Multi-Objective Decision Making: Towards Improvement of Accuracy, Interpretability and Design Autonomy in Hierarchical Genetic Fuzzy Systems

Myriam Regattieri Delgado*
myriam@dca.fee.unicamp.br

Fernando Von Zuben
vonzuben@dca.fee.unicamp.br

Fernando Gomide
gomide@dca.fee.unicamp.br

Department of Computer Engineering and Industrial Automation
State University of Campinas
Campinas, SP 13083-970, Brazil

Abstract—**This paper presents fuzzy modeling as a multi-objective decision making problem considering accuracy, interpretability and autonomy as goals. The proposed approach assumes that these goals can be handled via corresponding single-objective ϵ-constrained decision making problems whose solution is produced by a hierarchical evolutionary process. The fitting, generalization, and interpretation characteristics of the resulting fuzzy models are discussed using a classification problem.**

I. Introduction

Rule-based models are considered a linguistic representation of the mental model of a certain system by means of experience [1]. Fuzzy rule-based systems and models use fuzzy set theory to express expert knowledge in various domains [2]. In many applications, humans may be unable to extract knowledge out of massive amounts of numerical data, specially when the knowledge required is not easily available. This leads to the development of computer techniques to extract and represent knowledge in a fuzzy rule-based system. Data-driven fuzzy modeling, or fuzzy modeling (FM) for short, has attracted interest of many researches [3–7].

In FM, a fundamental aspect is the trade-off between two important criteria: accuracy and interpretability. The issue of accuracy is critical when models are used, for example, in control systems where the predicted value is fed back. Small errors will be propagated and reflected as errors in the long term behavior [1]. This kind of situation requires accurate fuzzy models, and precise fuzzy modeling is the only goal.

Although important, accuracy is not the only critical aspect in FM. In the last years, model interpretability has gained special attention [1, 5, 7–9]. When the focus is interpretability, linguistic fuzzy modeling generates systems for which the language is easily interpretable by human beings. Despite this useful appeal, there is no well-established definition of interpretability for fuzzy models [5]. Here, interpretability is identified by four characteristics: visibility, simplicity, consistency, and compactness.

Besides accuracy and interpretability, another important aspect in FM is design autonomy. Here autonomy means the degree of involvement of the system designer with the modeling procedure adopted. More precisely, the fewer the parameters defined *a priori* by the designer, the higher the degree of design autonomy.

Recently, numerous works and applications combining fuzzy set theory and evolutionary computation have appeared, and there is an increasing concern about the integration of these two areas. In particular, a great number of works explore the use of genetic algorithms (GAs) to design fuzzy systems. These hybrid approaches are named Genetic Fuzzy Systems (GFS) [10]. A GFS is basically a fuzzy system augmented by a learning process based on GAs [11]. GFS brought considerable attention of researchers from many areas [3, 10, 12, 13].

Genetic algorithms enrich optimization tools for fuzzy systems, particularly when the most significant design decisions can be encoded into a genetic-type representation - a chromosome. However, when complex design decisions must be made it would certainly be more appropriate to optimize a larger set of parameters encoded at different levels, producing a hierarchical genetic fuzzy system (HGFS) [12].

Hierarchical evolutionary rule-based fuzzy modeling will be addressed here, focusing on accuracy, interpretability and design autonomy issues. The solution assumes that these multiple goals can be treated by translating the multi-objective problem into single-objective ϵ-constrained problems [14]. This evolutionary approach uses the same strategy to evolve fuzzy systems as the one suggested by Delgado *et al.* [12], based on HGFS. In this case, the evolutionary parameters are adjusted not only to improve the models performance (accuracy), but also to guarantee the interpretability of the resulting fuzzy models. The hierarchical approach induces a minimum designer intervention by means of automatic tuning of many critical parameters of Mamdani or Takagi-Sugeno (TS) fuzzy models. In this paper, we emphasize TS models only, because the parameters of resulting fuzzy model can be easily identified using the input-output data. Moreover, TS models generally represents the behavior of complex nonlinear systems with a small number of rules, although they do not have a clear semantics due to the functional nature of the consequent.

*On leave from CEFET-PR. Supported by CAPES/PICDT.

0-7803-7280-8/02/$10.00 ©2002 IEEE

II. TAKAGI-SUGENO FUZZY MODELS

Takagi-Sugeno (TS) fuzzy models [15] are powerful inference systems with a reduced number of fuzzy rules. TS models have the consequent of the rules characterized by parametric functional relationships involving the input variables.

