A Fuzzy Residuated Approach to Case-Based Reasoning

Sandra Sandri

Instituto Nacional de Pesquisas Espaciais - INPE
CP 515, 12201-970 - São José dos Campos - SP, Brazil

Abstract. This paper addresses the use of residuated implication operators to create a fuzzy resemblance relation between cases so as to model the CBR basic principle "the more similar two problem descriptions are, the more similar are their solutions". We describe how this fuzzy relation can be exploited to identify case clusters, based of a finite number of level cuts from that relation, that are in turn used to solve a new problem. The paper proposes some formal results that characterize the sets of clusters obtained from the various level-cuts of the resemblance relation.

Keywords: case-based reasoning, residuated implication, similarity relations.

1 Introduction

Case based reasoning (CBR) [10,1] proposes to solve a problem using a principle that can be stated as "similar problems have similar solutions" [1]: it is based on a two-step procedure that exploits a base of already solved problems; the couple made of a problem and its solution is called a case. The first step of this procedure consists in retrieving problems in the base that are similar to the considered problem: it determines the cases in the base that are relevant to solving the problem at hand. The second step consists in reusing the solutions of these relevant problems, adapting them to the considered problem.

In this paper, we mainly focus on the first step and propose to address this task by combining two principles: on the one hand, the exploitation of a cluster decomposition of the solved problems and on the other hand, the definition of a fuzzy relation between solved problems. We use a weighted hypergraph as formalization of this fuzzy relation and present its exploitation in this framework.

Cluster decomposition in this CBR context has been introduced by [8]; it is based on a binary similarity measure between cases, called Case Resemblance Relation, that takes into account both their resemblances in the problem description space and in the solution description space. This measure defines a binary relation between cases; the corresponding graph of cases is then exploited and decomposed to identify clusters of similar cases.

Here we generalize this method to the case where the similarity measure between cases, that aggregates the similarities in the problem and the solution spaces, is not binary but takes values in $[0, 1]$, leading to a fuzzy relation instead of a crisp one. The problem is then to extract clusters from this fuzzy relation. To deal with this problem,

S. Greco et al. (Eds.): IPMU 2012, Part I, CCIS 297, pp. 51–60, 2012.

we propose to first obtain the relevant level-cuts from the fuzzy relation, thus creating a set of crisp relations. We then extract a set of clusters from the graph induced by the crisp relation.

We propose a principled definition of such similarity measures, based on the formulation of the CBR principle as a gradual rule. Indeed, the basic principle "similar problems have similar solutions" [1] can alternatively be formulated as "the more similar the problem descriptions are, the more similar the solution descriptions are", which belongs to the framework of gradual rules [2,9].

Fuzzy gradual rules, initially introduced by [5], have been studied from different points of view (see [3] for a survey): on one hand, they are interpreted as a fuzzy generalization of association rules, where the presence of one attribute must imply the presence of a second attribute, for each data point individually, and with a fuzzy residuated implication [5]. A more recent approach proposes to interpret fuzzy gradual rules as attribute co-variations, i.e. a global gradual tendency across the data, taking into account all points simultaneously ([9], see also [3]). In this paper we are interested in the first approach based on a fuzzy implication interpretation. More precisely, we interpret the measure used by [8] as the particular case where the Rescher-Gaines implication is used and we propose to generalize the approach to other operators, leading to fuzzy relations between cases.

The use of fuzzy sets theory in CBR has been addressed in the past by many authors. Which kinds of fuzzy rules should be used to model the CBR hypothesis is discussed, e.g., in [7]. The authors address the question of which subset of a case base should be used to derive a solution to a problem, given a set of reflexive and symmetric fuzzy relations for the problem and solution spaces. They claim that gradual rules should not be used when there exists two cases whose solutions are more similar than their problem descriptions, and possibility or certainty rules should be used instead. They also point out to the problem of ill-specified similarity relations, and propose the use of modifiers, for both gradual and non-gradual rules. In [6], the authors address fuzzy similarity-based models in case-based reasoning, giving distinct interpretations for the CBR hypothesis, to different constraints imposed on the relationship between similarities in the problem description and solution spaces Attaching individual attribute weight vectors to cases has been shown to be advantageous when similarity relations are used in CBR [11,12]. However, the cost of learning the weight vectors can be prohibitive.

