小さすぎると学習を十分に行うことができず、ネットワーク構造が大きすぎると一般的に汎化能力が劣化するとされていること

Fig. 4 Mutation operator

から、構造の複雑さに関するペナルティー項を付加する。したがって、ネットワークの評価は (3) 式により行う。

$$f = \begin{cases} -Nk \log E - C, & \text{for } E < 1/8 \\ -Nk \log E - aC, & \text{for } E \geq 1/8 \end{cases} \quad (3)$$

$$E = \frac{1}{Nk} \frac{1}{2} \sum_{p=1}^N \sum_{j=1}^k (t_{pj} - o_{pj})^2 \quad (4)$$

$$C = 2.0w + 1.5u \quad (5)$$

$$\text{if } E < \frac{1}{2} E_t^2 \text{ then } E = \frac{1}{2} E_t^2 \quad (6)$$

ここで、 E は自乗誤差、 C はネットワーク構造の複雑さ、 a は整数、 N は学習パターン数、 k は出力ユニット数、 w は結合の数、 u は中間ユニットと出力ユニットの数を示す。また、 t_{pj}, o_{pj} はそれぞれ教師信号と出力信号であり、 E_t は誤差に対するしきい値を示す。

(3) 式の第1項は誤差に関する項であり、第2項は構造の複雑さに関するペナルティー項である。すなわち、第1項より誤差が小さく、第2項より構造の小さいネットワークがより高い適応度を示すことになる。しかし、誤差が大きくても、構造が極めて小さいために高い適応度を示すことがある。そこで、誤差が一定値 ($1/8$) 以上の個体は、ペナルティー項の係数を a 倍大きくすることにより、誤差の小さいものの中で構造の小さいネットワークがより高い適応度を示すようにする。ここで、一定値は1パターン・1ユニット当たりの教師信号と出力信号の差を0.5として、(4) 式における E の値とした。このような値に設定したのは、ネットワークの出力値が0から1までの値であるので、1パターン・1ユニット

当たりの誤差が0.5以下であれば学習が進んでいると考えられるためである。

また、 a の値は初期値を2とし、 $E \geq 1/8$ の個体が集団中で最高適応度を示すとき、集団中に $E < 1/8$ の個体があれば、その個体が最高適応度を示すように a の値を大きくする。すなわち、(7)式に示すように a の値を変化する。

$$a = \left\lfloor \frac{f_{max} + Nk \log E_{max}}{-2.0w_{max} - 1.5u_{max}} \right\rfloor \quad (7)$$

ここで、 f_{max} は $E < 1/8$ である個体を示す最高適応度、 E_{max} は最高適応度を示す個体の自乗誤差、 w_{max} は最高適応度を示す個体の結合の数、 u_{max} は最高適応度を示す個体の中間ユニットと出力ユニットの数を示している。 $\lfloor x \rfloor$ は x の値を超えない最大の整数である。

逆に、 $E < 1/8$ の個体がなければ、 a の値は変化しない。このように a の値を可変にすることにより、学習ができていないのに構造が小さいために選択されることがなくなると期待される。

さらに、ある一定値以下の誤差に関しては、ネットワークは十分に学習できているとみなして、(6)式のように一定値以下の誤差は同じとする。これにより、より小さい構造を持つ個体の適応度が高くなっていくことになる。

また、(5)式に関しては、(3)式のペナルティ項として単なる結合の数だけでなく、ユニットの数とあわせて評価する方が良いという予備実験結果からこのように設定した。

3. 論理問題への適用

本節では、提案した方法を種々の論理問題に適用し、その問題を学習可能な最小構造を獲得できることを示す。

シミュレーション条件は以下のとおりである。シミュレーションは結合の大きさの初期値(± 0.2 の範囲)を10回変えて行う。なお、構造の探索は最大中間ユニット数を10個として、10個体用いて最終世代を100世代とする。各々の個体は、BP法における学習率(η)を0.3、慣性係数(α)を0.9として1世代当たり100回の学習をする。そして、 $E_t = 0.03$ 、 $p = 3\%$ 、交叉確率80%、突然変異確率3%とする。以降、特に指定しない限り、これらの条件を用いてシミュレーションを行う。

