Soft Computing (2018) 22:5021–5031
https://doi.org/10.1007/s00500-018-3239-1

FOCUS

# Multi-objective evolutionary algorithm for tuning the Type-2 inference engine on classification task

**Edward C. Hinojosa**[1] **· Heloisa A. Camargo**[2]

Published online: 21 May 2018


**Abstract**

Type-2 fuzzy systems have been investigated as an alternative formalism to deal with uncertainty when the classic Type-1 fuzzy systems do not offer the suitable flexibility for the representation of the information being modeled. The higher flexibility in representation comes with a higher complexity in the system modeling, mainly in the design of the Type-2 fuzzy sets and in the definition of the inference engine parameters. In this paper, we focus on the Type-2 fuzzy systems design, proposing a multi-objective evolutionary approach for tuning the Type-2 inference engine of a fuzzy rule-based classification system by means of automatically choosing the t-norm used in the inference process. The selection of the t-norm used plays an important hole, since different operators could lead to different results. In a preliminary version of this work, we have proposed an approach to design and optimize Type-2 fuzzy systems that includes the tuning of Type-2 fuzzy sets and the selection of rules. The additional tuning process proposed in this paper is an extension of the previous method in the sense that the same evolutionary procedure performs simultaneously the tuning of the inference mechanism and the tasks performed before. The evolutionary process is executed by means of a multi-objective genetic algorithm with three objectives that aim to balance the accuracy and interpretability of the system generated: the accuracy, the number of rules and the number of conditions in the rules. The proposed method has been compared with a state-of-the-art method proposed in the literature, presenting good results.

**Keywords** Type-2 fuzzy inference system · Fuzzy rule-based classification systems · Tuning · Multi-objective evolutionary algorithms

## 1 Introduction

Several computational intelligence techniques have demonstrated their usefulness and suitability in real-world problems where complexity and imperfect information are present. Fuzzy rule-based systems (FRBS), in particular, are recognized as consolidated techniques to deal with imprecision and uncertainty, besides being attractive for providing a linguistic description of the system in the form of rules.

Communicated by C. Kahraman.

✉ Edward C. Hinojosa
ehinojosa@unsa.edu.pe

Heloisa A. Camargo
heloisa@dc.ufscar.br

[1] National University of San Agustin, Arequipa, Peru

[2] Federal University of São Carlos, São Carlos, SP, Brazil

Despite being advantageous over the classic counterpart concerning the modeling of imprecision, the design of fuzzy systems requires the definition of accurate fuzzy memberships for the fuzzy sets that stand for linguistic values (Zadeh 1975), what is not always possible. In this sense, the concept of Type-2 fuzzy sets has gained importance and received attention as a means to deal with the type of uncertainty that Type-1 fuzzy sets cannot handle, such as vagueness and low reliability of linguistic terms (Karnik and Mendel 1998; Karnik et al. 1999; Castillo and Melin 2008).

According to Castillo and Melin (2012), experimental evidence of better performance of Type-2 fuzzy systems (T2FS) over Type-1 fuzzy systems (T1FS) has been reported in an expressive number of papers, as well as applications in intelligent control, pattern recognition, intelligent manufacturing, time series prediction and others. Despite their growing success, a general methodology to design and optimize T2FS is not available. In the Type-1 case, evolutionary techniques are the most often applied to automatically generate the system components. There are a large variety of approaches that generate and/or optimize the Data Base, the Rule Base of the FRBS of both, in sequence or simultaneously. Most relevant works have been reviewed and classified in Fazzolari et al. (2013), Herrera (2008), Cordón (2011) and Cordón et al. (2001).

While allowing a more powerful structure for knowledge representation and reasoning, T2FS pose additional challenges concerning their design and execution for the high number of parameters to be defined. To surpass the problem of difficult design and costly operations of Type-2 fuzzy sets, the notion of interval fuzzy sets, based on the concepts of lower and upper membership functions, was proposed (Liang and Mendel 2000). The simplified representation facilitated the generation and stimulated the use of T2FS, but the complexity remains higher than the one encountered in the T1FS.

Following the successful results regarding T1FS, evolutionary techniques have been shown to be robust and suitable enough to handle the additional complexity of T2FS. A thorough review on methods of bio-inspired computation, including genetic algorithms (GA), used to optimize T2FS can be found in Castillo and Melin (2012).

The work presented here focuses on the design and tuning of Type-2 fuzzy rule-based classification systems (FRBCS), by means of multi-objective genetic algorithms (MOGA). FRBCS are FRBS composed of rules of a particular format, dedicated to perform the classification task. According to the review reported in Castillo and Melin (2012), there are very few papers that focus on T2FS for the classification problem, contrary to what happens on the realm of Type-1 fuzzy systems.

### 1.1 Objectives

The objective of this work is to present a multi-objective evolutionary approach that optimizes Type-2 FRBCS, called Multi-Objective Genetic Type-2 Classifier Optimization (MOG-T2CO), including the tuning of Type-2 fuzzy sets parameters, the selection of rules and conditions and the tuning of the inference engine. In a previous work (Hinojosa and Camargo 2018), we presented a preliminary version of the method described here, where the first step is the generation of the Data Base (DB) by defining Type-2 fuzzy sets uniformly distributed on the dataset domains, and of the Rule Base (RB), by constructing Type-2 fuzzy classification rules using an adapted method inspired on the Wang and Mendel method (Wang and Mendel 1992). The generation of the initial knowledge base (KB) is followed by a multi-objective genetic process that simultaneously tunes the Type-2 fuzzy sets parameters and selects rules and conditions from the initial RB. The purpose of the Type-2 fuzzy sets tuning is to improve the system accuracy and, that of the selection process, is to improve the system complexity, reducing the number of rules and conditions in the RB and making the system more easily understandable.

In this paper, we introduce a more elaborated version of the method presented in Hinojosa and Camargo (2018), expanding the genetic process with the tuning of the inference engine as well. The tuning of the inference engine consists of selecting the t-norm operator to be used in the inference of the particular simplified inference method adopted in this work. The chromosome coding and the genetic operators used before are expanded to include information about the t-norms that can be selected by the genetic process. This proposal aims at facilitating the optimization of T2FS providing an automatic means of defining another parameter of the systems that otherwise has to be defined manually.

The evolutionary process developed here is, as in the preliminary version of the method, based on the Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Deb et al. 2002), modeled with three objectives. The first objective is related to the accuracy of the Type-2 FRBCS, measured as the classification rate; the second and third objectives are related to the interpretability of the systems, measured by the number of rules and number of conditions in the rule base, respectively.

Although the tuning of the inference mechanism is not an original idea by itself, having already been investigated in the context of T1FS (Crockett et al. 2006; Hinojosa and Carmago 2012), it has not been addressed in any of the related works analyzed here in the context of T2FS. In the next section, a selection of similar works and their relation with this proposal are presented and discussed in more detail.