The present work is a step in a global project of allowing weight vectors learning to become feasible also in large case bases, by first finding case base fragments in which the weight learning algorithm can be applied. Then, we can derive the answer to any problem presented to the case base using the weighted fragments. The approach proposed in [8], extended here, comes down to extracting the fragments that are consistent with the CBR principle considering the original relations. The extension aims at providing a larger set of possible frameworks to deal with specific applications.

This paper is organized as follows. In Section 2 we give some basic definitions and notation and in Section 3 we describe the original crisp approach. In Section 4 we present our extended fuzzy approach and discuss some of its properties. Section 5 finally brings the conclusions.

2 Basic Definitions and Notations

In this section, we recall some basic definitions that are used in the rest of the paper and provide some notation, successively concerning residuated implications, similarity relations, hypergraphs and imprecise partitions.

Given a left-continuous t-norm $\top$, a *residuated implication operator* $\rightarrow_\top$ is defined as $x \rightarrow_\top y = \sup_{z \in [0,1]} \top(x, z) \leq y$. Some well-known examples include

- the Gödel implication, residuum of $\top = \min$, defined as $x {\rightarrow_\top}_G y = 1$, if $x \leq y$, and y, otherwise;
- the Goguen implication, defined as $x {\rightarrow_\top}_\Pi y = 1$, if $x \leq y$, and y/x, otherwise.

The Rescher-Gaines implication operator, defined as $x {\rightarrow_\top}_{RG} y = 1$ if $x \leq y$, and 0, otherwise, is not a residuated operator itself but is the point-wise infimum of all residuated implications.

A *similarity relation* S on a domain U is a binary fuzzy relation, i.e. a mapping $S : U \times U \rightarrow [0, 1]$ that is both reflexive and symmetric. Some authors require it also satisfies the t-norm transitivity property, but we do not take it into consideration here as it does not play a role in our framework.

The set of similarity relations on a given domain U forms a lattice (not linearly ordered) with respect to the point-wise ordering (or fuzzy-set inclusion) relationship. The top of the lattice is the similarity S_{top} which makes all the elements in the domain maximally similar: $S_{top}(x, y) = 1$, for all $x, y \in U$. The bottom of the lattice S_{bot} is the classical equality relation: $S_{bot}(x, y) = 1$, if $x = y$, and 0, otherwise.

Particularly useful are families of parametric similarity relations $\mathcal{S} = \{S_0, S_{+\infty}\} \cup \{S_\beta\}_{\beta \in I \subseteq (0,+\infty)}$ that are such that: (i) $S_0 = S_{bot}$, (ii) $S_{+\infty} = S_{top}$, and (iii) $\beta < \beta'$, then $S_\beta \prec S_{\beta'}$, where $S \prec S'$ means $\forall x, y \in U, S(x, y) \leq S'(x, y)$ and $\exists x_0, y_0 \in U, S(x_0, y_0) < S'(x_0, y_0)$.

A *hypergraph* is a generalization of a non-directed graph, where edges can connect any number of vertices. Formally, it can be represented as a pair, $H = (N, E)$, where N is a set containing the vertices (nodes) and E is a set of non-empty subsets of N, called hyperedges. The set of hyperedges E is thus a subset of $2^N \backslash \emptyset$, where 2^N is the power set of N. An "ordinary graph" is then a hypergraph in which all hyperedges have at most two elements.

Given a hypergraph $H = (N, E)$, a hyperedge $h \in E$ is said to be maximal when $\nexists\, h' \in E$, such that $h \subseteq h'$ and $h' \neq h$. Each hyperedge in E is a clique, therefore, the set of maximal hyperedges is the set of maximal cliques of E.

Let B be a subset of a domain U. We propose the definition of an *imprecise partition* of B as a set $\mathcal{I}_B = \{B_1, ..., B_k\}$, $B_i \subseteq B$, $B_i \neq \emptyset$, such that

- $\bigcup_{i=1,k} B_i = B$ and
- $\nexists\, B_i, B_j \in B$, such that $B_i \subseteq B_j$ and $i \neq j$.