3.1 排他的論理和問題

まず、排他的論理和問題(XOR問題)¹⁾に適用する。この問題に対して、得られた誤差および適応度の最大値(max)・最小値(min)・平均値(avg)の変化をFig. 5に示す。なお、Fig. 5以降の図で示される適応度は線形スケーリングされる前の値である。

これらより、学習が進むとともに、適応度が高くなっていくことがわかる。その結果、最終世代においてFig. 6(a)に示されるような遺伝子型が得られた。この遺伝子型をデコーディングするとFig. 6(b)のような構造になる。これは、XOR問題を学習可能な最小構造である。Fig. 6(a)は得られた遺伝子型の一例ではあるが、10回のすべてのシミュレーションに

(a) Error

(b) Fitness

Fig. 5 Evolution process for XOR problem

Top
0 111 1 111 0 000 0 001 0 111
0 011 0 001 1 011 0 110 0 100 11
Bottom

(a) Genotype

(b) Phenotype

Fig. 6 The evolved network for XOR problem

においてFig. 6(b)のような最適構造が得られた。

3.2 4変数論理問題

つぎに、次式の4変数論理問題に適用する。

$$g = (a \cup c) \cap (b \cup d) \quad (8)$$

また、論理式とは全く無関係な入力、すなわち、ノイズ入力 n を加えてシミュレーションを行う。ここで、これらの入力値は全てバイナリ値である。なお、ノイズは“0”と“1”とをランダムに入力した。

その結果、10回中9回のシミュレーションにおいて、パラメータ数(結合数と中間・出力ユニット数との和)が7の構

Fig. 7 The evolved network for the logic problem with noise

Fig. 8 The number of parameters in the evolved network for ADD problem

Fig. 9 The evolved network for ADD problem

造が選択され、他の1回はパラメータ数が8の構造が選択された。パラメータ数が7の構造はすべて Fig. 7 に示すような構造であった。これは、最適な構造であるとともに、ノイズユニットを削除していることがわかる。

3.3 2進ビット加算問題

極小解の多いことで知られている2進ビット加算問題(ADD問題)¹⁾に適用する。この問題は2つの2ビットの和をネットワークに学習させる問題である。一般的にADD問題に対しては、入力ユニット4個、中間ユニット2個、出力ユニット3個からなり、密結合(結合数は26個)を持つネットワークがよく用いられる。すなわち、パラメータ数は31個となる。

10回のシミュレーションにおいて、選択されたネットワークのパラメータ数を Fig. 8 に示す。なお、この結果は0.1以下の誤差が得られた9回の内訳である。

また、誤差が0.1以下で構造が最小のネットワークは、Fig. 9 に示すような中間ユニット数が2個で、結合数が15個(すなわち、パラメータ数は20個)からなる構造であった。

Table 1 The number of parameters for various initial number of weights

The number of hidden units	The number of parameters
2	31 (31)
5	49 (55)
10	84 (95)

3.4 考察

3.4.1 GAによる構造選択の利点

GAによる構造選択の有効性を評価するために、代表的な構造学習法である忘却アルゴリズム²⁾をADD問題に適用する。シミュレーションは、忘却付き学習を30000回行った後、選択的忘却付き学習を30000回行う。そして、初期中間ユニット数2, 5, 10個の各場合に対して、結合の大きさの初期値を20回変えて学習を行った。なお、忘却量は 1×10^{-6} , 1×10^{-5} , 1×10^{-4} 、選択的忘却付き学習の際の忘却の対象とする結合の大きさに対するしきい値は0.001, 0.01, 0.1, 1.0として、各条件のもとで学習を行った。