Assume a TS fuzzy model composed of m fuzzy rules $R_j, j = 1, \cdots, m$, of the form:

$$R_j : \text{ If } X_1 \text{ is } A_1^j \cdots \text{ and } X_n \text{ is } A_n^j \text{ then } Y \text{ is } g_j(\mathbf{w}_j, \mathbf{x})$$

where $\mathbf{x} = [x_1 \cdots x_n]^T$ is the input vector (T means transpose), and the Q-dimensional vector $\mathbf{w}_j$ contains the parameters of function $g_j(.)$. The most common type of functions used is either constant ($Q = 1$) or linear ($Q = n + 1$). In this paper, nonlinear TS functions ($Q = \frac{n(n-1)}{2} + 2n + 2$) are considered as shown below:

$$\begin{aligned} g_j(\mathbf{w}_j, \mathbf{x}) &= w_{j0} + w_{j1}x_1 + \cdots + w_{jn}x_n + w_{j(n+1)}x_1x_1 + \\ &+ \cdots + w_{j(2n)}x_1x_n + w_{j(2n+1)}x_2x_2 + \cdots + w_{j(2n+2)}x_2x_n + \\ &+ \cdots + w_{j(\frac{n(n-1)}{2}+2n)}x_nx_n + w_{j(\frac{n(n-1)}{2}+2n+1)}x_1x_2\ldots x_n \, . \end{aligned} \tag{1}$$

The accuracy of the fuzzy model can be improved if numerically efficient estimation algorithms are used to define the consequent parameters. Rule antecedent parameters and rule-base structure are the ones to be found by the evolutionary algorithm (see Section IV). After these parameters are found, the elements of vector $\mathbf{w}_j \in \Re^Q$, where $\mathbf{w}_j = [w_{j0} \cdots w_{jQ-1}]^T$, $j = 1, ..., m$, are computed using least squares optimization. Either local or global optimization techniques can be used. The advantages and disadvantages of these techniques are discussed in [3,16]. Delgado *et al.* [16] compared local and global approaches to estimate the consequent parameters in a closed form. The use of more complex consequent functions introduces a large number (Q) of parameters. If Q is higher than necessary, then the parameters of the rules consequent are automatically pruned to avoid redundancy. An efficient pruning procedure is detailed in Delgado *et al.* [17]. This is the pruning procedure adopted here.

III. MULTI-OBJECTIVE DECISION MAKING AND GENETIC FUZZY SYSTEMS

Fuzzy modeling requires the consideration of multiple criteria in the design process. Recently, multiple goals in fuzzy system design has gained more interest as discussed in [3, 5, 13, 18–20].

Generally speaking, a multi-objective optimization problem can be formulated as:

$$\begin{aligned} \min \quad & \mathbf{h}(\mathbf{x}) \, , \\ \text{s.t.} \quad & \mathbf{x} \in \Omega \end{aligned}$$

where $\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), h_2(\mathbf{x}), \cdots, h_r(\mathbf{x})]^T$ is the vector of objectives, $h_i(.) : \Omega \to \Re, i = 1, \cdots, r$, and $\Omega \subseteq \Re^n$ is the subset of feasible solutions for which $\mathbf{h}$ is defined. In multi-objective optimization, the set of efficient solutions, also called non-dominated or Pareto-optimal, is composed of all those elements of the input space for which the corresponding vector of objectives cannot be improved in any of its components without degradation in another component.

Multi-objective optimization problems can be solved by different techniques [14]. Some techniques adopt the transformation of the original problem into a set of single-objective problems. The two most common ways to obtain the non-dominated solutions for the original problem are: 1) weighted method; 2) ϵ-constrained method. The former assumes an association to aggregate all objectives settled as a (often linear) combination of all criteria to define a single function to be optimized. The later comprises the solution of single-objective problems subject to the set of ϵ-constraints associated with the other objectives, except the one taken as reference.

In FM, more complex multi-objective decision making problems emerge. For these problems, the objectives are heterogeneous in the sense that they are not a function of the same set of variables. But the multi-objective decision theory is still a source of inspiration for many approaches. Genetic algorithms [21] appear as useful alternatives to carry out search for feasible solutions and Pareto-optimal solutions, resulting in multi-objective genetic fuzzy systems. Based on the weighted approaches, many methods adopt an aggregation formula (weighted sum) of the different criteria that are translated into a fitness measure of a solution of the original problem [3, 5, 13]. However, it is often hard to choose the appropriate weights. Experiments show that small changes in the weights may guide to completely different results. Alternatively, Pareto-optimal solutions can be generated and the decision-maker may choose the preferred solution [18–20].

