A Multi-Objective Evolutionary Algorithm for Rule Selection and Tuning on Fuzzy Rule-Based Systems

Rafael Alcalá Jesús Alcalá-Fdez María José Gacto Francisco Herrera

Abstract— **Recently, Multi-Objective Evolutionary Algorithms have been also applied to improve the difficult trade-off between interpretability and accuracy of Fuzzy Rule-Based Systems. It is know that both requirements are usually contradictory, however, a Multi-Objective Genetic Algorithm can obtain a set of solutions with different degrees of trade-off.**

This contribution presents a Multi-Objective Evolutionary Algorithm to obtain linguistic models with improved accuracy and the least number of possible rules. In order to minimize the number of rules and the system error, this model performs a rule selection and a tuning of the membership functions of an initial set of candidate linguistic fuzzy rules.

I. Introduction

Many automatic techniques have been proposed in the literature to extract a proper set of fuzzy rules from numerical data. However, most of these techniques usually try to improve the performance associated to the prediction error without inclusion of any interpretability measure, an essential aspect of Fuzzy Rule-Based Systems (FRBSs). In the last years, the problem of finding the right trade-off between interpretability and accuracy, in spite of the original nature of fuzzy logic, has arisen a growing interest in methods which take both aspects into account [1]. Of course, the ideal thing would be to satisfy both criteria to a high degree, but since they are contradictory issues generally it is not possible.

Recently, Multi-Objective Evolutionary Algorithms (MOEAs) [4], [9] have been also applied to improve the difficult trade-off between interpretability and accuracy of FRBSs, obtaining linguistic models not only accurate but also interpretable. Since this problem presents a multi-objective nature the use of these kinds of algorithms to obtain a set of solutions with different degrees of accuracy and interpretability is an interesting way to work. Most of these works apply MOEAs to obtain Mamdani FRBSs [6], [17], [18], [19], [20], [24] since they are much more interpretable than Takagi-Sugeno ones [22], [31], [32].

This contribution briefly reviews the state of the art in this recent topic, analyzing the most representative works of the specialized literature in order to point out the most important aspects that should been taken into account to deal with these kinds of problems. All these works try to obtain the complete Pareto (set of non-dominated solutions with different trade-off) by selecting or learning the set of rules better representing the example data, i.e., improving the system accuracy and decreasing the FRBS complexity but not considering learning or tuning of the Membership Function (MFs) parameters. In this way, this work presents a specific MOEA to obtain simpler and still accurate linguistic fuzzy models by applying rule selection and a tuning of the system parameters, which represents a more complex search space and therefore needs of different considerations respect to the works in the existing literature.

In order to do that, this contribution is arranged as follows. Next section presents a brief study of the existing MOEAs for general purpose which usually are modified or directly applied to obtain FRBSs with good interpretability-accuracy trade-off. Section III briefly analyzes the state of the art on the use of MOEAs to get the desired trade-off in different application areas of FRBSs. In Section IV, we present an algorithm to perform linguistic rule selection together with a tuning of MFs by using one of the most known MOEAs. Section V shows an experimental study of this method in a complex but interesting problem. Finally, Section VI points out some conclusions and further research lines.

II. Multi-Objective Evolutionary Algorithms

Evolutionary algorithms simultaneously deal with a set of possible solutions (the so-called population) which allows to find several members of the Pareto optimal set in a single run of the algorithm. Additionally, they are not too susceptible to the shape or continuity of the Pareto front (e.g., they can easily deal with discontinuous and concave Pareto fronts).

The first hint regarding the possibility of using evolutionary algorithms to solve a multi-objective problem appears in a Ph.D. thesis from 1967 [26] in which, however, no actual MOEA was developed (the multi-objective problem was restated as a single-objective problem and solved with a genetic algorithm). David Schaffer is normally considered to be the first to have designed a MOEA during the mid-1980s [27]. Schaffer's approach, called Vector Evaluated Genetic Algorithm (VEGA) consists of a simple genetic algorithm with a modified selection mechanism. However, VEGA had a number of problems from which the main one had to do with its inability to retain solutions with acceptable performance, perhaps above average, but not outstanding for any of the objective functions.

