Information Sciences 181 (2011) 5194–5213

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

A formal model for mining fuzzy rules using the *RL* representation theory

M. Delgado [a], M.D. Ruiz [a,*], D. Sánchez [b], J.M. Serrano [c]

[a] Dept. of Computer Science and A.I., University of Granada, Spain
[b] European Centre for Soft Computing, Mieres, Spain
[c] Dept. of Computer Science, University of Jaén, Spain

ARTICLE INFO

Article history:
Received 19 October 2010
Received in revised form 31 January 2011
Accepted 16 July 2011
Available online 27 July 2011

Keywords:
Data mining
Fuzzy association rules
Restriction levels
GUHA model
Absence of items
Negative rules

ABSTRACT

Data mining techniques managing imprecision are very useful to obtain meaningful and interesting information for the user. Among some other techniques, fuzzy association rules have been developed as a powerful tool for dealing with imprecision in databases and offering a good representation of found knowledge. In this paper we introduce a formal model for managing the imprecision in fuzzy transactional databases using the restriction level representation theory, a recent representation of imprecision that extends that of fuzzy sets. This theory introduces some new operators, keeping the usual crisp properties even when negation is involved.

The model allows us to mine fuzzy association rules in a straightforward way, extending the accuracy measures from the crisp case. In addition, we introduce several ways of representing and summarizing the obtained results, in order to offer new and very interesting semantics. As an application, we present how to extract fuzzy association rules involving both the presence and the absence of items using the proposed model, and we also perform some experiments with real fuzzy transactional datasets.

1. Introduction

As a tool for knowledge discovery and representation in ambiguous, ill-defined scenarios, fuzzy association rules have been developed and applied in numerous situations since their appearance in the 90s. First works used the theory of fuzzy subsets proposed by Zadeh [45], summarizing imprecise values with a clear semantic into groups and, later, using them for representing the different types of imprecision found in a stored database. Some advantages of the use of fuzzy sets are, on the one hand, softening bounds and, on the other hand, giving a formal representation for the most semantically significant and meaningful knowledge for the user. In [20,21], we can find a description of some of the most important works in the field of fuzzy association rules.

This work is intended to set a logical basis for the representation and evaluation of fuzzy association rules. Its foundation comes from the GUHA (General Unary Hypothesis Automaton) model [17] presented in [26] for association rules and also analyzed and complemented in [12]. It is based on two basic notions: a four fold contingency table (4*ft*-table for short) representing the frequencies of appearance of the itemsets involved in the rule, and the 4*ft*-quantifier which represents the measure used in the evaluation of association rules. This formalization using the GUHA model offers a good framework to study the properties of the validation measures and sets a starting point for developing new approaches in association rule

* Corresponding author.
E-mail addresses: mdelgado@ugr.es (M. Delgado), mdruiz@decsai.ugr.es (M.D. Ruiz), daniel.sanchezf@softcomputing.es (D. Sánchez), jschica@ujaen.es (J.M. Serrano).

0020-0255/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2011.07.029

mining. In particular we are interested in its development for fuzzy rules keeping the main properties of crisp rules. This has been carried out using the restriction level representation (RLR) theory [30].

The RLR theory is suitable when we have to operate with vague concepts and we need, or want to keep, the ordinary boolean properties. It is also useful when translating crisp procedures and for fuzzy properties formulation. In our case, we are interested in preserving the properties of the crisp measures used for the rule evaluation when extending them for mining fuzzy association rules. Specially, we use the RLR theory for generalizing the GUHA model for association rules to incorporate the case of fuzzy rules. The GUHA model gives us the suitable formalization for extending every interestingness measure expressed in terms of the frequencies of appearance of the itemsets involved in the rule through the RLR theory. A first proposal was presented in [10]. We showed how to use the GUHA model for representing information about the frequencies when mining fuzzy rules. This paper is intended to be an update to that work. We present new approaches for validating the results and we develop an algorithm for mining fuzzy rules based on the model philosophy and on the RLR theory obtaining very promising results.

In addition, some recent works in association rule mining cover the problem of discovering associations which involve not only the joint occurrence of items but also the absence of them [7,32,33,44]. The main problem is that the upward closure property of frequent itemsets is not satisfied when the absence is considered. The usual apriori-based algorithms cannot be easily adapted to mine rules containing absent items. Nevertheless, some approaches have been developed on crisp association rule mining [42,43,38], proposing new methods by imposing some specific requirements. And, as far as we know, there exist very few approaches for extending the problem of considering absent items when mining fuzzy rules [18,19]. In this work, we present how to deal with this special type of fuzzy rules by using our developed model. We also apply the proposed algorithm for mining fuzzy rules with absent items over both real and synthetic fuzzy databases.

The paper is organized as follows: Section 2 reviews the concepts necessary for the comprehension of this paper: concepts about crisp and fuzzy association rules, a brief summary of the GUHA model for the representation of crisp association rules and the approach for representing imprecise properties by means of restriction levels introduced in [30]. In Section 3 we develop our proposal for representing and evaluating fuzzy association rules using a combination of the GUHA model and the RLR theory. Section 4 shows how the model can be applied when mining fuzzy rules with absent items. In Section 5 an algorithm for mining fuzzy rules following the model philosophy is described and analysed. Section 6 shows the experimental results using our algorithm for mining fuzzy rules involving both the presence and the absence of items in some real and synthetic databases. We finish with some conclusions and possible lines for future research.

2. Background concepts

2.1. Association rules

Given a set I ("set of items") and a database D constituted by set of transactions, each one being a subset of I, association rules [1] are "implications" of the form $A \to B$ that relate the presence of itemsets (sets of items) A and B in transactions of D, assuming $A,B \subseteq I$, $A \cap B = \emptyset$ and $A,B \neq \emptyset$.

The ordinary measures proposed in [1] to assess association rules are *confidence* (the conditional probability $P(B|A)$) and *support* (the joint probability $P(A \cup B)$). An alternative framework was proposed in [5,6] where the accuracy is measured by means of Shortliffe and Buchanan's certainty factors [34], in the following way:

Definition 1 [9]. Let $\text{supp}(B)$ be the support of the itemset B, and let $\text{Conf}(A \to B)$ be the confidence of the rule. The *certainty factor* of the rule, denoted as $CF(A \to B)$, is defined as

$$CF(A \to B) = \begin{cases} \frac{\text{Conf}(A\to B)-\text{supp}(B)}{1-\text{supp}(B)} & \text{Conf}(A \to B) > \text{supp}(B), \\ \frac{\text{Conf}(A\to B)-\text{supp}(B)}{\text{supp}(B)} & \text{Conf}(A \to B) \leqslant \text{supp}(B). \end{cases} \quad (1)$$

The certainty factor yields a value in the interval $[-1,1]$ and measures how our belief that B is in a transaction changes when we are told that A is in that transaction. Positive values indicate that our belief increases, negative values mean that our belief decreases, and 0 means no change. Certainty factor has better properties than confidence, and helps to solve some of its drawbacks [6,9]. In particular, it helps to reduce the number of rules obtained by filtering those rules corresponding to statistical independence or negative dependence.

Some works [7,32,33,44] also deal with association rules where the absence of items in the transactions is taken into account. The formalization of this idea have been made by means of a negation operator which represents the complement of the occurrence of an item in a transactional database, that is, its absence. Formally, in these situations the set of items I has both positive $(i_1,\ldots,i_m)$ and negative $(\neg i_1,\ldots,\neg i_m)$ items, where $\neg i_k$ means that i_k is not present in a transaction. Therefore, a *negative association rule* is a rule that contains a negative item (i.e. a rule for which either its antecedent and/or its consequent can contain a conjunction of both present and absent items).

This idea must not be misled with other approaches where negative associations are mined by assessing their accuracy with a particular interestingness measure. That is the case of Au and Chan [4], who define a negative association if the adjusted difference is lower than a fixed value. Tsumoto presents in [40] another type of negative association. A negative rule is the contrapositive of a rule that is supported by all the positive examples in the database. Excluding these works, which use the value of the interestingness measure, a negative association rule will be that incorporating the negation of one or several items, where the negation represents the absence of the item in the rule.

2.2. Fuzzy association rules

In [9], the model for association rules is extended in order to manage fuzzy values in databases. The approach is based on the definition of fuzzy transactions as fuzzy subsets of items.

Definition 2 [9]. Let $I = \{i_1, \ldots, i_m\}$ be a finite set of items. A fuzzy transaction is a non empty fuzzy subset $\tilde{\tau} \subseteq I$.

For every item $i \in I$ and every transaction $\tilde{\tau}$, an item i will belong to $\tilde{\tau}$ with grade[1] $\tilde{\tau}(i)$ where $\tilde{\tau}(i)$ is a real number in the interval [0,1]. Let $A \subseteq I$ be an itemset. The membership grade of A to the fuzzy transaction $\tilde{\tau}$ is defined as $\tilde{\tau}(A) = \min_{i \in A} \tilde{\tau}(i)$.

According Definition 2 a crisp transaction is a special case of fuzzy transaction where every item in the transaction has membership grade equal to 1 or 0 depending on if they are in the transaction or not.

Example 1. Consider the set of items $I = \{i_1, i_2, i_3, i_4, i_5\}$ and the set of fuzzy transactions given by Table 1. In particular, we can see that $\tilde{\tau}_6$ is a crisp transaction. Some membership grades could be: $\tilde{\tau}_1(\{i_3, i_4\}) = 0.9$, $\tilde{\tau}_1(\{i_2, i_3, i_4\}) = 0.2$ and $\tilde{\tau}_2(\{i_1, i_2\}) = 1$.

Definition 3 [9]. Let I be a set of items, $\widetilde{D}$ a set of fuzzy transactions and $A, B \subset I$ two disjoint itemsets, i.e. $A \cap B = \emptyset$. A fuzzy association rule $A \rightarrow B$ is completely satisfied in $\widetilde{D}$ if and only if, $\tilde{\tau}(A) \leqslant \tilde{\tau}(B)$ for all $\tilde{\tau} \in \widetilde{D}$, that is, the membership grade of B is higher than the membership grade of A for all fuzzy transactions $\tilde{\tau}$ in $\widetilde{D}$.

This definition holds the meaning of crisp association rules because if we need $A \subseteq \tilde{\tau}$ to be satisfied, we also need to satisfy $B \subseteq \tilde{\tau}$. In our case this can be translated to $\tilde{\tau}(A) \leqslant \tilde{\tau}(B)$. In this way, since a crisp transaction is a special case of fuzzy transaction, a crisp association rule will be a special case of fuzzy association rule.

2.3. Formal model for crisp association rules

The logic model we are going to use is based in a method developed in the sixties by Hájek et al. [17]. This method is named GUHA (General Unary Hypotheses Automaton) and it features a good logic and statistical base which contributes to a better understanding of two important aspects of association rules: their nature and the properties of the interest measures used for their evaluation. In the following we will refer to this model as GUHA model although the used notation is not the same as that developed by the authors nor that found in posterior works [26,28].

The starting point is a binary database D where rows and columns represent transactions and items respectively. As a particular case, the items could be pairs of the form $\langle attribute, value \rangle$ or $\langle attribute, interval \rangle$, and the value of a transaction t_k in the position j will be 1 if the item i_j is satisfied in that transaction or 0 otherwise.

For this model, an *itemset* will be an aggregation of items using the usual logic connectors: $\wedge, \vee, \neg$. An association rule will be an expression of the type $\varphi \approx \psi$ where φ and ψ represent itemsets derived from D, and the symbol $\approx$ called 4*ft-quantifier* is an evaluation or condition for the fulfilment of the association rule which will depend on the interest measure used and on the *four fold table*, 4*ft-table* for short, associated to the itemsets φ and ψ. An example of association rule could be $i_1 \wedge i_3 \approx i_2 \wedge \neg i_4$.