Each $B_i \in B$ is called an *imprecise class*. An imprecise partition does not allow one class to be contained inside another one but, contrary to precise partitions, it does allow non-empty intersections between classes. Let $\mathcal{I}'_B$ and $\mathcal{I}''_B$ be two imprecise partitions of B. $\mathcal{I}'_B$ is said to be *finer* than $\mathcal{I}''_B$ ($\mathcal{I}'_B \preccurlyeq \mathcal{I}''_B$), when $\forall h' \in \mathcal{I}'_B, \exists h'' \in \mathcal{I}''_B$ such that $h' \subseteq h''$. Reciprocally, $\mathcal{I}''_B$ is said to be *coarser* than $\mathcal{I}'_B$.

3 Original Crisp Framework for Case Resemblance Hypergraph

In the following we describe the crisp approach, taking many of the definitions and notations from [8].

3.1 Basic Definitions

A case is defined as an ordered pair $c = (p, o) \in P \times O$ where p is the description of the solved problem and o the description of its solution. $P = P_1 \times ... \times P_n$ and O are respectively the problem description and the solution spaces.

Let $S_{out} \subseteq O^2$ denote the overall similarity relation on O. Let $S_{in} \subseteq P^2$ denote the overall similarity relation on the problem space. S_{in} may obtained by using a suitable aggregation function (e.g. means, t-norms, t-conorms, OWA operators, etc) applied on the set of similarity relations $\{S_1, ..., S_n\}$, each of which corresponding to a description variable.

3.2 Obtaining Clusters from a Case Base

Let $c_a = (p_a, o_a)$ and $c_b = (p_b, o_b)$ denote two cases in C. The crisp case resemblance relation S_{res} is defined as

$$S_{res}(c_a, c_b) = \begin{cases} 1, \text{if } 0 < S_{in}(p_a, p_b) \leq S_{out}(o_a, o_b) \\ 0, \text{otherwise} \end{cases} \tag{1}$$

Based on this relation, the case set can be organized through a decomposition in clusters based on this resemblance relation. Several (possibly intersecting) clusters of cases can be obtained from S_{res}.

A hypergraph $H = (C, E), E \subseteq C^2$, is said to *be compatible with* S_{res} if and only if it obeys the following conditions:

- $\forall c_a, c_b \in C$, if $S_{res}(c_a, c_b) = 1$, then $\exists h \in E$, such that $\{c_a, c_b\} \subseteq h$.
- $\forall c_a, c_b \in C$, if $S_{res}(c_a, c_b) = 0$, then $\nexists h \in E$, such that $\{c_a, c_b\} \subseteq h$.

A notable hypergraph $H = (C, E)$ compatible with S_{res} is the one in which E contains the maximal cliques of S_{res}, thus constituting an imprecise partitioning of case base C.

3.3 Computing a Solution to a New Problem According to a Cluster

Given a case base, a similarity relation for each variable, global similarity measures S_{in} and S_{out}, and a hypergraph $H = (C, E)$ compatible with S_{res}, we want to derive an appropriate solution o^* for a new problem description p^*.

First of all, we gather the clusters in H, containing cases whose problem descriptions are somewhat similar to p^*, denoted $E^* = \{h \in E \mid \exists c_i = (p_i, o_i) \in h, S_{in}(p_i, p^*) > 0\}$. For each $h = \{c_1, ..., c_r\} \in E^*$, we then compute its corresponding solution for p^*, denoted by o^*_h, using a suitable aggregation function. Weighted aggregation operators (eg. weighted means, weighted t-norms, etc) can be used when the the solution variables are numerical, using the similarity between each p_i and p^* as weights, considering the cases (p_i, o_i) in h. For non-numerical variables, a weighted voting method can be used.

3.4 Determining Cluster Strength in Relation to a New Problem

In [8] the final solution o^* is selected from O^*, the set of solutions for p^* from the clusters in E^*. To guide this selection, for each cluster in E^*, we calculate the strength of that cluster in relation to p^*. We take the solution from the cluster whose descriptions are the most strongly related to p^*.

The *cluster strength* of $h = \{c_1, ..., c_r\}$, $c_i = (p_i, o_i)$, in relation to a problem p^* can be obtained applying a suitable aggregation operator (e.g. means, t-norms, etc) on the $S_{in}(p_i, p^*)$ values.