After VEGA, the researchers designed a first generation of MOEAs characterized by its simplicity where the main lesson learned was that successful MOEAs had to combine a good mechanism to select non-dominated individuals (perhaps, but not necessarily, based on the concept of Pareto optimality) combined with a good mechanism to maintain diversity (fitness sharing was a choice, but not the only one). The

R. Alcalá, J. Alcalá-Fdez, M.J. Gacto and F. Herrera are with the Department of Computer Science and A.I., University of Granada, 18071 Granada, Spain (email: {alcala, jalcala, herrera}@decsai.ugr.es and mjgacto@ugr.es).
Supported by the Spanish CICYT Project TIN2005-08386-C05-01.

1-4244-1210-2/07/$25.00 ©2007 IEEE.

TABLE I

CLASSIFICATION OF MOEAS

Reference	MOEA	1^{st} Gen.	2^{nd} Gen.
[14]	MOGA	√	
[15]	NPGA	√	
[28]	NSGA	√	
[3]	micro-GA		√
[11]	NPGA 2		√
[10]	NSGA-II		√
[23]	PAES		√
[7], [8]	PESA & PESA-II		√
[33], [34]	SPEA & SPEA2		√

most representative MOEAs of this generation are the following: Nondominated Sorting Genetic Algorithm (NSGA) [28], Niched-Pareto Genetic Algorithm (NPGA) [15] and Multi-Objective Genetic Algorithm (MOGA) [14].

A second generation of MOEAs started when elitism became a standard mechanism. In fact, the use of elitism is a theoretical requirement in order to guarantee convergence of a MOEA. Many MOEAs have been proposed during the second generation (which we are still living today). However, most researchers will agree that few of these approaches have been adopted as a reference or have been used by others. In this way, the Strength Pareto Evolutionary Algorithm 2 (SPEA2) [34] and the Nondominated Sorting Genetic Algorithm II (NSGA-II) [10] can be considered as the most representative MOEAs of the second generation, also being of interest some others as the Pareto Archived Evolution Strategy (PAES) [23]. Table I shows a resume of the most representative MOEAs of both generations.

Finally, we have to point out that nowadays NSGA-II is the paradigm within the MOEA research community since the powerful crowding operator that this algorithm uses usually allows to obtain the widest Pareto sets in a great variety of problems, which is a very appreciated property in this framework. In this way, the question is: “Is NSGA-II the best MOEA to get the desired interpretability-accuracy trade-off of FRBSs?”. Next section presents the state-of-the-art on the use of the MOEAs to get this difficult trade-off in order to see how different researchers have affronted this problem.

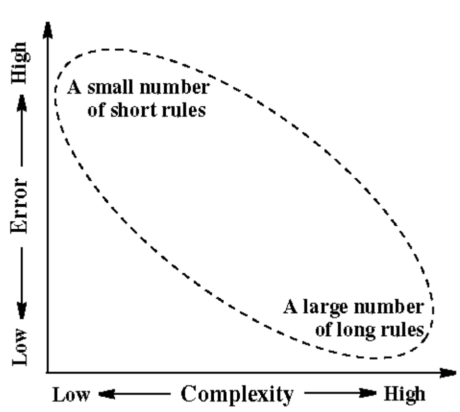

Fig. 1. Trade-off between the error and the interpretability of rule sets

III. USE OF MOEAS TO GET THE INTERPRETABILITY-ACCURACY TRADE-OFF OF FRBSS

As mentioned, MOEAs generate a family of equally valid solutions, where each solution tends to satisfy a criterion to a higher extent than another. For this reason, MOEAs have been also applied to improve the difficult trade-off between interpretability and accuracy of FRBSs, where each solution in the pareto front represents a different trade-off between interpretability and accuracy (see Figure 1).

The most continuous and prolific research activity in the application of MOEAs to Mamdani FRBS generation for finding the accuracy-interpretability trade off has been certainly performed by Ishibuchi’s group. Earlier works [17] were devoted to the application of simple MOEAs of the first generation to perform a rule selection on an initial set of classification rules involving *“don’t care”* conditions and considering two different objectives (classification accuracy and number of rules). Then, a third objective was also included in order to minimize the length of the rules by rule selection [18] or rule learning [18]. In [20], they apply a better MOEA, the Multi-Objective Genetic Local Search [16] (MOGLS), following the same approach for rule selection with three objectives. And finally, two algorithms based on a MOEA of the second generation (NSGA-II) have been proposed respectively for rule selection [24] and rule learning [21] considering the same concepts. In the literature, we can also find some papers of other researchers in this topic. For instance in [6], Cordon et al. use MOGA for jointly performing feature selection and fuzzy set granularity learning with only two objectives.