For any pair of itemsets φ and ψ the so called 4*ft-table* may be constructed from the database D as follows:

$\mathcal{M}$	ψ	$\neg\psi$
φ	a	b
$\neg\varphi$	c	d

This table will be noted by $\mathcal{M} = 4ft(\varphi, \psi, D) = \langle a, b, c, d \rangle$ where a, b, c and d will be non-negative integers satisfying that a is the number of objects (i.e. the number of rows of D) which contain at the same time the itemsets φ and ψ, b the number of objects satisfying φ and not ψ, and analogously for c and d. It is obvious that the inequality $a + b + c + d > 0$ is always satisfied.

When the fulfilment of the condition imposed by a quantifier $\approx$ comes from a four fold table, we will say that $\approx$ is a 4*ft*-quantifier. The association rule $\varphi \approx \psi$ will be true in the database D (or in the matrix $\mathcal{M}$) if and only if the condition(s) associated to the 4*ft*-quantifier $\approx$ are satisfied for the four fold table $4ft(\varphi, \psi, D)$.

Depending on the type of 4*ft*-quantifier we can express different kinds of associations between the itemsets φ and ψ. In [26] we can find some examples. The classical framework of support and confidence can be modeled by means of the *implication* 4*ft*-quantifier, $\Rightarrow_I$, as follows:

$$\Rightarrow_I(a, b, c, d) = \frac{a}{a+b} \quad (2)$$

[1] For sake of simplicity we note $\tilde{\tau}(i)$ as $\mu_{\tilde{\tau}}(i)$ where $\mu_{\tilde{\tau}} : I \rightarrow [0, 1]$ is the membership function associated to the fuzzy set $\tilde{\tau}$ on the referential I = {set of items}.

Table 1
Set of fuzzy transactions $\widetilde{D}_1$.

	i_1	i_2	i_3	i_4	i_5
$\tilde{\tau}_1$	1	0.2	1	0.9	0.9
$\tilde{\tau}_2$	1	1	0.8	0	0
$\tilde{\tau}_3$	0.5	0.1	0.7	0.6	0
$\tilde{\tau}_4$	0.6	0	0	0.5	0.5
$\tilde{\tau}_5$	0.4	0.1	0.6	0	0
$\tilde{\tau}_6$	0	1	0	0	0

imposing the following conditions:

$$\Rightarrow_I(a,b,c,d) \geqslant minconf \quad \text{and} \quad \frac{a}{n} \geqslant minsupp, \tag{3}$$

where $0 < minconf, minsupp < 1$ are the thresholds for the minimum confidence and minimum support respectively, and $n = a + b + c + d$ is the total number of transactions in the database D. By means of these impositions we collect two important facts from the association rule extraction. The first one is the computation of the 4*ft*-quantifier value (calculated from the 4*ft*-table) which measures the strength of the association, and the second one is the fulfilment of the previous conditions (imposed by the thresholds) that transforms the association rule $\varphi \approx \psi$ in a logic predicate that can be true or false depending on the satisfiability of the inequalities given by the thresholds. So, an association rule will be satisfied in D, noted by $Val(\varphi \approx \psi) = true$ if and only if, the conditions associated to quantifier $\approx$ are satisfied for the corresponding 4*ft*-table: $\mathcal{M} = 4ft(\varphi, \psi, D)$.

We have recently analyzed some of its important properties in the ambit of association rule extraction [12], emphasizing the following:

- It unifies the representation and the evaluation of association rules by means of two concepts: 4*ft*-table and 4*ft*-quantifiers[2] where 4*ft* stands for four fold table. The 4*ft*-table collects the information necessary to manage the association rule, and the 4*ft*-quantifier will represent the measurement for assessing the association rule.
- By analyzing the type of 4*ft*-quantifier which is defined in terms of the 4*ft*-table, we can study the properties of the association we will obtain using that 4*ft*-quantifier. This is deeply developed in [12] where we establish the existent relation between the principles for a good interestingness measure and the type of 4*ft*-quantifier.
- The model can be generalized to take into account several types of associations, not only association rules [28]. In particular, we have generalized the model to consider sets of rules with a significant meaning for the user like exception and anomalous rules [11].

In [12] we also explain in depth the existing relation between the 4*ft*-quantifiers and the interestingness measures used in the evaluation and validation of association rules. In fact, the certainty factor (Definition 1) can be seen as a 4*ft*-quantifier which fulfils all the principles for a good interestingness measure [12].

2.4. Representation by restriction levels

The basic idea of the RLR theory [30] is that vague properties defined on a set of objects X can be described by a collection of crisp representatives each one being a crisp realization under a certain restriction. The so called restriction levels (RL) are represented by values in the unit interval meaning possible levels of relaxation of the property where 1 corresponds to the most restrictive, 0 means no restriction at all and the restriction level 0.5 is halfway between being totally strict and no strict at all.

Definition 4 [30]. A *RL-set* Λ is a finite set of restriction levels $\Lambda = \{\alpha_1, \ldots, \alpha_m\}$ verifying that $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_m > \alpha_{m+1} = 0$, $m \geqslant 1$.

The *RL*-set of an atomic property represented by means of a fuzzy set A is defined as follows.

Definition 5 [30]. Let A be a fuzzy set defined on the referential X. Then the *RL*-set associated to A is given by:

$$\Lambda_A = \{A(x) \mid x \in support(A)\} \cup \{1\}, \tag{4}$$

where $A(x)$ is the membership grade of x to the fuzzy set A, and $support(.)$ denotes the support of a fuzzy set.

The employed *RL*-set to represent an imprecise property is obtained by the union of the *RL*-sets associated to the atomic properties defining that property.

[2] Do not mislead with the concept of fuzzy quantifier.

A *RL-representation* associated to an imprecise property in X is defined by a pair (Λ, ρ) where Λ is a *RL*-set and $\rho : \Lambda \to \mathcal{P}(X)$ is a function which applies each restriction level into a crisp realization in this level. For example, the *RL*-representation of an imprecise atomic property defined by a fuzzy set A will be the pair (Λ_A, ρ_A), where Λ_A is given by Eq. (4) and $\rho_A(\alpha) = A_\alpha = \{x \in X | A(x) \geqslant \alpha\}$ for all $\alpha \in \Lambda_A$.

Given an imprecise property P represented by (Λ_P, ρ_P), the set of crisp representatives of P is the set [30] $\Omega_P = \{\rho_P(\alpha) | \alpha \in \Lambda_P\}$.

Definition 6 [30]. Let (Λ, ρ) be a *RL*-representation with $\Lambda = \{\alpha_1, \ldots, \alpha_m\}$ verifying that $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_m > \alpha_{m+1} = 0$. Let $\alpha \in (0,1]$ and $\alpha_i, \alpha_{i+1} \in \Lambda$ satisfying that $\alpha_i > \alpha > \alpha_{i+1}$. Then we define

$$\rho(\alpha) = \rho(\alpha_i). \tag{5}$$

If we look to this definition, this extension for values that there are not in the *RL*-set of the function ρ, is the natural extension if we think of a fuzzy set A and its α-cuts. Using this definition the concept of equivalence between two *RL*-representations is straightforward.

Definition 7 [30]. Let (Λ, ρ) and (Λ', ρ') be two *RL*-representations on X. We will say that both representations (and the corresponding properties) are *equivalent*, noted by $(\Lambda, \rho) \equiv (\Lambda', \rho')$, if and only if, $\forall \alpha \in (0,1]$

$$\rho(\alpha) = \rho(\alpha'). \tag{6}$$

Summarizing, only a finite *RL*-set is necessary for defining a RLR, but the representation extends to any other *RL* in $(0,1]$.

The usual boolean operations are extended to RLRs by applying them on the representatives of the same *RL* of the arguments independently. In particular, we present here the logic operations of disjunction, conjunction and negation. The basic ideas of how they are defined can be found in [30].

Definition 8. Let P, Q be two imprecise properties with *RL*-representations (Λ_P, ρ_P), (Λ_Q, ρ_Q). Then, $P \wedge Q$, $P \vee Q$ and $\neg P$ are imprecise properties represented by $(\Lambda_{P\wedge Q}, \rho_{P\wedge Q})$, $(\Lambda_{P\vee Q}, \rho_{P\vee Q})$ and $(\Lambda_{\neg P}, \rho_{\neg P})$ respectively, where $\Lambda_{P\wedge Q} = \Lambda_{P\vee Q} = \Lambda_P \cup \Lambda_Q$, $\Lambda_{\neg P} = \Lambda_P$ and, for all $\alpha \in (0,1]$,

$$\begin{aligned} \rho_{P\wedge Q}(\alpha) &= \rho_P(\alpha) \cap \rho_Q(\alpha), \\ \rho_{P\vee Q}(\alpha) &= \rho_P(\alpha) \cup \rho_Q(\alpha), \\ \rho_{\neg P}(\alpha) &= \overline{\rho_P(\alpha)}, \end{aligned} \tag{7}$$

where $\overline{Y}$ is the usual complement of a crisp set Y.

Basic boolean properties that cannot be verified simultaneously by any standard fuzzy set theory (FST) hold simultaneously for RLRs. We want to remark that fuzzy sets are closed with respect to some of these RLR operations in the sense that the corresponding RLR yields the usual nested α-cut representation and hence the result is a fuzzy set. However, this is not necessarily true when negation is employed. In fact, given B a non-crisp fuzzy set, the RLR of $\neg B$ is not a fuzzy set as next example shows (see [30] for a more complete explanation and examples).

Example 2. Let A, B be the following two fuzzy sets: $A = 1/x_1 + 0.8/x_2 + 0.5/x_3 + 0.4/x_4$, $B = 0.9/x_1 + 0.6/x_3 + 0.5/x_4$ defined over the referential $X = \{x_1, \ldots, x_5\}$. We take their associated *RL*-sets (Λ_A, ρ_A), (Λ_B, ρ_B) and we perform some operations involving negation obtaining that $(\Lambda_{A\wedge\neg B}, \rho_{A\wedge\neg B})$ is not the *RL*-set of any fuzzy set because the crisp sets are not nested when decreasing the restriction level α_i (see Table 2).

2.5. Interpretation in terms of evidence

Given a *RL*-representation (Λ_A, ρ_A) for an atomic property A, the values of Λ_A can be interpreted as values of possibility for a possibility measure defined for all $\rho_A(\alpha_i) \in \Omega_A$ as

$$Pos(\rho_A(\alpha_i)) = \alpha_i. \tag{8}$$

Table 2
RL-sets associated to A, B, $\neg B$ and $A \wedge \neg B$.

α_i	$\rho_A(\alpha)$	$\rho_B(\alpha)$	$\rho_{\neg B}(\alpha)$	$\rho_{A\wedge\neg B}(\alpha)$
1	$\{x_1\}$	$\emptyset$	X	$\{x_1\}$
0.9	$\{x_1\}$	$\{x_1\}$	$\{x_2,x_3,x_4,x_5\}$	$\emptyset$
0.8	$\{x_1,x_2\}$	$\{x_1\}$	$\{x_2,x_3,x_4,x_5\}$	$\{x_2\}$
0.6	$\{x_1,x_2\}$	$\{x_1,x_3\}$	$\{x_2,x_4,x_5\}$	$\{x_2\}$
0.5	$\{x_1,x_2,x_3\}$	$\{x_1,x_3,x_4\}$	$\{x_2,x_5\}$	$\{x_2\}$
0.4	$\{x_1,x_2,x_3,x_4\}$	$\{x_1,x_3,x_4\}$	$\{x_2,x_5\}$	$\{x_2,x_5\}$

Following this interpretation we define a basic probability assignment in the usual way:

Definition 9. Let (Λ, ρ) be a RL-representation with crisp representatives Ω. The associated probability distribution m: $\Omega \to [0,1]$ is

$$m(Y) = \sum_{\alpha_i \mid Y=\rho(\alpha_i)} \alpha_i - \alpha_{i+1}. \tag{9}$$

The basic probability assignment m_F gives us information about how representative is each crisp set of the property F in Ω_F. For each $Y \in \Omega_F$, the value $m_F(Y)$ represents the proportion to which the available evidence supports claim that the property F is represented by Y. From this point of view, a *RL*-representation can be seen as a basic probability assignment in the sense of the theory of evidence, *plus a structure indicating dependencies between the possible representations of different properties*.

