

# ファジィ数遺伝子を用いたファジィ ES によるニューロエボリューション

## Evolving Neural Networks by means of Fuzzy Evolution Strategy with Fuzzy-valued Genotypes

和田 哲也 山下 彬 松瀬 高志 岡田 英彦

Tetsuya Wada, Akira Yamashita, Takashi Matsuse and Hidehiko Okada

### 1. はじめに

これまでに、ユニット間結合強度やユニットしきい値が実数ではなくファジィ数に拡張された階層型ニューラルネットワークが提案されている[1]. さらに、学習用データを用いて、このファジィニューラルネットワーク (FNN) の結合強度としきい値を学習させるために、従来のバックプロパゲーションアルゴリズムを拡張した手法も提案されている[1]. しかし、ファジィニューラルネットワークの教師なし学習法は提案されていなかった. これに対して、近年、進化計算手法を用いたニューラルネットワークの教師なし学習法が研究されており[2], ニューロエボリューションと呼ばれている.

本研究では、解の遺伝子表現 (genotype) の値としてファジィ値を用いることが可能な進化計算手法を提案している[3]. 本提案手法の有効性評価の一環として、提案手法の一種であるファジィ ES を、前記 FNN のニューロエボリューションに適用した.

### 2. ファジィニューラルネットワーク

本研究で用いるニューラルネットワークは、文献[1]で提案された FNN である. この FNN は、3 階層のフィードフォワード型であり、結合強度としきい値がファジィ数の値をとる. FNN の構造を図 1 に示す. なお、簡単化のため、FNN の出力層ユニット数は 1 とする.

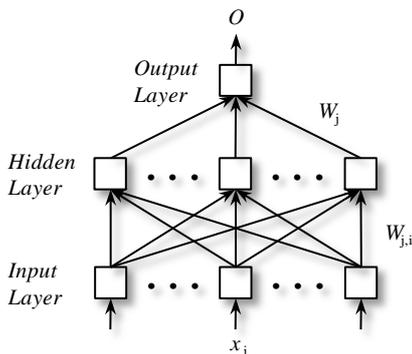


図 1: ファジィニューラルネットワークの構造[1]

この FNN は、実数入力ベクトル  $x$  を受け取り、ファジィ数出力値  $O$  を、式(1)-(5)に基づいて計算する.

入力層 :

$$o_i = x_i. \tag{1}$$

中間層 :

$$Net_j = \sum_i W_{j,i} o_i + \Theta_j, \tag{2}$$

$$O_j = f(Net_j). \tag{3}$$

出力層 :

$$Net = \sum_j W_j O_j + \Theta, \tag{4}$$

$$O = f(Net). \tag{5}$$

式(1)-(5)において、 $x_i, o_i$  は実数であり、 $Net_j, Net, W_{j,i}, W_j, \Theta_j, \Theta, O_j, O$  はファジィ数である. 簡単化のため、文献[1]ではこのファジィ数として対称三角型ファジィ数が用いられており、本稿でも同様とする. 対称三角型ファジィ数のメンバーシップ関数は、図 2 のように、中心  $c$  と幅  $w$ 、もしくはサポート集合の下限  $L$  と上限  $U$  によって規定できる.

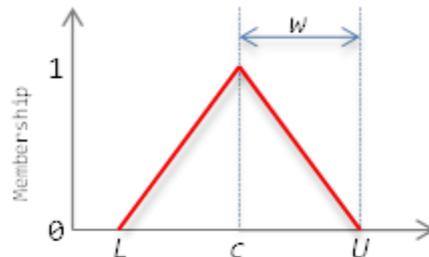


図 2: 対称三角型ファジィ数のメンバーシップ関数

ユニット内部関数  $f(x)$  にはシグモイド関数  $f(x) = 1/(1+e^{-x})$  を用いる. この  $f(x)$  によるファジィ数入力値からファジィ数出力値へのマッピングを図 3 に示す.