その結果、誤差が0.1以下となるときに得られた最小構造のパラメータ数を各初期中間ユニット数の場合について Table 1 に示す。なお、括弧内の値は学習前のネットワークのパラメータ数である。このように、初期中間ユニット数が2個の場合には結合は削除されず、また5, 10個の場合には1割ほど削除されている。しかし、提案方法による場合ではパラメータ数20個のネットワークが選択されているので、これに比べて大きな構造のネットワークとなっている。

3.4.2 学習回数について

1世代あたりの学習回数に対する検討として、学習回数を10, 100, 1000回として、ADD問題に対してシミュレーションを行なった。その結果、10回という少ない学習回数では与えられた問題に対して学習可能な構造を持つネットワークであっても、1世代の間に誤差が十分に減少しなかった。そして、最終的に得られるネットワークとして十分に誤差が減少していないネットワークがパラメータ数の少なさで選択される傾向があった。一方、学習回数を多くするとこのような問題はなくなるが、計算量が増大する。このようなことから、学習回数は各世代で学習可能な構造と不可能な構造を区別できる最小限の回数が良いと考えられ、今回のシミュレーションでは100回とした。

3.4.3 適応度について

与えられた問題を学習可能な最小構造を選択するための基準となる適応度関数について、以下の2点について検討する。

- 適応度関数を誤差の値により場合分けを行う必要性
- a の値を変動させることによる効果

まず、適応度関数を誤差の値により場合分けを行う必要性に関して検討を行う。適応度関数を次式で固定し、XOR問題に適用する。

$$f = -Nk \log E - C \quad (9)$$

適用した結果、10回のシミュレーション中6回において学

(a) Error

(b) Fitness

Fig. 10 Evolution process for XOR problem with the fitness function in Eq. (9)

習可能な最小構造が選択された。しかし、他の4回では誤差が大きくて学習できていないが、学習可能な最小構造よりも小さな構造が選択されていた。その場合の誤差と適応度の変化の一例を **Fig. 10** に示す。

Fig. 10 より46世代目に誤差の小さな構造が出現しているが、適応度が-4.85と集団中の平均適応度よりも低いために淘汰されている。平均適応度が高くなっているのは、学習できていないにもかかわらず構造が小さいために高い適応度を示す個体が多数存在したためである。したがって、適応度関数を誤差の値により場合分けを行う必要がある。

つぎに、(7) 式のように a の値を変動することによる効果を検討する。 a の値を変動させる場合と a の値を2に固定する場合について、XOR問題に対してシミュレーションを10回行った。それぞれの場合で選択されたネットワークのパラメータ数を **Fig. 11** に示す。このように a の値を固定すると、より大きい構造を選択する傾向であることがわかる。つぎに、 a の値を可変にしたときの誤差と適応度の変化を **Fig. 12** に示す。また、 a を固定させたときに得られたネットワークのパラメータ数が10の場合の誤差と適応度の変化を **Fig. 13** に示す。

Fig. 12 において、44世代目に適応度が減少している。これは、ネットワーク構造が大きい(パラメータ数が10個)が誤差の小さい個体が出現したが、それが最高適応度とならなかったために a の値が2から6へと変動したためである。それにより、その個体は最高適応度を示すようになっている。その後、そのネットワークの学習が進むことにより適応度が

Fig. 11 The number of parameters in the evolved network for XOR problem with variable a or fixed a

(a) Error

(b) Fitness

Fig. 12 Evolution process for XOR problem with variable a

上昇している。そして、83世代目に最小構造のパラメータ数7個のネットワークが現れ、それが90世代以降の適応度の上昇につながっている。

一方、**Fig. 13** では44世代目に誤差の小さい個体が出現しているが、ネットワーク構造が大きいために次世代以降に生き残れなかった。63世代目にネットワークのパラメータ数が10個の個体を選択されるが、それ以降さらに小さなネットワークは現れなかった。