At this point, we can see that all the methods mentioned were applied to classification problems for rule selection or rule learning, without learning or tuning the MFs that were initially fixed. Most of the works in this topic only consider quantitative measures of the system complexity in order to improve the interpretability of such systems, rarely considering qualitative measures. Moreover, MOEAs considered were slight modifications of MOEAs proposed for general use (MOGA, NSGA-II, etc.) or based on them. Notice that, although NSGA-II improves the results respect to other MOEAs, since to cross non-dominated rule sets with very different numbers of rules and different rule structures (forced by the NSGA-II crowding operator) usually gives a bad accuracy, this MOGA needed of an adaptation to favor the cross of similar solutions in order to also get good results for the accuracy objective [24]. The problem is that, although to directly apply this powerful algorithm to this problem gets a wider Pareto front with several good solutions, to improve the accuracy objective is more difficult than simplifying the fuzzy models, by which the Pareto front finally obtained still becomes sub-optimal respect to the accuracy objective.

On the other hand, there are a few works in the framework of fuzzy modeling for regression problems. In [19], authors show how a simple MOGA can be applied to a three-objective optimization problem (again not considering learning or tuning of parameters). Some applications of MOEAs have been also discussed in the literature to improve the difficult trade-off between accuracy and interpretability of Takagi-Sugeno models [29]. In [22], [31], [32], accuracy, interpretability and compactness have been considered as objectives to obtain interpretable and very accurate Takagi-Sugeno models. However, since Takagi-Sugeno models have a linear function in the consequent part of each fuzzy rule, they are close to accuracy representing another type of trade-off with less interpretable models [19]. For this reason, the type of rule most used to achieve the trade-off between accuracy and complexity are the fuzzy rules with linguistic terms in both the antecedent and consequent parts, i.e., Mamdani rules [25].

IV. A MOEA for Rule Selection and Tuning of Membership Functions

As we explain in the previous section most works in the field of fuzzy systems are applied to classification problems by learning or selecting rules, not considering tuning of the MF parameters. The main reason of this fact is that a tuning of parameters implies a lost of the interpretability to some degree. However, it is known that this way to work greatly improves the performance of the linguistic models so obtained, being another alternative to improve the interpretability-accuracy trade-off. For this reason, we would like to show an example of application that focus the research in the linguistic fuzzy modeling area, in order to evaluate the performance of MOEAs in a field which is still less explored, and with the objective of inject some ideas or recommendations in this open topic (improvement of the interpretability of very accurate models).

The proposed algorithm will perform rule selection from a given fuzzy rule set together with a parametric tuning of the MFs. To do that, we apply one of the most used MOEAs of the second generation, SPEA2 [34], considering two different objectives, system error and number of rules.

In the next subsections, we present the SPEA2 algorithm applied for linguistic fuzzy modeling. At first, the main components of the algorithm are proposed and then the main steps and characteristic are described.

A. Main Components of the Algorithm

As mentioned, we use the well-known SPEA2 to perform rule selection and tuning of MFs with the aim of improving the desired trade-off between interpretability and accuracy. In the following, the components needed to apply this algorithm in this concrete problem are explained. They are coding scheme, initial gene pool, objectives and genetic operators:

