fuzzy classification rules for multiobjective genetic fuzzy rule selection. Section IV shows an experimental study of this method on a set of 15 well-known datasets. Finally, Section V points out some conclusions.

II. Preliminaries: Fuzzy Rule-Based Classifiers Structure and Inference

Let us assume that we have m training (i.e., labeled) patterns $\vec{x}_p = (x_{p1}, ..., x_{pn})$, $p = 1, 2, ..., m$ from M classes in an n-dimensional pattern space where x_{pi} is the attribute value of the pth pattern for the ith attribute ($i = 1, ..., n$). For the simplicity of explanation, we assume that all the attribute values have already been normalized into real numbers in the unit interval [0, 1]. Thus the pattern space of our classification problem is an n-dimensional unit-hypercube [0, 1]^n.

For our n-dimensional pattern classification problem, we use fuzzy rules of the following type:

$$R_q : \text{ If } x_1 \text{ is } A_{q1} \text{ and ... and } x_n \text{ is } A_{qn} \text{ then Class } C_q \text{ with } CF_q, \quad (1)$$

where R_q is the label of the qth fuzzy rule, $\vec{x} = (x_1, ..., x_n)$ is an n-dimensional pattern vector, A_{qi} is an antecedent fuzzy set ($i = 1, ..., n$), C_q is a class label, and CF_q is a rule weight. We denote the antecedent fuzzy sets of R_q as a fuzzy vector $\vec{A}_q = (A_{q1}, A_{q2}, ..., A_{qn})$.

Fourteen fuzzy sets are initially considered in four fuzzy partitions with different granularities. Figure 1 depicts these partitions. In addition to those 14 fuzzy sets, we also use the domain interval [0, 1] itself as an antecedent fuzzy set in order to represent a *don't care* condition.

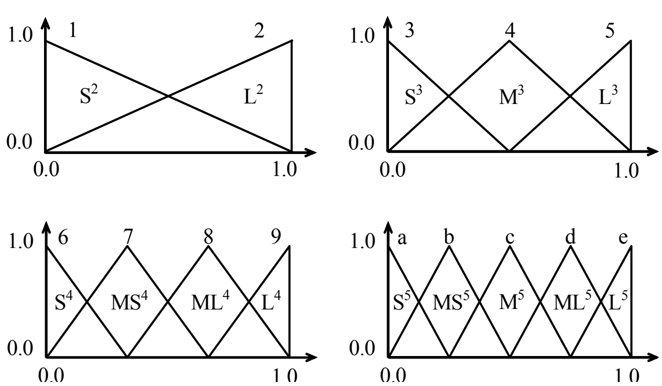

Fig. 1. The fourteen antecedent fuzzy sets considered.

Let S be a set of fuzzy rules of the form in (1). When an input pattern $\vec{x}_p$ is to be classified by S, first we calculate the compatibility grade of $\vec{x}_p$ with the antecedent part $\vec{A}_q = (A_{q1}, A_{q2}, ..., A_{qn})$ of each fuzzy rule R_q in S using the product operation as,

$$\mu_{\vec{A}_q}(\vec{x}_p) = \mu_{A_{q1}}(x_{p1}) \cdot ... \cdot \mu_{A_{qn}}(x_{pn}), \quad (2)$$

where $\mu_{A_{qi}}(\cdot)$ is the membership function of the antecedent fuzzy set A_{qi}. Then a single winner rule R_w is identified using the compatibility grade and the rule weight of each fuzzy rule as

$$\mu_{\vec{A}_w}(\vec{x}_p) \cdot CF_w = \max\{\mu_{\vec{A}_q}(\vec{x}_p) \cdot CF_q \mid R_q \in S\}. \quad (3)$$

The input pattern $\vec{x}_p$ is classified as the consequent class C_w of the winner rule R_w. When multiple fuzzy rules with different consequent classes have the same maximum value in (3), the classification of $\vec{x}_p$ is rejected. If there is no compatible fuzzy rule with $\vec{x}_p$, its classification is also rejected.

III. An Algorithm for Generating Single Granularity-Based Fuzzy Classification Rules

As we have already explained, multiobjective genetic fuzzy rule selection has been based on a previously fixed granularity [10], [11] (five linguistic terms in all the attributes) or multiple granularities [12]. Based on this last approach [12], in this section we propose a mechanism to generate single granularity-based fuzzy classification rules, a nearer to the interpretability approach. The proposed procedure is as follows:

- *Step 1:* Rule extraction with multiple granularities.
- *Step 2:* Specification of single granularity for each attribute based on the extracted rules.
- *Step 3:* Rule extraction with selected single granularities.
- *Step 4:* Multiobjective genetic fuzzy rule selection.

The original multiple granularities based procedure [12] is composed of Steps 1 and 4. Steps 2 and 3 are additional procedures. In Step 1, we extract a fixed short number of rules for each class based on well-known data mining rule evaluation measures [2] and multiple granularities. In Step 2, we select a single granularity for each attribute based on the extracted rules. Then, we extract the final set of candidate rules for each class by using the selected single granularities in Step 3. Step 4 is the same as the original one to perform multiobjective genetic fuzzy rule selection. The next subsections present detailed explanations of these steps.

A. Rule Extraction with Multiple Granularities (Step 1)

Since 14 antecedent fuzzy sets in Figure 1 and an additional *don't care* fuzzy set [0, 1] are used for each attribute of the n-dimensional classification problem, the total number of possible fuzzy rules is 15^n. Among these possible rules, we examine only short fuzzy rules with a small number of antecedent conditions (i.e., short fuzzy rules with many *don't care* conditions) to generate an initial set of candidate rules. In this work, we specify the maximum number of antecedent conditions as three for datasets with less than 30 attributes and two for datasets with more than or equal to 30 attributes.

The consequent class C_q and the rule weight CF_q of each fuzzy rule R_q are specified from training patterns compatible with its antecedent part $\vec{A}_q = (A_{q1}, A_{q2}, ..., A_{qn})$ in the