3.5 Using Attribute Weights

Weights can be attached to cases, so that cases considered more important for a given application have higher weights. Weight vectors can also be assigned to the description variables in a case: one can use either the same weight vector for all cases, or assign individual weight vectors to each case, so that more significant attributes inside a case receive higher weights (see [14] for an individual weight vector learning algorithm).

A weighted version of S_{in} can also be used to compute the clusters themselves. If individual weight vectors are used, the resulting relation is possibly asymmetric and one has to make it symmetric before applying the rest of the formalism (see [8]).

3.6 Example

Table 1 illustrates S_{in} and S_{out} for a simple case base as well as its computed S_{res} relation. As the relations are symmetric by definition, only the upper halves are shown. The maximal hypergraph for S_{res} is calculated as $H = (C, E)$, with $E = \{$\{a, b, e, f\}, \{a, d, e, f\}, \{a, g\}, \{d, h\}, \{g, h\}, \{c, d, e, f\}, \{c, g\}$\}$.

Table 1. Illustrative relations S_{in} (table i), S_{out} (table ii) and the resulting S_{res} (table iii)

i	p_a	p_b	p_c	p_d	p_e	p_f	p_g	p_h
p_a	1	0.80	0.00	0.07	0.49	0.19	0.07	0.96
p_b		1	0.00	0.00	0.35	0.02	0.00	0.76
p_c			1	0.51	0.20	0.47	0.33	0.00
p_d				1	0.37	0.80	0.81	0.07
p_e					1	0.49	0.38	0.45
p_f						1	0.63	0.19
p_g							1	0.08
p_h								1

ii	o_a	o_b	o_c	o_d	o_e	o_f	o_g	o_h
o_a	1	1	1	1	1	1	0.47	0.27
o_b		1	1	1	1	1	0.39	0.19
o_c			1	1	1	1	0.77	0.57
o_d				1	1	1	0.43	0.23
o_e					1	1	0.35	0.15
o_f						1	0.35	0.15
o_g							1	0.80
o_h								1

iii	c_a	c_b	c_c	c_d	c_e	c_f	c_g	c_h
c_a	1	1	0	1	1	1	1	0
c_b		1	0	0	1	1	0	0
c_c			1	1	1	1	1	0
c_d				1	1	1	0	1
c_e					1	1	0	0
c_f						1	0	0
c_g							1	1
c_h								1

Note that we have cases in the example above which would not be related through S_{res}. For example, $S_{res}(c_b, c_h) = 0$ because $S_{in}(p_b, p_h) = .76 > S_{out}(o_b, o_h) = .19$. However, using Gödel residuated operator we have $F_G(c_b, c_h) = .19$, so these cases will be considered as resembling each other on all level cuts $F_{G,\alpha}$, for $\alpha \in (0, .19]$.

4 Generalized Fuzzy Framework

In this section, we propose to extend the approach presented in the previous section, to a generalized fuzzy framework. More precisely, we propose to fuzzify the notion of case resemblance relation S_{res} described in Section 3.2. We then examine how the subsequent processing steps can be modified to perform case-based reasoning from a fuzzy case resemblance relation.

4.1 Fuzzy Case Resemblance Relation

As exposed in the introduction, we propose to modify the similarity measure between cases, so as to model a gradual formalization of the basic CBR principle, as "the more similar two problem descriptions are, the more similar are their solutions".

We consider a residuated implication operator ϕ (see Section 2) and define a *Fuzzy Case Resemblance Relation* (FCRR) as the mapping

$$\begin{array}{ll} F_\phi : & C^2 \rightarrow [0,1] \\ & (c_a, c_b) \mapsto F_\phi(c_a, c_b) = \begin{cases} 0, & \text{if } S_{in}(p_a, p_b) = 0 \\ \phi(S_{in}(p_a, p_b), S_{out}(o_a, o_b)), & \text{otherwise} \end{cases} \end{array} \tag{2}$$

Note that (as with S_{res}), the first condition in the definition of F_ϕ is necessary, otherwise two cases would be considered completely similar while having completely dissimilar problem descriptions.