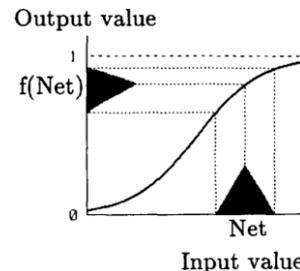


図 3: 中間層・出力層ユニットにおける入出力関係[1]

式(2)-(5)におけるファジィ演算は拡張原理[4]に基づいて行われ、ファジィ数の  $\alpha$  カット集合である区間値を用いた

演算は、次の式(6)-(8)に基づいて行われる[5]。2つの閉区間を、 $A=[a^L, a^U], B=[b^L, b^U]$ とする。

$$\begin{aligned} A+B &= [a^L, a^U] + [b^L, b^U] \\ &= [a^L+b^L, a^U+b^U]. \end{aligned} \quad (6)$$

$$\begin{aligned} k \cdot A &= k[a^L, a^U] \\ &= [ka^L, ka^U], \text{ if } k \geq 0, \text{ or } [ka^U, ka^L], \text{ if } k < 0. \end{aligned} \quad (7)$$

$$\begin{aligned} A \cdot B &= [a^L, a^U] \cdot [b^L, b^U] \\ &= [\min(a^L b^L, a^L b^U, a^U b^L, a^U b^U), \\ &\quad \max(a^L b^L, a^L b^U, a^U b^L, a^U b^U)]. \end{aligned} \quad (8)$$

図1のFNNのフィードフォワード計算の詳細は文献[1]を参照されたい。FNNの入力層ユニット数を $n$  ( $i=1,2,\dots,n$ )、中間層ユニット数を $m$  ( $j=1,2,\dots,m$ )とすると、 $W_{ji}$ の数は $mn$ 、 $W_j$ の数は $m$ であり、結合強度の数は $mn+m$ である。さらに、ユニットしきい値の数は $m+1$ である。したがって、FNNに含まれる結合強度およびしきい値の総数は $mn+2m+1$ である。本研究が提案するファジィESは、この $mn+2m+1$ 個のファジィ数を genotype の値として用いる。genotype を  $V=(V_1, V_2, \dots, V_D)$ と表すと、 $D=mn+2m+1$ であり、 $V_1, V_2, \dots, V_D$ は、対応する結合強度もしくはしきい値がとるファジィ数の値である。 $V_1, V_2, \dots, V_D$ は、それぞれ、上限と下限、もしくは中心と幅によって規定できる。 $V_i=[v_i^L, v_i^U]$ もしくは $V_i=(v_i^C, v_i^W)$ であり、 $v_i^L, v_i^U, v_i^C, v_i^W$ は $V_i$ の上限、下限、中心、幅を表す。

### 3. ファジィES

提案するファジィESのプロセス構成は、従来のESと共通である(図3)。ただし、解の初期化、適合度評価、子個体生成の方法が、genotype のファジィ数を扱うことができるように拡張される。

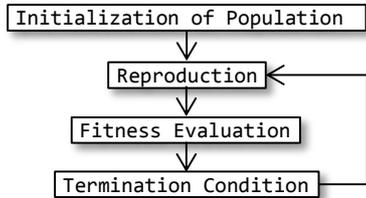


図4: ファジィESのプロセス構成

#### 3.1 個体の初期化

個体の集団サイズを $P$ とする。初期化プロセスでは、各個体の genotype  $V_1, V_2, \dots, V_P$ をランダムに初期化する。 $V_p=(V_{p,1}, V_{p,2}, \dots, V_{p,D})$ であり、 $V_{p,d}$ は対称三角型ファジィ数の中心・幅もしくはサポート集合の下限・上限である。

( $p=1,2,\dots,P, d=1,2,\dots,D$ )。すなわち、 $V_{p,d}=(V_{p,d}^C, V_{p,d}^W)$ もしくは $V_{p,d}=[V_{p,d}^L, V_{p,d}^U]$ と表される。前者をCWモデル、後者をLUモデルと呼ぶことにする。 $V_p$ はニューラルネットの結合強度もしくはしきい値のため、 $V_{p,d}$ の初期値とし

ては、絶対値があまり大きくない値が望ましいと考えられる。そこで、 $V_{p,d}$ の初期値を、平均0の正規乱数、もしくは中心0の一様乱数を用いて決定する。 $V_{p,d}=[V_{p,d}^L, V_{p,d}^U]$ の場合、乱数を2つ生成し、その小さい(大きい)ほうを $V_{p,d}^L$ ( $V_{p,d}^U$ )に代入する。一方、 $V_{p,d}=(V_{p,d}^C, V_{p,d}^W)$ の場合、乱数を2つ生成し、その一方を $V_{p,d}^C$ に、もう一方の絶対値を $V_{p,d}^W$ にそれぞれ代入する。