In this paper, FM will be viewed as a multi-objective decision making problem, for which accuracy, interpretability and design autonomy are the goals, and feasible solutions are achieved from ϵ-constrained optimization problems. The aim here is to minimize the error criterion to improve accuracy, but subject to visibility, simplicity, compactness and consistency constraints. Visibility is associated with two different characteristics: the γ-completeness and the α-overlapping. The former, known as coverage property [1], fixes a minimum degree (γ) of overlapping in the universe partition, that guarantees granulations without gaps [4, 22]. The later, also called distinguishability property [1], fixes a maximum degree (α) of overlapping in the universe partition [12]. The simplicity of each rule is evaluated by its length, measured by the number of features or variables minus the number of irrelevant features identified by *don't care* conditions [13]. Consistency and compactness are associated with the rule base. Compactness is measured by the total of fuzzy rules in the rule base. A consistent rule-base presumes the absence of conflicting rules, i.e., rules with the similar antecedents, but with very different consequent parameters.

Therefore, the corresponding ϵ-constrained problem may be expressed as follows:

$$\min_{\mathbf{v}\in\Omega(\epsilon)} \quad \text{MSE} = \frac{1}{N}\sum_{p=1}^{N}\left(y_p - \hat{y}_p\right)^2$$

$$s.t. \quad \left\{ \begin{array}{rcccl} k\text{-consistency} & = & \epsilon_1 & = & \text{0-consistency} \\ \text{TotRules} & \leq & \epsilon_2 & = & \text{MaxRules} \\ \text{TotMF} & \leq & \epsilon_3 & = & \text{MaxGranularity} \\ \alpha\text{-overlapping} & \leq & \epsilon_4 & = & \alpha\text{max} \\ \gamma\text{-completeness} & \geq & \epsilon_5 & = & \gamma_{\text{min}} \end{array} \right.$$

where $\mathbf{v}$ is the decision variable, and the vector $\epsilon = \{\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_5\}$, defined in $\varepsilon = \{\epsilon : \Omega(\epsilon) \neq \emptyset\}$ is associated with interpretability requirements: consistency (ϵ_1), compactness (ϵ_2), simplicity (ϵ_3), visibility (ϵ_4 and ϵ_5).

The formulation above is a single objective ϵ-constrained version of the original multi-objective problem. If a solution provided by genetic algorithms is a unique solution of the ϵ-constrained problem or if it solves all the ϵ-constrained problems when we consider each constrain as a reference objective function, than this solution is Pareto-optimal. For genetic operators, the violation of visibility conditions needs repairing procedures to shift the membership functions. Repairing algorithms represent a major advantage of evolutionary techniques to adjust membership function parameters when contrasted with classical optimization techniques.

Next section details the hierarchical evolutionary process, pointing out how the constraints are satisfied during evolution.

IV. The Hierarchical Evolutionary Approach

The hierarchical evolutionary process is structured in modules that consider the membership functions (or partition set) at the first level, the population of individual rules at the second level, the population of sets of rules at the third level and the population of fuzzy systems at the fourth level [12]. This FM approach may be applied to evolve Takagi-Sugeno fuzzy models or Mamdani fuzzy models. The coding scheme makes use of real and integer encoding (depending on the level) and four populations (each one associated with one level) evolve interactively.

A. Encoding Scheme

The chromosomes representing the partition set (level I) use real values encoding. The chromosomes representing individual rules and rule-base individuals (level II and III, respectively) use integer encoding. Fuzzy systems individuals (level IV) use real values to encode t-norm parameters p_t and integer codes for the remaining alleles. The chromosomes at all levels have hierarchical relationships, as summarized in Fig. 1.

A partition set individual (level I) contains all the membership functions defined in the universe of the variables involved, and is represented by a real code chromosome. The chromosome is formed by the concatenation of all the partition sets associated with each variable. Each member of the population of individual rules (level II in Fig. 1) represents a fuzzy proposition. This population accepts different combinations of membership functions as identified by their indexes (the order in the