2.6. RL-numbers

On the basis of *RL*-representations and operations, we introduced in [29] the *RL*-numbers as a representation of imprecise quantities. This approach offers two main advantages: (1) *RL*-numbers are representations of imprecise quantities that can be easily obtained by extending usual crisp measurements to fuzzy sets. (2) Arithmetic and logical operations on *RL*-numbers are straightforward and unique extensions of the operations on crisp numbers, verifying the usual properties of crisp arithmetic and logical operations. In addition, the imprecision does not necessarily increase through operations, and it can even diminish. The following definitions and properties are from [29]:

Definition 10. A *RL*-real number is a pair $(\Lambda, \mathcal{R})$ where Λ is a *RL*-set and $\mathcal{R} : (0,1] \to \mathbb{R}$.

We shall note $\mathbb{R}_{RL}$ the set of *RL*-real numbers. The *RL*-real number R_x is the representation of a (precise) real number x iff $\forall \alpha \in \Lambda_{R_x},\ \mathcal{R}_{R_x}(\alpha) = x$. We shall denote such *RL*-real number as R_x or, equivalently, x, since in the crisp case, the set Λ_{R_x} is not important. Operations are extended as follows:

Definition 11. Let $f : \mathbb{R}^n \to \mathbb{R}$ and let $R_1, \ldots, R_n$ be *RL*-real numbers. Then $f(R_1, \ldots, R_n)$ is a *RL*-real number with

$$\Lambda_{f(R_1,\ldots,R_n)} = \bigcup_{1 \leqslant i \leqslant n} \Lambda_{R_i} \tag{10}$$

and, $\forall \alpha \in \Lambda_{f(R_1,\ldots,R_n)}$

$$\mathcal{R}_{f(R_1,\ldots,R_n)}(\alpha) = f(\mathcal{R}_{R_1}(\alpha), \ldots, \mathcal{R}_{R_n}(\alpha)). \tag{11}$$

It is obvious that operations defined in this way are consistent extensions of crisp operations. We want to remark that operations not defined for certain combinations of real values are not defined for *RL*-numbers that verify that combination in at least one restriction level. This is the case of division by 0, i.e., R/R' is defined if and only if $0 \notin \Omega_{R'}$.

Dubois and Prade introduce the notion of *gradual element*, in particular, gradual number in [15] as missing concepts in the theory of fuzzy sets and their use has been mainly investigated for the representation of vague quantities. In [30] we have done an extensive study about the similarities and differences between the use of fuzzy, gradual or *RL* numbers concluding that they are different but complementary. Fuzzy numbers are fuzzy concepts defined on the domain of the numbers, vague intervals define restrictions on the numbers, whilst gradual/*RL*-numbers are true vague numbers. All are useful, but for different purposes; they have the same usefulness as intervals and numbers, respectively, in the crisp case. More specifically we propose the following:

- *RL*-numbers are the correct choice for extending measures (cardinality, probability of fuzzy events, etc.) to the case of fuzzy information represented by either fuzzy sets, in particular, or RLRs in general. Any such measure on a fuzzy set should yield a *RL*-number when the (possibly fuzzy) set/event we want to measure is well known (possibly by a representation as a fuzzy set).
- Fuzzy numbers (intervals) are an useful, correct, intuitive way to define fuzzy restrictions with semantics of fuzzy interval like "around" or "approximately between x and y". This is useful in two different situations: (1) when the imprecise quantity we want to represent is ill-known and (2) when providing meaningful information to a human (fuzzy intervals are better suited than *RL*-numbers for this purpose).

3. Formal model for mining fuzzy association rules

3.1. Introduction and related works

Fuzzy rules represent in a comprehensive way the information obtained from a database. Nevertheless, their evaluation by means of appropriate quality measures is not a straightforward step [13], specially if we assume the semantics associated to the fuzzy rule. At this respect several approaches have been developed:

Sudkamp establishes quality measures for fuzzy rules basing on a classification for a transaction into three types: examples, counterexamples and irrelevant examples [36], and he proposes to compute the measures of *confirmation* and *confidence* for fuzzy rules depending on the membership degree of a transaction into each type.

Dubois et al. [14] propose a method to obtain the quality measures for fuzzy rules in a systematic way basing on a classification of transactions into examples, counterexamples and irrelevant examples associated to a rule. In the fuzzy case, the corresponding partition will be fuzzy but its definition will depend on the semantics we want to capture in the rule. According to this approach, several types of fuzzy rules can be defined depending on the chosen focus: a set conjunction approach or an implication approach. Both approaches coincide in the crisp case but must be differentiated in the fuzzy one.

De Cock et al. also propose a study of the fuzzy rule semantics by defining the associated positive and negative examples of a fuzzy rule [8].

We want to establish a formalization for fuzzy association rules which enables to extend in a straightforward way the measures used to assess their accuracy. Moreover, we want to keep the main properties of such measures from the crisp to the fuzzy case. In particular, we use the model presented in Section 2.3 and the RLR theory, which brings us the possibility to use every crisp mining algorithm for the fuzzy case by means of a parallelization process (see Section 5).

The key idea of our approach has some similarities with the previous cases since the formal model gives a classification of the transactions into four different types (examples supporting $A \wedge B$, $A \wedge \neg B$, $\neg A \wedge B$ and $\neg A \wedge \neg B$). The main difference is that we do not measure in these sets. We first consider the tuples pertaining to each set in each restriction level and then we aggregate the measure for each level depending on the difference between the levels using the basic probability assignment in formula (9). In this way, in each restriction level we have a value for the satisfiability of the rule which will be then aggregated in every level, being possible even to decrease its global value, when the interest measure range oscillates in the interval $[-1,1]$ (this is the case of the certainty factor).

3.2. Formal model for fuzzy rules. Our approach

This section is devoted to present the generalization of the formal model in Section 2.3 for fuzzy rules using the RLR theory. Then we will present a framework for extending the interestingness measures for their validation from the crisp to the fuzzy case by some examples.

Let $A,B \in I$ be two itemsets in a fuzzy database $\widetilde{D}$. We consider that $\widetilde{\Gamma}_A$ and $\widetilde{\Gamma}_B$ are the fuzzy sets defined in $\widetilde{D}$ as $\widetilde{\Gamma}_A(\tilde{\tau}) = \tilde{\tau}(A)$ and $\widetilde{\Gamma}_B(\tilde{\tau}) = \tilde{\tau}(B)$ respectively. Following the RLR theory presented in Section 2.4 their respective *RL*-representations noted by $(\Lambda_{\widetilde{A}}, \rho_{\widetilde{A}})$, $(\Lambda_{\widetilde{B}}, \rho_{\widetilde{B}})$, are given by Eqs. (4) and (5).

Note that the set $\{A \wedge B, A \wedge \neg B, \neg A \wedge B, \neg A \wedge \neg B\}$ is a partition of the fuzzy database $\widetilde{D}$ (in the same way as $\{\varphi \wedge \psi, \varphi \wedge \neg\psi, \neg\varphi \wedge \psi, \neg\varphi \wedge \neg\psi\}$ is a partition in the crisp case). So, we can take their associated fuzzy sets defined in $\widetilde{D}$ joint with their respective *RL*-representations noted by $(\Lambda_{\widetilde{A}\wedge\widetilde{B}}, \rho_{\widetilde{A}\wedge\widetilde{B}})$, $(\Lambda_{\widetilde{A}\wedge\neg\widetilde{B}}, \rho_{\widetilde{A}\wedge\neg\widetilde{B}})$, and so on. We remark that the obtained *RL*-sets will contain the same set of restriction levels, that is, all of them will be equal to $\Lambda_{\widetilde{A}} \cup \Lambda_{\widetilde{B}}$.

For every $\alpha \in \Lambda_Y$, $\rho_Y(\alpha)$ is a crisp set, then we can compute its cardinality noting it, as usual, by $|\rho_Y(\alpha)|$. To generalize the concepts of four fold table and 4*ft*-quantifier, we define for every restriction level $\alpha_i \in \Lambda_{\widetilde{A}} \cup \Lambda_{\widetilde{B}}$ the associated 4*ft*-table $\mathcal{M}_{\alpha_i} = 4ft(\widetilde{\Gamma}_A, \widetilde{\Gamma}_B, \widetilde{D}, \alpha_i)$ as follows:

$\mathcal{M}_{\alpha_i}$	$\widetilde{\Gamma}_B$	$\widetilde{\Gamma}_{\neg B}$
$\widetilde{\Gamma}_A$	a_i	b_i
$\widetilde{\Gamma}_{\neg A}$	c_i	d_i

where a_i, b_i, c_i and d_i are non-negative integers such that $a_i = \left|\rho_{\widetilde{A}\wedge\widetilde{B}}(\alpha_i)\right|$, $b_i = \left|\rho_{\widetilde{A}\wedge\neg\widetilde{B}}(\alpha_i)\right|$ and analogously with c_i and d_i. The 4*ft*-quantifiers are computed for every restriction level using the 4*ft*-table in the particular level. It is worth to mention that for all $\alpha_i \in \Lambda_{\widetilde{A}} \cup \Lambda_{\widetilde{B}}$ the following equality is fulfilled

$$a_i + b_i + c_i + d_i = n = |\widetilde{D}|. \tag{12}$$

Until now we have obtained for each restriction level α_i a 4*ft*-table $\mathcal{M}_{\alpha_i}$ and the value of the 4*ft*-quantifier in that level, noted $\approx(a_i, b_i, c_i, d_i)$.

Using the representation the formal model offers us, we are able to generalize every interest measure from the crisp to the fuzzy case, in particular, every 4*ft*-quantifier by means of the basic probability distribution (see Definition 9) as follows:

$$\sum_{\alpha_i \in \Lambda_{\widetilde{A}} \cup \Lambda_{\widetilde{B}}} (\alpha_i - \alpha_{i+1})(\approx (a_i, b_i, c_i, d_i)). \tag{13}$$

The following theorem shows that the formal model for fuzzy rules is a good generalization of the crisp case.

Theorem 1. *Let A and B be two itemsets in a crisp database D. Then the fuzzy formal model previously defined coincides with the crisp one presented in Section* 2.3.

Proof. To prove it, we first define the itemsets A and B seen as fuzzy itemsets: $\widetilde{\Gamma}_A(t) = t(A),\ \widetilde{\Gamma}_B(t) = t(B) \in [0,1]$ where t is a transaction in D and $t(A), t(B) \in \{0,1\}$ is the following indicator function:

$$t(A) = \begin{cases} 1 & \text{if } A \in t, \\ 0 & \text{if } A \notin t. \end{cases} \tag{14}$$

Aanalogously for B. The associated RL-sets to $\widetilde{\Gamma}_A$ and $\widetilde{\Gamma}_B$ are $\Lambda_{\widetilde{A}} = \Lambda_{\widetilde{B}} = \{1\}$. And their associated RL-representations are: $(\Lambda_{\widetilde{A}}, \rho_{\widetilde{A}})$ and $(\Lambda_{\widetilde{B}}, \rho_{\widetilde{B}})$ where

$$\begin{aligned} \rho_{\widetilde{A}}(1) &= A_1 = \{t \in D \mid t(A) \geqslant 1\} = \{t \in D \mid A \in t\}, \\ \rho_{\widetilde{B}}(1) &= B_1 = \{t \in D \mid t(B) \geqslant 1\} = \{t \in D \mid B \in t\}. \end{aligned} \tag{15}$$

Similarly, we compute the RL-representations for the sets $\widetilde{\Gamma}_A \wedge \widetilde{\Gamma}_B,\ \widetilde{\Gamma}_A \wedge \neg\widetilde{\Gamma}_B,\ \neg\widetilde{\Gamma}_A \wedge \widetilde{\Gamma}_B,\ \neg\widetilde{\Gamma}_A \wedge \neg\widetilde{\Gamma}_B$. We can see that the 4*ft*-table for the restriction level $\alpha = 1$ coincides with the 4*ft*-table for the crisp itemsets A and B (see Section 2.3)

$\mathcal{M}_1 \equiv \mathcal{M}$	B	$\neg B$
A	a_1	b_1
$\neg A$	c_1	d_1

With respect to the interest measures for fuzzy rules, it is immediate to see that the generalization of 4*ft*-quantifiers defined in Eq. (13) coincides with the one defined for the crisp case:

$$\sum_{\alpha_i \in \Lambda_{\widetilde{A}} \cup \Lambda_{\widetilde{B}}} (\alpha_i - \alpha_{i+1})(\approx (a_i, b_i, c_i, d_i)) = (1-0)(\approx (a_1, b_1, c_1, d_1)). \qquad \square \tag{16}$$

In the following we present the fuzzy extension for the particular cases of support and confidence measures.