#### 3.2 子個体の生成

ある親個体を $V_a$ とし、この親個体から生成される子個体を $V_z$ とする。集団から $V_a$ を選ぶ方法は従来のESと同一の方法を用いる。

LUモデルの場合は、

$$\bullet V_a=(V_{a,1}, V_{a,2}, \dots, V_{a,D}), V_{a,i}=[v_{a,i}^L, v_{a,i}^U]$$

$$\bullet V_z=(V_{z,1}, V_{z,2}, \dots, V_{z,D}), V_{z,i}=[v_{z,i}^L, v_{z,i}^U]$$

である。子個体の $v_{z,i}^L$ と $v_{z,i}^U$ は次式によって決定する。

$$v_{z,i}^L = v_{a,i}^L + rand, \quad (9)$$

$$v_{z,i}^U = v_{a,i}^U + rand. \quad (10)$$

ただし、 $rand$ は乱数を表しており、平均0の正規乱数や中心0の一様乱数を用いる。また、 $v_{z,i}^L > v_{z,i}^U$ となった場合は、 $v_{z,i}^L \leq v_{z,i}^U$ が満たされるように修正する必要がある。この修正方法として、例えば次の4通りがあげられる：(1)  $v_{z,i}^L < v_{z,i}^U$ , (2)  $v_{z,i}^U < v_{z,i}^L$ , (3)  $tmp < v_{z,i}^L, v_{z,i}^L < v_{z,i}^U, v_{z,i}^U < tmp$ , (4)  $v_{z,i}^L < (v_{z,i}^L + v_{z,i}^U)/2, v_{z,i}^U < v_{z,i}^L$ .

CWモデルの場合は、

$$\bullet V_a=(V_{a,1}, V_{a,2}, \dots, V_{a,D}), V_{a,i}=(v_{a,i}^C, v_{a,i}^W)$$

$$\bullet V_z=(V_{z,1}, V_{z,2}, \dots, V_{z,D}), V_{z,i}=(v_{z,i}^C, v_{z,i}^W)$$

である。子個体の $v_{z,i}^C$ と $v_{z,i}^W$ は、LUモデルの場合と同様に、次式によって決定する。

$$v_{z,i}^C = v_{a,i}^C + rand, \quad (11)$$

$$v_{z,i}^W = v_{a,i}^W + rand. \quad (12)$$

$v_{z,i}^W < 0$ となった場合は、 $v_{z,i}^W \geq 0$ が満たされるように修正する必要がある。この修正方法として、例えば $v_{z,i}^W < 0$ とすることが挙げられる。

#### 3.3 個体の適合度評価

genotype  $V_p$ に対応するFNNの個体の適合度を評価するために、そのFNNを対象問題に適用して性能を測定する。後述する実験では、対象問題はファジィ関数 $y=F(x)$ の近似である。ただし、 $F(x)$ はFNNおよびファジィESのプロセスに対して与えられず、さらに、入力 $x$ に対するFNNの目標出力値も陽には与えられない。あるサンプリングされた複数の入力値に対して、FNNの出力値と関数 $F(x)$ の値の誤差の総和だけがファジィESに対して与えられる。ファジィESの適合度評価プロセスにおいては、この誤差の総

和を個体の適合度評価値として用い、集団内の個体をランキングする。

#### 4. 評価実験

FNN の対象問題としてファジィ関数の近似問題を取り上げ、本研究で提案したファジィ ES による FNN の教師なし学習を実験的に試みる。簡単化のため、ファジィ関数  $F(x)$  の入力  $x$  は  $0 \leq x \leq 1$  の実数とする。これにより、FNN の入力層ユニット数は 1 である。

集団内の個体をランクづけするために、各個体の適合度を次のように決定した。ある個体に対応する FNN  $\hat{F}$ 、 $0 \leq x \leq 1$  の範囲で一様ランダムにサンプリングされた入力値を与え ( $x_r$  とする)、その入力値に対する FNN の出力ファジィ値を得る ( $O_r$  とする)。一方、その入力値に対する目標関数  $F$  の値も求める ( $F(x_r)$  とする)。この  $O_r$  と  $F(x_r)$  の誤差  $e_r$  を、文献[1]と同様に、 $e_r = \sum_h (h * (([O_r]_h^L - [F(x_r)]_h^L)^2 + ([O_r]_h^U - [F(x_r)]_h^U)^2))$  と定義する。ここで、 $[A]_h$  はファジィ数  $A$  の  $h$  レベル集合を表し、 $[A]_h^L$  および  $[A]_h^U$  はその  $h$  レベル集合である区間の下限および上限を表している。後述する実験では、 $\Sigma$  の計算に  $h \in \{0.2, 0.4, \dots, 1.0\}$  を用いた。 $x_r$  をランダムに 100 回サンプリングして誤差  $e_r$  の総和を求め、個体の適合度として用いる。前記の通り、ファジィ ES に対してフィードバックされる値はこの誤差  $e_r$  の総和のみであり、 $F(x_r)$ 、 $e_r$  のいずれの値もファジィ ES には与えられない。