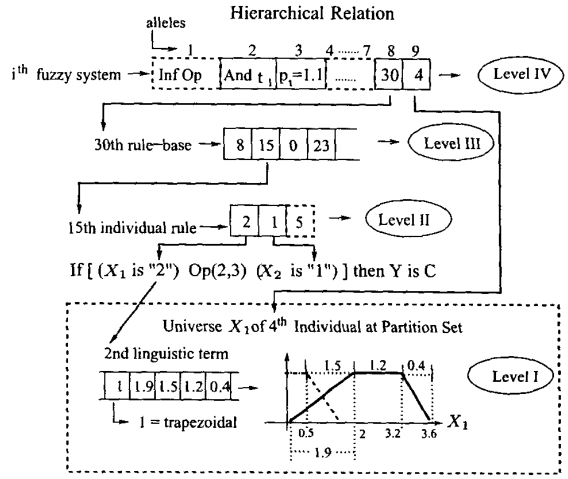

Fig. 1. Encoding and hierarchical relations among individuals at the different levels (rules with two antecedent variables are assumed)

partition set). Null values indicate *don't care* conditions, and the genetic operators prioritizes simpler individual rules (rules with more *don't care* conditions). The population of set of rules (level III) has individuals formed by indexes that identify the corresponding rules. The length of the chromosome, fixed by the constraint ϵ_2, determines the maximum number of fuzzy rules but smaller rule-bases are always aimed at first. Each individual at level IV represents a fuzzy system. At this level, the code of each chromosome associates a specific set of rules (allele at site 8) and a partition set (allele at site 9), with a subset of operators used to define the inference mechanism (alleles at site 1 to 7). When Takagi-Sugeno fuzzy models are chosen, alleles at site 4 to 7 (which are associated with Mamdani fuzzy models and identify the rule-semantic, rule aggregation operator and defuzzification method) are not considered. In this case, only alleles at sites 2,3,8 and 9 are evolved. For example, the *i-th* fuzzy system, depicted at level IV in Fig. 1, uses the *30-th* set of fuzzy rules at level III and the *4-th* partition set individual at level I. Then , each individual rule that composes its rule-base aggregates the antecedent part by the t-norm $\mathbf{t_1}$ given by:

$$a\,\mathbf{t}_1\,b = \frac{1}{1 + \sqrt[p_t]{\left(\frac{1-a}{a}\right)^{p_t} + \left(\frac{1-b}{b}\right)^{p_t}}}, \tag{2}$$

with an associated parameter $p_t = 1.1$ (see Klement et al. [23] for details on $\mathbf{t}$-norm taxonomy). And more, the *15-th* individual rule of the *i-th* fuzzy system could be given by:

If (X_1 is "2") $\mathbf{t}_1$ (X_2 is "1") then Y is $g_{15}(\mathbf{w}_{15}, \mathbf{x})$, for TS fuzzy models; and
If (X_1 is "2") $\mathbf{t}_1$ (X_2 is "1") then Y is "5", for Mamdani fuzzy models.

In Fig. 1, for the *i-th* fuzzy system, the linguistic terms indexed by "2" and "1", and eventually "5" for Mamdani models, are defined in the *4-th* chromosome at level I. In the case of TS fuzzy systems, functions $g(.)$ of the input variables are used at the consequent part of the rules, as shown in Equation (1) of Section II.

B. Hierarchical Evolutionary Algorithm

The design process uses a GA strategy to produce improved fuzzy model parameters along generations. The main steps of the hierarchical evolutionary algorithm are summarized as follows:

1. Start with Generation = 1;
2. Initialize populations for each module;
3. If Takagi-Sugeno fuzzy modeling
 (a) Compute the optimal parameter for the consequent of each individual at level IV (see Section II);
4. Calculate the fitness (ft) of each individual at all population levels as follows:
 (a) Fuzzy System (level IV): $ft_{\text{FS}}(i)$ is based on the fuzzy system performance;
 (b) Rule-Base (level III): $ft_{\text{RB}(k)} = max\left(ft_{\text{FS}(b)}, \cdots, ft_{\text{FS}(d)}\right)$, where $b, \cdots, d$ are the fuzzy systems of which the rule-base (k) is part.
 (c) Indiv. Rule (level II): $ft_{\text{I}(j)} = mean\left(ft_{\text{RB}(m)}, \cdots, ft_{\text{RB}(p)}\right)$, where $m, \cdots, p$ are the rule-bases of which the individual rule (j) is part.
 (d) Part. Set (level I): $ft_{\text{PS}(q)} = max\left(ft_{\text{FS}(x)}, \cdots, ft_{\text{FS}(z)}\right)$, where $x, \cdots, z$ are the fuzzy systems of which the partition set (q) is part.
5. If the stop condition does not hold, do:
 (a) From level IV to level I apply the evolutionary operations (selection, crossover and mutation) to form a new population;
 (b) Generation = Generation + 1;
 (c) Return to step 3;