Definition 12 (*Itemset support*). Let $A \subseteq I$ be an itemset and $(\Lambda_{\widetilde{A}}, \rho_{\widetilde{A}})$ the associated RL-representation to the fuzzy set $\widetilde{\Gamma}_A$ in $\widetilde{D}$. The support of A in the fuzzy transactional database $\widetilde{D}$ is defined as follows

$$\text{Fsupp}(A) = \sum_{\alpha_i \in \Lambda_{\widetilde{A}}} (\alpha_i - \alpha_{i+1}) \left(\frac{\left|\rho_{\widetilde{A}}(\alpha_i)\right|}{|\widetilde{D}|} \right). \tag{17}$$

Taking the right part of the formula (17) and using the 4*ft*-table associated to itemsets A and B we can see that for the restriction level α_i

$$\frac{\left|\rho_{\widetilde{A}}(\alpha_i)\right|}{|\widetilde{D}|} = \frac{a_i + b_i}{a_i + b_i + c_i + d_i} \tag{18}$$

is the itemset support of A when the database is crisp. In addition, $a_i + b_i + c_i + d_i$ is constant for every restriction level being the number of fuzzy transactions of $\widetilde{D}$ that will be noted by n as in the crisp case.

The support and the confidence of a fuzzy rule $A \to B$ are defined in a similar way.

Definition 13 (*Rule support*). Let $A, B \subseteq I$ be two disjoint itemsets and $(\Lambda_{\widetilde{A}}, \rho_{\widetilde{A}}),\ (\Lambda_{\widetilde{B}}, \rho_{\widetilde{B}})$ the RL-representations associated to the fuzzy sets $\widetilde{\Gamma}_A$ and $\widetilde{\Gamma}_B$ in $\widetilde{D}$. Then, the support of the fuzzy rule $A \to B$ in $\widetilde{D}$ is defined as

$$\text{FSupp}(A \to B) = \text{Fsupp}(A \wedge B) = \sum_{\alpha_i \in \Lambda_{\widetilde{A}} \cup \Lambda_{\widetilde{B}}} (\alpha_i - \alpha_{i+1}) \left(\frac{\left|\rho_{\widetilde{A} \wedge \widetilde{B}}(\alpha_i)\right|}{|\widetilde{D}|} \right). \tag{19}$$

Definition 14 (*Rule confidence*). Let $A, B \subseteq I$ be two disjoint itemsets and $(\Lambda_{\widetilde{A}}, \rho_{\widetilde{A}}), (\Lambda_{\widetilde{B}}, \rho_{\widetilde{B}})$ the RL-representations associated to the fuzzy sets $\widetilde{\Gamma}_A$ and $\widetilde{\Gamma}_B$ in $\widetilde{D}$. Then, the confidence of the fuzzy rule $A \to B$ in $\widetilde{D}$ is defined as

$$\text{FConf}(A \to B) = \sum_{\alpha_i \in \Lambda_{\widetilde{A}} \cup \Lambda_{\widetilde{B}}} (\alpha_i - \alpha_{i+1})(\Rightarrow_I (a_i, b_i)) = \sum_{\alpha_i \in \Lambda_{\widetilde{A}} \cup \Lambda_{\widetilde{B}}} (\alpha_i - \alpha_{i+1}) \left(\frac{\left|\rho_{\widetilde{A} \wedge \widetilde{B}}(\alpha_i)\right|}{\left|\rho_{\widetilde{A}}(\alpha_i)\right|} \right). \tag{20}$$

Previous definition has the inconvenience of presenting indeterminations of the form "$\frac{0}{0}$" when $\left|\rho_{\widetilde{A}}(\alpha_i)\right| = 0$. This happens when there does not exist transactions satisfying at the same time the antecedent and the consequent. So, to preserve the Definition 3 of fuzzy rule we will take the value 1 for that indetermination.

Table 3
Fuzzy sets $\widetilde{\Gamma}_A$ and $\widetilde{\Gamma}_B$.

	$\tilde{\tau}_1$	$\tilde{\tau}_2$	$\tilde{\tau}_3$	$\tilde{\tau}_4$	$\tilde{\tau}_5$	$\tilde{\tau}_6$
$\widetilde{\Gamma}_A$	1	0.8	0.5	0	0.4	0
$\widetilde{\Gamma}_B$	0.9	0	0.6	0.5	0	0

Table 4
Associated *RL*-representations to $\widetilde{\Gamma}_A$, $\widetilde{\Gamma}_B$, $\neg\widetilde{\Gamma}_A$ and $\neg\widetilde{\Gamma}_B$.

α_i	$\rho_{\widetilde{A}}$	$\rho_{\widetilde{B}}$	$\rho_{\neg\widetilde{A}}$	$\rho_{\neg\widetilde{B}}$
1	$\{\tilde{\tau}_1\}$	$\emptyset$	$\{\tilde{\tau}_2, \tilde{\tau}_3, \tilde{\tau}_4, \tilde{\tau}_5, \tilde{\tau}_6\}$	$\widetilde{D}_1$
0.8	$\{\tilde{\tau}_1, \tilde{\tau}_2\}$	$\{\tilde{\tau}_1\}$	$\{\tilde{\tau}_3, \tilde{\tau}_4, \tilde{\tau}_5, \tilde{\tau}_6\}$	$\{\tilde{\tau}_2, \tilde{\tau}_3, \tilde{\tau}_4, \tilde{\tau}_5, \tilde{\tau}_6\}$
0.6	$\{\tilde{\tau}_1, \tilde{\tau}_2\}$	$\{\tilde{\tau}_1, \tilde{\tau}_3\}$	$\{\tilde{\tau}_3, \tilde{\tau}_4, \tilde{\tau}_5, \tilde{\tau}_6\}$	$\{\tilde{\tau}_2, \tilde{\tau}_4, \tilde{\tau}_5, \tilde{\tau}_6\}$
0.4	$\{\tilde{\tau}_1, \tilde{\tau}_2, \tilde{\tau}_3, \tilde{\tau}_4\}$	$\{\tilde{\tau}_1, \tilde{\tau}_3, \tilde{\tau}_4\}$	$\{\tilde{\tau}_5, \tilde{\tau}_6\}$	$\{\tilde{\tau}_2, \tilde{\tau}_5, \tilde{\tau}_6\}$
0.2	$\{\tilde{\tau}_1, \tilde{\tau}_2, \tilde{\tau}_3, \tilde{\tau}_4\}$	$\{\tilde{\tau}_1, \tilde{\tau}_3, \tilde{\tau}_4\}$	$\{\tilde{\tau}_5, \tilde{\tau}_6\}$	$\{\tilde{\tau}_2, \tilde{\tau}_5, \tilde{\tau}_6\}$

It is worth to mention that the defined measures for the support and the confidence in the fuzzy case are the same as those proposed in [9] where it is used a semantic approximation based on the evaluation of quantified sentences using the *GD*-method and the relative fuzzy quantifier $Q(x) = x$ to evaluate the sentences.

In the following, we present two examples. The former shows in a detailed way, by means of a toy fuzzy database, the process of computing the support and the confidence of fuzzy rules using the proposed formal model. The last example shows how to extend the certainty factor measure presented in [6] which is a stronger measure than confidence, and that we will use later in the experimental evaluation.

Example 3. Let $I = \{i_1, i_2, \ldots, i_6\}$ and $\widetilde{D}_1$ be the fuzzy database in Table 1. We take the itemsets $A = \{i_1, i_3\}$ and $B = \{i_4\}$ whose associated fuzzy sets $\widetilde{\Gamma}_A$ and $\widetilde{\Gamma}_B$ are in Table 3 with *RL*-representations $(\Lambda_{\widetilde{A}}, \rho_{\widetilde{A}})$ and $(\Lambda_{\widetilde{B}}, \rho_{\widetilde{B}})$. In Tables 4 and 5 we show the resulting *RL*-representations from applying the negation to both fuzzy sets and their conjunction for the *RL*-set $\Lambda = \{1, 0.8, 0.6, 0.4, 0.2\}$. Now, we follow with the computation of support and confidence of some fuzzy rules in $\widetilde{D}_1$ using the proposed model. First, we compute the 4*ft*-tables $\mathcal{M}_{\alpha_i} = 4ft(\widetilde{\Gamma}_A, \widetilde{\Gamma}_B, \widetilde{D}_1)$ for every restriction level in Λ (Table 6)[3] and then, using Eqs. (19) and (20) we calculate the support and the confidence of the rule $A \rightarrow B$ being 0.266 and 0.5 respectively.

Taking the itemsets $C = \{i_1, i_5\}$ and $E = \{i_4\}$ the support of the fuzzy rule $C \rightarrow E$ is 0.2, and its confidence is equal to 1. This is due to $\tilde{\tau}_i(C) \leqslant \tilde{\tau}_i(E)$ for all $i = 1, 2, \ldots, 6$ following Definition 3. If we use the formula for the confidence given in (20) we get the same result but having into account that when the indetermination "$\frac{0}{0}$" appears we consider it as 1, as we commented after Definition 14. Other association rules found in $\widetilde{D}_1$ are given in Table 7.

Several works have pointed out some drawbacks of the support/confidence framework to assess association rules (see for instance [35]). In general, the same problems occur when dealing with fuzzy association rules. In this paper we shall employ the certainty factor for the validation of association rules (see Definition 1).

The following example shows how to generalize the certainty factor measure that, seen as a 4*ft*-quantifier, is noted by $\equiv_{CF}$ (due to its good properties the certainty factor belongs to the class of equivalence 4*ft*-quantifiers, see [12]).

Example 4. Let $A, B \subset I$ be two disjoint itemsets in D and let $4ft(A, B, D) = \langle a, b, c, d\rangle$ be their associated 4*ft*-table. The 4*ft*-quantifier $\equiv_{CF}$ which represents the certainty factor is given as follows:

$$\equiv_{CF}(a,b,c,d) = \begin{cases} \frac{ad-bc}{(a+b)(b+d)} & \text{if } ad > bc, \\ 0 & \text{if } ad = bc, \\ \frac{ad-bc}{(a+b)(a+c)} & \text{if } ad < bc. \end{cases} \quad (21)$$

To generalize this quantifier to the fuzzy case we only have to consider the 4*ft*-table for $\Lambda_{\widetilde{A}} \cup \Lambda_{\widetilde{B}}$ (the union of the *RL*-sets of $\widetilde{\Gamma}_X$ and $\widetilde{\Gamma}_Y$) for each restriction level and then compute:

$$\equiv_{CF}(A \rightarrow B) = \sum_{\alpha_i \in \Lambda_{\widetilde{A}} \cup \Lambda_{\widetilde{B}}} (\alpha_i - \alpha_{i+1})(\equiv_{CF}(a_i, b_i, c_i, d_i)), \quad (22)$$

where the *RL*-representations of both itemsets must be normalized to fulfil Definition 3. In particular, for the fuzzy database $\widetilde{D}_1$, and the previous itemsets, we show the associated certainty factors in Table 8.