本実験では例題として 3 種類の目標関数を用いた。これらの関数は次の通りである。なお、 $F(x)$  のサポート集合を  $[F(x)]_0$ 、このサポート集合 (区間) の下限、上限を  $[F(x)]_0^L$ 、 $[F(x)]_0^U$  とする。目標関数の出力値はすべて対称三角型ファジィ数であり、 $[F(x)]_1$  は区間  $[F(x)]_0$  の中点 (つまり  $[F(x)]_1 = ([F(x)]_0^L + [F(x)]_0^U) / 2$ ) である。

例題(1)

$$[F(x)]_0^L = 0.1 \sin(2\pi x) + 0.1x^2 + 0.3$$

$$[F(x)]_0^U = 0.1 \sin(2\pi x) + 0.1x^2 + 0.6$$

例題(2)

$$[F(x)]_0^L = 0.1 \sin(2\pi x) - 0.1x + 0.4$$

$$[F(x)]_0^U = 0.1 \sin(2\pi x) + 0.1x + 0.6$$

例題(3)

$$[F(x)]_0^L = 0.1 \sin(2\pi x) + 0.1x + 0.3$$

$$[F(x)]_0^U = 0.1 \sin(2\pi x) - 0.1x + 0.6$$

これらのファジィ関数を図 5-7 に示す。図中の 3 つの曲線が、下から順に、 $[F(x)]_0^L$ 、 $[F(x)]_1$ 、 $[F(x)]_0^U$  を表している。

例題(1)の実験結果を図 8-13 に示す。図 8-10 (図 11-13) は LU (CW) モデルを用いた場合の結果である。FNN およびファジィ ES を表 1 の通りに設定したとき、近似誤差最小の解が得られた。図 8・11 は、各世代の最良個体による誤差が世代交代に伴って減少する様子を表している。また図 9・12 は、目標関数、初期世代の最良個体による関数および、全 100,000 個体中の最良個体による関数をそれぞれ

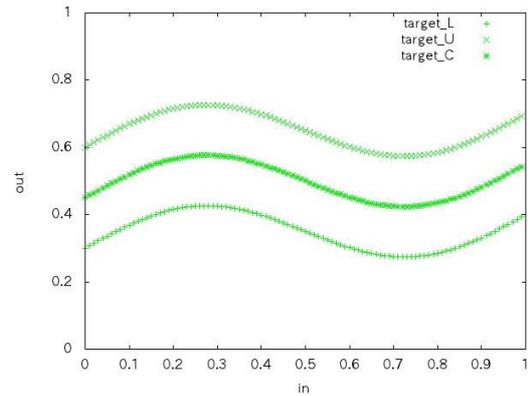


図 5: 本実験において FNN が近似する目標ファジィ関数(1)

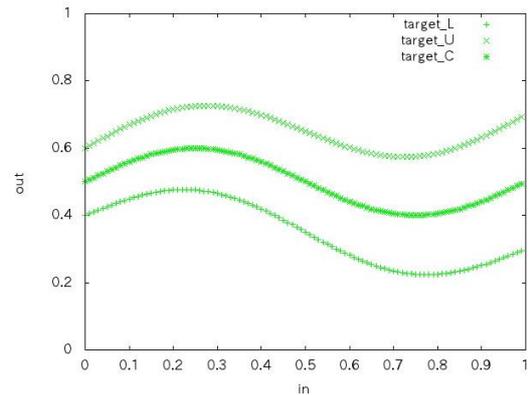


図 6: 本実験において FNN が近似する目標ファジィ関数(2)

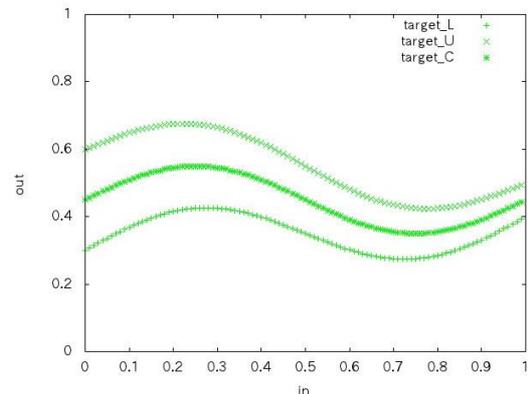


図 7: 本実験において FNN が近似する目標ファジィ関数(3)

