

# Multi-Objective Evolutionary Design of Fuzzy Rule-Based Systems\*

**Hisao Ishibuchi**

Department of Industrial Engineering  
Osaka Prefecture University  
Sakai, Osaka, Japan  
hisaoi@ie.osakafu-u.ac.jp

**Takashi Yamamoto**

Department of Industrial Engineering  
Osaka Prefecture University  
Sakai, Osaka, Japan  
yama@ie.osakafu-u.ac.jp

**Abstract** - This paper clearly demonstrates advantages of our evolutionary multiobjective optimization approach to the design of fuzzy rule-based classification systems over single-objective methods. The main advantage of our approach is that a large number of tradeoff (i.e., non-dominated) fuzzy rule-based systems can be obtained by its single run with respect to conflicting objectives: accuracy maximization and complexity minimization. By analyzing the obtained fuzzy rule-based systems, the decision maker can understand the tradeoff between these two objectives. Such understanding is of great help when the decision maker chooses a final compromise fuzzy rule-based system. In the case of single-objective methods, only a single fuzzy rule-based system is obtained based on the pre-specified preference of the decision maker. We compare four formulations of genetic algorithm-based rule selection through computational experiments on well-known benchmark data sets. The four formulations have two objectives, their weighted sum, three objectives, and their weighted sum, respectively.

**Keywords:** Pattern classification, fuzzy systems, rule selection, genetic algorithms, multiobjective optimization.

## 1 Introduction

In 1990s, various approaches were proposed to improve the accuracy of fuzzy rule-based systems. Those approaches were often based on evolutionary computation [6], [19] and neural networks [1], [20], [25]. While the accuracy of fuzzy rule-based systems can be significantly improved by those approaches, the complexity is also increased. That is, the interpretability is usually degraded by the learning of fuzzy rule-based systems. For obtaining fuzzy rule-based systems with high interpretability, some recent studies took into account the tradeoff between the accuracy of fuzzy rule-based systems and their complexity [3]-[5], [11]-[18], [22]-[24]. A prevailing strategy for handling the tradeoff in those studies is to combine an accuracy measure and a complexity measure into a single

scalar objective function so that standard optimization techniques can be employed. Only a few studies handled the design of fuzzy rule-based systems in the framework of multiobjective optimization where a large number of non-dominated fuzzy rule-based systems are obtained.

In this paper, we intend to demonstrate advantages of multiobjective approaches over single-objective ones to the design of fuzzy rule-based classification systems. As in our former studies [12]-[16], we use the following three-objective formulation to find a large number of non-dominated fuzzy rule-based classification systems:

$$\text{Maximize } f_1(S) \text{ and minimize } f_2(S), f_3(S), \quad (1)$$

where  $S$  is a fuzzy rule-based classification system,  $f_1(S)$  is the number of correctly classified training patterns by  $S$ ,  $f_2(S)$  is the number of fuzzy rules in  $S$ , and  $f_3(S)$  is the total rule length of fuzzy rules in  $S$ . Since the number of antecedent conditions of each rule is referred to as the rule length,  $f_3(S)$  is the same as the total number of antecedent conditions of fuzzy rules in  $S$ . In the three-objective formulation in (1), the first objective is viewed as accuracy maximization while the second and third objectives correspond to complexity minimization.

To examine the effect of the third objective (i.e., the minimization of the total rule length), we also use the following two-objective formulation:

$$\text{Maximize } f_1(S) \text{ and minimize } f_2(S). \quad (2)$$

Advantages of these multiobjective formulations are demonstrated through computational experiments on some benchmark data sets from UC Irvine Machine Learning Repository. The three-objective formulation in (1) is compared with the single-objective approach based on the following weighted scalar objective function:

$$\text{Maximize } w_1 f_1(S) - w_2 f_2(S) - w_3 f_3(S). \quad (3)$$

Moreover the two-objective formulation in (2) is compared with the single-objective approach based on the following weighted scalar objective function:

$$\text{Maximize } w_1 f_1(S) - w_2 f_2(S). \quad (4)$$

\* 0-7803-8566-7/04/\$20.00 © 2004 IEEE.