このように、 a の値を可変にすることにより限られた世代数の中で、与えられた問題を学習可能な最小構造を選択する可能性が高くなることがわかる。

以上より、本方法において用いている適応度関数は構造の最小化に有効であることが確認された。

4. コンプレッサの音響診断への適用

複雑で大規模な問題として、コンプレッサの音響診断データ¹⁷⁾に提案方法を適用する。ここでは、予め抽出された特

Fig. 13 Evolution process for XOR problem with fixed α

微量を入力として異常内容を識別する異常識別ネットワークの構造選択を対象とする。このネットワークは3層の階層型ネットワークである。シミュレーション条件は、最大中間ユニット数を50個、BP法における学習率・慣性係数ともに0.1とする以外は前節と同じ値を用いる。なお、異常識別ネットワークは32次元の特徴量を入力として、4種類のパルス状態を識別するように構成される。すなわち、入力ユニット数は32個、出力ユニット数は4個となる。このようなネットワークに対して、240個の入力データにより学習を行いながら、構造を選択していく。また、世代毎に最高適応度を示す個体に対して240個の未知データを与えて識別率の高いネットワークを選択する。これを最終世代まで続ける。そして、最終世代で選択されたネットワークに対して、別の480個の未知データを与えて識別率を算出する。

その結果、中間ユニット数が14個、結合数が319個の構造が選択され、識別率81.5%であった。

さらに、得られた構造とその識別率を評価するために、3.4節と同様に階層型ネットワークに忘却アルゴリズムを適用する。なお、学習前のネットワークの中間ユニット数は14個と50個とし、それぞれ密な結合を持つネットワークを用いる。また、忘却付き学習を30000回、選択的忘却アルゴリズムを30000回行なう。なお、忘却量は 1×10^{-6} , 1×10^{-5} , 1×10^{-4} , 選択的忘却付き学習における忘却の対象とする結合の大きさに対するしきい値は0.001, 0.01, 0.1, 1.0, 学習率・慣性係数はともに0.1として、各条件の下で結合の大きさの初期値を30回変更してシミュレーションを行なった。各中間ユニット

Table 2 Discrimination accuracy with BP with forgetting and GA

	BP with forgetting	GA
Number of initial hidden units	50	14
Number of weights	785	348
Accuracy [%]	79.2	80.8

数で最も高い識別率が得られたネットワークの結合数とその識別率をTable 2に示す。また、比較のため提案方法により構造選択した場合において得られた識別率もTable 2に示す。

このように、提案方法により構造選択されたネットワークを忘却アルゴリズムにより選択されたネットワークと比較すると、汎化能力において大きな差は見られない。一方、初期中間ユニット数で差があるが、学習後のパラメータ数（結合数）は、提案方法において選択された中間ユニット数に等しい14個の場合でも提案方法より約1割ほど多い結果となっている。

5. おわりに

階層型ニューラルネットワークをパターン認識に用いる際の最適な構造選択を目的として、遺伝的アルゴリズムの適用を検討した。新たなネットワークのコーディング方法および構造の複雑さに関するペナルティ項を導入した適応度関数を提案した。そして、提案方法の有効性を評価するために種々の論理問題と実際的な問題としてコンプレッサの音響診断に適用した。得られた結果より、階層型ニューラルネットワークの構造選択において提案方法の有効性が確認された。

しかし、適応度関数の係数などは試行錯誤により決定したものであり、今後その理論的根拠などについて検討していく必要がある。さらに、本研究では計算量の点から世代間で学習した結合の大きさを引き継ぐことにしているが、結合の大きさを引き継がない場合との比較を今後していく必要があろう。一方、今回の実験では個体数を10個と比較的少ない個体に設定しているが、これでは探索の初期段階でほとんど類似個体になっている可能性も考えられる。このような初期収束を防止するため、今後は個体数を増やす、あるいは多様性を維持するための仕組みを導入していく必要があろう。