[3] Theoretically we can compute the 4*ft*-tables for each restriction level in $\Lambda_A \cup \Lambda_B$ but in practice we will fixed the *RL*-set into a finite set of levels in the unit interval because the possible different levels that can appear in a database could be very large.

Table 5
RL-representations associated to the possible conjunctions between the fuzzy sets $\widetilde{\Gamma}_A$, $\widetilde{\Gamma}_B$, $\neg\widetilde{\Gamma}_A$ and $\neg\widetilde{\Gamma}_B$.

α_i	$\rho_{\widetilde{A}\wedge\widetilde{B}}$	$\rho_{\widetilde{A}\wedge\neg\widetilde{B}}$	$\rho_{\neg\widetilde{A}\wedge\widetilde{B}}$	$\rho_{\neg\widetilde{A}\wedge\neg\widetilde{B}}$
1	$\emptyset$	$\{\tilde{\tau}_1\}$	$\emptyset$	$\{\tilde{\tau}_2,\tilde{\tau}_3,\tilde{\tau}_4,\tilde{\tau}_5,\tilde{\tau}_6\}$
0.8	$\{\tilde{\tau}_1\}$	$\{\tilde{\tau}_2\}$	$\emptyset$	$\{\tilde{\tau}_3,\tilde{\tau}_4,\tilde{\tau}_5,\tilde{\tau}_6\}$
0.6	$\{\tilde{\tau}_1\}$	$\{\tilde{\tau}_2\}$	$\{\tilde{\tau}_3\}$	$\{\tilde{\tau}_4,\tilde{\tau}_5,\tilde{\tau}_6\}$
0.4	$\{\tilde{\tau}_1,\tilde{\tau}_3,\tilde{\tau}_4\}$	$\{\tilde{\tau}_2\}$	$\emptyset$	$\{\tilde{\tau}_5,\tilde{\tau}_6\}$
0.2	$\{\tilde{\tau}_1,\tilde{\tau}_3,\tilde{\tau}_4\}$	$\{\tilde{\tau}_2\}$	$\emptyset$	$\{\tilde{\tau}_5,\tilde{\tau}_6\}$

Table 6
$4ft(\mathcal{M}_{\alpha_i}, A, B, \widetilde{D})$ where $\alpha_i \in \Lambda$.

	a_i	b_i	c_i	d_i
$\mathcal{M}_1$	0	1	0	5
$\mathcal{M}_{0.8}$	1	1	0	4
$\mathcal{M}_{0.6}$	1	1	1	3
$\mathcal{M}_{0.4}$	3	1	0	2
$\mathcal{M}_{0.2}$	3	1	0	2

Table 7
Some fuzzy rules in database $\widetilde{D}_1$ considering $\Lambda = \{1, 0.8, 0.6, 0.4, 0.2\}$.

Rule	Support	Confidence
$\{i_1, i_2\} \to \{i_3\}$	0.167	0.8
$\{i_4\} \to \{i_5\}$	0.2	0.767

Table 8
Certainty factor for some fuzzy rules in $\widetilde{D}_1$ taking $\Lambda = \{1, 0.8, 0.6, 0.4, 0.2\}$.

Rule	Support	Confidence	Certainty factor
$\{i_1, i_3\} \to \{i_4\}$	0.266	0.5	0.33
$\{i_1, i_5\} \to \{i_4\}$	0.2	1	1
$\{i_1, i_2\} \to \{i_3\}$	0.167	0.8	0.6
$\{i_4\} \to \{i_5\}$	0.2	0.767	0.48

We want to remark that in the case of the itemsets $A = \{i_1, i_5\}$ and $B = \{i_4\}$ we have to normalize their *RL*-representations (to divide all the restriction levels by the highest one, in this case we divide by 0.9) to obtain the corresponding value 1 for the certainty factor, since if $\tilde{\tau}_i(A) \leqslant \tilde{\tau}_i(B)$ for all $i = 1, \ldots, 6$ the association rule is totally accurate.

3.3. Alternatives for the validation of fuzzy rules using the formal model

The proposed formalization for fuzzy rules by restriction levels, allows the user to choose several options in order to manage the obtained results. In the previous section we have dealt with the standard option, i.e. we evaluate the accuracy of fuzzy rules by a crisp number given by the value of the associated 4*ft*-quantifier (see Eq. (13)) but it is also possible to obtain a detailed view of the results via restriction levels. Then, depending on the user interests several types of presentation for the obtained results can be chosen. We elucidate the following ones:

1. Summarizing the results by a crisp number.
2. Obtaining a crisp result for each restriction level.
3. Summarizing the results using a *RL*-number.

The first option corresponds to the value of the 4*ft*-quantifier obtained by its generalization for fuzzy rules via Eq. (13). This is the classic approach that have been taken into account in most of the approaches until now, where those fuzzy rules that exceed the imposed thresholds for the support and the confidence (or other measure) are extracted from the data. In Examples 3 and 4 we used this criteria, we showed the mined rules choosing the crisp value associated to the 4*ft*-quantifier. Next sections present two new alternatives for showing the results to have different perspectives for managing the information from the set of extracted rules.

3.3.1. A set of crisp results

The proposed formal model parallelizes the process for obtaining fuzzy rules, as for each restriction level we compute the value associated to the corresponding crisp rule where the level is seen as the evidence we have for the fulfilment of the rule. Therefore, by imposing fixed thresholds in each restriction level we obtain a set of crisp rules at each level. In this way, the user could be interested in obtaining the set of crisp rules satisfied with evidence greater than a fixed α, proceeding thence in different ways.

- We can restrict the mining process in only those levels greater than α, that is, we search those rules exceeding the imposed thresholds at levels α_i greater or equal than the fixed level α defined by the user. This choice will accelerate the mining process as we take only those rules whose evidence is $\geqslant\alpha$, but we lose information because rules with less evidence but near α are discarded.
- We mine all rules in each restriction level satisfying the imposed thresholds, and we also compute the value of the 4*ft*-quantifier which summarizes the accuracy of the fuzzy rule. Having both, we can center our attention in those levels greater or equal than the fixed α but having the information of the total accuracy of the rule in the rest of the levels.

Example 5. Continuing the previous examples, we consider the fuzzy database $\widetilde{D}_1$ in Table 1. If we set $\alpha = 0.6$ we are saying that we are interested in those rules with evidence greater than 0.6. The first option could be searching only those rules satisfying the thresholds for support and confidence/CF in levels greater than 0.6. If we look at Table 9 we can see that rule $\{i_1,i_3\} \to \{i_4\}$ does not have a high value in levels 0.6, 0.8 and 1 so, it would not be interesting for the user. Rule $\{i_1,i_2\} \to \{i_3\}$ would be extracted at levels 0.6 and 0.8 but not at level 1.

The second option also gives to the user a summary measure for each rule that says if it would be extracted or not in global. In our example the rule $\{i_1,i_3\} \to \{i_4\}$ is not interesting, and the rule $\{i_1,i_2\} \to \{i_3\}$ will be mined if the chosen 4*ft*-quantifier is greater or equal than the fixed threshold, in our example if *minconf* = 0.6 or *minCF* = 0.8.

It is worth to mention that this second option also gives the possibility to explore at each level the set of mined rules, having at the same time the summary value associated to each rule (see Table 9).

3.3.2. Summarizing results using RL

In order to manage the obtained results using restriction levels we present two different choices. The first one is to present for each mined rule its associated values in each restriction level by considering *RL*-numbers. Imagine that the fuzzy rule $A \to B$ is mined in $\widetilde{D}$ because its support and confidence given in (19) and (20) exceeds the minsupp and minconf thresholds for fuzzy rules. If so, we can show to the user its associated support and confidence values in each considered restriction level by the following *RL*-numbers:

$$\mathrm{RLSupp}(A \to B) = \mathrm{Supp}_1(A \to B)/1 + +\mathrm{Supp}_{\alpha_2}(A \to B)/\alpha_2 + \cdots + \mathrm{Supp}_{\alpha_m}(A \to B)/\alpha_m, \tag{23}$$

$$\mathrm{RLConf}(A \to B) = \mathrm{Conf}_1(A \to B)/1 + +\mathrm{Conf}_{\alpha_2}(A \to B)/\alpha_2 + \cdots + \mathrm{Conf}_{\alpha_m}(A \to B)/\alpha_m, \tag{24}$$

where $\mathrm{Supp}_\alpha(A \to B)$ and $\mathrm{Conf}_\alpha(A \to B)$ represent respectively the support and the confidence of $A \to B$ at level α. The RLCF of a rule could be defined analogously.

This option provides to the user the measurement values of each rule at each restriction level. On the other hand, if we are more interested in knowing the mined rules at each restriction level, we offer the following methodology.

Another interesting issue of our method is that, in terms of the restriction levels evidence, we can summarize the obtained set of fuzzy association rules too. We obtain all crisp rules in each restriction level and then we reduce the number of existent rules by discarding those less interesting according to the associated thresholds at each restriction level. After this pruning, we can interpret the resulting set of association rules in terms of evidence by using *RL* obtaining an expression of the following type

$$Ruleset = \{i_1 \to i_3, i_2 \to i_4\}/0.6 + \{i_1 \to i_3, i_2 \to \neg i_4, i_1 \to \neg i_2\}/0.5 + \{i_1 \to i_2\}/0.4, \tag{25}$$

where, between others, (25) shows that rules $i_1 \to i_3$ and $i_2 \to i_4$ exceed the imposed thresholds with evidence 0.6. It could be also helpful for the user ranking the obtained rules at each level by the relevance measure used, for instance by certainty factor, obtaining in this case that rule $i_1 \to i_3$ is more certain that $i_2 \to i_4$ with evidence 0.6. And of course, if the user is

Table 9
Measurement for rules $\{i_1,i_3\} \to \{i_4\}$ and $\{i_1,i_2\} \to \{i_3\}$ respectively seen by restriction levels joint with their summary values.

α	Supp	Conf	CF	Supp	Conf	CF
1	0	0	0	0	0	−1
0.8	0.166...	0.5	0.4	0.166...	1	1
0.6	0.166...	0.5	0.25	0.166...	1	1
0.4	0.5	0.75	0.5	0.166...	1	1
0.2	0.5	0.75	0.5	0.333...	1	1
Summary	0.266	0.5	0.33	0.167	0.8	0.6

interested in knowing the associated values for the accuracy measures they could be showed at the same time by only modifying the previous expression by the following one

$$Ruleset = \{i_1 \rightarrow i_3(0.2, 0.9), i_2 \rightarrow i_4(0.1, 0.8)\}/0.6 + \{i_1 \rightarrow i_3(0.3, 0.9), i_2 \rightarrow \neg i_4(0.1, 0.6), i_1 \rightarrow i_2(0.15, 0.7)\}/0.5 + \{i_1 \rightarrow i_2(0.25, 0.9)\}/0.4, \quad (26)$$

where the numbers in parentheses indicate the values for the accuracy measures used of the corresponding crisp rule in that level, for instance (support, CF).

Let us notice that previous expressions are not actually fuzzy sets, but still can be very helpful in order to interpret the set of results, as we relate the relevance degree and the basic probability assignment (formula (13)) for each rule or set of rules.

4. Application for managing the absence of items

Most of works in data mining search the joint occurrence of items in the database. This is the starting point in this area although some other approaches have been developed employing new techniques. A new interesting research line currently in progress is that of searching new and useful relations considering not only the presence but also the absence of items. At this respect we can distinguish several approaches that we can divide into two different groups. The first class contains those approaches that formalize the concept of absence by means of the complementary or the negation of items. In this case the set of items I contains both positive and negative items $(i_1, \neg i_1, \ldots, i_m, \neg i_m)$ where $\neg i_k$ means that i_k is not present in a transaction. The approaches belonging to the second class are those that search a group of rules with a predefined meaning considering for that the absence of items too. In this class we can stress some important types of rules like exception rules (a set of three or two rules where the second one involves the absence of an item meaning that the second rule is an exception to the first one) [37] or anomalous rules [11] (a set of three rules meaning that the second rule represents an anomalous behavior that deviates from the usual one, represented by the first rule).