This paper is organized as follows. Section 2 summarizes the related works and highlights the relation with the proposal presented in this paper. Section 3 describes the basic concepts of FRBCS and Type-2 fuzzy sets necessary to understand the proposal. Section 4 presents the process used in this work for generating the initial RB. The method proposed here to optimize T2FS is introduced in Sect. 5. In Sect. 6, the experimental study and the results obtained are described. Finally, in Sect. 7, we point out some conclusions and future works.

## 2 Related work

In this section, we describe a summary of the representative work that shows some similarity with the study developed here and have, in some way, motivated the elaboration of this research.

In a work by Chua and Tan (2008), a method for genetically evolving Type-2 fuzzy rule-based classifiers was proposed. The results show that the performance of a Type-2 fuzzy classifier is better or at least comparable to a Type-1 fuzzy classifier. Although obtaining a successful result in an application to detect the presence of a human in a vehicle, the genetic algorithm used has only one objective and the interpretability issue was not considered.

In the work of Cai et al. (2007), a novel fuzzy neural network combining a Type-2 fuzzy logic system and a genetic algorithm based on a Takagi–Sugeno–Kang fuzzy neural network, is presented. An implementation using interval fuzzy sets was adopted, and the results were compared against a number of other traditional neuro-fuzzy classifiers, showing superior performance regarding classification rate. However, the fuzzy neural model does not allow any optimization of the interpretability aspects.

In the work presented in Sanz et al. (2013), a method called IVTURS-FARC is proposed to tune interval-valued fuzzy sets parameters and select rules, using a MOGA, after generating an initial Rule Base (RB), using the FARC-HD method (Alcalá-Fdez et al. 2011a).

Interval-valued fuzzy reasoning method (IV-FRM) with TUning and Rule Selection (IVTURS) is a linguistic fuzzy rule-based classification method based on a new completely interval-valued fuzzy reasoning method. The paper contains an extensive study of the reasoning method based on Interval-valued fuzzy sets (IVFS), and the method developed there combines, as we do in our proposal, the tuning of IVFS and rules selection. According to Alcalá-Fdez et al. (2011a), the aforementioned combination improves both the interpretability and the accuracy of the final fuzzy system. The analysis of the results has shown that IVTURS outperforms two of the high performing fuzzy classification algorithms found in the literature, FARC-HD (Alcalá-Fdez et al. 2011a) and FURIA (Hühn and Hüllermeier 2009). Despite presenting good accuracy results compared to the state-of-the-art methods, the results concerning the complexity of the system in terms of number of rules and conditions have not been reported, preventing the readers to evaluate how much the evolutionary process improves the systems interpretability. Besides, the method used to generate the initial rule base is based on the generation of all possible association rules. Although the size of the rules is limited, the process as a whole has a high computational cost.

A genetic approach to tune Type-2 membership functions, named lateral displacement and expansion/compression (LDEC), was proposed in Shukla and Tripathi (2014). In this paper, $\alpha$ and $\beta$ parameters are calculated to adjust the parameters of interval, such that $\alpha$ tuning deals with lateral displacement, whereas $\beta$ tuning carries out compression/expansion operation. Even though the authors mention that the interpretability and accuracy are taken into account, the fitness function evaluates only the classification error, indicating that the complexity of the system is not improved during the genetic process. Another drawback of this proposal is that the experiments use only one dataset.

In the work described in Lucca et al. (2015), the authors introduce a family of Choquet-based non-associative aggregation functions for application in the fuzzy reasoning method proposed by Barrenechea et al. (2013) for fuzzy rule-based classification systems. Even though the FRBCS addressed in this paper are T1FS, the study presented supports the assumption adopted in our work that changes in the parameters of the reasoning process lead the system to different results, and that special attention must be directed to the definition of these parameters.

The works presented in Sanz et al. (2013) and in Lucca et al. (2015) are the ones that address some kind of optimization regarding the inference process. However, their focus is on the aggregation of the final results of each individual rule, and not in the inference process itself. Thus, none of the papers described here perform the tuning of the inference engine by choosing the t-norm used in the inference step, as ours does. As stated in the previous discussion, the specific fuzzy reasoning method used can give different results in special in T2FS. The standard strategy is to previously select a particular operator and build the inference system with it, repeating all the experiments in case a decision to experiment another one is made. This is an advantage offered by our approach to the construction of T2FS, for yet another parameter definition can be included in the same evolutionary process without compromising the computational cost.

Considering our discussion on the related work, the main positive impacts of the method proposed here can be summarized as follows:

- Presenting a combined evolutionary process that simultaneously optimize parameters of different components of the T2FS such as the inference engine (t-norm operator), the DB (Type-2 fuzzy sets parameters) and the RB (number of rules and conditions) allowing to improve the accuracy and interpretability of the system in a balanced way;
- Using a simple and fast method to generate the initial rule base, as opposed to the ones used in similar work;
- Providing means for a more flexible and extensive search for T2FS that are good enough for each particular application, for allowing the automatic selection of the operator used in the rules inference and freeing the systems designer from manually making this selection and risking to repeat experiments when the decision is to change the operator.

## 3 Fuzzy rule-based classification systems

A FRBS has five components: fuzzification interface, inference engine, knowledge base (RB and DB) and defuzzification interface. A typical FRBS is depicted in Fig. 1.

FRBCS are FRBS learned with the specific purpose of performing the classification task. The main difference between

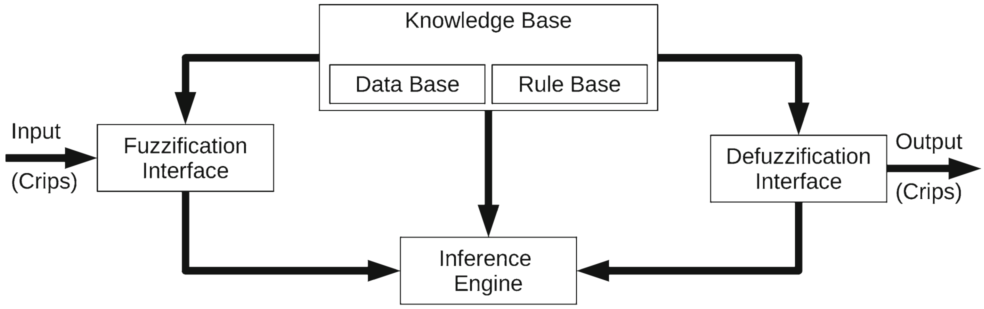


**Fig. 1** Main elements of a FRBS

a FRBS and FRBCS is that in FRBCS a defuzzification interface is not necessary. The output of the inference engine is a crisp output, which is the class of the object being classified.

In this work, we used a classic fuzzy reasoning method in the inference engine. This method classifies an example using the rule that has the highest compatibility degree with the example. To define the class $C_j \in C = \{C_1, C_2, \ldots, C_m\}$ of an example $e_q$ represented by $n$ features $e_q = \{a_{q1}, a_{q2}, \ldots, a_{qn}\}$, this method applies the following steps:

1. Calculate the compatibility degree between example $e_q$ and all fuzzy rules in the RB. Generally, the evaluation of the compatibility degree uses a t-norma.
2. Find rule $R_i \in$ RB with the highest compatibility degree with the example $e_q$.
3. Assign the class $C_j$ (consequent of class $R_i$) to example $e_q$.