- **Coding scheme and initial gene pool**
  A double coding scheme for both *rule selection* (C_S) and *tuning* (C_T) is used:
  $$C^p = C_S^p C_T^p$$
  In the $C_S^p = (c_{S1}, \ldots, c_{Sm})$ part, the coding scheme consists of binary-coded strings with size m (with m being the number of initial rules). Depending on whether a rule is selected or not, values '1' or '0' are respectively assigned to the corresponding gene. In the C_T part, a real coding is considered, being m^i the number of labels of each of the n variables comprising the data base,
  $$C_i = (a_1^i, b_1^i, c_1^i, \ldots, a_{m^i}^i, b_{m^i}^i, c_{m^i}^i), \;\; i = 1, \ldots, n \;\; ,$$
  $$C_T^p = C_1 C_2 \ldots C_n \;\; .$$
  The initial population is obtained with all individuals having all genes with value '1' in the C_S part. And in the C_T part the initial data base is included as first individual. The remaining individuals are generated at random within the corresponding variation intervals. Such intervals are calculated from the initial data base. For each MF, $C_i^j = (a^j, b^j, c^j)$, the variation intervals are calculated in the following way:
  $$[I_{a^j}^l, I_{a^j}^r] = [a^j - (b^j - a^j)/2, a^j + (b^j - a^j)/2]$$
  $$[I_{b^j}^l, I_{b^j}^r] = [b^j - (b^j - a^j)/2, b^j + (c^j - b^j)/2]$$
  $$[I_{c^j}^l, I_{c^j}^r] = [c^j - (c^j - b^j)/2, c^j + (c^j - b^j)/2]$$
- **Objectives**
  Two objectives are minimized for this problem: the number of rules (interpretability) and the Mean Squared Error (accuracy),
  $$\text{MSE} = \frac{1}{2 \cdot |E|} \sum_{l=1}^{|E|} (F(x^l) - y^l)^2,$$
  with $|E|$ being the size of a data set E, $F(x^l)$ being the output obtained from the FRBS decoded from the said chromosome when the l-th example is considered and y^l being the known desired output. The fuzzy inference system considered to obtain $F(x^l)$ is the *center of gravity weighted by the matching* strategy as defuzzification operator and the *minimum t-norm* as implication and conjunctive operators.
- **Genetic Operators**
  The crossover operator depends on the chromosome part where it is applied: the BLX-0.5 [13] in the C_T part and the HUX [12] in the C_S part.
  Finally, four offspring are generated by combining the two from the C_S part with the two from the C_T part (the two best replace to their parent). The mutation operator changes a gene value at random in the C_S and C_T parts (one in each part) with probability P_m.

B. SPEA2 Based Approach

Considering the components defined and the descriptions of the authors in [34], the SPEA2 algorithm consists of the next steps:

Input: N (population size),
$\overline{N}$ (external population size),
T (maximum number of generations).
Output: A (non-dominated set).

1) Generate an initial population P_0 and create the empty external population $\overline{P}_0 = \emptyset$.
2) Calculate fitness values of individuals in P_t and $\overline{P}_t$.

3) Copy all non-dominated individuals in $P_t \cup \overline{P}_t$ to $\overline{P}_{t+1}$. If $|\overline{P}_{t+1}| > \overline{N}$ apply truncation operator. If $|\overline{P}_{t+1}| < \overline{N}$ fill with dominated in $P_t \cup \overline{P}_t$.
4) If $t \geq T$, return A and stop.
5) Perform binary tournament selection with replacement on $\overline{P}_{t+1}$ in order to fill the mating pool.
6) Apply recombination (BLX-HUX) and mutation operators to the mating pool and set P_{t+1} to the resulting population. Go to step 2 with $t = t + 1$.

V. Experiments

In this section, we present an example on the use of MOEAs to obtain linguistic models with a good trade-off between interpretability and accuracy in a real-world problem [5] with 4 input variables that consists of estimating the maintenance costs of medium voltage lines in a town. To do that, we also compare the proposed algorithm with the paradigm of MOEAs, NSGA-II [10], by also considering the same components described in section IV-A in order to show the good behavior of SPEA2 in this specific framework. Methods considered for the experiments are briefly described in Table II. WM method is considered to obtain the initial rule base to be tuned. T and S methods perform the tuning of parameters and rule selection respectively. TS indicates tuning together with rule selection in the same algorithm. All of them consider the accuracy of the model as the sole objective. MOEAs considered (SPEA2 and NSGA-II) perform rule selection from a given fuzzy rule set together with the parametric tuning of the MFs considering two objectives, system error and number of rules.