## 2 Design of fuzzy rule-based systems

Let us assume that we have  $m$  training patterns  $\mathbf{x}_p = (x_{p1}, \dots, x_{pn})$ ,  $p = 1, 2, \dots, m$  from  $M$  classes where  $x_{pi}$  is the attribute value of the  $p$ -th pattern for the  $i$ -th attribute ( $i = 1, 2, \dots, n$ ). For simplicity of explanation, we assume that the pattern space is normalized into the  $n$ -dimensional unit hyper-cube  $[0, 1]^n$ , i.e.,  $x_{pi} \in [0, 1]$ . In this section, we explain our three-objective approach [16] to the design of fuzzy rule-based classification systems from the given training patterns. Our approach consists of two stages: heuristic extraction of candidate fuzzy rules and genetic rule selection where an evolutionary multiobjective optimization algorithm is used to find a large number of non-dominated rule sets with respect to the three objectives in (1).

### 2.1 Fuzzy classification rules

For our  $n$ -dimensional  $M$ -class pattern classification problem, we use fuzzy rules of the following form:

$$\text{Rule } R_q: \text{If } x_1 \text{ is } A_{q1} \text{ and } \dots \text{ and } x_n \text{ is } A_{qn} \\ \text{then Class } C_q \text{ with } CF_q, \quad (5)$$

where  $R_q$  is the label of the  $q$ -th rule,  $\mathbf{x} = (x_1, \dots, x_n)$  is an  $n$ -dimensional pattern vector,  $A_{qi}$  is an antecedent fuzzy set (i.e., linguistic value such as *small* and *large*),  $C_q$  is a class label, and  $CF_q$  is a rule weight. Fuzzy rules of this form were used for classification problems in [11]-[16].

We define the compatibility grade of each training pattern  $\mathbf{x}_p = (x_{p1}, \dots, x_{pn})$  with the antecedent part  $\mathbf{A}_q = (A_{q1}, \dots, A_{qn})$  using the product operator as

$$\mu_{\mathbf{A}_q}(\mathbf{x}_p) = \mu_{A_{q1}}(x_{p1}) \cdot \dots \cdot \mu_{A_{qn}}(x_{pn}), \quad (6)$$

where  $\mu_{A_{qi}}(\cdot)$  is the membership function of  $A_{qi}$ . To determine the consequent class  $C_q$ , we calculate the confidence of the fuzzy rule " $\mathbf{A}_q \Rightarrow \text{Class } h$ " for each class as an extension of its non-fuzzy version [2] as

$$c(\mathbf{A}_q \Rightarrow \text{Class } h) = \frac{\sum_{\mathbf{x}_p \in \text{Class } h} \mu_{\mathbf{A}_q}(\mathbf{x}_p)}{\sum_{p=1}^m \mu_{\mathbf{A}_q}(\mathbf{x}_p)}. \quad (7)$$

The consequent class  $C_q$  is specified by identifying the class with the maximum confidence:

$$c(\mathbf{A}_q \Rightarrow \text{Class } C_q) = \max_{h=1,2,\dots,m} \{c(\mathbf{A}_q \Rightarrow \text{Class } h)\}. \quad (8)$$

On the other hand, the rule weight  $CF_q$  is specified as follows:

$$CF_q = c(\mathbf{A}_q \Rightarrow \text{Class } C_q) - \sum_{\substack{h=1 \\ h \neq C_q}}^m c(\mathbf{A}_q \Rightarrow \text{Class } h). \quad (9)$$

When we classify a new pattern  $\mathbf{x}_p$  by a fuzzy rule-based classification system  $S$ , we use a single winner-based fuzzy reasoning method. A single winner rule  $R_w$  is chosen from the rule set  $S$  for the new pattern  $\mathbf{x}_p$  as

$$\mu_{\mathbf{A}_w}(\mathbf{x}_p) \cdot CF_w = \max \{ \mu_{\mathbf{A}_q}(\mathbf{x}_p) \cdot CF_q \mid R_q \in S \}. \quad (10)$$

### 2.2 Heuristic extraction of candidate rules

We use 14 antecedent fuzzy sets in Figure 1 to generate candidate rules. That is, we simultaneously use four fuzzy partitions of different granularities for each attribute. We also use *don't care* as an antecedent fuzzy set. Thus the total number of possible combinations of the antecedent fuzzy sets is  $(14+1)^n$  for  $n$ -dimensional classification problems.

