



ELSEVIER

Information Sciences 136 (2001) 53–67

INFORMATION
SCIENCES

AN INTERNATIONAL JOURNAL

www.elsevier.com/locate/ins

Fuzzy modeling using genetic algorithms with fuzzy entropy as conciseness measure

T. Suzuki ^{a,*}, T. Kodama ^b, T. Furuhashi ^b, H. Tsutsui ^c

^a *Department of Information Electronics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan*

^b *Department of Information Engineering, Mie University, 1515 Kamihama-cho, Tsu 514-8507, Japan*

^c *Research and Development Headquarters, Yamatake Corporation, 1-12-1, Kawana, Fujisawa-shi 251-8522, Japan*

Abstract

In this paper, a fuzzy modeling method using genetic algorithms (GAs) with a conciseness measure is presented. This paper introduces De Luca and Termini's fuzzy entropy to evaluate the shape of a membership function, and proposes another measure to evaluate the deviation of a membership function from symmetry. A combined measure is then derived from these two measures, and a new conciseness measure is defined for evaluation of the shape and allocation of the membership functions of a fuzzy model. Numerical results show that the new conciseness measure is effective for fuzzy modeling formulated as a multi-objective optimization problem. © 2001 Published by Elsevier Science Inc.

Keywords: Fuzzy modeling; Fuzzy entropy; Genetic algorithms; Conciseness measure; Linguistic interpretability

1. Introduction

Fuzzy models [1] have been constructed by knowledge acquisition from experts, but the knowledge acquisition through interviews has often been

* Corresponding author.

E-mail addresses: suzuki@bioele.nuee.nagoya-u.ac.jp (T. Suzuki), kodama@pa.info.mie-u.ac.jp (T. Kodama), furuhashi@pa.info.mie-u.ac.jp (T. Furuhashi), tsutsui@atc.yamatake.co.jp (H. Tsutsui).

difficult because they seldom have explicit knowledge that can be represented with fuzzy if–then type rules.

Many methods, that automatically derive if–then type fuzzy rules from numerical data, have been proposed to overcome the problem of knowledge acquisition. Since tuning of both the antecedent and consequent part of fuzzy rules can be formulated as an optimization problem, evolutionary algorithms have been applied to solve this problem. However, automatically derived fuzzy models are not often linguistically interpretable, as recognized in the literature [2–4].

In this paper, fuzzy modeling using genetic algorithms (GAs) using a new conciseness measure is presented. Conciseness is a criterion that represents the linguistic interpretability of fuzzy models, and it is defined by referring to the shape and allocation of the membership functions of a fuzzy model.

The conciseness of fuzzy models has been evaluated by the number of fuzzy rules, the number of membership functions [5], or the degree of freedom term of Akaike's information criterion (AIC) [6,7]. However, fuzzy models that have the same number of membership functions cannot be distinguished with these measures.

This paper introduces De Luca and Termini's fuzzy entropy [8] for evaluation of the shape of a membership function. De Luca and Termini proposed fuzzy entropy as a measure of fuzziness. They used Shannon's function, and defined a measure that became largest at the grade of membership of 0.5. Several authors have attempted to quantify fuzziness and proposed fuzzy entropy [9,10]. This measure has been applicable to various industrial applications, e.g., image processing [11], fuzzy clustering [12], etc.

De Luca and Termini's fuzzy entropy, however, cannot evaluate the deviation of a membership function. This paper proposes a new measure for the deviation of a membership function from symmetry, and derives a combined measure of their fuzzy entropy and this proposed measure. Then a new conciseness measure is defined for evaluating the shape and allocation of the membership functions of a fuzzy model.

Since the new conciseness measure is in conflict with the accuracy of fuzzy models, fuzzy modeling using GAs with these two criteria is formulated as a multi-objective optimization problem [13–16].

The rest of this paper is organized as follows. Section 2 describes fuzzy modeling using GAs with the new conciseness measure. Next, the conciseness of fuzzy models are discussed in Section 3 in an illustrative way, followed by Section 4 that defines a new conciseness measure. And in Section 5, numerical results show that the conciseness measure is in conflict with the accuracy of fuzzy models, and then is applied to fuzzy modeling using GAs. Finally, concluding remarks are given in Section 6.