TABLE II

Methods Considered for Comparison

Méthod	Ref.	Description
WM	[30]	Wang & Mendel algorithm
WM+T	[2]	Tuning of Parameters
WM+S	[2]	Rule Selection
WM+TS	[2]	Tuning and Rule Selection
SPEA2	[34]*	Tuning and Rule Selection with SPEA2
NSGA-II	[10]*	Tuning and Rule Selection with NSGA-II

* based on that algorithm

In the next subsections, we describe this real-world problem and finally we show the results obtained.

A. Problem Description and Experiments

Estimating the maintenance costs of the medium voltage electrical network in a town [5] is a complex but interesting problem. Since a direct measure is very difficult to obtain, it is useful to consider models. These estimations allow electrical companies to justify their expenses. Moreover, the model must be able to explain how a specific value is computed for a certain town. Our objective will be to relate the *maintenance costs of the medium voltage lines* with the following four variables: *sum of the lengths of all streets in the town*, *total area of the town*, *area that is occupied by buildings*, and *energy supply to the town*. We will deal with estimations of minimum maintenance costs based on a model of the optimal electrical network for a town in a sample of 1,059 towns.

To develop the different experiments, we consider a *5-folder cross-validation model*, i.e., 5 random partitions of data each with 20%, and the combination of 4 of them (80%) as training and the remaining one as test. For each one of the 5 data partitions, the tuning methods have been run 6 times, showing for each problem the averaged results of a total of 30 runs. In the case of methods with multi-objective approach (SPEA2 and NSGA-II), the averaged values are calculated considering the most accurate solution from each Pareto obtained. In this way, the multi-objective algorithms are compared with several single objective based methods. This way to work differs with the previous works in the specialized literature (see section III) in which one or several Pareto fronts are presented and an expert should after select one solution. Our main aim following this approach is to compare the same algorithm by only considering an accuracy objective (WM+TS) with the most accurate solution found by the multi-objective ones in order to see if the Pareto fronts obtained are not only wide but also optimal (similar solutions to that obtained by WM+TS should be included in the final Pareto).

The initial linguistic partitions are comprised by *five linguistic terms* with equally distributed triangular shape MFs. The values of the input parameters considered by the MOGAs are shown in the next: population size of 200, external population size of 61 (in the case of SPEA2), 50000 evaluations and 0.2 as mutation probability per chromosome. Different sizes of population were probed showing not very different results but presenting the best performance around 200 individuals.

B. Results and Analysis

The results obtained by the analyzed methods are shown in table III, where $\#R$ stands for the number of rules, MSE$_{tra}$ and MSE$_{tst}$ respectively for the averaged error obtained over the training and test data, σ for the standard deviation and t for the results of applying a *test t-student* (with 95 percent confidence) in order to ascertain whether differences in the performance of the multi-objective approach are significant when compared with that of the other algorithms in the table. The interpretation of this column is:

- $\star$ represents the best averaged result.
- $+$ means that the best result has better performance than that of the corresponding row.

TABLE III

Results obtained by the studied methods

Method	#R	MSE$_{tra}$	σ_{tra}	t	MSE$_{tst}$	σ_{tst}	t
WM	65	57605	2841	+	57934	4733	+
WM+T	65	18602	1211	+	22666	3386	+
WM+S	40.8	41086	1322	+	59942	4931	+
WM+TS	41.9	14987	391	+	18973	3772	=
SPEA2	**33**	**13272**	1265	*	**17533**	3226	*
NSGA-II	41.0	14488	965	=	18419	3054	=

Analysing the results showed in table III we can highlight the two following facts:

- NSGA-II obtains the same accuracy and the same number of rules than the models obtained with WM+TS (single objective-based approach) considering the most accurate result of each obtained Pareto. Therefore, we could consider that this algorithm gets good solutions, from the most accurate ones (with more complexity) to the most simple ones (with the worst accuracy).
- The SPEA2 method shows a reduction of the MSE_{tra} and produces more or less the same MSE_{tst} respect to the models obtained by only considering the accuracy objective (WM+TS). Moreover, a considerable number of rules have been removed from the initial FRBS, obtaining simpler models with a similar performance. In this way, the most accurate models obtained by SPEA2 considering a multi-objective approach get a better trade-off between interpretability and accuracy than those obtained by a single objective based algorithm (which theoretically should obtain the most accurate results).