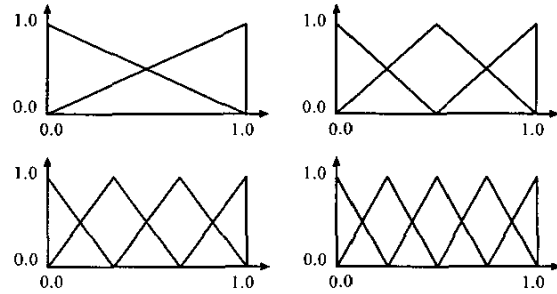


Figure 1. Fourteen fuzzy sets from four fuzzy partitions.

It is impractical to examine all the  $15^n$  combinations of the antecedent fuzzy sets for high-dimensional classification problems. In this paper, we only examine short fuzzy rules with only a few antecedent conditions (i.e., with many *don't care* conditions). The number of antecedent conditions excluding *don't care* conditions of each fuzzy rule is referred to as the rule length. Candidate rule extraction is performed in a greedy manner for each class using a heuristic rule evaluation measure. We use the following rule evaluation measure:

$$f_{\text{SLAVE}}(R_q) = s(\mathbf{A}_q \Rightarrow \text{Class } C_q) \\ - \sum_{\substack{h=1 \\ h \neq C_q}}^m s(\mathbf{A}_q \Rightarrow \text{Class } h), \quad (11)$$

where  $s(\cdot)$  is the support measure of fuzzy rules, which is defined as follows:

$$s(\mathbf{A}_q \Rightarrow \text{Class } h) = \frac{1}{m} \sum_{\mathbf{x}_p \in \text{Class } h} \mu_{\mathbf{A}_q}(\mathbf{x}_p). \quad (12)$$

The heuristic rule evaluation measure in (11) is a modified version of a rule evaluation criterion used in an iterative fuzzy GBML (genetics-based machine learning) algorithm called SLAVE [10].

### 2.3 Genetic rule selection

Let us assume that  $N$  fuzzy rules have been extracted as candidate rules using the SLAVE criterion in (11). A subset  $S$  of the  $N$  candidate rules is handled as an individual in genetic rule selection, which is represented by a binary string of the length  $N$  as

$$S = s_1 s_2 \cdots s_N, \quad (13)$$

where  $s_j = 1$  and  $s_j = 0$  mean that the  $j$ -th candidate rule is included in  $S$  and excluded from  $S$ , respectively.

We use a well-known high-performance evolutionary multiobjective optimization algorithm called NSGA-II [7], [8] to find a large number of non-dominated rule sets from the candidate rules with respect to the three objectives in (1). The NSGA-II algorithm is also employed in the two-objective case in (2). On the other hand, we use a standard single-objective genetic algorithm with a single elite solution in the case of the single-objective formulations in (3) and (4).

We use two problem-specific heuristic tricks in the NSGA-II algorithm and the standard single-objective genetic algorithm. One is biased mutation where a larger probability is assigned to the mutation from 1 to 0 than that from 0 to 1. This heuristic is used to efficiently decrease the number of fuzzy rules in each rule set. The other is the removal of unnecessary fuzzy rules. Since we use the single winner-based method for classifying each pattern, some fuzzy rules in  $S$  may be chosen as winner rules for no patterns. We can remove those fuzzy rules without degrading the number of correctly classified training patterns (i.e.,  $f_1(S)$ ). At the same time, the removal of such an unnecessary fuzzy rule improves the number of fuzzy rules (i.e.,  $f_2(S)$ ) and the total rule length (i.e.,  $f_3(S)$ ). Thus we remove all fuzzy rules that are not selected as winner rules for any training patterns from the rule set  $S$ . The removal of unnecessary fuzzy rules is performed for each rule set after  $f_1(S)$  is calculated and before  $f_2(S)$  and  $f_3(S)$  are calculated.