A typical fuzzy classification rule in the RB can be expressed by:

$$R_i\text{: }\textbf{IF } V_1 \textbf{ IS } T_{1l_1} \textbf{ AND } V_2 \textbf{ IS } T_{2l_2} \textbf{ AND } \ldots \textbf{ AND } V_n \textbf{ IS } T_{nl_n} \textbf{ THEN } \text{Class} = C_j$$

where $R_i$, fuzzy rule with identifier $i$; $V_1, V_2, \ldots, V_n$, linguistic variables or features of the set of examples considered in the problem; $T_{1l_1}, T_{2l_2}, \ldots, T_{\mathrm{nl}_n}$], linguistic terms used to represent the feature values; $C_j$, class of rule $R_i$.

In this work, the linguistic terms are represent by Type-2 fuzzy sets. The Type-2 fuzzy sets theory was proposed by Zadeh (1975) for modeling the uncertainties inherent to the definition of the membership functions of antecedents and consequents in a fuzzy inference system. The basic idea of Type-2 fuzzy sets is that, for each specific value of $x$ in the domain of the fuzzy set, there is not a single value for the membership function, as is the case in Type-1 fuzzy sets, but instead, the membership function takes on a set of values with different membership degrees. The definition of such sets is then based on concepts of primary membership function and secondary membership function.

According to Mendel and John (2002), an interval Type-2 fuzzy set, $\tilde{A}$ on $X$, is defined by a Type-2 membership function, $0 \leq \mu_{\tilde{A}}(x, u) \leq 1$, where $x \in X$ and $J_x \subseteq [0, 1]$, i.e.,

$$\tilde{A} = \left\{((x, u), \mu_{\tilde{A}}(x, u)) | \forall x \in X, \forall u \in J_x \subseteq [0, 1]\right\}. \quad (1)$$

A Type-2 fuzzy set can be defined with different forms. Figure 2 illustrates a triangular Type-2 fuzzy set, which is the form of fuzzy sets used in this work. The uncertainty of a Type-2 fuzzy set $\tilde{A}$ is represented by a region called the Footprint of Uncertainty (FOU($\tilde{A}$)). The FOU describes the domain that supports all the secondary degrees of a Type-2 fuzzy set and allows the representation of Type-2 fuzzy sets in two dimensions instead of three. The FOU of a Type-2 fuzzy set $\tilde{A}$ is delimited by two Type-1 fuzzy sets, the lower membership function (LMF) and the upper membership function (UMF).

In Fig. 2, the value $\underline{\mu}_{\tilde{A}(x)}$ for $x$ is defined by lower membership function (LMF) for $x$ and the value $\overline{\mu}_{\tilde{A}(x)}$ for $x$ is defined by upper membership function (UMF). The uncertainty of $\tilde{A}$ is represented by the Footprint of Uncertainty (FOU($\tilde{A}$)).

## 4 Fuzzy rule learning process

Before the selection and tuning process described in Sect. 4 is applied, the initial RB, from which the rules in the final rule base will be selected, must be generated. The initial RB is generated by a fuzzy rule learning process based on an extend version of the Wang and Mendel algorithm (Wang and Mendel 1992), adapted to the use of Type-2 fuzzy sets. The fuzzy rule learning process consists of three steps: DB generation, rules generation and elimination of redundant and inconsistently rules.

First, we predefined the DB with Type-2 membership functions uniformly distributed adopting a process similar to the one used in Trk et al. (2014) and considering the minimum and maximum input values for each attribute in the dataset of

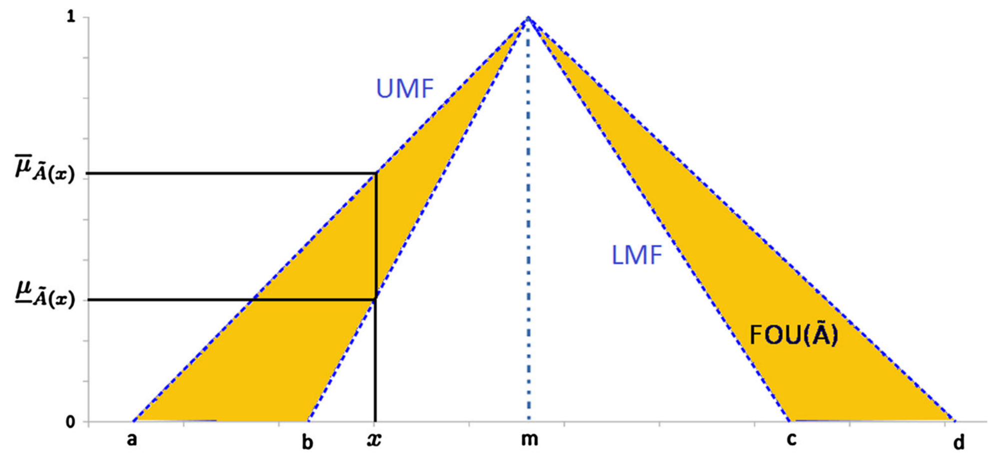


**Fig. 2** Interval Type-2 fuzzy set

examples (min and max). Figure 3 shows an example of five Type-2 fuzzy sets representing values of a linguistic variable where each one have five parameters $\{a_i, b_i, m_i c_i, d_i\}$ where $b_i - a_i = m_i - b_i = c_i - m_i = d_i - c_i$.

Second, for each example in $E = \{e_1, e_2, \ldots, e_p\}$ labeled with a class from the set of classes $C = \{C_1, C_2, \ldots, C_m\}$, where each $e_q \in E$ is defined by a set of $n$ features $e_q = \{a_{q_1}, a_{q_2}, \ldots, a_{q_n}\}$, the values of $\underline{\mu}_{\tilde{A}(a_{q_j})}$ and $\overline{\mu}_{\tilde{A}(a_{q_j})}$ are calculated for each Type-2 fuzzy set in the domain of feature $j$.

After that, the linguistic term with maximum value for $\underline{\mu}_{\tilde{A}(a_{q_j})} + \overline{\mu}_{\tilde{A}(a_{q_j})}$ is included as a condition in the fuzzy rule. This calculation is repeated for each one of the features to form a rule with $n$ conditions. For each $e_q \in E$ a fuzzy rule is generated.

Third, for each fuzzy rule in the RB with the same antecedent (conflicting and redundant) the fuzzy rule with the highest degree is selected to remain in the RB and the other ones are eliminated. The degree of each rule is calculated by applying a t-norm operator among all membership degrees in the UMF and all membership degrees in the LMF of the Type-2 fuzzy sets in the rule separately and then summing these two results.

An example of the calculation of fuzzy rules degrees and classification method is illustrated in Fig. 4 for the particular case of a RB with two fuzzy rules and three antecedents in each fuzzy rule, where $\lambda^i_{C_i}(x')$ is the rule firing of the rule $i$ with class $C_i$ for the input $x'$.