These results are due to the large search space that involves this problem. There are some initial rules that should be removed since they do not cooperate in a good way with the remaining ones. Even in the case of only considering an accuracy-based objective, the large search space that supposes the tuning of parameters makes very difficult to remove these kinds of rules since bad rules are tuned together with the remaining ones searching for their best cooperation. The use of a multi-objective approach favors a better selection of the ideal number of rules, preserving some rule configurations until the rule parameters are evolved to dominate solutions including bad rules.

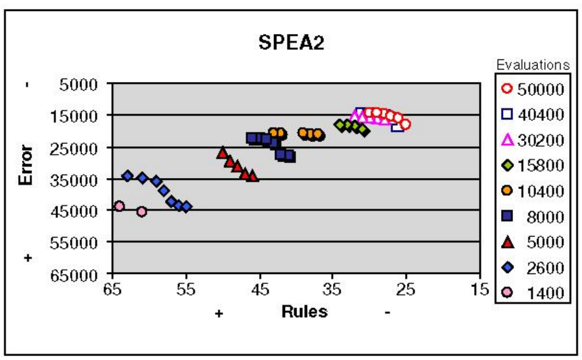

Fig. 2. Evolution of the Pareto fronts of SPEA2

On the other hand, NSGA-II tries to obtain a wider Pareto front by crossing very different solutions based on its crowding operator. However, in this problem, it is difficult to obtain accurate solutions by favoring the crossing of solutions with very different rule configurations (those in the Pareto), which try to obtain the best accuracy by learning very different parameters for the MFs. In our opinion, this is the main reason by which this algorithm does not work as well as SPEA2 in this particular problem.

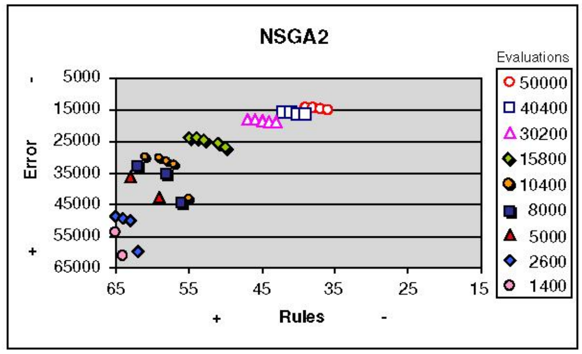

Fig. 3. Evolution of the Pareto fronts of NSGA2

In Figures 2 and 3, we can see the Pareto evolution for each multi-objective algorithm. We can observe as the Pareto moves along without having a wide extension, even in the case of NSGA-II. In this way, although SPEA2 implements a truncation operator that is similar to the crowding operator in NSGA-II, this operator is never used since the number of non-dominated solutions is always very lower than the size of the external population (which is completed with the best solutions that do not belong to the Pareto and that can be close or similar to those in the Pareto). This favors the evolution of the parameters of the MFs by allowing the crossing of solutions with more or less the same subset of rules and different parameters. All these facts suggest the design of more specific algorithms in order to get even better solutions for these kinds of problems, probably, solutions with a better performance that considering a single objective and with a minor number of rules.

VI. Concluding Remarks

In this work we have analyzed the use of MOEAs to improve the trade-off between interpretability and accuracy of FRBSs. A brief revision of the state of the art in this topic has been performed. From this study we can point out the following facts:

- Most of the contributions in this topic were made in the framework of fuzzy classification, considering Mamdani FRBSs.
- Most of the works only consider quantitative measures of the system complexity to determine the FRBS interpretability.
- None of the works considered a learning or tuning of the MFs, only performing a rule learning or selection.
- The MOEAs considered were slight modifications of MOEAs proposed for general use (MOGA, NSGA-II, etc.) or specifically developed for this concrete and difficult problem. It is due to the special nature of this problem, in which to improve the accuracy objective is more difficult than simplifying the fuzzy models, by which the Pareto front finally obtained still becomes sub-optimal respect to the accuracy objective. This specially occurs in algorithms as NSGA-II [24], since to cross non-dominated rule sets with very different

numbers of rules and different rule structures (forced by the NSGA-II crowding operator) usually gives a bad accuracy, by which this MOGA needs of an adaptation to favor the cross of similar solutions in order to also get good results for the accuracy objective.