The initialization phase (step 1) comprises random and deterministic generation of population at each level. At the partition set (level I), real code chromosomes are generated to uniformly distribute the membership functions over the associated universes. At level II, different fuzzy propositions encoded by integer chromosomes are randomly generated. *Don't care* conditions represented by null values are introduced to attend the simplicity criterion. At the third level, each integer chromosome, that represents a set of fuzzy rules, is randomly generated. Rule exclusions are possible to attend the compactness criterion. At level IV, alleles are randomly initialized (except alleles at site associated with parameter p_t). Alleles at sites 2 and 4 associated with antecedent aggregation and rule semantic (this later is specific to Mamdani Models) are generated to cover all the possible norm operators. Alleles at site 3 and 5 associated with t-norm parameters p_t are initialized with the value $p_t = 2.0$. Alleles at site 8 and 9 are randomly generated and define which rule-base and partition set will be used by the current fuzzy system.

After all the population levels have been initialized, the next step involves the optimization (only in the case of TS models) of the consequent parameters of each fuzzy system at level IV.

Fitness calculation (step 4) is performed after defining all fuzzy system parameters. This process starts at level IV and finishes at level I. At level IV, the fitness is evaluated by decoding the chromosome representation, and then measuring the fitness function that depends on the performance of each fuzzy system, when applied to the problem under consideration.

At step 5, either the stop condition (maximum number of generations or error criterion) is verified or the algorithm continues. The evolutionary operators work downward, that is, selection, crossover and mutation are applied from the top level (IV) to the bottom level (I).

C. Evolutionary Operators

Selection is the first evolutionary operator applied and uses the tournament technique [21] to select 80% of the individuals. The remaining 20% of the population is chosen by a technique that favors diversity in the population. This diversity criterion focus the choice of the "most diverse" chromosomes when compared with the one with the highest fitness. The measure of diversity is either the Euclidean distance for real-coded chromosomes or the Hamming distance for integer codes. The selection process is combined with an elitist strategy. The second evolutionary operator is the 1-point crossover. The crossover point is randomly chosen. Mutation, the last operator, is applied to real and integer encoding parameters.

At the partition level, two visibility conditions must be fulfilled by the set of membership functions to achieve the interpretability requirements: 1) the γ-*completeness*, and 2) the maximum degree α of overlapping. These conditions state that, given a value x of one of the inputs within the operation range, we can always find a linguistic term A such that $\mu_A(x) \geq \gamma$, and an unique linguistic term B such that $\mu_B(x) \geq \alpha$. This assumption means a minimum (γ) and maximum (α) overlapping degrees among the membership functions during evolution. Figure 2 shows examples of interpretability analysis in the partition set.

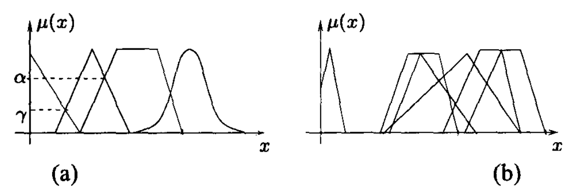

Fig. 2. Interpretability analysis: (a) an interpretable partition set; (b) a non-interpretable partition set

The genetic operators change the shape and location of membership functions but the minimum and maximum overlapping degrees are always fulfilled. In the case of visibility restriction violation, a repairing procedure is used to attend the γ-completeness or α-overlapping criteria, shifting the membership functions when it is necessary.

At the second level, crossover and mutation are applied to try different combinations of linguistic terms of each proposition. Mutation changes the current value by a new one chosen from the set $\{0, 1, ..., L_i\}$, where $L_i \leq \epsilon_3$ means the maximum number of linguistic terms for the i-th variable. Null values identify *don't care conditions*.

At the rule-base level, crossover and mutation change the integer indexes associated with individual rules. At this level, alleles (representing individual rules) with lower fitness have higher probability of being changed by mutation. The new values are chosen among the values of the set $\{0, 1, ..., S_{II}\}$, where S_{II} means the size of the population at level II. Null values are possible (indicating a rule elimination). The length of the chromosome is defined by the constraint ϵ_2. Null values mean the exclusion of the corresponding rule in the rule-base.

At the fuzzy system level, crossover produces combination of two parents. The mutation of integer alleles changes a value by a new one chosen among all the possibilities. For example, the new value of alleles at site 2 (see Fig. 1) is chosen from $\{1, \cdots, 9\}$, where the indexes represent the t-norms accepted for antecedent aggregation [2]; for alleles at site 8, the new values are chosen from $\{0, 1, ..., S_{III}\}$; and for alleles at site 9 are chosen from $\{0, 1, ..., S_I\}$, where S_{III} and S_I are the population sizes at levels III and I, respectively. Uniform mutation operator is applied to alleles at sites associated with parameter p_t because it achieved better results when compared with non-uniform mutation. The possibility of adopting different inference operators gives flexibility to the resulting fuzzy system and improves the degree of design autonomy.