参考文献

- 1) D. E. Rumelhart, G. E. Hinton and R. J. Williams: Learning Internal Representations by Error Propagation, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, MIT Press, I, 318/362 (1986)
- 2) 石川眞澄: 忘却を用いたコネクションリストモデルの構造学習アルゴリズム, 人工知能学会誌, 5-5, 595/603 (1990)
- 3) Y. L. Cun, J. S. Denker and S. A. Solla: Optimal Brain Damage, D. S. Touretzky eds., Advances in Neural Information Processing Systems, 2, 598 / 605 (1990)
- 4) 栗田多喜夫: 情報量基準による3層ニューラルネットの隠れ層のユニット数の決定法, 電子情報通信学会論文誌 D-II, J73-DII-2, 1872/1878 (1990)
- 5) A. S. Weigend, D. E. Rumelhart and B. A. Huberman: Generalization by Weight-Elimination with Application to Forecasting, R. P. Lippmann et al. eds., Advances in Neu-

- ral Information Processing Systems, **3**, 875 / 882 (1991)
- 6) 和田, 川人: 新しい情報量基準と Cross Validation による汎化能力の推定, 電子情報通信学会論文誌 D-II, **J74-D-II** - 7, 955 / 965 (1991)
- 7) R. Reed: Pruning Algorithm - A Survey -, IEEE Trans. on Neural Networks, **4** - 5, 740 / 747 (1994)
- 8) D. E. Goldberg: Genetic Algorithm in Search, Optimization, and Machine Learning, Addison-Wesley (1989)
- 9) 北野宏明: 遺伝的アルゴリズム, 産業図書, (1993)
- 10) H. Kitano: Designing Neural Networks Using Genetic Algorithm with Graph Generation System, Complex System **4**, 461/476 (1990)
- 11) H. Kitano: Neurogenetic Learning: an Integrated Method of Designing and Training Neural Networks using Genetic Algorithm, Physica, **D-75**, 225/238 (1994)
- 12) G. F. Miller, P. M. Todd and S. U. Hegde: Designing Neural Network using Genetic Algorithm, Proc. of the 3rd International Conference on Genetic Algorithms, 379/384 (1989)
- 13) W. Schiffmann, M. Joost and R. Werner: Synthesis and Performance Analysis of Multilayer Neural Network Architectures, Technical Report, 16, University of Koblenz (1992)
- 14) V. Maniezzo: Genetic Evolution of the Topology and Weight Distribution of Neural Networks, IEEE Trans. on Neural Networks, **5**-1, 39/53 (1994)
- 15) S. T. Harp, T. Samad and A. Guha: Towards the Genetic Synthesis of Neural Networks, Proc. of the 3rd International Conference on Genetic Algorithms, 360/369 (1989)
- 16) 田中, 古賀, 谷野: ニューラルネットワークの構造決定における遺伝的アルゴリズムの応用, システム制御情報学会論文誌, **7-12**, 528/530 (1994)
- 17) 小谷, 植田, 松本, 金川: ニューラルネットワークによる回転機の音響診断法, 計測自動制御学会論文集, **31-3**, 382/390 (1995)

赤 澤 堅 造 (正会員)

1965年大阪大学工学部電気工学科卒業, 71年同大学院博士課程単位取得退学. 同年大阪大学工学部電気工学科助手, 講師, 助教授を経て, 90年神戸大学工学部計測工学科教授. 現学科名は情報知能工学科. 医用生体工学の研究, 特に筋ダイナミクス, 手足の運動制御の機構, 筋電制御義手, 生体情報工学などの研究に従事. 日本ME学会論文賞, 学会賞受賞. 電気学会, 日本ME学会, 計測自動制御学会, IEEE, バイオメカニズム学会各会員. (工学博士).

[著 者 紹 介]

芝 直 樹

1994年, 神戸大学工学部計測工学科卒業. 同年同大学院修士課程情報知能工学専攻入学. 在学中, 遺伝的アルゴリズム・ニューラルネットワークに関する研究に従事. 96年三菱電機(株)入社

小 谷 学 (正会員)

1983年, 神戸大学大学院修士課程計測工学専攻修了. 同年(株)神戸製鋼所入社. 89年神戸大学工学部計測工学科助手を経て現在情報知能工学科助教授. 音響診断, ニューラルネットワークを用いたパターン認識などの研究に従事. 電子情報通信学会, システム制御情報学会, 日本神経回路学会などの会員(工学博士).