2. Fuzzy modeling using GAs with conciseness measure

This section describes the details of fuzzy modeling using GAs with a new conciseness measure. This measure is defined in Section 4 that follows the discussion of the conciseness of fuzzy models in Section 3. In this paper, a very simple procedure of fuzzy modeling using GAs is described to clarify the feasibility of the new conciseness measure. The goal of the fuzzy modeling is to obtain fuzzy models which are concise enough for human beings, not to mention that they should have accurate model output, thus the fuzzy modeling is considered as a multi-objective optimization problem.

2.1. Fuzzy model

A single-input single-output fuzzy model with simplified fuzzy inference [17] is used in this paper. The output y of a fuzzy model with the input x is given by

$$y = \sum_{i=1}^{N_m} \mu_i(x) \cdot c_i, \quad (1)$$

where $\mu_i(x)$ and c_i ($i = 1, \dots, N_m$) are grades of membership in the antecedent parts and singletons in the consequent parts, respectively. N_m is the number of membership functions.

Fuzzy models are identified from a set of data D and the singleton in the consequent part of a rule is given by

$$c_i = \sum_{k=1}^{N_d} y_k \cdot \mu_i(x_k), \quad (2)$$

where N_d is the number of input–output pairs (x_k, y_k) ($k = 1, \dots, N_d$).

2.2. Membership functions

The following conditions for allocating membership functions are used to limit the search space of fuzzy modeling:

(a) For all $x \in X$, membership functions $\mu_i(x)$ ($i = 1, \dots, N_m$) satisfy

$$\sum_{i=1}^{N_m} \mu_i(x) = 1. \quad (3)$$

(b) For all x that are not crest points, i.e. x s.t. $\mu_i(x) = 1$, there are exactly two fuzzy sets for which $\mu_i(x) > 0$ and for all the others $\mu_i(x) = 0$.

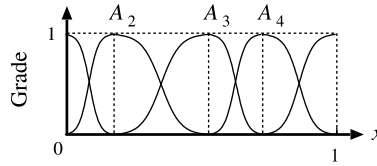


Fig. 1. Example allocation of membership functions that satisfies the conditions.

(c) Each membership function is similar with respect to the crest point $x = a$, in the sense that $\mu_{i_i}(x) = \mu_{r_i}(1 - \frac{1-a}{a}x)$, where $\mu_{i_i}(x) = \{\mu_i(x) \mid x \leq a, \mu_i(x) > 0\}$, $\mu_{r_i}(x) = \{\mu_i(x) \mid a \leq x, \mu_i(x) > 0\}$.

(d) All the membership functions are convex.

An example allocation of membership functions that satisfies these conditions is shown in Fig. 1. These conditions allow a fuzzy model to be determined with the positions of the crest points and the shape of membership functions.

2.3. Rank-based evaluation

A rank-based evaluation is used for finding Pareto-optimal solutions with the two criteria: conciseness and accuracy. Conciseness is measured by the conciseness measure defined in Eq. (13), and accuracy is measured by mean squared error given by

$$\frac{1}{N_t} \sum_{k=1}^{N_t} (y_k - \hat{y}_k)^2, \quad (4)$$

where N_t is the number of the test data, y is the output of the modeling target, and \hat{y} is the model output.

The rank of each chromosome i in the population R_i is given by

$$R_i = 1 + q_i, \quad (5)$$

where chromosome i is inferior to q_i chromosomes.

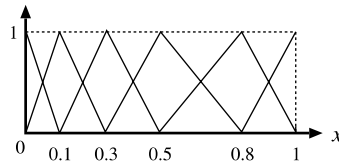
2.4. Chromosome encoding

Since simplified fuzzy inference is employed and the allocation of membership functions is restricted with conditions (a)–(d) in Section 2.2, the parameters of a fuzzy model are the following two parameters: the positions of the crest points of the membership functions and the shape of the membership functions.

The positions of the crest points and the shape of the membership functions of a fuzzy model are encoded into a chromosome. The length of all the chromosomes in the population is fixed, and each chromosome has $N_m + 1$

0.1	0.3	0.5	0.8	trimf
position of crest points				shape

(a)



(b)

Fig. 2. Example of a chromosome: (a) chromosome; (b) allocation of membership functions represented by the above chromosome.