## 5 Type-2 fuzzy classifier optimization

In this section, the method we propose for Type-2 fuzzy classifier optimization using a multi-objective genetic algorithm, called MOG-T2CO, is presented. The optimization proposed includes the tuning of Type-2 fuzzy sets parameters, tuning the inference engine and selecting rules and conditions. The MOGA used in this work is the NSGA-II algorithm (Deb et al. 2002). Since NSGA-II is a very well-known evolutionary algorithm, the basic procedure is not described in details. In the following, we describe the modeling of the main components of a MOGA such as the chromosome coding, the chromosome evaluation procedure and the objectives defined.

Each chromosome ($\mathrm{CR}_i$) represents an entire knowledge base and is composed of four distinct parts, organized in a structure similar to the one proposed in Shukla and Tripathi (2014), illustrated in (2).

$$\mathrm{CR}_i = \mathrm{CR}_{M_i} + \mathrm{CR}_{R_i} + \mathrm{CR}_{\mathrm{Co}_i} + \mathrm{CR}_{\mathrm{t\_norm}}. \tag{2}$$

The first part of the chromosome $\mathrm{CR}_{M_i}$ encodes the parameters of all the Type-2 fuzzy sets for each linguistic variable in the DB, and according to the specific membership function format adopted in this paper, triangular interval Type-2 membership functions, each function is represented with five parameters, as explained in the previous section. Each gene in this part of the chromosome encodes a parameter of a linguistic value as a real value. Figure 5 illustrates the general form of a $\mathrm{CR}_{M_i}$ , considering the five parameters that represent the Type-2 fuzzy sets $(a, b, m, c, d)$ for each linguistic value of each linguistic variable, where $v$ represents the number of linguistic variables.

In order to preserve the semantics of the fuzzy sets, we define the limits for the values that the corresponding gene for each parameter can assume in the chromosome. The upper and lower limits are defined in a way that two consecutive parameters cannot swap, as noticed in Eqs. (3)–(8). When new chromosomes are generated by the genetic operators, they are checked for validity considering theses limits. The limits are illustrated in Fig. 6.

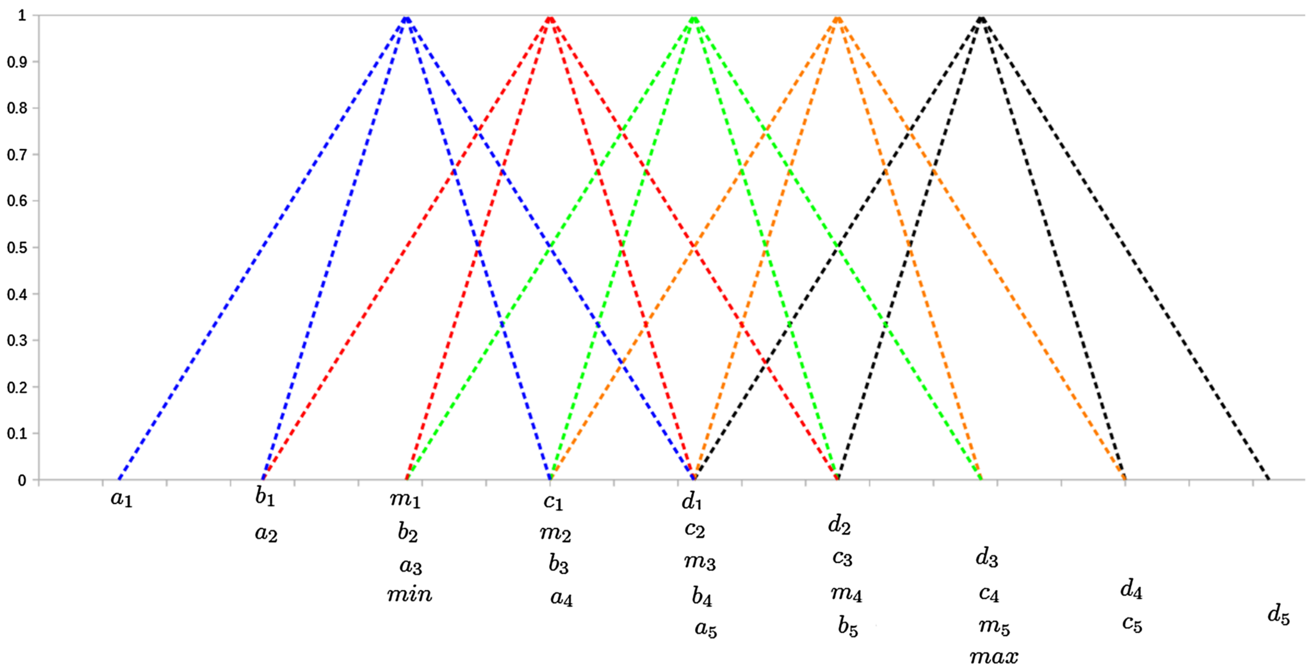


**Fig. 3** DB with Type-2 membership functions uniformly distributed

$$\begin{aligned}\text{Diff} &= (b-a)/2 \\ &= (m-b)/2 = (c-m)/2 = (d-c)/2 \end{aligned} \tag{3}$$

$$a_{\text{lower}} = a - \text{Diff};\ a_{\text{upper}} = a + \text{Diff} \tag{4}$$

$$b_{\text{lower}} = a_{\text{upper}};\ b_{\text{upper}} = b + \text{Diff} \tag{5}$$

$$m_{\text{lower}} = b_{\text{upper}};\ m_{\text{upper}} = m + \text{Diff} \tag{6}$$

$$c_{\text{lower}} = m_{\text{upper}};\ c_{\text{upper}} = c + \text{Diff} \tag{7}$$

$$d_{\text{lower}} = c_{\text{upper}};\ d_{\text{upper}} = d + \text{Diff}. \tag{8}$$

The fuzzy rules are encoded in the second part of the chromosome, $\text{CR}_R$. Note that the size of this part of the chromosome is equal to the number of rules in the rule base generated in the learning process described in Sect. 4, since each rule is coded in one gene. The upper and lower limits for each gene representing a rule are 1.0 and 0.0, respectively. A fuzzy rule is considered valid and part of the RB if its corresponding gene value is greater than 0.5. The inference engine considers only valid rule in the inference process. Figure 7 shows the general format of this part of the chromosome ($\text{CR}_R$), assuming that the initial rule base has $r$ rules.

The conditions of all rules in the RB are encoded in the third part of the chromosome, $\text{CR}_{\text{Co}}$, where each conditions is coded in one gene The conditions in the rules are also explicitly represented in the chromosome so the evolutionary process can select them, allowing the appearance of rules with less conditions. As in $\text{CR}_R$, the upper and lower limits for each gene in $\text{CR}_{\text{Co}}$ are 1.0 and 0.0, respectively, and a condition is considered valid and part of the rule if its corresponding gene value is greater than 0.5. Obviously, a condition is considered valid if, besides having a code value higher than 0.5, it belongs to a valid rule. Figure 8 illustrates the general format of the chromosome in its third part assuming that a dataset has $v$ linguistic variable and $r$ rules were generated for the initial RB. The size of this part of the chromosome is $v * r$.