V. Experiments and Results

The performance of the HGFS has been evaluated with the Iris classification data set. Popularized by Fisher [24], this three-class classification problem has 150 four-dimensional vectors representing 50 samples of each species: *Iris setosa* (Cl_1), *Iris versicolor* (Cl_2), and *Iris virginica* (Cl_3). The four dimensional pattern vector $x_p = (x_{1p}, x_{2p}, x_{3p}, x_{4p}), p = 1, \cdots, 150$, has been normalized within the range [0,1]. The attributes are as follows: x_{1p} is the sepal length, x_{2p} is the sepal width, x_{3p} is the petal length, and x_{4p} is the petal width.

The training (Tr) and testing (Ts) sets have 75 patterns each (25 patterns randomly chosen from each class). The purpose is twofold: to evaluate the generalization capabilities in a classification context, and to compare with the results provided by Castellano and Fanelli [25], because their approach named Compact Min Fuzzy (CMinFuzzy) has outperformed (in terms of trade-off between accuracy/compactness) many other approaches that solved the Iris classification problem.

Here, HGFS with non-linear consequents optimized by the global method is used to solve the multi-objective problem of fuzzy modeling based on ϵ-constrained method proposed in Sect. III. To test the benefits of this methodology, the proposed approach is compared with HGFS based on a weighted method, following the one proposed in [8] which assumes that the fitness of each fuzzy system is given by:

$$\text{fitness} = W_{error} \frac{1}{\text{RMSE}_{tr}} - W_{comp}\, m ,$$

where W_{error} and W_{comp} are positive weights, m define the total of fuzzy rules in the rule base, and RMSE_{tr} measures the root mean squared error for training data.

HGFS uses fixed population sizes in each module: 30 individuals in the population of fuzzy systems (level IV), 70 individuals in the population of set of fuzzy rules (level III) and 10 individuals in the population of partition set (level I). Since the population of individual rules (level II) contains all possible combinations of the linguistic terms, no evolution is necessary at this level. The evolutionary operators (selection, crossover and mutation) are those detailed in Sect. IV-C. Crossover and mutation genetic operators are performed with probability $P_C = 0.2$ and $P_M = 0.08$, respectively, for all the three levels (I, III and IV) that evolve. All these evolutionary parameters have been chosen after many tests using different values.

For comparison purposes, in the ϵ-constrained method the constraint ϵ_2 is defined as 4 resulting in a maximum of 4 fuzzy rules in each rule-base; the constraint $\epsilon_3 = 2$ fixed a maximum of 2 linguistic terms for each variable. For the weighted method, different weights have been tested, but the best results were achieved with $W_{erro} = 10$ and $W_{comp} = 0.1$. The ANFIS system assumed the same partition for all variables, that is, each with 2 Gaussian membership functions. This means a total of 16 fuzzy rules.

If we denote by f the fuzzy system evolved by HGFS or obtained by ANFIS, then the following assignment is devised to relate a value of $f(\mathbf{x}_p)$ with each class Cl_i: $Cl_1 \to f(\mathbf{x}_p) = 1$; $Cl_2 \to f(\mathbf{x}_p) = 2$; $Cl_3 \to f(\mathbf{x}_p) = 3$. Table I shows the classification errors for the test set obtained by Castellano and Fanelli [25], ANFIS, and HGFS, based on two optimization methods: ϵ-constrained method (ϵ-constrMet) and weighted method (WeigMet).

TABLE I

Iris data simulation results

Approach	Cycles	Rules	Misclassification	
			Training	Test
ANFIS	100	16	0	6
CMinFuzzy [25]	-	5	0	4
HGFS (WeigMet)	100	7	0	7
HGFS (ϵ-constrMet)	100	4	0	3

As Table I shows, the fuzzy classifier provided by the proposed approach (HGFS based on the ϵ-constrained method) outperforms the other approaches because it achieved the best compromise between classification performance (classification rate of 96% in the test set) and system complexity (a total of 4 fuzzy rules). The results emphasize the benefits of translating the multi-objective problem into a single-objective ϵ-constrained problem, for which the solution may be produced by a hierarchical evolutionary process based on genetic algorithms.