On the other hand, this contribution has presented an algorithm based on SPEA2 and a case of study on the use of MOEAs to obtain simpler and still accurate linguistic fuzzy models by also considering a tuning of the system parameters, which represents a more complex search space and therefore needs of different considerations respect to the works in the existing literature.

The results obtained have shown that the use of MOEAs can represent a way to obtain even more accurate and simpler linguistic models than those obtained by only considering performance measures. In this case (also performing a tuning of the parameters), the problem of crossing very different solutions with different number of rules and very different parameters becomes more important since to obtain a wide Pareto with the best solutions is practically impossible. Therefore, as further work, more specific algorithms should be proposed in order to get the best possible solutions.

References

[1] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena (Eds.), "Interpretability issues in fuzzy modeling," *Studies in Fuzziness and Soft Computing* 129 (Springer, Heidelberg, Germany, 2003).

[2] J. Casillas, O. Cordón, M.J. del Jesus, and F. Herrera, "Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction," *IEEE Trans. Fuzzy Systems* vol. 13, no. 1, pp. 13–29, 2005.

[3] C Coello, and G. Toscano, "A Micro-Genetic Algorithm for multiobjective optimization," *First Int. Conf. on Evolutionary Multi-Criterion Optimization*, LNCS 1993 (London, UK, 2001) pp. 126–140.

[4] C.A. Coello, D.A. Van Veldhuizen, and G.B. Lamont, "Evolutionary Algorithms for Solving Multi-Objective Problems", Kluwer Academic Publishers (2002).

[5] O. Cordón, F. Herrera, and L. Sánchez, "Solving electrical distribution problems using hybrid evolutionary data analysis techniques," *Applied Intelligence* vol. 10, pp. 5–24, 1999.

[6] O. Cordon, F. Herrera, M.J. del Jesus, and P. Villar, "A multiobjective genetic algorithm for feature selection and granularity learning in fuzzy-rule based classification systems," *Proc. IX IFSA World Congress and XX NAFIPS Int. Conf.* (Vancouver, Canada, 2001) pp. 1253–1258.

[7] D. Corne, J. Knowles, and M. Oates, "The Pareto Envelope-based Selection Algorithm for multiobjective optimization," *Proc. Parallel Problem Solving from Nature VI Conf.*, LNCS 1917 (Paris, France, 2000) pp. 839–848.

[8] D. Corne, N. Jerram, J. Knowles, and M. Oates, "PESA-II: Region based selection in evolutionary multiobjective optimization," *Proc. Genetic and Evolutionary Computation Conf.* (San Francisco, CA, 2001) pp. 283–290.

[9] K. Deb, "Multi-objective optimization using evolutionary algorithms," John Wiley & Sons (Chichester, UK, 2001).

[10] K. Deb, S. Agrawal, A. Pratab, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II, " *IEEE Trans. Evolutionary Computation* vol. 6, no. 2, pp. 182–197, 2002.

[11] M. Erickson, A. Mayer, and J. Horn, "The Niched Pareto Genetic Algorithm 2 applied to the design of groundwater remediation systems, " *Proc. First Int. Conf. on Evolutionary Multi-Criterion Optimization*, LNCS 1993 (London, UK, 2001) pp. 681–695.

[12] L.J. Eshelman, "The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination," *Foundations of genetic Algorithms* vol. 1, pp. 265–283, 1991.

[13] L.J. Eshelman, and J.D. Schaffer, "Real-coded genetic algorithms and interval-schemata," *Foundations of Genetic Algorithms* vol. 2, pp. 187–202, 1993.

[14] C.M. Fonseca, and P.J. Fleming, "Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization," *Proc. 5th Int. Conf. on Genetic Algorithms* (San Mateo, CA, 1993) pp. 416–423.

[15] J. Horn, N. Nafpliotis, and D.E. Goldberg, "A niched pareto genetic algorithm for multiobjective optimization," *Proc. First IEEE Conf. on Evolutionary Computation, IEEE World Congress on Computational Intelligence* vol. 1 (Piscataway, NJ, 1994) pp. 82–87.