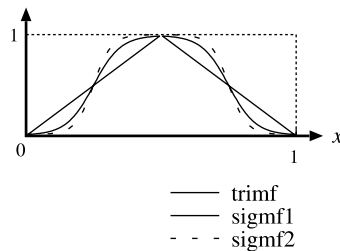


Fig. 3. Shapes of membership functions.

genes: N_m for storing the positions of the crest points and one for the shape. An example is shown in Fig. 2, where $N_m = 6$. Figs. 2(a) and (b) show a chromosome and the allocation of membership functions represented by the chromosome, respectively.

The positions of the crest points of the membership functions at both of the ends are always $x = 0$ and 1 , respectively. Only the positions of the intermediate crest points appear in a chromosome. The shape parameter is either “trimf”, “sigmf1” or “sigmf2”, and the shapes of them are shown in Fig. 3.

The number of the chromosomes in the population was fixed at N_c .

2.5. Genetic operators

Genetic operators are applied in the usual order: selection, crossover and mutation. The selection operation is based on rank which is assigned using the conciseness measure and the accuracy. The roulette wheel selection method is used to select N_c chromosomes. The probability for the roulette wheel selection is given by

$$\frac{P_i}{\sum_{i'} P_{i'}}, \quad (6)$$

where $P_i = 1/R_i$. For the crossover operation, the chromosomes are selected in pairs, and for each pair the two chromosomes are crossed over at a random position. Some chromosomes among the lowest ranked chromosomes are randomly selected and mutated.

3. Conciseness of fuzzy models

One of the most important features of a fuzzy model is its consistent representation of pattern knowledge and symbolic knowledge. In this paper, a pattern means a feature vector extracted from raw data, and a symbol means a label assigned to a pattern set which has a fuzzy border. And a fuzzy pattern set is defined by a membership function, which is effective for agreement between information contained in a continuous space and symbolic knowledge in a discrete space. Fuzzy inference, that uses min–max–center of gravity or product–sum–center of gravity for treating pattern–symbol pairs, is a good tool for interface between pattern processing in a continuous space and symbolic processing in a discrete space. Humans are clearly conscious of symbolic processing, and a fuzzy model using pattern–symbol pairs to describe its processing is comprehensible to humans.

The question here is about the quantitative measure of *comprehensibility* of fuzzy models. How comprehensible is a model with fuzzy membership functions more than that with crisp membership functions? How about the case with neural networks?

This paper studies the *conciseness* of fuzzy models. A concise fuzzy model is comprehensible. Let us try to grasp the input–output relationships from fuzzy if–then type rules. The following three fuzzy rules are assumed to be given:

If x is SMALL, then y is 1.0.

If x is MEDIUM, then y is 0.4.

If x is BIG, then y is 0.2.

The membership functions, SMALL, MEDIUM and BIG, in Fig. 4 are also assumed to be given.

The membership functions in Fig. 4(a) are crisp and equidistantly allocated on the universe of discourse. Those in Figs. 4(b) and (c) are triangular ones. While the membership functions in (b) are evenly allocated on the universe of discourse, those in (c) are unevenly allocated. The membership functions in Fig. 4(d) have parts with gentle slope where grades of membership are nearly equal to 0.5, and are not equidistantly allocated. The solid lines in Fig. 5 show

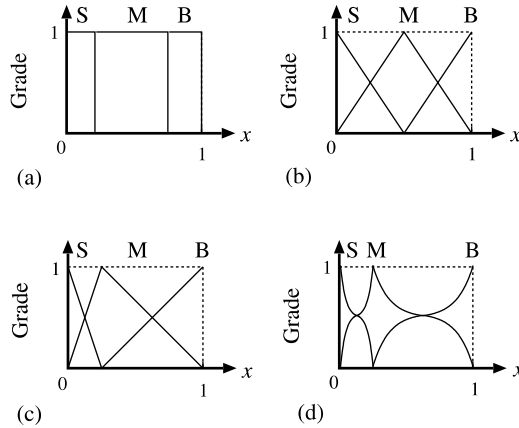


Fig. 4. Examples of membership functions: (a) crisp, equidistant allocation; (b) triangular, equidistant allocation; (c) triangular, not-equidistant allocation; (d) grade nearly equal to 0.5, not-equidistant allocation.