Finally, the fourth part $\text{CR}_{\text{t\_norm}}$ of the chromosome encodes a t-norm to be used in the inference engine adopted here, as shown in Fig. 4. The t-norm is encoded using only one gene, and each possible t-norm is associated to an interval of values. The upper and lower limits for each gene in $\text{CR}_{\text{Co}}$ are 1.0 and 0.0, respectively. We consider, at the current stage of our research, four possible t-norms,as shown in Table 1 with their respective range of values used in the chromosome coding.

To form the initial population of the MOGA, the first chromosome encodes the KB with the Type-2 fuzzy sets DB defined by uniform distribution, all fuzzy rules and all conditions obtained in the fuzzy learning process described in Sect. 4 and the Minimum t-norm. The other chromosomes in the initial population are encoded randomly, considering the upper and lower limits of each gene.

Each chromosome is evaluated by means of the calculation of three objectives. The first objective is defined as the error rate of the FRBCS coded in the chromosome. The error rate is calculated based on the results of classification of the training examples using the single winner inference rule. The winner fuzzy rule is defined among the valid rules, that is, those rules that the fuzzy corresponding gene value is greater than or equal to 0.5. The winner inference rule is the one with higher compatibility degree $D_{R_i}$.

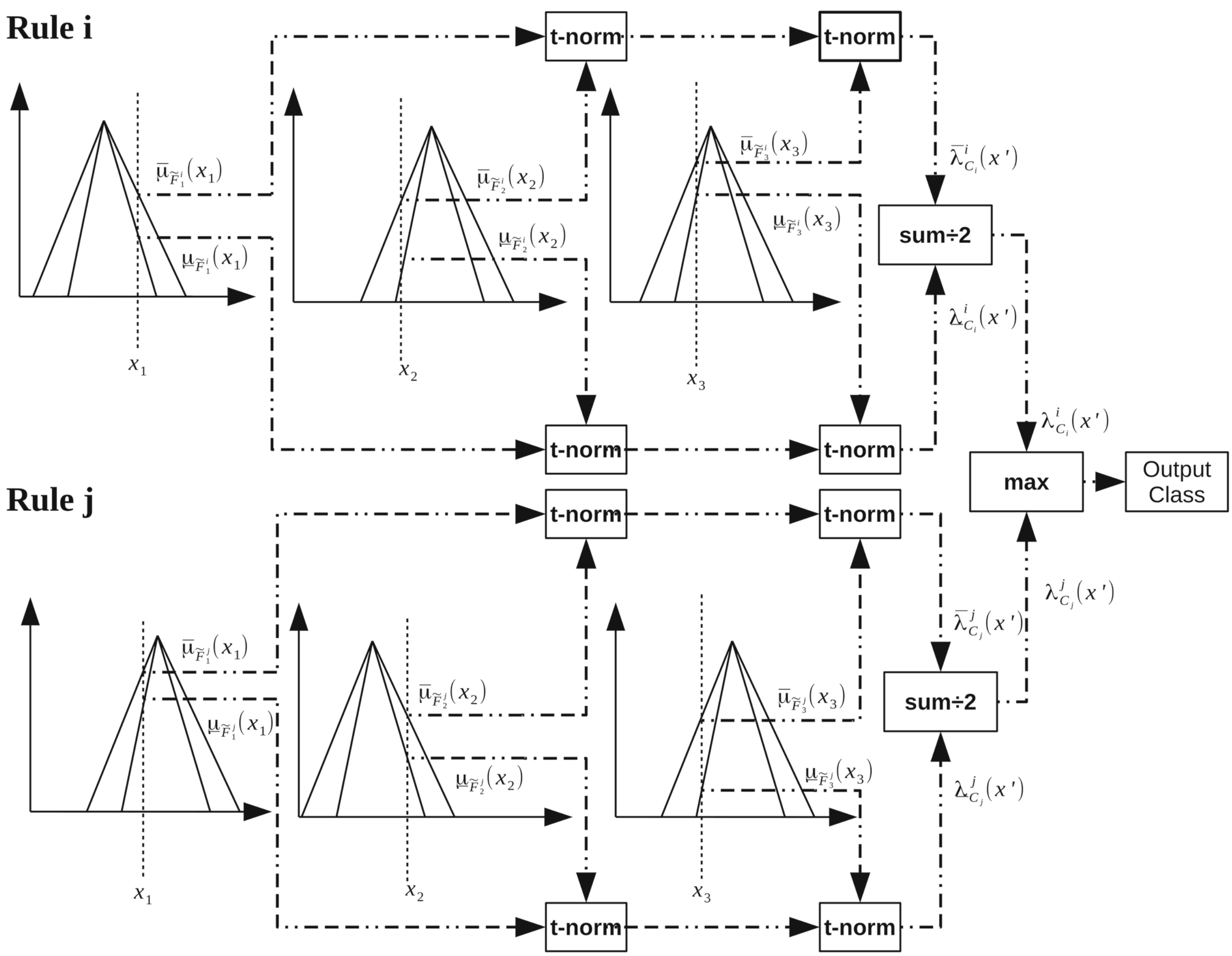


**Fig. 4** Fuzzification interface and inference engine

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $a_{1_1}$ | $b_{1_2}$ | $m_{1_3}$ | $c_{1_4}$ | $d_{1_5}$ | $a_{2_1}$ | $b_{2_2}$ | $m_{2_3}$ | $c_{2_4}$ | $d_{2_5}$ | … | $a_{v_1}$ | $b_{v_2}$ | $m_{v_3}$ | $c_{v_4}$ | $d_{v_5}$ |

**Fig. 5** Chromosome encoding for the membership functions parameters

The value of the second objective is calculated by counting genes with values greater than 0.5 in $\mathrm{CR}_M$. These genes represent the indexes of the fuzzy rules that are valid for the RB, that is, rules that are considered in the inference engine. Similar to the second objective, the third objective is calculated by counting genes with values greater than 0.5 in $\mathrm{CR}_{\mathrm{Co}}$ taking into consideration that this condition is in a fuzzy rule considered in the inference engine. These two objectives represent the total number of rules and total number of conditions in the RB, respectively, and are to be minimized, since the goal is to search for RB with low complexity.

New populations are generated using the genetic operators of selection, crossover and mutation. Tournament selection, based on the dominance of the solutions and crowding distances, is the method used for selecting a chromosome for crossover operator. Simulated binary crossover (SBX) (Deb and Kalyanmoy 2001) and polynomial mutation (Deb and Kalyanmoy 2001) are the crossover and mutation operator used in our proposal, respectively.

Based on these modeling elements, the evolutionary process will search for solutions that present a good accuracy and a good interpretability in terms of complexity.

An example of tuned Type-2 fuzzy sets on a linguistic variable domain is shown in Fig. 9.

## 6 Experiments and results

The method described in Sect. 5 was applied on ten well-known datasets extracted from KEEL repository (Alcalá-Fdez et al. 2011b), as shown in Table 2.

Table 3 shows the parameters used in NSGA-II algorithm.