In this section we explain how to apply our approach for mining fuzzy rules considering both the presence and the absence of items in the line of the first class of approaches. Nevertheless, we have also accomplished some progress in the second class when considering only crisp rules [11]. In the following, first we will review the efforts made until now in this field and then we will present our proposal for managing the absence of items using the developed formal model for fuzzy rules.

4.1. Related approaches

As we mention at the end of Section 2.1, there also exist some works that mine the so called negative rules without taking into account the negation of items. Examples of this type are the proposal in [4] where the rule is extracted if the adjusted difference measure is lesser than a fixed value, and the proposal of Tsumoto [40] that extracts negative rules coming from the contrapositive of a rule supported by all the positive examples in the data.

Some problems arise when considering the absence of items: the density of data is higher, the complexity may increase significantly in terms of the number of data items and some prune strategies used to restrict the search space and to guarantee the efficiency in classical AR mining algorithms cannot be used. This is due to the fact that the absence of items does not fulfil the upward closure property of frequent itemsets. For this reason, the authors have addressed the task of mining negative association rules from different points of view:

- Obtain negative associations using a predefined taxonomy or graph-based structure.
- Define a new measure to obtain stronger negative rules. This type of approaches take advantage of the defined measure and its properties to mine a set of stronger negative rules smaller than that obtained using only the support-confidence framework.
- Consider only those negative items whose positive are frequent in the mining process. In general, the negative items have very high support and they do not fulfil the upward closure property, so with this imposition the set of candidates to be in a rule is pruned, and therefore the number of extracted rules is substantially reduced.

The first point differs from the other two because it is based on a predefined domain knowledge which is not available in all cases. In this group we can find a novel approach presented by Savasere et al. [32,33] where the domain knowledge is given in the form of a taxonomy which is then used to mine the negative associations, and the approach of Yuan et al. [44] which employs a hierarchical graph-structured taxonomy containing classification information about the similarity between items.

In the second group we can highlight several approaches: Wu et al. [42] use a Piatetsky–Shapiro principle [25] for discarding uninteresting rules joint with a minimum interestingness threshold for a better pruning of the frequent items generated. Yan et al. [43] mine positive and negative fuzzy rules using the support-confidence framework and they incorporate in the process a new measure called *degree of implication* for measuring the relative fraction of transactions which are not negative examples of the rule. Teng et al. [38] focus in searching substitution rules (which contain negative items) employing the chi-square and the negative correlation measures.

Approaches belonging to the last group consider only those negative items whose positive is frequent in the mining process. To our knowledge this type of imposition was first used in [38] where the negative rules are mined among the set of frequent positive items, measuring their negative correlation and then considering those satisfying the support and confidence thresholds. This idea has been also employed in [2] by Antonie et al. where they first searched the frequent positive items and then used the rho correlation measure to divide the mining process in two parts: if two items X, Y are positively correlated, rules $X \to Y$ or $\neg X \to \neg Y$ can be found, and if X, Y are negatively correlated they search for the rules $\neg X \to Y$ or $X \to \neg Y$.

Han and Beheshti formalized this idea in [18,19] to mine fuzzy positive and negative rules by defining a valid negative rule $A \to \neg B$ such as the one that fulfils the following conditions:

(1) $A \cup B = \emptyset$,
(2) $\text{supp}(A) \geqslant minsupp$,
(3) $\text{supp}(B) \geqslant minsupp$,
(4) $\text{supp}(A \to \neg B) \geqslant minsupp$,
(5) $\text{Conf}(A \to \neg B) \geqslant minconf$.

The definitions for the rules $\neg A \to B$ and $\neg A \to \neg B$ are analogous.

Wang et al. [41] also propose some prune strategies by defining two new measures based on the fulfilment or not of (1)–(5). Nevertheless, in this work the authors propose an algorithm that uses X and Y and their negations directly, where I only contains positive items. In this case, the mined rules involve the disjunction of items, since for instance, if $X = i_1 \wedge i_2$ then $\neg X = \neg i_1 \vee \neg i_2$. We have to be careful in these situations where the disjunction of items appears, as the user may not be interested in such kind of rules.

4.2. Our approach for managing the absence of items in fuzzy rules

The absence of items have been formalized by considering the negation operator $\neg$ which, in this case, represents the complement of the occurrence of an item. In our model for fuzzy rules, we have considered the logic operators $\wedge$, $\vee$, $\neg$ between items that could be considered for mining them. In particular, the negation has an active role as it is used for computing the four frequencies involved in the 4*ft*-tables $\mathcal{M}_\alpha$. Then we have in the 4*ft*-tables all the information needed for both positive and negative items.

Let $I = \{i_1, \ldots, i_m, \neg i_1, \ldots, \neg i_m\}$ be a set of items and $\widetilde{D}$ be a set of fuzzy transactions, where an item i_j is satisfied in a transaction $\tilde{\tau} \in \widetilde{D}$ to a certain degree between 0 and 1, so this degree can be considered as the restriction level in which i_j is fulfilled. We want to remark that when dealing with a conjunction of two or more items, being negative at least one of them, the subsequent restriction level set may not correspond necessarily to a fuzzy set, as we commented before.

We can see that the obtained measures for negative items using the proposed model are the ones that we should expect. Let $A, B \subset I$ be two itemsets with associated 4*ft*-table for each restriction level $\mathcal{M}_{\alpha_i} = 4ft(\widetilde{\Gamma}_A, \widetilde{\Gamma}_B, \widetilde{D}) = \langle a_i, b_i, c_i, d_i \rangle$, then $4ft(\widetilde{\Gamma}_{\neg A}, \widetilde{\Gamma}_B, \widetilde{D}) = \langle c_i, d_i, a_i, b_i \rangle$:

$\mathcal{M}_{\alpha_i}$	B	$\neg B$
A	a_i	b_i
$\neg A$	c_i	d_i

	B	$\neg B$
$\neg A$	c_i	d_i
A	a_i	b_i

$$\text{supp}_{\alpha_i}(\neg A) = \frac{c_i + d_i}{a_i + b_i + c_i + d_i} = \frac{(a_i + b_i + c_i + d_i) - (a_i + b_i)}{a_i + b_i + c_i + d_i} = 1 - \text{supp}_{\alpha_i}(A), \tag{27}$$

$$(\Rightarrow_I)_{\alpha_i}(c_i, d_i, a_i, b_i) = \frac{c_i}{c_i + d_i} = \text{Conf}_{\alpha_i}(\neg A \to B). \tag{28}$$

In consequence, it is not necessary to compute extra frequencies when the negation of items is involved. Finally, let us notice that, due to the followed representation, and opposite to the fuzzy sets case, it is not possible to find itemsets containing both items i and $\neg i$ since items cannot be in a transaction and absent at the same time (as we deal with crisp transactions in each restriction level), as expected.

5. Algorithmic implementation using the formal model

The implementation is one of the crucial stages in the data mining area. Nowadays, research about new and fast algorithms for large databases is constantly growing [16,39].

The proposed model for fuzzy association rule mining provides a unified framework for the extraction of different kinds of association rules by only modifying the 4*ft*-quantifier. In particular we have applied the formal model to extract crisp exception and anomalous rules in [11] obtaining acceptable results. In the following, we present how to extract fuzzy rules by means of a parallelization of a particular crisp mining process. Then, any crisp mining algorithm could be used for mining fuzzy association rules by applying it in every restriction level in a straightforward way. Afterwards, we compute the summary measures and present the results using one of the several methods we have proposed in Section 3.3.

5.1. Algorithm and implementation issues

The algorithm we propose uses a simple variation of Apriori, modifying it for dealing with a set of items represented by means of BitSets. Previous works [23,27] have proposed another Apriori-based algorithm using a bit-string representation of items. In both cases, quite good results were obtained with respect to time. One advantage of using a bit-string representation of items is that it speeds up logical operations such as conjunction or cardinality. Instead of strings of bits, in our implementation we use the java class `java.util.BitSet` which contains the implementation of the object `BitSet` and some useful operations. The BitSet object stores a set of bits (zero or one) in each position. The main idea is to store the whole database into a vector of BitSets with size equal to the number of transactions and dimension equal to the number of items. For each transaction, a bit in the BitSet will take the value 1 if an item (or itemset when dealing with conjunction of items) is satisfied, or 0 if not. The general framework for mining fuzzy rules is described in Algorithm 1 whose main steps are:

1. A database preprocessing for transforming it into K sets of BitSets, one for each restriction level.
2. The mining process which consists in a variation of Apriori (considering the support of items) where we compute the 4*ft*-table and then the value for the 4*ft*-quantifier chosen by the user. This process is done in parallel for each restriction level.
3. We summarize the results following the preferred method by the user. In particular, this step takes the files generated in the previous step and then performs the suitable computations for showing the results using the user preferences.

Algorithm 1. Mining Fuzzy Rules

Input: Fuzzy transactional database $\widetilde{D}$, set of RLs, *minsupp*, *minCF*
Output: Set of fuzzy association rules.
1. **Database Preprocessing**
   1.1 Transform the fuzzy transactional database into k boolean databases (k being the number of RLs).
   1.2 Store the database into k vectors of BitSets, one for each RL.
2. **Mining Process For** each *RL* mine rules with support and CF greater than *minsupp* and *mincf*
   2.1 Search the set of candidates **If** i_k is a frequent item, add it to candidates. Store BitSet vector indexes associated to the frequent itemsets and their computed cardinalities.
   2.2 Search frequent l-itemsets using the previous set of candidates (C)
   2.3 Form the rules using C and compute their associated 4*ft*-table Store all the rules in a file joint with the 4*ft*-table.
3. **Summarizing the obtained results**
   3.1 Read all the found rules in every restriction level α
   3.2 Compute the corresponding summarized measures using $\mathcal{M}_\alpha$

In Algorithm 1 steps 1.1 and 1.2 are conjunctly processed in order to read the database only once. For each *RL* the created BitSet vector will have dimension equal to the number of items in the database and each element of the vector will contain a BitSet with the value one or zero in position i if the item appears or not in the ith transaction. For computing the itemsets frequencies we use the *cardinality* function implemented in the BitSet java class. Then the computation of the frequencies in the 4*ft*-table are easily done using the cardinalities of the antecedent X, the consequent Y and their conjunction $X \wedge Y$ as follows, that does not increase the computational cost of the algorithm:

$$\begin{aligned}
a &= cardinality(X \wedge Y),\\
b &= (a+b) - a = cardinality(X) - cardinality(X \wedge Y),\\
c &= (a+c) - a = cardinality(Y) - cardinality(X \wedge Y),\\
d &= (a+b+c+d) - a - b - c = |\widetilde{D}| - a - b - c.
\end{aligned}$$

In our experiments, we have considered support and the certainty factor measure given in Eq. (21), but other types of measures can also be implemented immediately by means of appropriate quantifiers collecting the type of association the user wants to extract from the data.

5.2. Complexity aspects

The proposed Algorithm 1 has three different parts, where the second one is the mining process. The overall complexity of our approach will depend on the complexity of each part. In the Apriori case the complexity of the second part is $O(n2^i)$, where n is the number of transactions and i the number of items. Our algorithm repeat the mining process for each restriction level. If we note with k the number of RLs, we have a theoretic complexity of $O(kn2^i)$. When necessary, the complexity can be reduced to $O(n2^i)$ by a suitable parallelization in every restriction level.

The third step is the most time consuming as it depends on the number of formed rules in each level. We first collect the different rules found in every level, and then we compute the measures used to resume the results.

Concerning space, the size of memory requirements for standard databases is acceptable. For a database with 61,810 transactions and 33 items the memory occupied by the vector of BitSets is around 1 MB. But as we need a different vector for each RL, the consumed memory will be also increased.