4.2 Adaptation of the Methodology to the Fuzzy Approach

Contrary to S_{res}, the new case resemblance relation F_ϕ is not necessarily crisp and thus requires an adaptation of both the cluster decomposition of the case base and the computation of the solution to a new problem.

The proposed adaptation of the crisp methodology relies on the α level cut decomposition of the fuzzy case resemblance relation. We propose to derive hypergraphs from a Crisp Case Resemblance Relation (CCRR), defined as

$$\forall \alpha \in (0,1], F_{\phi,\alpha}(c_i, c_j) = \begin{cases} 1, \text{if } F_\phi(c_i, c_j) \geq \alpha \\ 0, \text{otherwise.} \end{cases} \tag{3}$$

Even though an infinite number of level cuts can be derived from FCRR F_ϕ, we only need a finite number of CCRRs $F_{\phi,\alpha}$, one for each distinct α greater than 0 in F_ϕ.

The other issues mentioned in Section 3 can then be dealt with in a straightforward manner. It suffices to make $S_{res} = F_{\phi,\alpha}$, discarding definition (1), and apply the procedures given in the previous section.

4.3 Example

Table 2 shows the fuzzy case resemblance relations $F_{G,\alpha}$ and $F_{\Pi,\alpha}$, obtained by the application of Gödel and Goguen implications, respectively, on the data in Table 1. In

Table 2. CCRRs based on FCRRs F_G and F_Π, for S_{in} and S_{out} given in Table 1

F_G	c_a	c_b	c_c	c_d	c_e	c_f	c_g	c_h
c_a	1	1	0	1	1	1	1	0.27
c_b		1	0	0	1	1	0	0.19
c_c			1	1	1	1	1	0
c_d				1	1	1	0.43	1
c_e					1	1	0.35	0.15
c_f						1	0.35	0.15
c_g							1	1
c_h								1

F_Π	c_a	c_b	c_c	c_d	c_e	c_f	c_g	c_h
c_a	1	1	0	1	1	1	1	0.28
c_b		1	0	0	1	1	0	0.25
c_c			1	1	1	1	1	0
c_d				1	1	1	0.53	1
c_e					1	1	0.92	0.33
c_f						1	0.55	0.79
c_g							1	1
c_h								1

our example, the relevant values of α for F_G and F_Π are respectively $A_G = \{.15, .19, .27, .35, .43, 1\}$ and $A_\Pi = \{.25, .28, .33, .53, .55, .79, .92, 1\}$.

Table 3 presents the imprecise partitions $E_{G,\alpha}$ generated from the CCRRs $F_{\phi,\alpha}, \alpha \in A_\phi$, derived from FCRRs F_G and F_Π. For example, using Gödel residuated operator with $\alpha = .19$, we obtain hypergraph $H_{G,.19} = (C, E_{G,.19})$ from $F_{G,.19}$, with $E_{G.19} = \{\{a, b, e, f\}, \{a, b, h\}, \{a, d, e, f, g\}, \{a, d, g, h\}, \{c, d, e, f, g\}\}$.

Table 3. Maximal cliques obtained for each relevant $F_{G,\alpha}$ and $F_{\Pi,\alpha}$ level cut from Table 2

α	$E_{G,\alpha}$				
1	$\{\{a, b, e, f\}$,		$\{a, d, e, f\}, \{a, g\}$,	$\{d, h\}, \{g, h\}$,	$\{c, d, e, f\}, \{c, g\}\}$
.43	$\{\{a, b, e, f\}$,		$\{a, d, e, f\}, \{a, d, g\}$,	$\{d, g, h\}$,	$\{c, d, e, f\}, \{c, d, g\}\}$
.35	$\{\{a, b, e, f\}$,		$\{a, d, e, f, g\}$,	$\{d, g, h\}$,	$\{c, d, e, f, g\}\}$
.27	$\{\{a, b, e, f\}$,		$\{a, d, e, f, g\}$,	$\{a, d, g, h\}$,	$\{c, d, e, f, g\}\}$
.19	$\{\{a, b, e, f\}$,	$\{a, b, h\}$,	$\{a, d, e, f, g\}$,	$\{a, d, g, h\}$,	$\{c, d, e, f, g\}\}$
.15	$\{\{a, b, e, f, h\}$,		$\{a, d, e, f, g, h\}$,		$\{c, d, e, f, g\}\}$