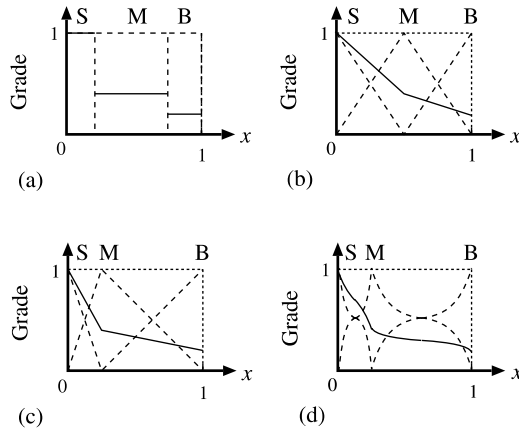


Fig. 5. Input–output relationships: (a) crisp, equidistant allocation; (b) triangular, equidistant allocation; (c) triangular, not-equidistant allocation; (d) grade nearly equal to 0.5, not-equidistant allocation.

the input–output relationships obtained from the membership functions in Fig. 4.

The question that arises here is which relationships are easier for us to image only from the above three fuzzy if–then type rules. From the three discrete rules, we can image the step-wise input–output relationships of model (a) most easily. It becomes more difficult with model (b) and more with model (c). The relationships of model (d) are the most difficult for us to image.

From the above observation, the conciseness of fuzzy models is defined in this paper as follows.

Definition 1 (*Conciseness of fuzzy model*). A fuzzy model is said to be more concise if the membership functions are more equidistantly allocated on the universe of discourse, and the shapes of membership functions are less fuzzy.

Definition 1 defines the conciseness of fuzzy models as the easiness for grasping the correspondence between the discrete fuzzy rules and the continuous values.

4. Fuzzy entropy

A quantitative measure of the conciseness of fuzzy models is examined in the following two subsections.

4.1. De Luca and Termini's fuzzy entropy

De Luca and Termini [8] defined fuzzy entropy of fuzzy set A as

$$d(A) = \int_{x_1}^{x_2} \{-\mu_A(x) \ln \mu_A(x) - (1 - \mu_A(x)) \ln(1 - \mu_A(x))\} dx, \quad (7)$$

where $\mu_A(x)$ is the membership function of fuzzy set A . If $\mu_A(x) = 0.5$ for all x on the support of A , then the fuzzy entropy of fuzzy set A is the maximum.

This fuzzy entropy can distinguish the shapes of membership functions, i.e., triangular shape, sigmoidal shape, etc., and coincides with the definition of the conciseness. Thus this entropy can be a candidate for a quantitative measure of the conciseness of fuzzy models.

In the case where conditions (a) and (b) in Section 2.2 are given, De Luca and Termini's entropy can be simplified. Assuming that two membership functions $\mu_A(x)$ and $\mu_B(x)$ are overlapping and for all $x \in [x_1, x_2]$ $\mu_A(x) + \mu_B(x) = 1$, then

$$d(A) + d(B) = -2 \int_{x_1}^{x_2} \{\mu_A(x) \ln \mu_A(x) + \mu_B(x) \ln \mu_B(x)\} dx. \quad (8)$$

Under conditions (a) and (b) in Section 2.2, we can use the following measure for evaluation of the shape of membership functions instead of Eq. (7):

$$d(A) = - \int_{x_1}^{x_2} \mu_A(x) \ln \mu_A(x) dx. \quad (9)$$

4.2. Measure for deviation of membership function

This paper defines a quantitative measure of the deviation of a membership function from symmetry. The membership function is assumed to satisfy conditions (a)–(d) in Section 2.2. This measure is defined by considering Eq. (9).

Definition 2 (*Measure for deviation of membership function*). The measure for the deviation of fuzzy set A from symmetry is given by

$$r(A) = \int_{x_1}^{x_2} \mu_C(x) \ln \frac{\mu_C(x)}{\mu_A(x)} dx, \tag{10}$$

where x_1 and x_2 are the left and right terminal points of the support of fuzzy set A , respectively; $\mu_A(x)$ is the membership function of fuzzy set A ; $\mu_C(x)$ is a symmetrical membership function with respect to the vertical line through the crest, which has the same support as that of fuzzy set A .

Fig. 6 illustrates an example of fuzzy sets A and C .