## 3 Computational experiments

We used seven data sets in Table 1 available from the UC Irvine Machine Learning repository (<http://www.ics.uci.edu/~mllearn/>). Data sets with missing values are marked by "\*" in the third column of Table 1. Since we did not use incomplete patterns with missing values, the number of patterns in the third column does not include those patterns. For comparison, experimental results by Elomaa and Rousu [9] are cited in Table 1. They applied six variants of the C4.5 algorithm [21] to 30 data sets. The performance of each variant was examined by ten iterations of the whole ten-fold cross-validation (10CV) procedure. We show in the last two columns of Table 1 the best and worst error rates on test patterns among the six variants reported in [9] for each data set.

Table 1. Data sets used in our computational experiments.

Data set	Attributes	Patterns	Classes	C4.5 in [9]	
				Best	Worst
Breast W	9	683*	2	5.1	6.0
Diabetes	8	768	2	25.0	27.2
Glass	9	214	6	27.3	32.2
Heart C	13	297*	5	46.3	47.9
Iris	4	150	3	5.7	7.5
Sonar	60	208	2	24.6	35.8
Wine	13	178	3	5.6	8.8

\* Incomplete patterns with missing values are not included.

We applied our two-stage rule selection method to the seven data sets in Table 1. All attribute values were normalized into real numbers in the unit interval [0, 1]. We generated 300 fuzzy rules for each class as candidate rules in a greedy manner using the SLAVE criterion in (11). That is, the best 300 candidate rules with the largest values of the SLAVE criterion were found for each class. Thus the total number of candidate rules was  $300M$  where  $M$  is the number of classes. In this heuristic procedure, we examined candidate rules of length 3 or less for the six data sets except for the sonar data with 60 attributes. Since the sonar data have a large number of attributes, we only examined short candidate rules of length 2 or less.

The NSGA-II algorithm was executed using the following parameter values:

Population size: 200 strings,

Crossover probability: 0.8 (uniform crossover),

Biased mutation probabilities:

$$p_m(0 \rightarrow 1) = 1/300M \text{ and } p_m(1 \rightarrow 0) = 0.1,$$

Stopping condition: 5000 generations.

The standard single-objective genetic algorithm was also executed using the same parameter values and the weight values  $w_1 = 10$ ,  $w_2 = 1$ ,  $w_3 = 1$ .

For evaluating the generalization ability of obtained rule sets, we used the 10CV technique as in [9]. First each data set was randomly divided into ten subsets of the same size. One subset was used as test patterns while the other nine subsets were used as training patterns. Our two-stage rule selection method was applied to training patterns to find non-dominated rule sets. The generalization ability of obtained rule sets was evaluated by classifying test patterns. This train-and-test procedure was iterated ten times so that all the ten subsets were used as test patterns. As in [9], we iterated the whole 10CV procedure ten times using different data partitions. Thus our method was executed 100 times in total for each data set.

In Table 2, we show experimental results by the two-objective formulation in (2) on the Wisconsin breast cancer data set (Breast W in Table 1). A number of non-dominated rule sets were obtained by a single run of the NSGA-II algorithm, which was executed 100 times during the 10 iterations of the whole 10CV procedures. The obtained rule sets were categorized into several groups

according to the number of fuzzy rules. Then the average error rates on training patterns and test patterns were calculated for each group of rule sets with the same number of fuzzy rules. The average rule length was also calculated for each group. These average results are summarized in Table 2. In the last column, the number of rule sets in each group is shown. That is, the last column shows the number of runs of the NSGA-II algorithm from which the corresponding rule set was obtained. We can see that from Table 2 that a number of rule sets were obtained by each run of the NSGA-II algorithm. The experimental results in Table 2 are depicted in Figure 2. We also show in Figure 2 the experimental results by the standard single-objective genetic algorithm based on the weighted scalar objective function of the two objectives.

Table 2. Experimental results on the Wisconsin breast cancer data set by the two-objective formulation.

Number of rules	Average length	Average error rates		Number of runs
		Training	Test	
1	2.16	35.41	35.63	100
2	2.00	3.19	4.27	100
3	2.06	2.59	4.34	100
4	2.02	2.25	5.35	92
5	2.00	2.06	4.99	93
6	1.93	1.84	4.28	84
7	1.92	1.73	3.29	54

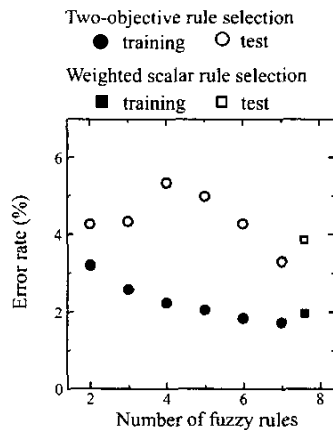


Figure 2. Experimental results in Table 2 on the Wisconsin breast cancer data set.