α	$E_{\Pi,\alpha}$			
1	$\{\{a, b, e, f\}$,	$\{a, d, e, f\}, \{a, g\}$,	$\{d, h\}, \{g, h\}$,	$\{c, d, e, f\}, \{c, g\}\}$
.92	$\{\{a, b, e, f\}$,	$\{a, d, e, f\}, \{a, e, g\}$,	$\{d, g\}, \{g, h\}$,	$\{c, d, e, f\}, \{c, e, g\}\}$
.79	$\{\{a, b, e, f\}$,	$\{a, d, e, f\}, \{a, e, g\}$,	$\{d, g\}, \{f, g\}, \{g, h\}$,	$\{c, d, e, f\}, \{c, e, g\}\}$
.55	$\{\{a, b, e, f\}$,	$\{a, d, e, f\}, \{a, e, f, g\}$,	$\{d, f, g\}, \{f, g, h\}$,	$\{c, d, e, f\}, \{c, e, f, g\}\}$
.53	$\{\{a, b, e, f\}$,	$\{a, d, e, f, g\}$,	$\{d, f, g, h\}$,	$\{c, d, e, f, g\}\}$
.33	$\{\{a, b, e, f\}$,	$\{a, d, e, f, g\}$,	$\{d, e, f, g, h\}$,	$\{c, d, e, f, g\}\}$
.28	$\{\{a, b, e, f\}$,	$\{a, d, e, f, g, h\}$,		$\{c, d, e, f, g\}\}$
.25	$\{\{a, b, e, f, h\}$,	$\{a, d, e, f, g, h\}$,		$\{c, d, e, f, g\}\}$

Note that, given a FCRR F_ϕ, many clusters appear in hypergraphs derived from different CCRRs $F_{\phi,\alpha}$. In the example shown here, we see that the total number of clusters using Gödel (respec. Goguen) operator is 29 (respec. 44) but only 16 (respec. 20) of them are distinct.

4.4 Some Properties of the Fuzzy Approach

Hypergraphs Hierarchy. The maximal hypergraphs $H_{\phi,\alpha} = (C, E_{\phi,\alpha})$, generated from the level cuts $F_{\phi,\alpha}$ of a given FCRR F_ϕ, are nested:

$$\text{if } \alpha \geq \alpha' \text{ then } E_{\phi,\alpha} \preccurlyeq E_{\phi,\alpha'}$$

(see Section 2 for the definition of $\preccurlyeq$). $E_{\phi,\alpha}$ is thus an imprecise partition finer than $E_{\phi,\alpha'}$.

The proof is simple, based on the fact that similarity relations $F_{\phi,\alpha}$ are such that $(\alpha < \alpha') \rightarrow (F_{\phi,\alpha'} \prec F_{\phi,\alpha})$ (see Section 2 for the definition of $\prec$). This property is illustrated in Table 3: any cluster from $E_{\phi,\alpha}$ is included in a coarser $E_{\phi,\alpha'}$, with $\alpha \geq \alpha'$.

Residuated Implication Operators. The residuated implication operators share some interesting characteristics in this framework. Let z be the number of distinct values appearing in F_ϕ, collected in $A = \{\alpha_1, ..., \alpha_z\}$, $\alpha_1 = 1$.

It holds that:

1. For all ϕ, $F_{\phi^\uparrow} = F_{\phi,1}$ = S_{res}, and thus $H_1 = H$. In other words, whatever residuated implication operator is used to calculate the FCRR for a case base, the finest imprecise partition generated from that FCRR coincides with the crisp case resemblance relation S_{res}. The proof is straightforward: by definition, $F_{\phi,1}(c_a, c_b) = 1$ if $S_{in}(p_a, p_b) > 0$ and $\phi(S_{in}(p_a, p_b), S_{out}(o_a, o_b)) = 1$. Due to the properties of residuated operators, the second condition only holds when $S_{in}(p_a, p_b) \leq S_{out}(o_a, o_b)$. The two conditions coincide with the definition of S_{res}, which completes the proof. It must also be underlined that the crisp case resemblance relation is a specific case of the fuzzy extension, that corresponds to the choice of the Rescher-Gaines implication: more formally, $F_{RG} = S_{res}$.
2. For all ϕ,

$$F_{\phi_\downarrow}(c_i, c_j) = F_{\phi,z} = \begin{cases} 1, \text{if} \min(S_{in}(p_a, p_b), S_{out}(o_a, o_b)) > 0 \\ 0, \text{otherwise.} \end{cases} \quad (4)$$