In the case where the shape of the membership functions in Fig. 6 is triangular, the measure $r_{tri}(A)$ is expressed as

$$r_{tri}(A) = s \left[\frac{|d|}{s} - \frac{1}{2} \ln \left\{ 2 \left(\frac{1}{2} + \frac{|d|}{s} \right) \right\} \right] \quad \left(0 < a \leq 1, -\frac{1}{2} \leq d \leq \frac{1}{2} \right), \tag{11}$$

where s is the width of support, and d is the deviation of the crest point of fuzzy set A from that of the isosceles triangular fuzzy set C . The value of this measure $r_{tri}(A)$ is monotonically increasing with the absolute value of d . Numerical calculation gives that $r(A)$ is also monotonically increasing with the absolute value of d with any shapes of membership functions that satisfy conditions (a)–(d) in Section 2.2. This measure, which evaluates the deviation of a membership

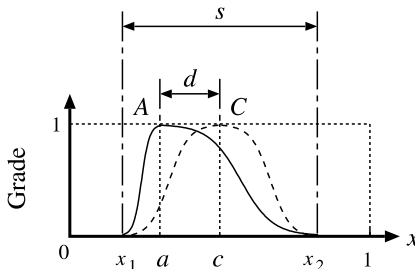


Fig. 6. Membership function A and symmetrical membership function C .

function, is another good candidate for the conciseness measure of fuzzy models. A combination of De Luca and Termini's fuzzy entropy in Eq. (9) and the deviation measure in Eq. (10) can evaluate the conciseness defined in Section 3.

4.3. Combined measure

One way of combining the two measures is summation. By summing the fuzzy entropy $d(A)$ in Eq. (9) and the measure for deviation of a membership function $r(A)$ in Eq. (10), a new measure $dr(A)$ is obtained

$$\begin{aligned} dr(A) &= d(A) + r(A) = - \int_{x_1}^{x_2} \mu_A(x) \ln \mu_A(x) + \int_{x_1}^{x_2} \mu_C(x) \ln \frac{\mu_C(x)}{\mu_A(x)} dx \\ &= - \int_{x_1}^{x_2} \mu_C(x) \ln \mu_A(x) dx. \end{aligned} \quad (12)$$

The fuzzy entropy $d(A)$ can evaluate the shape of a membership function. And if the shape is fixed, the measure $r(A)$ can evaluate the deviation of a membership function.

4.4. Conciseness measure

A new conciseness measure dr_{avr} is introduced to evaluate the shapes and allocations of N_m fuzzy sets A_i ($i = 1, \dots, N_m$) on the universe of discourse X on x -axis. The new conciseness measure dr_{avr} is defined as

$$dr_{avr} = \frac{1}{N_m - 2} \sum_{i=2}^{N_m-1} dr(A_i), \quad (13)$$

where $dr(A)$ is the combined measure in Eq. (12), which evaluates the shape and deviation of a membership function, N_m is the number of fuzzy sets A_i ($i = 1, \dots, N_m$) on the universe of discourse X on x -axis.

5. Numerical results

This section describes numerical results to show usefulness of the conciseness measure for fuzzy modeling. The following single-input/single-output function is used as a modeling target throughout this section:

$$f(x) = \begin{cases} 1 - 2x & (0 \leq x \leq 0.5), \\ -4x^2 + 8x - 3 & (0.5 < x \leq 1). \end{cases} \quad (14)$$

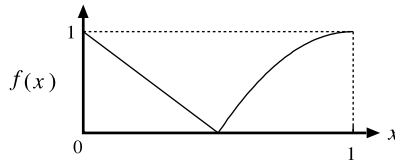


Fig. 7. Modeling target for numerical experiments ($f(x)$ in Eq. (14)).

Fig. 7 depicts this function. Conditions (a)–(d) in Section 2.2 was imposed and two membership functions were set to overlap everywhere.

5.1. Conciseness measure vs. accuracy of fuzzy models

To examine the relationships between the conciseness measure (dr_{avr}) and the accuracy, 1000 fuzzy models were randomly generated and their conciseness measure and accuracy were calculated. Among them, the fuzzy models near the Pareto front are shown in Fig. 8 with their conciseness measure and accuracy. In this case, the shape of membership functions was fixed to triangular and the number of membership functions of a fuzzy model was set at 6. Each dot in the figure corresponds to a fuzzy model that has a unique combination of the crest points of membership functions. From Fig. 8, it is observed that the conciseness measure and the accuracy are in conflict as indicated with the broken line.