表 1: FNN およびファジィ ES の設定

	LU モデル	CW モデル
中間層ユニット数	3	5
ES 戦略	(10,100)-ES	(10,100)-ES
世代数	1,000	1,000
式(9)-(12)の rand	[-0.1,0.1]	[-0.1,0.1]
	一様乱数	一様乱数

表している。さらに、図 10・13 は、入力  $x=0.5$  に対して目標関数が出力するファジィ数のメンバーシップ関数、お

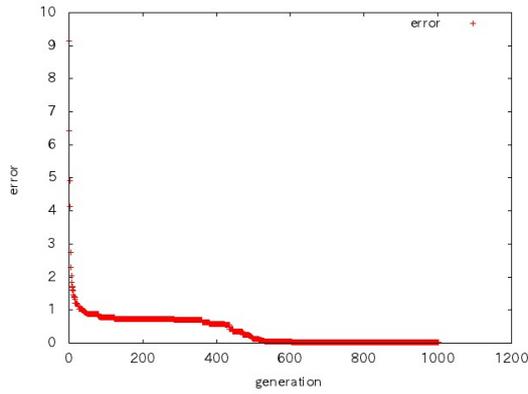


図 8: 世代交代に伴う誤差の変化

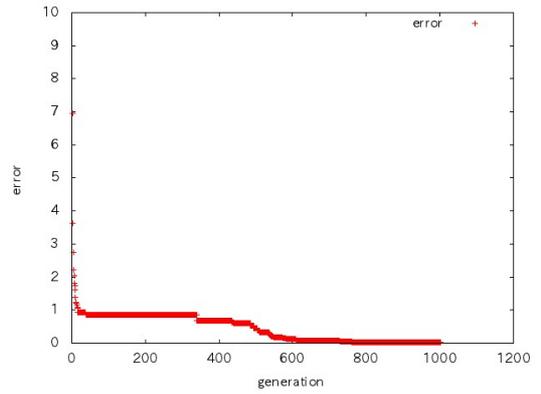


図 11: 世代交代に伴う誤差の変化

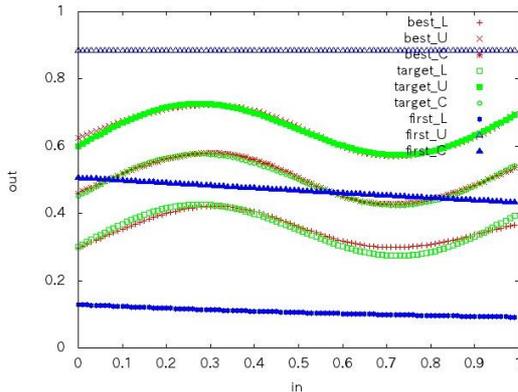


図 9: 目標関数および初期/最良個体による関数

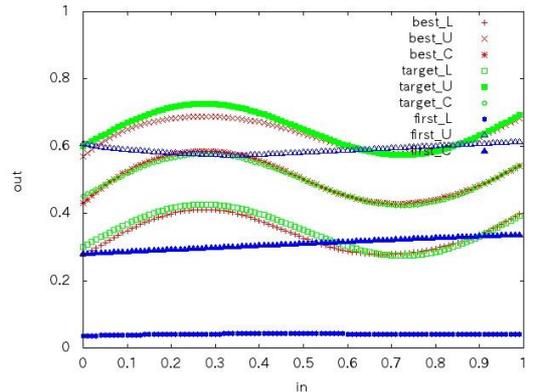


図 12: 目標関数および初期/最良個体による関数

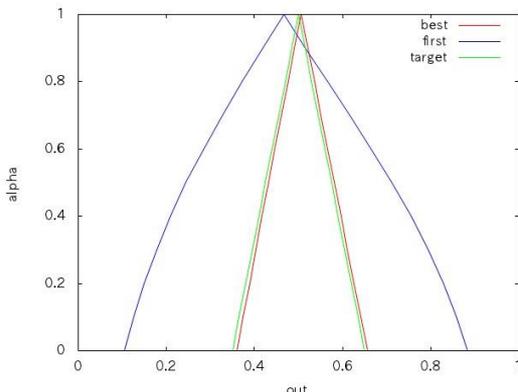


図 10:  $x=0.5$  に対する目標関数および初期/最良個体による関数の出力ファジィ数

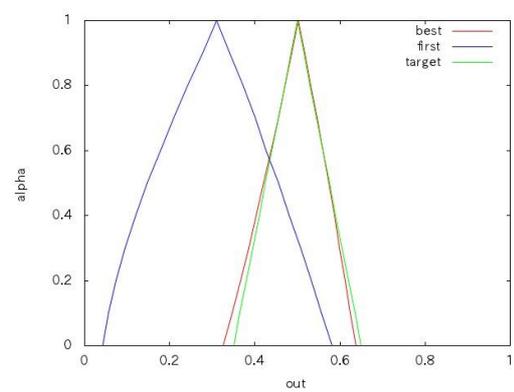


図 13:  $x=0.5$  に対する目標関数および初期/最良個体による関数の出力ファジィ数

よび、同じ入力  $x=0.5$  に対して初期個体/最良個体のニューラルネットワークが出力するファジィ数のメンバーシップ関数を表している。

次に、例題(2)の実験結果を例題(1)と同様に図 14-19 に示す。図 14-16 (図 17-19) は LU (CW) モデルを用いた場合の結果である。ファジィ ES を表 1 の通りに、また FNN の中間層ユニット数は LU/CW モデルとも 3 に設定したとき、近似誤差最小の解が得られた。

最後に、例題(3)の実験結果を例題(1)(2)と同様に図 20-25 に示す。図 20-22 (図 23-25) は LU (CW) モデルを用いた場合の結果である。ファジィ ES を表 1 の通りに、また FNN の中間層ユニット数は 5 (LU モデル) および 3 (CW モデル) に設定したとき、近似誤差最小の解が得られた。