   Therefore, given two FCRRs generated from the same case base with different residuated implication operators, their coarsest imprecise partitions also coincide. The proof is based on monotonicity properties of the t-norm operators used to obtain the residuated implication operators.

Properties (1) and (2) imply all residuated operators applied to a given case base generate at least two common crisp relations, the largest and the smallest ones compatible with the case base. As a consequence, they also generate the same corresponding set of maximal cliques. For example, in Table 3, we see that $E_{G_1} = E_{\Pi_1}$ and $E_{G.15} = E_{\Pi.25}$. Moreover, we have $E_{G_1} = E_{\Pi_1} = E$ (see Section 3.6).

In the example, it is easy to check that Gödel and Goguen operators lead to very different crisp relations and only coincide in the unavoidable ones. Nevertheless, they have many clusters in common and it is possible to conceive heuristics that would be able to take advantage of that fact, leading to a reduced computational effort, if more than one operator is to be used in a given application.

4.5 Experiments Using Related Approaches

In [11], the authors used similarity relations associated to the description and solution variables spaces and the weighting approach proposed in [14] (see also [2]) on a real-world classification problem. It was shown that weighting the attributes in each case in the training set tends to lead to better results than the non weighted counterparts. Arguably, the weighting is able to overcome possible discrepancies on the similarities relations and reality, in what regards a given problem. Indeed, different relations may induce the same order in the pairs of cases in what regards their similarity, but a particular valuation may be more consistent than another with the relations associated to the other variables.

Weighting is computationally expensive, which makes it impracticable to use in large case bases. However, a good compromise can be obtained if we first use the non-weighted fuzzy approach proposed here to generate clusters and then apply the weighting algorithm on each cluster, as if it were itself a (smaller) case base. This approach has been tested in [12] for the same data used in [11], with a choice of parameters for S_{out} that generated a single crisp relation. The experiments using clusters behaved in general better than those without clusters.

5 Conclusions

We extended here an approach to CBR retrieval and reuse, based on the determination of clusters of similar cases, using a residuated implication operator to create a fuzzy resemblance relation. In this extension, the resemblance between cases is no longer necessarily crisp, as in the original approach. To extract clusters from this fuzzy relation, we propose to first derive its relevant level cuts, thus creating a set of crisp relations, and then obtaining the clusters therefrom. The approach proposed in [8] is equivalent to the the gradual rules framework proposed in [7], based on Rescher-Gaines implication operator, which can be very restrictive when the similarity relations are not properly tuned. The approach proposed here provides a means to deal with this problem.

Results from the use of this strategy on a real-world experiment using with a single FCRR led to good results [12], which suggests that the extended approach proposed here is promising, as it allows a larger choice of reasonable (imprecise) partitioning of the case base for learning weights, making it possible to learn weight vectors in large case bases. It is interesting to note that the use of the extended approach for a given choice of residuated operator does not necessarily increase very much the overall cost of the procedure, because many clusters are naturally present in several hypergraphs.

The important drawback of using an approach based on a set of similarity relations is that the number of parameters required to build the relations might be large. On the other hand, for many applications, similarity relations can be easily obtained from the experts in the domain of the case base, either directly or by converting a fuzzy sets into a relation (see [13] for a conversion method). Last but not least, one can use learning algorithms to find the similarity relations parameters if experts are not available.

In any case, properly tuning parameters may be difficult and for this reason, the association of the fuzzy approach proposed here and weighting is promising, by compensating for the improperly tuned parameters. The other drawback is the computational cost. So far, we have studied the use of taking the clusters as the maximal cliques, a NP-complete problem. In practical terms, however, this is not usually not significant, mainly for two reasons. On the one hand, the set of cases compatible with a new problem is usually small. On the other hand, when a large number of cases is compatible with the problem at hand, these cases themselves can be clustered around the output values. Nevertheless, when the set of compatible cases is large, one can reduce it by using thresholds in the overall description similarity relation between the problem and the cases compatible with it.