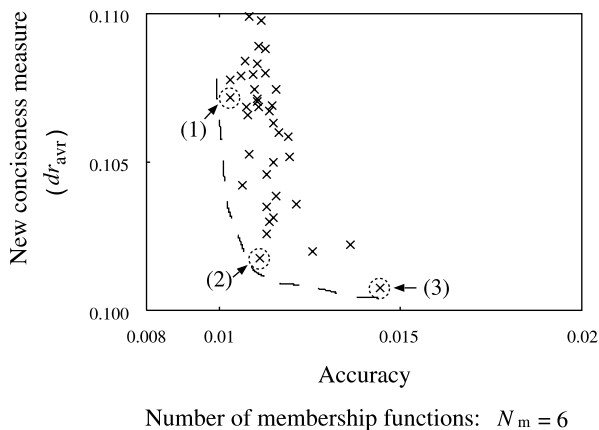


Fig. 8. New conciseness measure vs. accuracy of fuzzy models with triangular membership functions.

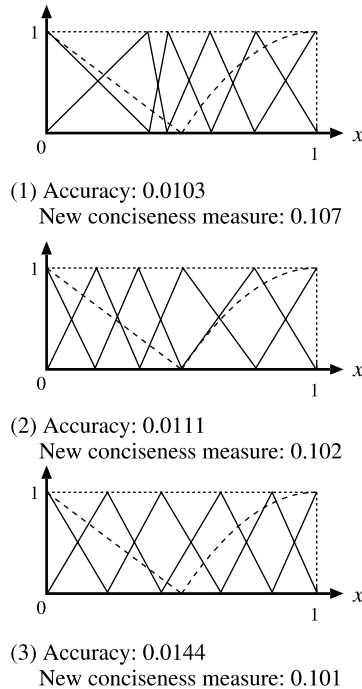


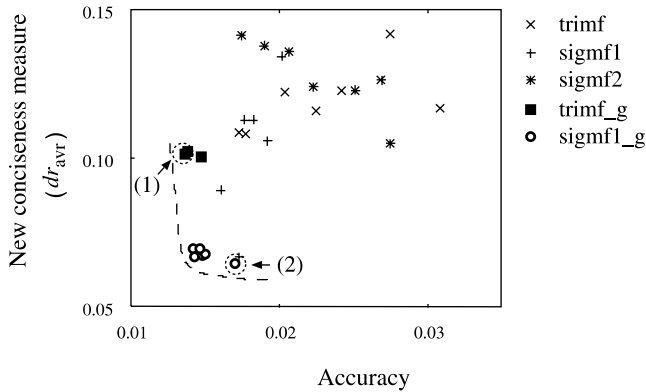
Fig. 9. Allocation of the membership functions of the fuzzy models (1), (2) and (3) in Fig. 8.

Figs. 9(1), (2) and (3) show the allocations of the membership functions of the fuzzy models (1), (2) and (3), which were on the Pareto front in Fig. 8, respectively. The less the conciseness measure was, the more equidistant the allocation of membership functions was.

The conciseness measure was also examined with fuzzy models that have various shapes of membership functions. 10 fuzzy models were taken from the Pareto front in Fig 8, and for each fuzzy model, the conciseness measure dr_{avr} and the accuracy were calculated by varying the shapes of membership functions. It was observed that the conciseness measure and the accuracy were in conflict when the shapes of membership functions were near triangular.

5.2. Fuzzy modeling using GAs with conciseness measure

Fuzzy modeling using GAs with the conciseness measure was done. The parameters for the numerical experiments were the following: the number of chromosomes N_c was set at 50; the shape parameter in a chromosome was either “trimf”, “sigmf1” or “sigmf2” in Fig 3; the crossover rate and mutation rate were set at 0.5 and 0.05, respectively.



Number of membership functions: $N_m = 6$

Fig. 10. Fuzzy models acquired after 10 generations of genetic operations.

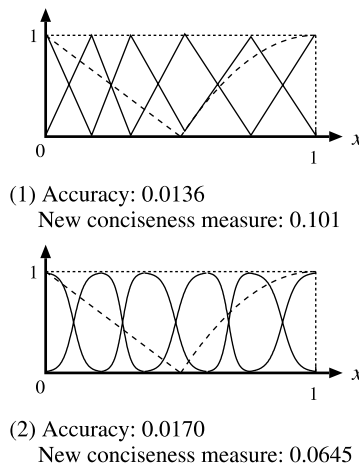


Fig. 11. Allocation of membership functions of acquired fuzzy models.

Figs. 10 and 11 show the results. The dots labeled “trimf”, “sigmf1” and “sigmf2” are the initial chromosomes, which were randomly generated, and the dots labeled “trimf_g” and “sigmf1_g” are the chromosomes after 10 generations of genetic operations.