[16] H. Ishibuchi, and T. Murata, "Multi-objective genetic local search algorithm," *Proc. third IEEE Int. Conf. Evolutionary Computation* (Japan, 1996) pp. 119–124.

[17] 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. 135–150, 1997.

[18] H. Ishibuchi, T. Nakashima, and T. Murata, "Three-objective genetics-based machine learning for linguistic rule extraction," *Information Sciences* vol. 136, pp. 109–133, 2001.

[19] H. Ishibuchi, and T. Yamamoto, "Interpretability Issues in Fuzzy Genetics-Based Machine Learning for Linguistic Modelling, in: J. Lawry, J.G. Shanahan, A.L. Ralescu (Ed.), " *Modelling with Words: Learning, Fusion, and Reasoning within a Formal Liguistic Representation Framework*, LNCS 2873, Springer-Verlag (Berlin, Heidelberg, 2003) pp. 209–228.

[20] 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, no. 1, pp. 59–88, 2004.

[21] H. Ishibuchi, and Y. Nojima, "Analysis of interpretability-accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning," *International Journal of Approximate Reasoning* vol. 44, no. 1, pp. 4–31, 2007.

[22] F. Jimenez, A.F. Gomez-Skarmeta, H. Roubos, and R. Babuska, "Accurate, Transparent, and Compact Fuzzy Models for Function Approximation and Dynamic Modeling through Multi-objective Evolutionary Optimization," *Proc. First Int. Conf. on Evolutionary Multi-Criterion Optimization*, LNCS 1993 (Zurich, Switzerland, 2001) pp. 653–667.

[23] J.D. Knowles, and D.W. Corne, "Approximating the non dominated front using the Pareto archived evolution strategy," *Evolutionary Computation* vol. 8, no. 2, pp. 149–172, 2000.

[24] K. Narukawa, Y. Nojima, and H. Ishibuchi, "Modification of evolutionary multiobjective optimization algorithms for multiobjective design of fuzzy rule-based classification systems," *Proc. 2005 IEEE Int. Conf. on Fuzzy Systems* (Reno, USA, 2005) pp. 809–814.

[25] E.H. Mamdani, and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International Journal of Man-Machine Studies* vol. 7, pp. 1–13, 1975.

[26] R.S. Rosenberg, "Simulation of genetic populations with biochemical properties." MA Thesis, Univ. Michigan, Ann Harbor, Michigan, 1967.

[27] J.D. Schaffer, "Multiple objective optimization with vector evaluated genetic algorithms," *Proc. First Int. Conf. on Genetic Algorithms* (Pittsburgh, USA, 1985) pp. 93–100.

[28] N. Srinivas, and K. Deb, "Multiobjective optimization using nondominated sorting in genetic algorithms," *Evolutionary Computation* vol. 2, pp. 221–248, 1994.

[29] T. Takagi, and M. Sugeno, "Fuzzy identification of systems and its application to modeling and control," *IEEE Trans. Systems, Man,* and *Cybernetics* vol. 15, no. 1, pp. 116–132, 1985.

[30] L.X. Wang, and J.M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Trans. Systems, Man,* and *Cybernetics* vol. 22, no. 6, pp. 1414–1427, 1992.

[31] H.L. Wang, S. Kwong, Y.C. Jin, W. Wei, and K.F. Man, "Multi-Objective Hierarchical Genetic Algorithm for Interpretable Fuzzy Rule-Based Knowledge Extraction, " *Fuzzy Sets and Systems* vol. 149, no. 1, pp. 149–186, 2005.

[32] H.L. Wang, S. Kwong, Y.C. Jin, W. Wei, and K.F. Man, "Agentbased evolutionary approach for interpretable rule-based knowledge extraction," *IEEE Trans. Systems, Man,* and *Cybernetics*Part C: Applications and Reviews vol. 35, no. 2, pp. 143–155, 2005.

[33] E. Zitzler, and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach," *IEEE Trans. Evolutionary Computation* vol. 3, no. 4, pp. 257–271, 1999.

[34] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization," *Proc. Evolutionary Methods for Design, Optimization and Control with App. to Industrial Problems* (Barcelona, Spain, 2001) pp. 95–100.