In Figure 2, we can observe a clear tradeoff between the number of fuzzy rules and the error rates on training patterns (see the closed circles). On the other hand, the relation between the number of fuzzy rules and the error rate on test patterns is unclear in Figure 2 (see the open circles). In Figure 3, we show experimental results by the two-objective formulation on the other data sets. The

tradeoff between the number of fuzzy rules and the error rates on training patterns is clearly observed for all data sets. On the other hand, the relation between the number of fuzzy rules and the error rates on test patterns is not clear. In Figure 3 (C), we observe the overfitting of obtained rule sets to training data (i.e., deterioration in the error rates on test patterns due to the increase in the number of fuzzy rules). We can also see from Figure 2 and Figure 3 that good results were not always obtained by the weighted scalar objective function.

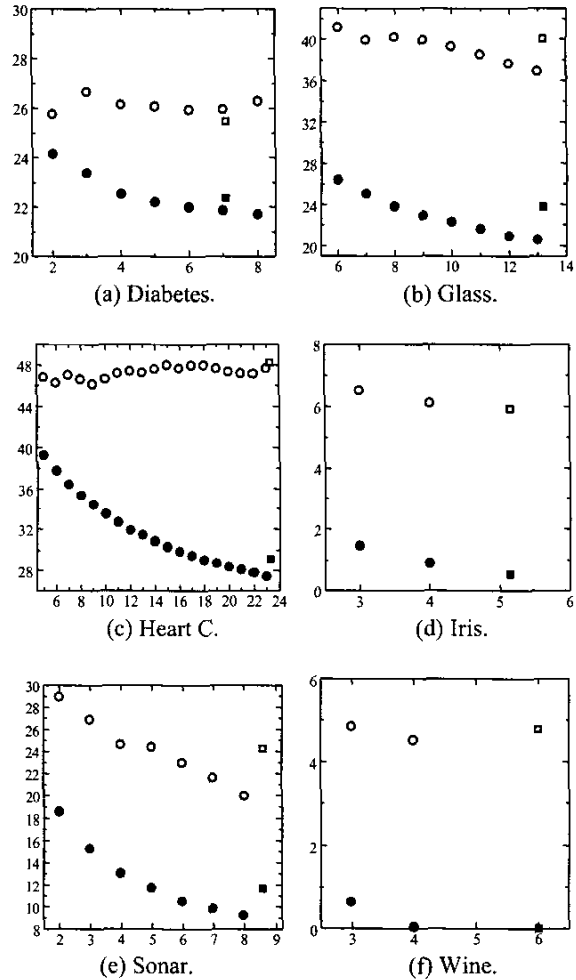


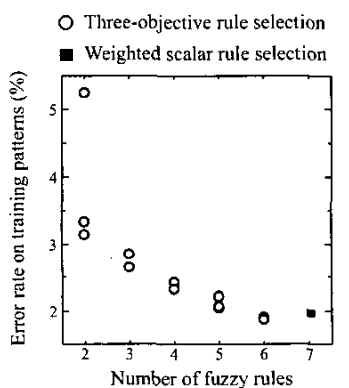
Figure 3. Experimental results by the two-objective approach on the other data sets. Each figure is depicted in the same manner as Figure 2.