The following section will show in practice the performance of our algorithm for the particular case of considering the absence of items also in the mining process. In addition we apply our algorithm to some real and synthetic fuzzy databases.

6. Experiments considering the absence of items

There exist few approaches taking into account fuzzy negative rules. Works by Yan et al. [43] and Han and Beheshti [19] seem to be the most important ones in this field, as far as we know. Our approach follows a totally new focus as it employs the *RLR* theory to obtain a new type of fuzzy rules where the negation of items does not always match with that defined in fuzzy sets theory.

We propose a simple modification of Algorithm 1 to consider the absence of items too. The main problem is that the absence of an item tends to be always frequent as its occurrence usually has low support. Han and Beheshti [19] proposed to take only negative items whose positive is frequent [19] to reduce the total number of extracted rules. In order to see if this condition is adequate or not we have taken into account two different approaches, one that only considers those items in $I' = \{i_1,\ldots,i_m,\neg i_1,\ldots,\neg i_m\}$ that are frequent and the other following the approach developed in [19] where we only consider those negative items whose positive is frequent. For the former, we use the Algorithm 1 considering that $i_k \in I'$ and for the latter, we consider that $i_k \in I$ where $I = \{i_1,\ldots,i_k\}$ and we also impose in step 2.1 that if i_k is frequent, we add i_k and $\neg i_k$ to the set of candidates. It is worth to mention that in the mining rule process a rule of the form $i_k \rightarrow \neg i_k$ cannot be found as $i_k \wedge \neg i_k$ is the null itemset and its cardinality is equal to zero and consequently its support too (remember that in a given restriction level it is impossible that both i_k and $\neg i_k$ appear simultaneously).

Concerning the complexity, as we consider the absence of items too, we increment the complexity to $O(kn2^{2i})$ where k is the number of restriction levels, n the number of transactions and i the number of "positive" items. Again, by a suitable parallelization in each restriction level we can reduce to $O(n2^{2i})$.

6.1. Experimentation results

In our experiments we took a fixed *RL*-set $\Lambda = \{1, 0.9,\ldots,0.2, 0.1\}$ with just 10 elements which is sufficient to obtain reliable fuzzy rules. In order to prove the performance of our algorithm we have carried out several experiments with the databases in Table 10. The first database has been used in [31] to obtain interesting information about olive agriculture in the south of Spain. The second database is Auto-mpg from the known UCI machine learning repository [3], where we have fuzzified continuous attributes using the following set of linguistic labels: *low, medium* and *high*. The following three *Fresas* databases were employed in [24] for query refinements in Information Retrieval (IR) where the difference between them is in the *binary* case if we consider that an item is present or not in a document, in this case in a transaction; in the *frequency* case, the item has a frequency grade in the unit interval where its frequency is divided by the maximum frequency in the collection of documents; and in the *idf* case we have considered the inverse document frequency measure mostly used in the IR field. The *Forest Cov-type* databases are reduced versions originating from the database used in [22] where we have excluded the binary attributes and we have fuzzyfied the remaining attributes using two different methods: into equi-depth intervals and the *k*-means clustering for reducing their numeric domains to linguistic ones employing the labels *low, medium, high*. The last database is a random choice of Soil database transactions in order to obtain a fuzzy database with a higher number of transactions.

Table 10
Databases description employed in the experiments.

Fuzzy database	Size	Positive items	Potential items
Soil	541	33	66
Auto-mpg	398	39	78
Fresas (binary)	99	837	1674
Fresas (frequency)	99	837	1674
Fresas (idf)	99	837	1674
Forest Cov-type (equidepth)	581,012	37	74
Forest Cov-type (*k*-means)	581,012	37	74
Synthetic from Soil DB	99,811	33	66

Table 11
Fuzzy rules obtained in Soil (left), Auto-mpg (middle) and Synthetic Soil (right) with only one item in both parts of the rule.

Approach	*ms*	*mc**mcf*	# Rules	sec	# Rules	sec	# Rules	sec
Conf without restriction	0.01	0.8	1106	10	1408	13	1106	10
		0.9	1022	7	1402	12	1020	8
	0.05	0.8	1083	10	1377	13	1083	10
		0.9	996	8	1371	16	994	8
CF without restriction	0.01	0.8	26	1	9	1	31	1
		0.9	22	1	7	1	27	1
	0.05	0.8	24	1	9	1	29	1
		0.9	21	1	7	1	26	1
Conf with restriction	0.01	0.8	72	1	13	1	72	1
		0.9	54	1	7	1	52	1
	0.05	0.8	45	1	13	1	44	1
		0.9	28	1	7	1	26	1
CF with restriction	0.01	0.8	26	1	6	1	31	1
		0.9	22	1	4	1	27	1
	0.05	0.8	24	1	6	1	29	1
		0.9	20	1	4	1	25	1

To illustrate the complexity in time we discussed before and to show the performance of both proposals, we have carried out an extensive battery of experiments in the mentioned databases in a 2.13 GHz Intel Core i3 notebook with 4 GB of main memory, running Windows 7.

We have run diverse experiments in order to study the performance of our proposal from the following points of view:

1. Suitability of using confidence or certainty factor measures when mining negative rules.
2. Differences between the use of the imposition proposed by Han and Beheshti [18] for the confidence and the certainty factor.
3. Study of time consumed in the third part of the algorithm depending on the number of mined rules.

For comparing the behavior of the algorithm employing the confidence and the certainty factor measures we have executed the algorithm with both measures in all databases imposing only one item in both parts of the rule (antecedent and consequent) for several threshold values as we show in Tables 11–13. We refer to the imposition proposed by Han and Beheshti [18] when we mention "with restriction" in Tables 11–13. In general we observed a high decrease in the number of extracted rules when employing the certainty factor. Let us notice that, in Table 13, the consumed time in the Conf approach without imposing the restriction is so high with respect to the one in the *CF* approach, that we have not considered to include it as it takes in these cases more than a day in their execution. Anyway, in this point we are more interested in comparing the different approaches, and it is sufficient to know that the Conf approach without imposing the restriction is nonviable as it extracts too many rules and the consumed time is also disproportionate.

In addition, relative to the second point, when carrying the same experiments imposing the condition proposed by Han and Beheshti [18], i.e. the absence of items must have their positive frequent, we found that this imposition reduces the total number of mined rules but it can leave behind some interesting rules, and sometimes by only imposing a measure stronger than confidence it is enough to obtain a manageable set of rules. In particular, some of the rules that have not been found when using that imposition but employing the certainty factor instead of confidence are the following:

"*IF* height = medium
THEN ¬depth = high (FSupp = 0.119 & $\equiv_{CF}$ = 0.781)

found in Soil database. A possible meaning is that when the height is medium it tends to occur that depth is medium or low, that is, it is not high. This rule could be interesting, but as the item depth = high is not frequent, when imposing the restriction of Han and Beheshti this rule was not found. In the Auto-mpg database, we obtain the following rule:

"*IF* ¬cylinders = average
THEN ¬max_cylinders = 5 (FSupp = 0.525 & $\equiv_{CF}$ = 0.999)

which shows a logical association rule, as ¬cylinders = average can be translated to cylinders = more than 3 or less than 8 and ¬max_cylinders = 5 is equivalent to max_cylinders = 3 or 4 or 6 or 8, so if the number of cylinders is not the average then the maximum number of cylinders tends not to be 5 (which is just the average of the set {3,4,5,6,8}).

However when considering rules with more than one item in the antecedent or in the consequent, the number of rules is highly increased if we do not impose the mentioned condition as Table 14 shows.[4] Summarizing the results to answer the second point, we observed that it would be convenient to use some pruning strategy (like the one in [18]) in order to obtain a

[4] The results obtained in the Auto-mpg database has similar values to those shown in Table 14 with quantities of the same order.

Table 12
Fuzzy rules obtained in `Forest equidepth` (left) and `Forest k-means` (right) with only one item in both parts of the rule.

Approach	*ms*	*mc**mcf*	# Rules	sec	# Rules	sec
Conf without restriction	0.01	0.8	1232	8	1297	9
		0.9	1228	8	1228	8
	0.05	0.8	1232	8	1297	8
		0.9	1228	8	1228	9
CF without restriction	0.01	0.8	4	1	8	1
		0.9	4	1	7	1
	0.05	0.8	4	1	8	1
		0.9	4	1	7	1
Conf with restriction	0.01	0.8	10	1	23	1
		0.9	6	1	6	1
	0.05	0.8	10	1	23	1
		0.9	6	1	6	1
CF with restriction	0.01	0.8	4	1	8	1
		0.9	4	1	7	1
	0.05	0.8	4	1	8	1
		0.9	4	1	7	1

Table 13
Fuzzy rules obtained in `Fresas binary` (left), `Fresas frequency` (middle) and `Fresas idf` (right) with only one item in both parts of the rule.

Approach	*ms*	*mc**mcf*	# Rules	sec	# Rules	sec	# Rules	sec
Conf without restriction	0.05	0.8	≈719,000		≈702,000		≈701,000	
		0.9	≈719,000		≈701,000		≈700,000	
	0.1	0.8	≈702,000		≈700,000		≈700,000	
		0.9	≈699,000		≈699,000		≈699,000	
CF without restriction	0.05	0.8	22,436	4724	756	2011	469	1442
		0.9	22,435	4754	749	1993	469	1477
	0.1	0.8	12,815	922	577	461	408	383
		0.9	12,814	928	570	457	408	375
Conf with restriction	0.05	0.8	1447	18	36	7	15	4
		0.9	1291	18	22	7	10	4
	0.1	0.8	53	1	13	1	8	1
		0.9	30	1	6	1	6	1
CF with restriction	0.05	0.8	611	5	15	2	10	1
		0.9	610	6	15	2	10	1
	0.1	0.8	21	1	6	1	6	1
		0.9	21	1	6	1	6	1

Table 14
Fuzzy rules obtained in the `Soil DB` with two items in the antecedent and/or the consequent of the rule.

Approach	*ms*	*mc*/*mcf*	Database	# Rules
Conf without restriction			Soil	≈280 000
	0.05	0.9	auto-mpg	≈555 000
			Forest equidepth	≈388 000
			Forest *k*-means	≈385 000
CF without restriction			Soil	16,940
	0.05	0.9	auto-mpg	≈44,000
			Forest equidepth	≈23,000
			Forest *k*-means	≈22,000
Conf with restriction			Soil	317
	0.05	0.9	auto-mpg	45
			Forest equidepth	81
			Forest *k*-means	20
CF with restriction			Soil	283
	0.05	0.9	auto-mpg	26
			Forest equidepth	65
			Forest *k*-means	5

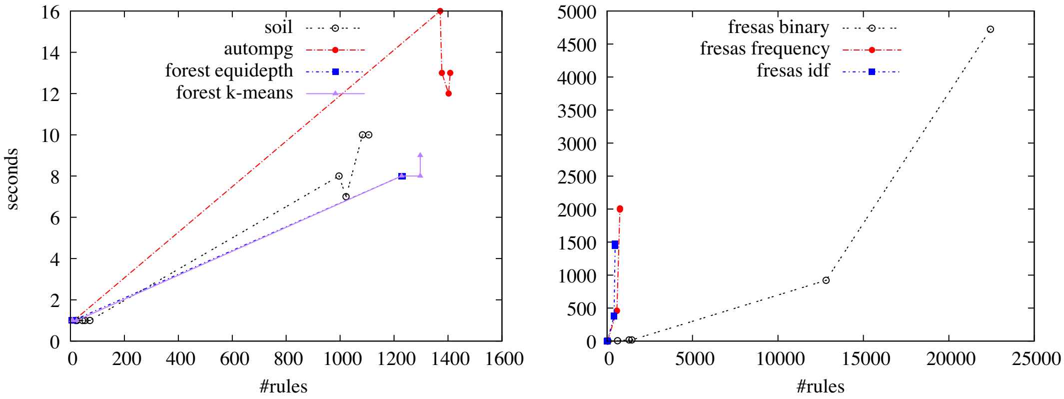

Fig. 1. Time in seconds depending of the number of extracted rules in several experiments with different threshold impositions.