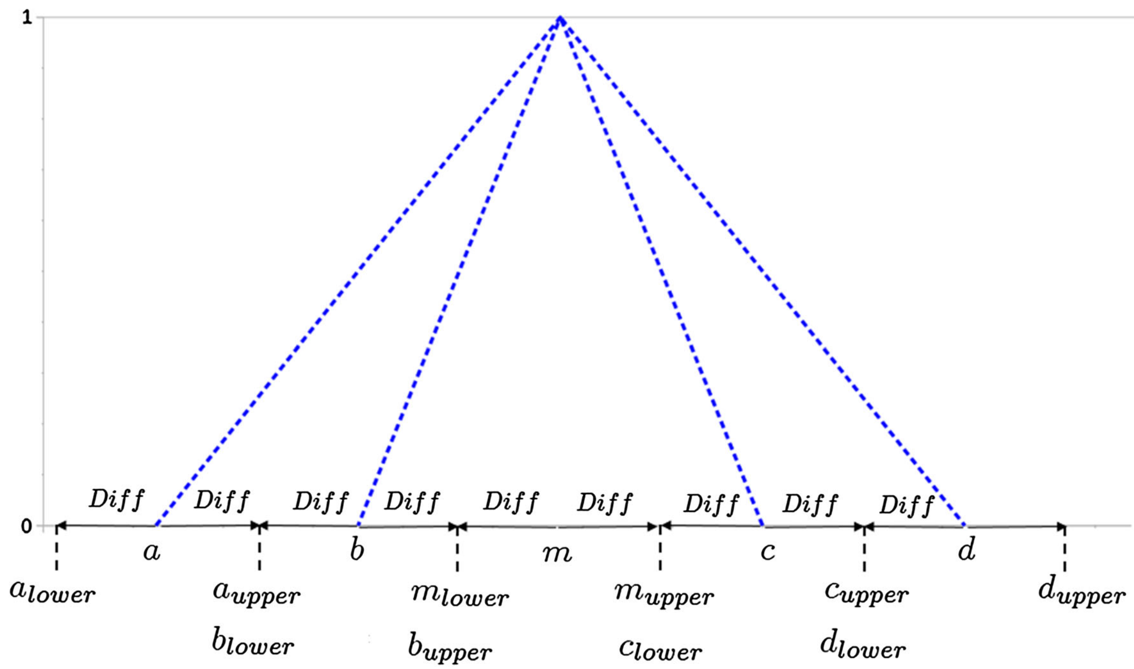


**Fig. 6** Type-2 membership functions limits

| | | | | |
|---|---|---|---|---|
| $0.0 \leq R_1 \leq 1.0$ | $0.0 \leq R_2 \leq 1.0$ | $0.0 \leq R_3 \leq 1.0$ | ... | $0.0 \leq R_r \leq 1.0$ |

**Fig. 7** Chromosome encoding for $\mathrm{CR}_R$

| | | | | | | |
|---|---|---|---|---|---|---|
| $0.0 \leq Co_{1_1} \leq 1.0$ | $0.0 \leq Co_{2_1} \leq 1.0$ | ... | $0.0 \leq Co_{v_1} \leq 1.0$ | ... | $0.0 \leq Co_{v_r} \leq 1.0$ |

**Fig. 8** Chromosome encoding for $\mathrm{CR}_{\mathrm{Co}}$

All experiments were run using fivefold cross-validation. Results after learning process and multi-objective inference engine tuning are shown in Table 4. Second and third columns show the mean error rate in training ($\mathrm{Tr}_{\mathrm{er}}$) and test ($\mathrm{Te}_{\mathrm{er}}$) dataset, respectively. Fourth and fifth columns show the quantity of fuzzy rules and conditions in the RB, respectively. Standard deviation is shown next to each result.

In order to obtain a more solid conclusion on the quality of these results, we compared them with the ones described in Lucca et al. (2015), which were generated by means of a method that uses a family of Choquet-based non-associative aggregation functions in the fuzzy inference reasoning. The method described in Lucca et al. (2015) was selected to be compared with our approach because it shares the general objective of our research, which is to improve the performance of fuzzy rule-based classification systems, searching for the best operator to be used in the inference process, among a set of candidates. In that work, the best operator found for each dataset was used in the comparison analysis. Following the same strategy, we also selected the highest classification rate for each dataset to be used in comparison with our proposal. The accuracy and standard deviation values for both methods are shown in Table 5, for 10 datasets. In both methods, the values appearing in the table were obtained from the average of each run of a fivefold cross-validation testing strategy. These results demonstrate that the method proposed in our work can offer a performance at least as good as a state-of-the-art method.

The work in Lucca et al. (2015) did not focused on the balance between accuracy and interpretability and did not

**Table 1** t_norms

| No. | t_norm | Equation | Range |
|---|---|---|---|
| 1 | Minimum | $\mathrm{Min}(x, y) = \min(x, y)$ | $0.00 \geq \mathrm{CR}_{\mathrm{t\_norm}} \leq 0.25$ |
| 2 | Algebraic product | $\mathrm{AP}(x, y) = x \cdot y$ | $0.25 > \mathrm{CR}_{\mathrm{t\_norm}} \leq 0.50$ |
| 3 | Einstein product | $\mathrm{EP}(x, y) = \frac{x \cdot y}{2-[x+y-x \cdot y]}$ | $0.50 > \mathrm{CR}_{\mathrm{t\_norm}} \leq 0.75$ |
| 4 | Hamacher product | $\mathrm{HP}(x, y) = \frac{x \cdot y}{x+y-x \cdot y}$ | $0.75 > \mathrm{CR}_{\mathrm{t\_norm}} \leq 1.00$ |