In Table 3, we show experimental results by the three-objective formulation in (1) on the Wisconsin breast cancer data set. Since the total rule length was taken into account in addition to the classification accuracy and the number of fuzzy rules, rule sets with the same number of fuzzy rules but the different total rule length were obtained

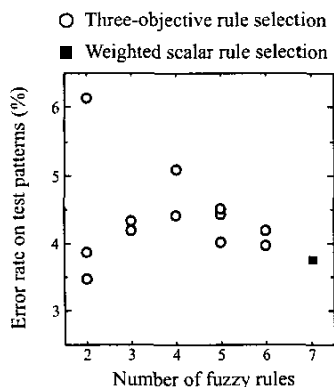
as non-dominated rule sets. As a result, more non-dominated rule sets were obtained by the three-objective formulation in Table 3 than the two-objective formulation in Table 2. The experimental results in Table 3 are depicted in Figure 4 where the experimental results by the weighted scalar objective function of the three objectives are also shown. As in Figure 2, the tradeoff between the number of fuzzy rules and the error rates on training patterns is clear in Figure 4 (a). On the other hand, the relation between the number of fuzzy rules and the error rate on test patterns is unclear in Figure 4 (b). A part of the experimental results in Table 3 are visualized in Figure 5 from a different viewpoint where the relation between the average rule length and the error rates is examined for some rule sets with the same number of rules (i.e., two rules in (a) and four rules in (b)). In Figure 5, we can observe a clear tradeoff between the average rule length and the error rates on training patterns. We can also observe the overfitting of rule sets to training patterns (i.e., the increase in the error rate on test patterns) due to the increase in the average rule length in Figure 5.

Table 3. Experimental results on the Wisconsin breast cancer data set by the three-objective formulation.

Number of rules	Average length	Average error rate		Number of runs
		Training	Test	
1	1.00	35.43	35.43	100
2	1.00	5.25	6.13	100
2	1.50	3.34	3.47	100
2	2.00	3.15	3.87	92
3	1.33	2.85	4.19	79
3	1.67	2.64	4.33	92
4	1.50	2.42	4.41	72
4	1.75	2.32	5.09	36
5	1.40	2.21	4.43	35
5	1.60	2.05	4.51	61
5	1.80	2.07	4.02	35
6	1.50	1.91	4.19	35
6	1.67	1.87	3.97	45

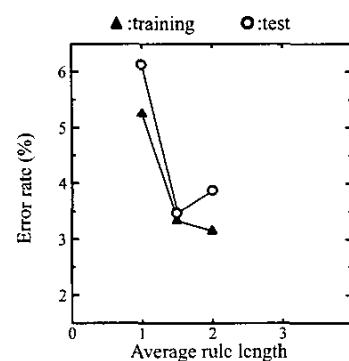


(a) Error rates on training patterns.

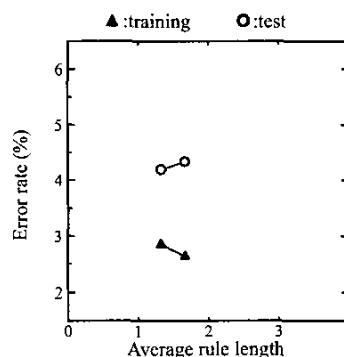


(b) Error rates on test patterns.

Figure 4. Experimental results by the three-objective formulation on the Wisconsin breast cancer data sets.



(a) Rule sets with two fuzzy rules.



(b) Rule sets with four fuzzy rules.

Figure 5. Relation between the average rule length and the average error rates on training and test patterns.

## 4 Conclusions

In this paper, we compared single-, two-, and three-objective approaches to the design of fuzzy rule-based systems for classification problems. The main advantage

of the three-objective approach is that a large number of non-dominated rule sets are obtained by its single run. The decision maker can understand the tradeoff between the accuracy and the complexity of fuzzy rule-based systems by the obtained rule sets. Another advantage is that the second and third objectives (i.e., the number of fuzzy rules and the total rule length) can work as a safeguard against the overfitting of fuzzy rule-based classification systems to training patterns. This advantage was also examined in our former studies [14], [15].