It is important to point out that some approaches presented in the literature achieve 100% of correct classification in Iris data when all the 150 available patterns were used in the training process. This means that the cardinalities of training set Tr and test set Ts are given by $S_{Tr} = 150$ and $S_{Ts} = 0$, respectively. When we consider all the available patterns in

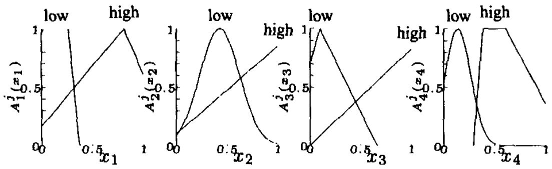

Fig. 3. The partition for variables x_1, x_2, x_3, and x_4

the training process, the HGFS was also able to correctly classify all the patterns. Although, since the aim here is to test the generalization capability of the proposed approach, the attention has been given to the simulations considering patterns distributed among training and test sets.

At the end of the evolutionary process, for the fuzzy system considering HGFS based on ϵ-constrained method, the resulting rule-base produced is detailed bellow:

R_1 : If x_3 is *low* and x_4 is *low* then $y = 0.86 - 0.3x_1 + 0.19x_2 + 0.31x_3 + 0.09x_4 - 0.14x_1^2 - 0.23x_2^2 - 2.86x_3^2$;
R_2 : If x_1 is *low* and x_2 is *low* and x_3 is *high* then $y = 1.15 - 1.97x_1 - 6.06x_2 + 1.82x_3 + 3.41x_4 + 5.18x_2^2$;
R_3 : If x_1 is *high* and x_3 is *low* and x_4 is *high* then $y = 9.7 - 0.17x_1 + 1.97x_2 + 2.65x_3 - 33.7x_4 + 0.36x_1^2 - 3.44x_2^2 - 0.8x_3^2 + 31x_4^2$;
R_4 : If x_1 is *high* and x_2 is *high* and x_4 is *low* then $y = 1.52 - 0.49x_1 + 0.2x_2 + 3.83x_3 + 0.32x_4 + 0.29x_1^2 - 0.42x_2^2 - 2.42x_3^2 - 0.24x_4^2 + 0.33x_1x_2x_3x_4$.

The antecedent aggregation '**and**' evolved is given by Equation (2), with $p_t = 1.66$. The linguistic terms *low* and *high* are shown in Fig. 3. As it can be noted, a good level of accuracy is achieved and all constraints, needed to ensure the interpretability criterion, are fulfilled: visibility, simplicity (*don't care* conditions in all the fuzzy rules), compactness (only 4 fuzzy rules), and consistency.

VI. Conclusions

In this paper, accuracy, interpretability and autonomy were envisioned from the multi-objective decision making framework and associated single-objective ϵ-constrained problems. Therefore, in addition to precision, interpretability of the rule-based model is achieved through a set of constraints imposed on the search process. Interpretability is based on visibility, simplicity, compactness and consistency goals. The use of HGFS to solve the ϵ-constrained problem improves performance in terms of accuracy, satisfies interpretability constraints, and provides an automatic adjustment mechanism for a number of critical parameters, which increases autonomy by minimizing user intervention.

Acknowledgments

Myriam Regattieri Delgado acknowledges CEFET/PR and CAPES/PICDT, Fernando Von Zuben the CNPq grant 300910/96-7, and Fernando Gomide the CNPq grant 300729/86-3, for their support.