As future work, we intend to investigate alternative mechanisms to obtain hypergraphs, apart from maximal clusters and to study how to reduce the number of crisp relations derived from the FCRR, so that a smaller number of hypergraphs would have to be created. We already know that the minimal set of crisp relations should include the two extreme possibilities, as discussed in Section 5, that can be obtained from any residuated operator. We believe that analyzing the crisp relations obtained from distinct residuated operators will allow us to derive a good heuristic to select some of these crisp relations to derive the associated hypergraph. Finally, we intend to study more deeply the relation between the approach presented and the mining of fuzzy gradual rules, as presented in [9].

Acknowledgments. The author is indebted to Marie-Jeanne Lesot for invaluable help in preparing this manuscript, to the paper anonymous reviewers for useful comments and suggestions and to FAPES for grant No 2012/02077-8.

References

1. Aamodt, A., Plaza, E.: Case-based reasoning: foundational issues, methodological variations and system approaches. AI Commun. 7(1), 39–59 (1994)
2. Armengol, E., Esteva, F., Godo, L., Torra, V.: On learning similarity relations in fuzzy case-based reasoning. Trans. on Rough Sets, 14–32 (2004)
3. Bouchon-Meunier, B., Laurent, A., Lesot, M.-J., Rifqi, M.: Strengthening fuzzy gradual rules through "all the more" clauses. In: Proc. FuzzIEEE 2010 (WCCI 2010), pp. 2940–2946 (2010)
4. Dubois, D., Prade, H.: Possibility theory: an approach to computerized processing of uncertainty. Plenum Press (1988)
5. Dubois, D., Prade, H.: Gradual inference rules in approximate reasoning. Information Sciences 61(1-2), 103–122 (1982)
6. Esteva, F., Garcia-Calves, P., Godo, L.: Fuzzy similarity-based models in case-based reasoning. In: Proc. FuzzIEEE 2002, vol. 2, pp. 1348–1353 (2002)
7. Hüllermeier, E., Dubois, D., Prade, H.: Fuzzy rules in case-based reasoning. In: Proc. AFIA 1999, pp. 45–54 (1999)
8. Fanoiki, T., Drummond, I., Sandri, S.: Case-based reasoning retrieval and reuse using case resemblance hypergraphs. In: Proc. FuzzIEEE 2010 (WCCI 2010), pp. 1–7 (2010)
9. Hüllermeier, E.: Implication-Based Fuzzy Association Rules. In: Siebes, A., De Raedt, L. (eds.) PKDD 2001. LNCS (LNAI), vol. 2168, pp. 241–252. Springer, Heidelberg (2001)
10. Kolodner, J.: Cased-based reasoning. Morgan Kaufmann (1993)
11. Martins-Bedé, F.T., Godo, L., Sandri, S., Dutra, L.V., Freitas, C.C., Carvalho, O.S., Guimarães, R.J.P.S., Amaral, R.S.: Classification of Schistosomiasis Prevalence Using Fuzzy Case-Based Reasoning. In: Cabestany, J., Sandoval, F., Prieto, A., Corchado, J.M. (eds.) IWANN 2009, Part I. LNCS, vol. 5517, pp. 1053–1060. Springer, Heidelberg (2009)
12. Sandri, S., Mendonça, J., M.-Bedê, F., Guimarães, R., Carvalho, O.: Weighted fuzzy similarity relations in case-based reasoning: a case study in classification. In: WCCI 2012 (to appear, 2012)
13. Sandri, S., Toledo Martins-Bedé, F.: Order Compatible Fuzzy Relations and Their Elicitation from General Fuzzy Partitions. In: Liu, W. (ed.) ECSQARU 2011. LNCS, vol. 6717, pp. 640–650. Springer, Heidelberg (2011)
14. Torra, V.: On the learning of weights in some aggregation operators: the weighted mean and OWA operators. Math. and Soft Comp. 6 (2000)