これらの図に示した結果から、教師なし学習にも関わらず、最良個体の関数は目標関数をより精度で近似できていることがわかる。

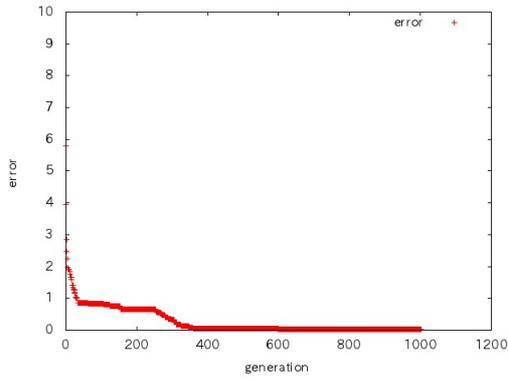


図 14: 世代交代に伴う誤差の変化

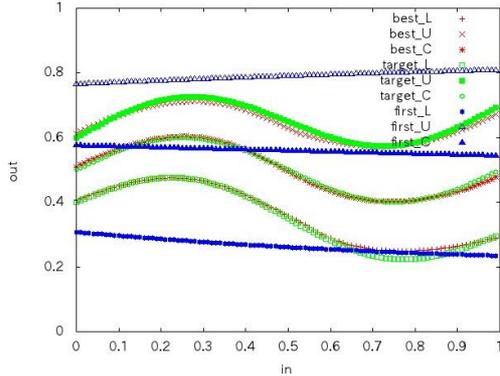


図 15: 目標関数および初期/最良個体による関数

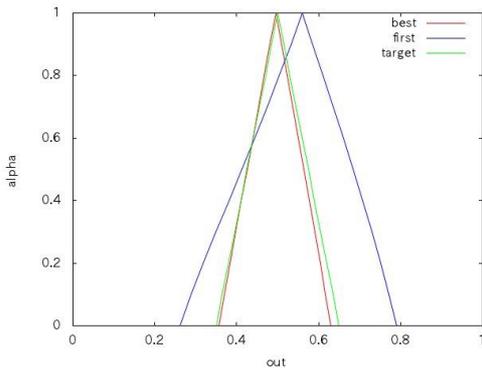


図 16:  $x=0.5$  に対する目標関数および初期/最良個体による関数の出力ファジィ数

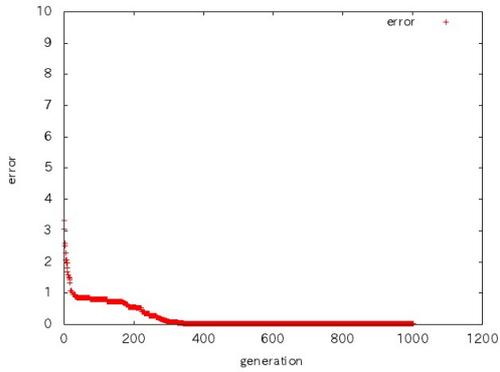


図 17: 世代交代に伴う誤差の変化

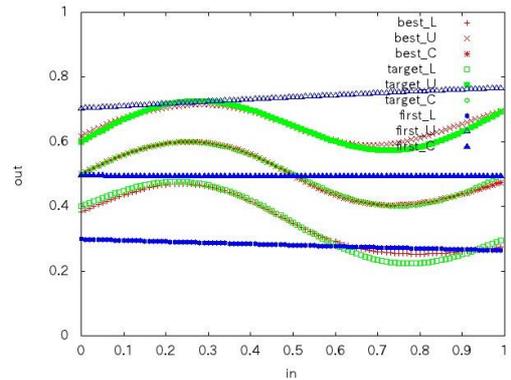


図 18: 目標関数および初期/最良個体による関数

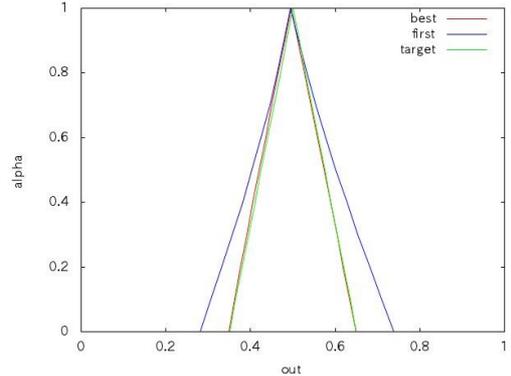


図 19:  $x=0.5$  に対する目標関数および初期/最良個体による関数の出力ファジィ数

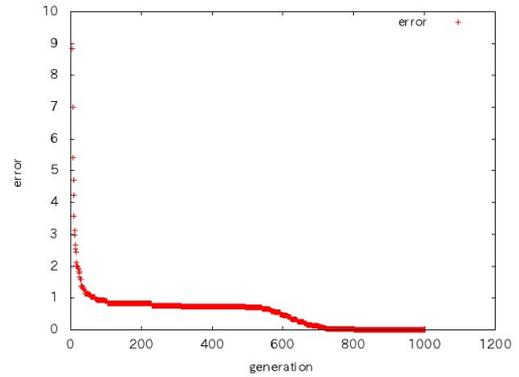


図 20: 世代交代に伴う誤差の変化

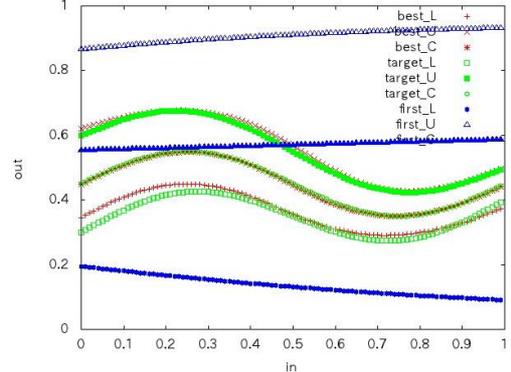


図 21: 目標関数および初期/最良個体による関数

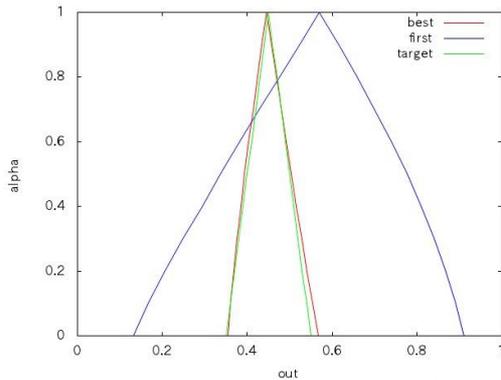


図 22:  $x=0.5$  に対する目標関数および初期/最良個体による関数の出力ファジィ数

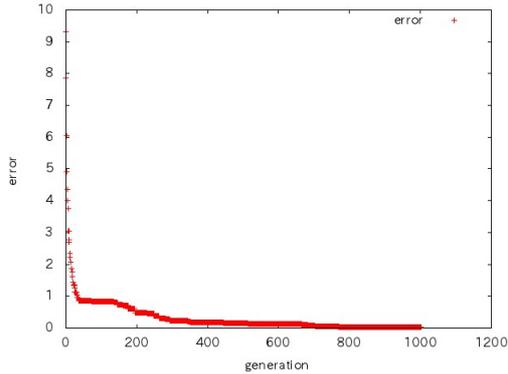


図 23: 世代交代に伴う誤差の変化

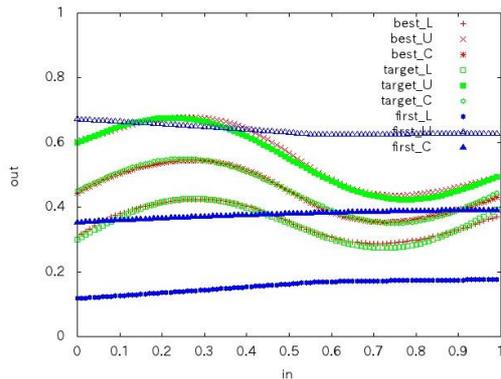