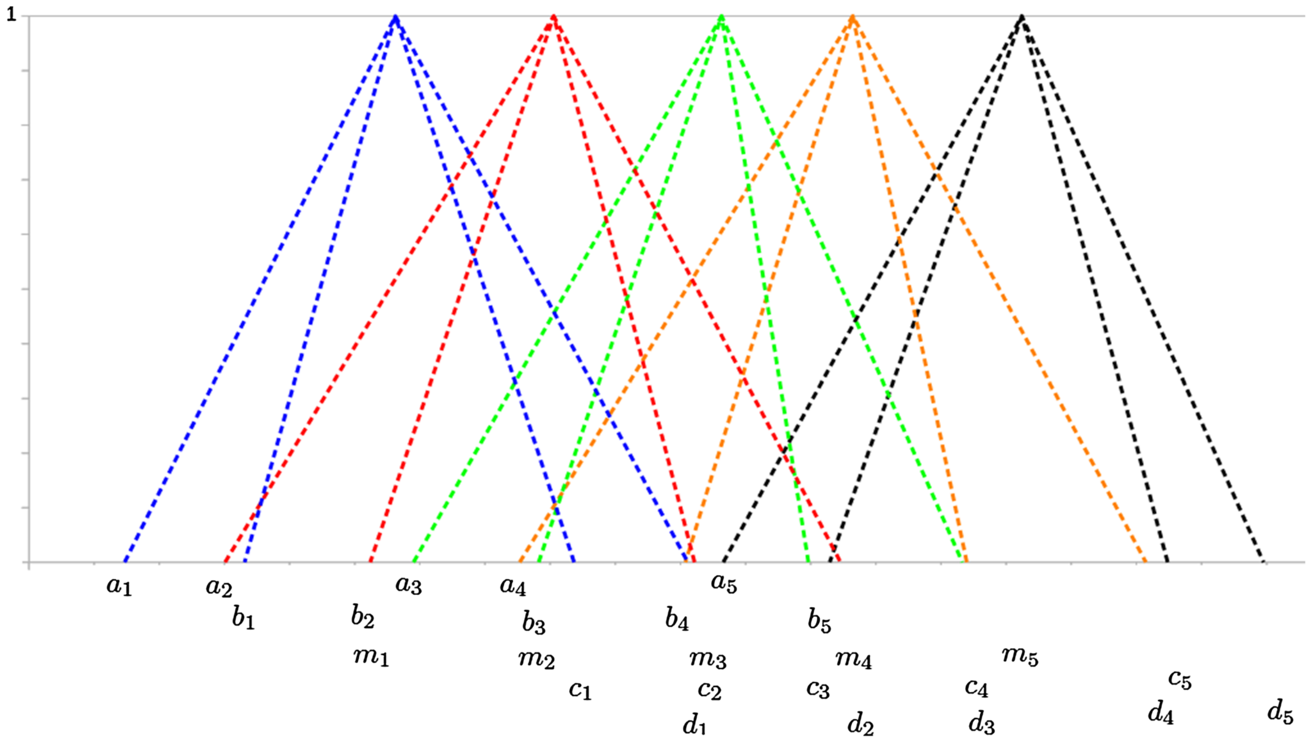


**Fig. 9** Tuned Type-2 fuzzy sets in a linguistic variable

**Table 2** Datasets used in the experiments

| No. | Dataset | #Attributes | #Examples | #Classes |
|---|---|---|---|---|
| 1 | balance | 4 | 625 | 3 |
| 2 | banana | 2 | 5300 | 2 |
| 3 | ecoli | 7 | 336 | 8 |
| 4 | glass | 9 | 214 | 7 |
| 5 | iris | 4 | 150 | 3 |
| 6 | led7digit | 7 | 500 | 10 |
| 7 | newthyroid | 5 | 215 | 3 |
| 8 | pima | 8 | 768 | 2 |
| 9 | wine | 13 | 178 | 3 |
| 10 | wisconsin | 9 | 683 | 3 |

**Table 3** Parameters of NSGA-II algorithm

| Parameter | Value |
|---|---|
| Size of the population | 100.0 |
| Crossover probability | 1.0 |
| Mutation probability | 0.1 |
| Number of generations | 2000.0 |

present results on the number of rules and conditions of the generated models. For this reason, the comparison analysis done in our work includes only the accuracy rates.

Finally, we carried out a Wilcoxon signed-rank test (Wilcoxon 1945) to compare both methods. Table 6 demonstrates that the results obtained using MOG-T2CO are comparable to the ones obtained by the Choquet method, since neither one is significantly better than the other (considering a $\rho < 0.05$).

## 7 Conclusions

This work presents a multi-objective genetic approach to tune T2FS, which includes selecting the t-norm used in the inference, tuning the Type-2 fuzzy sets and selecting rules and conditions from an initial rule base. The main contribution of this research is to provide a method capable of handling the complexity involved in the generation and optimization of the T2FS for classification, using a dataset. The method proposed here optimizes, in a coherently combined way, different components of the T2FS simultaneously: the inference engine, the data base and the rule base. It also takes advantage of the combined optimization to improve both the systems accuracy and interpretability. By tuning the Type-2 fuzzy sets parameters and the t-norm used in the inference process, the accuracy of the system is improved. By selecting rules and conditions, the complexity of the system also improves. This way, at the end of the evolutionary process, it is possible to obtain a set of systems with a good balance between accu-

**Table 4** Results obtained by MOG-T2CO

| Dataset | $Tr_{er}$ | $Te_{er}$ | #Rules | #Conditions |
|---|---|---|---|---|
| balance | $0.0000 \pm 0.0000$ | $0.1760 \pm 0.0316$ | $500.00 \pm 0.00$ | $2000.00 \pm 0.00$ |
| banana | $0.2860 \pm 0.0460$ | $0.2904 \pm 0.0424$ | $6.60 \pm 2.42$ | $8.40 \pm 4.63$ |
| ecoli | $0.1362 \pm 0.0264$ | $0.2321 \pm 0.0069$ | $129.60 \pm 14.92$ | $804.40 \pm 113.59$ |
| glass | $0.3504 \pm 0.0887$ | $0.3691 \pm 0.0159$ | $30.40 \pm 15.40$ | $156.20 \pm 113.98$ |
| iris | $0.0350 \pm 0.0111$ | $0.0467 \pm 0.0267$ | $4.00 \pm 1.26$ | $5.40 \pm 1.62$ |
| led7digit | $0.3945 \pm 0.0481$ | $0.4180 \pm 0.0694$ | $47.80 \pm 8.89$ | $251.00 \pm 63.47$ |
| newthyroid | $0.0279 \pm 0.0170$ | $0.0233 \pm 0.0147$ | $26.80 \pm 6.79$ | $97.80 \pm 30.22$ |
| pima | $0.2821 \pm 0.0134$ | $0.2649 \pm 0.0199$ | $230.20 \pm 23.45$ | $1122.80 \pm 238.97$ |
| wine | $0.0366 \pm 0.0150$ | $0.0500 \pm 0.0111$ | $53.20 \pm 8.57$ | $372.40 \pm 84.14$ |
| wisconsin | $0.0190 \pm 0.0099$ | $0.0365 \pm 0.0206$ | $145.20 \pm 27.82$ | $808.60 \pm 271.93$ |

**Table 5** Comparison of MOG-T2CO and Choquet method

| Dataset | MOG-T2CO | Choquet method |
|---|---|---|
| balance | $0.1760 \pm 0.0316$ | $0.1008 \pm 0.0091$ |
| banana | $0.2904 \pm 0.0424$ | $0.3734 \pm 0.0092$ |
| ecoli | $0.2321 \pm 0.0069$ | $0.2500 \pm 0.0924$ |
| glass | $0.3691 \pm 0.0159$ | $0.4047 \pm 0.0279$ |
| iris | $0.0467 \pm 0.0267$ | $0.0533 \pm 0.0298$ |
| led7digit | $0.4180 \pm 0.0694$ | $0.3580 \pm 0.0563$ |
| newthyroid | $0.0233 \pm 0.0147$ | $0.0651 \pm 0.0400$ |
| pima | $0.2649 \pm 0.0199$ | $0.2455 \pm 0.0161$ |
| wine | $0.0500 \pm 0.0111$ | $0.0611 \pm 0.0444$ |
| wisconsin | $0.0365 \pm 0.0206$ | $0.0409 \pm 0.0099$ |

**Table 6** Results of Wilcoxon signed-rank test between MOG-T2CO and Choquet method

| | Value |
|---|---|
| $N$ | 10 |
| Estimated median | $-0.0081$ |
| $\rho$ | 0.6100 |

racy and interpretability, with respect to complexity. As an additional advantage, the semantics of fuzzy sets are also preserved, for the tuning process limits the value that can be generated in the corresponding part of the chromosome.