## References

- [1] S. Abe, M.-S. Lan, and R. Thawonmas, "Tuning of a fuzzy classifier derived from data," *International Journal of Approximate Reasoning*, vol. 14, no. 1, pp. 1-24, January 1996.
- [2] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. I. Verkamo, "Fast discovery of association rules," in U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy (eds.), *Advances in Knowledge Discovery and Data Mining*, pp. 307-328, AAAI Press, Menlo Park, 1996.
- [3] J. Casillas, O. Cordon, F. Herrera, and L. Magdalena (eds.), *Accuracy Improvements in Linguistic Fuzzy Modeling*, Springer-Verlag, 2003.
- [4] J. Casillas, O. Cordon, F. Herrera, and L. Magdalena (eds.), *Interpretability Issues in Fuzzy Modeling*, Springer-Verlag, 2003.
- [5] L. Castillo, A. Gonzalez, and R. Perez, "Including a simplicity criterion in the selection of the best rule in a genetic fuzzy learning algorithm," *Fuzzy Sets and Systems*, vol. 120, no. 2, pp. 309-321, June 2001.
- [6] O. Cordon, F. Herrera, F. Hoffman, and L. Magdalena, *Genetic Fuzzy Systems*, World Scientific, Singapore, 2001.
- [7] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*, John Wiley & Sons, Chichester, 2001.
- [8] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, April 2002.
- [9] T. Elomaa and J. Rousu, "General and efficient multisplitting of numerical attributes," *Machine Learning*, vol. 36, no. 3, pp. 201-244, September 1999.
- [10] A. Gonzalez and R. Perez, "SLAVE: A genetic learning system based on an iterative approach," *IEEE Trans. on Fuzzy Systems*, vol. 7, no. 2, pp. 176-191, April 1999.
- [11] H. Ishibuchi, T. Murata, and I. B. Turksen, "Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems," *Fuzzy Sets and Systems*, vol. 89, no. 2, pp. 135-149, July 1997.
- [12] H. Ishibuchi, T. Nakashima, and T. Murata, "Three-objective genetics-based machine learning for linguistic rule extraction," *Information Sciences*, vol. 136, no. 1-4, pp. 109-133, August 2001.
- [13] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, "Selecting fuzzy if-then rules for classification problems using genetic algorithms," *IEEE Trans. on Fuzzy Systems*, vol. 3, no. 3, pp. 260-270, August 1995.
- [14] H. Ishibuchi and T. Yamamoto, "Effects of three-objective genetic rule selection on the generalization ability of fuzzy rule-based systems," *Lecture Notes in Computer Science*, vol. 2632, pp. 608-622, April 2003.
- [15] H. Ishibuchi and T. Yamamoto, "Evolutionary multiobjective optimization for generating an ensemble of fuzzy rule-based classifiers," *Lecture Notes in Computer Science*, vol. 2723, pp. 1077-1088, July 2003.
- [16] H. Ishibuchi and T. Yamamoto, "Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining," *Fuzzy Sets and Systems*, vol. 141, pp. 59-88, January 2004.
- [17] Y. Jin, "Fuzzy modeling of high-dimensional systems: Complexity reduction and interpretability improvement," *IEEE Trans. on Fuzzy Systems*, vol. 8, no. 2, pp. 212-221, April 2000.
- [18] Y. Jin, W. von Seelen, and B. Sendhoff, "On generating FC<sup>3</sup> fuzzy rule systems from data using evolution strategies," *IEEE Trans. on Systems, Man and Cybernetics - Part B: Cybernetics*, vol. 29, no. 4, pp. 829-845, December 1999.
- [19] C. L. Karr and E. J. Gentry, "Fuzzy control of pH using genetic algorithms," *IEEE Trans. on Fuzzy Systems*, vol. 1, no. 1, pp. 46-53, February 1993.
- [20] D. Nauck and R. Kruse, "A neuro-fuzzy method to learn fuzzy classification rules from data," *Fuzzy Sets and Systems*, vol. 89, no. 3, pp. 277-288, August 1997.
- [21] J. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, CA, 1993.
- [22] H. Roubos and M. Setnes, "Compact and transparent fuzzy models and classifiers through iterative complexity reduction," *IEEE Trans. on Fuzzy Systems*, vol. 9, no. 4, pp. 516-524, August 2001.
- [23] M. Setnes, R. Babuska, and B. Verbruggen, "Rule-based modeling: Precision and transparency," *IEEE Trans. on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, vol. 28, no. 1, pp. 165-169, February 1998.
- [24] M. Setnes and H. Roubos, "GA-based modeling and classification: Complexity and performance," *IEEE Trans. on Fuzzy Systems*, vol. 8, no. 5, pp. 509-522, October 2000.
- [25] V. Uebele, S. Abe, and M. -S. Lan, "A neural-network-based fuzzy classifier," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 25, no. 2, pp. 353-361, February 1995.