References

[1] J. Espinosa and J. Vandewalle, "Constructing fuzzy models with linguistic integrity from numerical data-AFRELI algorithm," *IEEE Transactions on Fuzzy Systems*, Vol. 8:5, pp. 591–600, 2000.
[2] W. Pedrycz and F. Gomide, An Introduction to Fuzzy Sets: Analysis and Design. Cambridge: MIT Press, 1998.
[3] F. Hoffmann and O. Nelles, "Structure identification of TSK-fuzzy systems using genetic programming," in Proceedings of IPMU 2000, Madrid, Spain, pp. 438–445, July 2000.
[4] J.-S. Jang, "ANFIS: Adaptive-network-based fuzzy inference systems," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 23:3, pp. 665–685, 1993.
[5] Y. Jin, "Fuzzy modeling of high-dimensional systems: Complexity reduction and interpretability improvement," *IEEE Transactions on Fuzzy Systems*, Vol. 8:2, pp. 212–221, 2000.
[6] C. L. Karr, "Genetic algorithms for fuzzy controllers," *AI Expert*, Vol. 6:2, pp. 26–33, 1991.
[7] K. Nozaki, H. Ishibuchi, and H. Tanaka, "A simple but powerful heuristic method for generating fuzzy rules from numerical data," *Fuzzy Sets and Systems*, Vol. 86, pp. 251–270, 1997.
[8] H. Ishibuchi, D. Takeuchi, and T. Nakashima, "Ga-based approaches to linguistic modeling of nonlinear functions," in Proceedings of 9thIFSA World Congress and 20thNAFIPS International Conference, Vancouver, CA, July 2001.
[9] M. Setnes, R. Babuška, and H. B. Verbruggen, "Rule-based modeling: Precision and transparency," *IEEE Transactions on Systems, Man, and Cybernetics-Part C.*, Vol. 28:1, pp. 165–169, 1998.
[10] F. Herrera and L. Magdalena, "Introduction: Genetic fuzzy systems," *International Journal of Intelligent Systems*, Vol. 13, pp. 887–890, 1998.
[11] O. Cordón, F. Herrera, F. Gomide, F. Hoffmann, and L. Magdalena, "Ten years of genetic fuzzy systems: Current framework and new trends," in Proceedings of 9thIFSA World Congress and 20thNAFIPS International Conference, Vancouver, CA, pp. 1241–1246, July 2001.
[12] M. R. Delgado, F. Von Zuben, and F. Gomide, "Hierarchical genetic fuzzy systems," *Information Sciences - Special Issue on Recent Advances in Genetic Fuzzy Systems*, Vol. 136:1–4, pp. 29–52, 2001.
[13] H. Ishibuchi and T. Nakashima, "Genetic-algorithm-based approach to linguistic approximation of nonlinear functions with many input variables," in Proceedings of FUZZ-IEEE'99, Seoul, Korea, pp. 779–784, Aug. 1999.
[14] V. Chankong and Y. Haimes, Multiobjective Decision Making - Theory and Methodology. Elsevier, 1983.
[15] T. Takagi and M. Sugeno, "Derivation of fuzzy control rules from human operator's control actions," in Proceedings of the IFAC Symp. on Fuzzy Information, Knowledge Representation and Decision Analysis, Marseilles, France, pp. 55–60, 1983.
[16] M. R. Delgado, F. Von Zuben, and F. Gomide, "Local and Global Estimation of Takagi-Sugeno Consequent Parameters in Genetic Fuzzy Systems," in Proceedings of 9thIFSA World Congress and 20thNAFIPS International Conference, Vancouver, CA, pp. 1247–1252, July 2001.
[17] M. R. Delgado, F. Von Zuben, and F. Gomide, "Optimal Parameterization of Evolutionary Takagi-Sugeno Fuzzy Systems," in Proceedings of 8th Information Processing and Management of Uncertainty in Knowledge-Based Systems Conference (IPMU), Madrid, Spain, pp. 650–657, July 2000.
[18] O. Cordón, M. del Jesus, F. Herrera, and P. Villar, "A multiobjective genetic algorithm for feature selection and granularity learning in fuzzy-rule-based classification systems," in Proceedings of 9thIFSA World Congress and 20thNAFIPS International Conference, Vancouver, CA, pp. 1253–1258, July 2001.
[19] L. Gacôgne, "Research of Pareto set by genetic algorithm, application to multicriteria optimization of fuzzy controller," in Proceeding of 5th European Congress on Intelligent Techniques and Soft Computing, EUFIT'97, Aachen, Germany, pp. 837–845, Sept. 1997.
[20] A. Gómez-Skarmeta, F. Jiménez, and J. Ibánez, "Pareto-optimality in fuzzy modeling," in In 6th European Congress on Intelligent Techniques and Soft Computing, EUFIT'98, Aachen, Germany, pp. 694–700, Sept. 1998.
[21] Z. Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs. Springer-Verlag, 1996.
[22] A. Gonzales and R. Perez, "Completeness and consistency conditions for learning fuzzy rules," *Fuzzy Sets and Systems*, Vol. 96, pp. 37–51, 1998.
[23] E. Klement, R. Mesiar, and E. Pap, Triangular Norms. Kluwer Academic Publishers, 2000.
[24] R. A. Fisher, "The use of multiple measurements in taxonomic problems," in Annual Eugenics, Vol. 7, pp. 179–188, 1936.
[25] G. Castellano and A. M. Fanelli, "A staged approach for generation and compression of fuzzy classification rules," in Proceedings of FUZZ-IEEE'2000, San Antonio, USA, pp. 42–47, May 2000.