The most evident practical implication of the study is that the system designer can avoid using the standard procedure of manually defining the t-norm operator to be used in the inference process. On the contrary, this selection can be left to the optimization process and different t-norms can easily be tested.

An analysis of the current status of the method allows us to identify a few limitations that deserve to be investigated further. Among these, we can say that the definition of the initial fuzzy sets and fuzzy rules remains an issue. Even though experiments show that the tuning of T2FS allows delivering better results in comparison with the non-tuned version, tuning is limited by the initial definition of the fuzzy set. In the future, different informative methods can be used in the definition of Type-2 fuzzy sets. Several design decisions have been made to make this research possible. For instance, only a small set of t-norms was considered in the tuning process. Running experiments with a larger number of t-norms can open unknown possibilities of analysis regarding the systems behavior.

Finally, it is important to notice that results are sensitive to the particular MOGA used in the evolutionary tuning. As part of the future work, we also intend to expand the experiments with the use of others MOGA, specially those that are based on different elements with respect to the one used in this research, such as MOEA/D (Martinez and Coello 2014) or SPEA2 (Zitzler et al. 2001).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

Alcalá-Fdez J, Alcala R, Herrera F (2011) A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning. IEEE Trans Fuzzy Syst 19(5):857–872. https://doi.org/10.1109/TFUZZ.2011.2147794

Alcalá-Fdez J, Fernández A, Luengo J, Derrac J, García S, Sánchez L, Herrera F (2011) Keel data-mining software tool: data set repository, integration of algorithms and experimental analysis framework. J Mult-Valued Log Soft Comput 17(2–3):255–287

Barrenechea E, Bustince H, Fernandez J, Paternain D, Sanz JA (2013) Using the Choquet integral in the fuzzy reasoning method of fuzzy rule-based classification systems. Axioms 2(2):208–223. https://doi.org/10.3390/axioms2020208

Cai A, Quek C, Maskell DL (2007) Type-2 GA-TSK fuzzy neural network. In: 2007 IEEE Congress on evolutionary computation, pp 1578–1585. https://doi.org/10.1109/CEC.2007.4424661

Castillo O, Melin P (2008) Type-2 fuzzy logic: theory and applications. Studies in fuzziness and soft computing. Springer, Berlin

Castillo O, Melin P (2012) Optimization of type-2 fuzzy systems based on bio-inspired methods: a concise review. Inf Sci 205:1–19. https://doi.org/10.1016/j.ins.2012.04.003

Chua TW, Tan WW (2008) Genetically evolved fuzzy rule-based classifiers and application to automotive classification. In: Simulated evolution and learning. Springer, Berlin, pp 101–110

Cordón O (2011) A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: designing interpretable genetic fuzzy systems. Int J Approx Reason 52(6):894–913

Cordón O, Herrera F, Hoffmann F, Magdalena L (2001) Genetic fuzzy systems: evolutionary tuning and learning of fuzzy knowledge bases. In: Advances in fuzzy systems—applications and theory, vol 19. World Scientific Publishing Co. Pte. Ltd, Singapore

Crockett KA, Bandar Z, Fowdar J, O'Shea J (2006) Genetic tuning of fuzzy inference within fuzzy classifier systems. Expert Syst 23(2):63–82. https://doi.org/10.1111/j.1468-0394.2006.00325.x

Deb K, Kalyanmoy D (2001) Multi-objective optimization using evolutionary algorithms. Wiley, New York

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6(2):182–197

Fazzolari M, Alcala R, Nojima Y, Ishibuchi H, Herrera F (2013) A review of the application of multiobjective evolutionary fuzzy systems: current status and further directions. IEEE Trans Fuzzy Syst 21(1):45–65

Herrera F (2008) Genetic fuzzy systems: taxonomy, current research trends and prospects. Evol Intell 1(1):27–46

Hinojosa E, Carmago HA (2012) Multiobjective genetic optimization of fuzzy partitions and t-norm parameters in fuzzy classifiers. In: 2012 Brazilian symposium on neural networks, pp 154–159. https://doi.org/10.1109/SBRN.2012.45

Hinojosa CE, Camargo HA (2018) A multi-objective evolutionary algorithm for tuning Type-2 fuzzy sets with rule and condition selection on fuzzy rule-based classification system. Springer, Berlin pp 389–399

Hühn J, Hüllermeier E (2009) Furia: an algorithm for unordered fuzzy rule induction. Data Min Knowl Discov 19(3):293–319. https://doi.org/10.1007/s10618-009-0131-8

Karnik NN, Mendel JM (1998) Introduction to type-2 fuzzy logic systems. In: 1998 IEEE international conference on fuzzy systems proceedings. In: IEEE World Congress on Computational Intelligence (Cat. No.98CH36228), vol 2, pp 915–920

Karnik NN, Mendel JM, Liang Q (1999) Type-2 fuzzy logic systems. IEEE Trans Fuzzy Syst 7(6):643–658. https://doi.org/10.1109/91.811231

Liang Q, Mendel JM (2000) Interval type-2 fuzzy logic systems: theory and design. IEEE Trans Fuzzy Syst 8(5):535–550. https://doi.org/10.1109/91.873577

Lucca G, Dimuro GP, Mattos V, Bedregal B, Bustince H, Sanz JA (2015) A family of Choquet-based non-associative aggregation functions for application in fuzzy rule-based classification systems. In: 2015 IEEE international conference on fuzzy systems (FUZZ-IEEE), pp 1–8

Martinez SZ, Coello CAC (2014) A multi-objective evolutionary algorithm based on decomposition for constrained multi-objective optimization. In: 2014 IEEE Congress on evolutionary computation (CEC), pp 429–436

Mendel JM, John RIB (2002) Type-2 fuzzy sets made simple. IEEE Trans Fuzzy Syst 10(2):117–127. https://doi.org/10.1109/91.995115

Sanz JA, Fernández A, Bustince H, Herrera F (2013) IVTURS: a linguistic fuzzy rule-based classification system based on a new interval-valued fuzzy reasoning method with tuning and rule selection. IEEE Trans Fuzzy Syst 21(3):399–411. https://doi.org/10.1109/TFUZZ.2013.2243153

Shukla PK, Tripathi SP (2014) A new approach for tuning interval type-2 fuzzy knowledge bases using genetic algorithms. J Uncertain Anal Appl 2(1):4

Trk S, John R, Özcan E (2014) Interval type-2 fuzzy sets in supplier selection. In: 2014 14th UK workshop on computational intelligence (UKCI), pp 1–7

Wang LX, Mendel JM (1992) Generating fuzzy rules by learning from examples. IEEE Trans Syst Man Cybern 22(6):1414–1427

Wilcoxon F (1945) Individual comparisons by ranking methods. Biom Bull 1(6):80–83. https://doi.org/10.2307/3001968

Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning. J Inf Sci 8:199

Zitzler E, Laumanns M, Thiele L (2001) SPEA2: improving the strength Pareto evolutionary algorithm, Technical report