From Fig. 10, it is observed that the fuzzy models are distributed on the Pareto front at the 10th generation as indicated with the broken line. Figs. 11(1) and (2) show the allocations of the membership functions of the fuzzy models (1) and (2), which were on the Pareto front in Fig. 10, respectively. A trade-off between the conciseness measure and the accuracy enabled the successful search for variety of concise fuzzy models with good accuracy.

6. Conclusions

This paper presented a fuzzy modeling method using the new measure for the conciseness of fuzzy models. This paper defined the conciseness of fuzzy models, and quantified the conciseness by introducing fuzzy entropy. De Luca and Termini's fuzzy entropy could evaluate the shapes of membership functions, but their entropy could not distinguish similar shaped membership functions. This paper defined a measure for deviation of a membership function from symmetry. This is another measure for the conciseness of fuzzy models. With De Luca and Termini's measure and the measure for deviation, a combined measure was derived. Based on the combined measure, a conciseness measure was defined to evaluate the shape and allocation of membership functions of a fuzzy model. This new conciseness measure was in conflict with the accuracy of fuzzy models in the case where the membership functions were near triangular. Numerical results showed that the new conciseness measure was effective for fuzzy modeling formulated as a multi-objective optimization problem.

References

- [1] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, *IEEE Trans. Syst., Man, Cybern.* 15 (1985) 116–132.
- [2] M. Setnes, R. Babuška, H.B. Verbruggen, Rule-based modeling: precision and transparency, *IEEE Trans. Syst., Man, Cybern.* C 28 (1) (1998) 165–169.
- [3] J.V. de Oliveira, Semantic constraints for membership function optimization, *IEEE Trans. Syst., Man, Cybern.* A 29 (1) (1999) 128–138.
- [4] M. Setnes, H. Roubos, Ga-fuzzy modeling and classification: complexity and performance, *IEEE Trans. Fuzzy Syst.* 8 (5) (2000) 509–522.
- [5] S. Matsushita, T. Furuhashi, et al., Selection of input variables using genetic algorithm for hierarchical fuzzy modeling, in: *Proceedings of the 1996 First Asia-Pacific Conference on Simulated Evolution and Learning*, 1996, pp. 106–113.
- [6] H. Akaike, Information theory and an extension of the maximum likelihood principle, in: *2nd International Symposium on Information Theory*, 1973, pp. 267–281.
- [7] H. Nomura, S. Araki, I. Hayashi, N. Wakami, A learning method of fuzzy reasoning by delta rule, in: *Proceedings of Intelligent System Symposium*, 1992, pp. 25–30.
- [8] A. De Luca, S. Termini, A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory, *Inf. Control* 20 (1972) 301–312.
- [9] N.R. Pal, J.C. Bezdek, Measuring fuzzy uncertainty, *IEEE Trans. Fuzzy Syst.* 2 (2) (1994) 107–118.
- [10] N.R. Pal, On quantification of different facets of uncertainty, *Fuzzy Sets Syst.* 107 (1999) 81–91.
- [11] S.D. Zeno, L. Cinque, Image thresholding using fuzzy entropies, *IEEE Trans. Syst., Man, Cybern.* B 28 (1) (1998) 15–23.
- [12] S. Miyamoto, M. Mukaidono, Fuzzy c -means as a regularization and maximum entropy approach, in: *Proceedings of the 7th International Fuzzy Systems Association World Congress (IFSA'97)*, vol. II, 1997, pp. 86–92.

- [13] J.D. Schaffer, Multiple objective optimization with vector evaluated genetic algorithms, in: Proceedings of the First International Conference on Genetic Algorithms and Their Applications, 1985, pp. 93–100.
- [14] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Reading, MA, 1989.
- [15] S.J. Louis, G.J.E. Rawlins, Pareto optimality, ga-easiness and deception, in: Proceedings of the Fifth International Conference on Genetic Algorithms, 1993, pp. 118–123.
- [16] C.M. Fonseca, P.J. Fleming, Genetic algorithms for multiobjective optimization: formulation, discussion and generalization, in: Proceedings of the Fifth International Conference on Genetic Algorithms, 1993, pp. 416–423.
- [17] M. Mizumoto, Fuzzy control under various approximate reasoning methods, in: Proceedings of the Second IFSA Congress, 1987, pp. 143–146.