Table 15
Experiment results for an Apriori algorithm with *minsupp* = 0 and *minCF* = 0.1.

Database	Rules	sec
Soil	4224	2.5
Auto-mpg	5928	4.2
Forest equidepth	5328	125.6
Forest *k*-means	5328	127.6
Fresas binary	2,798,928	376667.3
Fresas frequency	2,798,928	386246.9
Fresas idf	2,798,928	389938.3

manageable set of rules in some cases, for instance for a high number of items (like the *Fresas* databases). In other cases it would be sufficient to use a strong accuracy measure with appropriate properties like the certainty factor.

Regarding the last point, we have seen that the most consuming time part of our algorithm is its third part which summarizes the obtained results. It strongly depends on the total number rules that must be processed for obtaining the summary. In Fig. 1 we show some of the times produced in our experimentation in function of the number of rules obtained in the summary. When the number of rules to be processed exceeds 100000 in every level, the time is highly increased. As, in real world cases, the user is usually interested in obtaining a reduced set of rules for being able to manage, then in these cases we think that the obtained times are reasonable.

We have also performed some experiments in a Xeon Core2Duo (4 kernels) computer with 8 GB of RAM memory running under Fedora 12 (64 bits architecture). We have implemented a basic Apriori algorithm to obtain the results for each restriction level and afterwards we resume the results. We considered the support and the certainty factor accuracy measures with associated thresholds 0 and 0.1 respectively in order to study the time requirements in a different algorithm. Results are shown in Table 15.

7. Conclusions

We have presented an extension of the GUHA model to formalize fuzzy association rules using the RLR theory. This proposal offers a unified view for managing fuzzy rules via restriction levels. It also gives a process for extending the validation measures used in the extraction process from the crisp to the fuzzy case (Eq. (13)). In addition, we have developed several alternatives for managing the obtained results using our model that allows the user to explore the extracted rules from different points of view. In particular, we have applied the model for managing the absence of items, proving that it does not entail extra computations, and the involved measures are the logical extensions when the operator $\neg$ is used.

We have proposed a general algorithm following the model philosophy and using the BitSet computation which accelerates the logical computations between itemsets. This algorithm has been applied for mining fuzzy rules involving both the presence and the absence of items obtaining reasonably time performance in some real and synthetic fuzzy databases. We want to remark that although we have offered an algorithmic approach, the proposed formalization provides a method for using every crisp rule mining algorithm for mining fuzzy rules by only parallelizing the mining process in each restriction level and then summarizing the obtained results.

We have also compared our approach with that of Han and Beheshti [18] which imposes that the negation of an item must have a frequent positive to consider it a candidate. We obtained that the imposition of a stronger measure than confidence, the certainty factor in our case, could be an alternative to that imposition when rules with single items in antecedent and consequent are considered. The more items appear in one of the rule sides, the more the imposition drastically reduces the number of extracted rules. At this respect improvements in this field are needed, opening a line for future studies: to provide stronger measures, different impositions or distinct methods to obtain a suitable set of rules involving the "negation" of items.

Also in this way, we think that a promising line is to study a predefined group of rules involving the absence of items having significant semantics for the user. Some efforts have been already done when mining exceptions or anomalous rules [11]. Another interesting type of rule is that proposed in [38] which has a substitutive meaning by using the negation in the rule consequent.

Acknowledgements

The research reported in this paper was partially supported by the Andalusian Government (Junta de Andalucía) under project P07-TIC03175 and from the Spanish Ministry for Science and Innovation by the project Grants TIN2006-15041-C04-01 and TIN2009-08296.

References

[1] R. Agrawal, T. Imielinski, A. Swami, Mining associations between sets of items in massive databases, in: ACM-SIGMOD International Conference on Data, 1993, pp. 207–216.
[2] M.L. Antonie, O.R. Zaïane, Mining positive and negative association rules: an approach for confined rules, in: European PKDD Conference, 2004, pp. 27–38.
[3] A. Asuncion, D.J. Newman, UCI machine learning repository, 2007.
[4] W.H. Au, K.C.C. Chan, An effective algorithm for discovering fuzzy rules in relational databases, in: Proceedings of the 7th IEEE Internat Conference on Fuzzy Systems, 1998, pp. 1314–1319.
[5] F. Berzal, I. Blanco, D. Sánchez, M.A. Vila, A new framework to assess association rules, LNCS 2189 (2001) 95–104.
[6] F. Berzal, M. Delgado, D. Sánchez, M.A. Vila, Measuring accuracy and interest of association rules: a new framework, Intelligent Data Analysis 6 (3) (2002) 221–235.
[7] S. Brin, R. Motwani, J.D. Ullman, S. Tsur, Dynamic itemset counting and implication rules for market basket data, SIGMOD Record 26 (2) (1997) 255–264.
[8] M. De Cock, C. Cornelis, E.E. Kerre, Elicitation of fuzzy association rules from positive and negative examples, Fuzzy Sets and Systems 149 (1) (2005) 73–85.
[9] M. Delgado, N. Marín, D. Sánchez, M.A. Vila, Fuzzy association rules: general model and applications, IEEE Transactions on Fuzzy Systems 11 (2) (2003) 214–225.
[10] M. Delgado, M.D. Ruiz, D. Sánchez, A restriction level approach for the representation and evaluation of fuzzy association rules, in: Proceedings of the IFSA-EUSFLAT, Lisbon, Portugal, 2009, pp. 1583–1588.
[11] M. Delgado, M.D. Ruiz, D. Sánchez, New approaches for discovering exception and anomalous rules, International Journal of Uncertanity, Fuzziness and Knowledge-Based Systems 19 (2) (2011) 361–399.
[12] M. Delgado, M.D. Ruiz, D. Sánchez, Studying interest measures for association rules through a logical model, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 18 (1) (2010) 87–106.
[13] D. Dubois, E. Hüllermeier, H. Prade, A note on quality measures for fuzzy association rules, in: Proceedings IFSA03, 10th International Fuzzy Systems Association World Congress, No. 2715 in Lecture Notes in Artificial Intelligence, 2003, p. 677.
[14] D. Dubois, E. Hüllermeier, H. Prade, A systematic approach to the assessment of fuzzy association rules, Data Mining Knowledge Discovery 13 (2) (2006) 167–192.
[15] D. Dubois, H. Prade, Gradual elements in a fuzzy set, Soft Computing 12 (2008) 165–175.
[16] D. Dudek, RMAIN: association rules maintenance without reruns through data, Information Sciences 179 (24) (2009) 4123–4139.
[17] P. Hájek, I. Havel, M. Chytil, The GUHA method of automatic hypotheses determination, Computing 1 (1966) 293–308.
[18] J. Han, M. Beheshti, Discovering both positive and negative fuzzy association rules in large transaction databases, Journal of Advanced Computational Intelligence and Intelligent Informatics 10 (3) (2006) 287–294.
[19] J. Han, M. Beheshti, Mining fuzzy association rules: Interestingness measure and algorithm, in: IEEE International Conference on Granular Computing, 2006, pp. 659–662.
[20] T-P Hong, Y-C Lee, Fuzzy Sets and Their Extensions: Representation, Aggregation and Models, Chapter: An Overview of Mining Fuzzy Association Rules, Springer, 2008. pp. 397–410.
[21] E. Hüllermeier, Fuzzy sets in machine learning and data mining, Applied Soft Computing 11 (2) (2011) 1493–1505.
[22] H. Liu, F. Hussain, C.L. Tan, M. Dash, Discretization: an enabling technique, Data Mining and Knowledge Discovery 6 (2002) 393–423.
[23] E. Louie, T.Y. Lin, Finding association rules using fast bit computation: machine-oriented modeling, LNAI 1932 (2000) 486–494.
[24] M.J. Martín-Bautista, D. Sánchez, J. Chamorro-Martínez, J.M. Serrano, M.A. Vila, Mining web documents to find additional query terms using fuzzy association rules, Fuzzy Sets and Systems 148 (1) (2004) 85–104.
[25] G. Piatetsky-Shapiro, Discovery, analysis and presentation of strong rules, Knowledge Discovery in Databases (1991) 813–818.
[26] J. Rauch, Logic of association rules, Applied Intelligence 22 (2005) 9–28.
[27] J. Rauch, M. Šimunek, An alternative approach to mining association rules, Studies in Computational Intelligence (SCI) 6 (2005) 211–231.
[28] Jan Rauch, Classes of association rules: an overview, Studies in Computational Intelligence 118 (2008) 315–337.
[29] D. Sánchez, M. Delgado, M.A. Vila, RL-numbers: an alternative to fuzzy numbers for the representation of imprecise quantities, in: IEEE International Conference on Fuzzy Systems, 2008, pp. 2058–2065.
[30] D. Sánchez, M. Delgado, M.A. Vila, J. Chamorro-Martínez, On a non-nested level-based representation of fuzziness. Fuzzy Sets and Systems (2011), in press, doi:10.1016/j.fss.2011.07.022.
[31] D. Sánchez, J.M. Serrano, M.A. Vila, M. Delgado, G. Calero, J. Sánchez, V.M. Aranda, Building a fuzzy logic information network and a decision-support system for olive cultivation in Andalusia, Spanish Journal of Agricultural Research 6 (2008) 252–263.
[32] A. Savasere, E. Omiecinski, S. Navathe, An efficient algorithm for mining association rules in large databases, in: Proceedings of the 21st Conference on Very Large Databases, Zürich, Switzerland, 1995, pp. 432–444.
[33] A. Savasere, E. Omiecinski, S. Navathe, Mining for strong negative associations in a large database of customer transactions, in: ICDE, IEEE Computer Society, 1998, pp. 494–502.
[34] E. Shortliffe, B. Buchanan, A model of inexact reasoning in medicine, Mathematical Biosciences 23 (1975) 351–379.
[35] C. Silverstein, S. Brin, R. Motwani, Beyond market baskets: generalizing association rules to dependence rules, Data Mining and Knowledge Discovery 2 (1998) 39–68.

[36] T. Sudkamp, Examples, counterexamples, and measuring fuzzy associations, Fuzzy Sets and Systems 149 (1) (2005) 57–71.
[37] E. Suzuki, Discovering interesting exception rules with rule pair, in: Proceedings of the Workshop on Advances in Inductive Rule Learning at PKDD-04, 2004, pp. 163–178.
[38] W. Teng, M. Hsieh, M. Chen, On the mining of substitution rules for statistically dependent items, in: Proceedings of the ICDM, 2002, pp. 442–449.
[39] Y.J. Tsaya, T.J. Hsua, J.R. Yub, FIUT: a new method for mining frequent itemsets, Information Sciences 179 (11) (2009) 1724–1737.
[40] S. Tsumoto, Automated discovery of positive and negative knowledge in clinical databases, IEEE Engineering in Medicine and Biology Magazine 19 (4) (2000) 56–62.
[41] H. Wang, X. Zhang, G. Chen, Mining a complete set of both positive and negative association rules from large databases, in: Proceedings of the PAKDD 2008 Lecture Notes in Artificial Intelligence, 5012, 2008, pp. 777–784.
[42] X. Wu, C. Zhang, S. Zhang, Mining both positive and negative association rules, in: Proceedings of ICML, 2002, pp. 658–665.
[43] P. Yan, G. Chen, C. Cornelis, M. De Cock, E. Kerre, Mining positive and negative fuzzy association rules, Knowledge-Based Intelligent Information and Engineering Systems (KES'2004) 3213 (2004) 270–276.
[44] X. Yuan, B. Buckles, Z. Yuan, J. Zhang, Mining negative association rules, in: Proceedings of ISCC, 2002, pp. 623–629.
[45] L.A. Zadeh, Fuzzy sets, Information and Control 8 (1965) 338–353.