図 24: 目標関数および初期/最良個体による関数

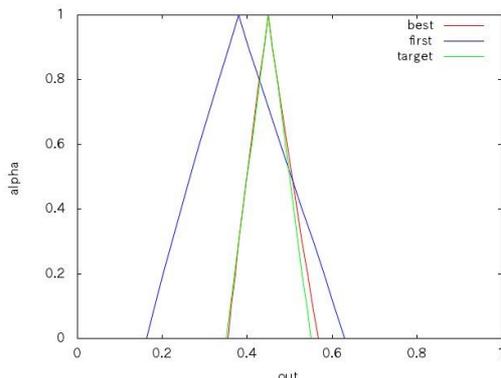


図 25:  $x=0.5$  に対する目標関数および初期/最良個体による関数の出力ファジィ数

## 5. 実験結果の考察

FNN のパラメータ (中間層ユニット数), ファジィ ES のパラメータ (全個体数統一のもとでの集団サイズと世代数など), 目標関数を様々に変えて実験を進めている. これまでの実験の結果から, 以下のような知見が得られてきている.

- 望ましい設定のもとでは, LU モデルより CW モデルの方が, 解の進化が良好である.
- $(\mu+\lambda)$ -ES より  $(\mu,\lambda)$ -ES の方が近似誤差の小さい解を得やすい.
- $(\mu,\lambda)$ -ES において,  $\lambda$  を  $\mu$  より相対的に大きくとる方が, 解の進化が良好である.
- 式(9)-(12)の rand には,  $[-1.0, 1.0]$  一様乱数より  $[-0.1, 0.1]$  一様乱数を用いるほうが, 解の進化が良好である.
- 生成個体数 100,000 の条件のもとで, 個体数を相対的に多くした場合より, 世代数を相対的に多くした方が, 解の進化が良好である.

## 6. まとめ

本稿では, 本研究で提案したファジィ ES を用いて FNN の教師なし学習を行う方法を示した. 本稿における FNN の学習目標は, 明示的には与えられない目標ファジィ関数を近似することである. 実験の結果, 提案手法を用いて FNN のファジィ結合強度とファジィしきい値をチューニングすることで, 教師なし学習にも関わらず, 良好な精度で FNN に目標関数を近似させられることがわかった. また, 様々な設定のもとでの実験の結果, 望ましい設定のもとでは LU モデルより CW モデルの方が解の進化が良好であることなどがわかった.

今後の課題として, 提案手法に関する 5 章の知見の一般性をさらなる評価実験によって確かめてゆくこと, 本研究で提案している他のファジィ進化計算手法 (ファジィ GA, ファジィ DE, ファジィ PSO など[3]) との比較, 他の問題への提案手法の適用に基づいた有効性評価などがある.

## 参考文献

- [1] H. Ishibuchi, H. Tanaka and H. Okada, Fuzzy neural networks with fuzzy weights and fuzzy biases, Proc. of IEEE International Conferences on Neural Networks, pp.1650-1655, 1993.
- [2] X. Yao, Evolving artificial neural networks, Proceedings of the IEEE, 87(9), 1423-1447, 1999.
- [3] 岡田, 解の遺伝子表現に区間値やファジィ値を用いた区間・ファジィ進化計算手法の提案, 情報処理学会関西支部平成 24 年度支部大会, 2012.
- [4] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning - I, II, and III, Information Sciences, Vol.8, pp.199-249, pp.301-357 and vol.9, pp.43-80, 1975.
- [5] G. Alefeld and J. Herzberger, Introduction to Interval Computation, Academic Press, 1983.