

Reduction of the size of L -fuzzy contexts. A tool for differential diagnoses of diseases

Cristina Alcalde^a and Ana Burusco^{b,c}

^aDepartment of Applied Mathematics. University of the Basque Country UPV/EHU, Plaza Europa 1, 20018 San Sebastian, Spain; ^b Department of Statistics, Computer Science and Mathematics, Public University of Navarra, Campus de Arrosadía, 31006 Pamplona, Spain; ^cInstitute of Smart Cities. Public University of Navarra, Campus de Arrosadía, 31006 Pamplona, Spain

ARTICLE HISTORY

Compiled February 14, 2019

ABSTRACT

Information extraction from an L -fuzzy context becomes a hard problem when we work with a large set of objects and/or attributes. The goal of this paper is to present two different and complementary techniques to reduce the size of the context. First, using overlap indexes, we will establish rankings among the elements of the context that will allow us to determine those that do not provide relevant information and eliminate them. Second, by means of Choquet integrals, we will aggregate some objects or attributes of the context in order to jointly use the provided information. One interesting application of the developed theory consists on helping in the differential diagnoses of diseases that share a large number of symptoms and, therefore, that are difficult of distinguish.

KEYWORDS

L -fuzzy context; L -fuzzy concept; Choquet integral; overlap indexes; differential diagnosis

1. Introduction

The L -fuzzy concept analysis [Burusco and Fuentes-González (1994, 1998); Burusco and Fuentes-González (2000)] is a theory that studies the information arising from an L -fuzzy context using the L -fuzzy concepts as tools. An L -fuzzy concept is a pair of L -fuzzy sets that can be interpreted as a group of elements (objects) that shares some characteristics (attributes). The set of these L -fuzzy concepts has the structure of complete lattice.

When the cardinality of this L -fuzzy concept lattice is large, the obtained result may not be manageable. One of the factors that determines the size of the L -fuzzy concept lattice is the cardinality of the lattice L . The other is the size of the L -fuzzy context. In most situations, not all objects or attributes are of equal importance from the point of view of the concepts. In fact, some objects or attributes could be eliminated (eliminating their corresponding row or column) without loss of relevant information, thus reducing the size of the context.

Over the past, several researchers have developed models in order to reduce the size of this lattice. In [Belohlávek and Vychodil (2005, 2012)], the authors used hedges to control the size of the concept lattice. A different approach is developed by these authors in [Belohlávek and Vychodil (2006, 2013)]. Also Wei and Qi [Wei and Qi (2010)], and Medina [Medina (2012)] have published works from the point of view of the attributes for fuzzy oriented concept lattices or using T-implications as Li and Zhang do [Li and Zhang (2013)].

Other methods to reduce the complexity of the lattice using fuzzy similarity [Belohlávek (2000)] or block relations [Konecny and Krupka (2011)] have also been developed.

None of the previous works use aggregation operators for the purpose of studying this task.

Sometimes, the L -fuzzy context values are independent and we can use usual aggregations as weighted means [Calvo and Mesiar (2003b,a)], OWA operators [Fodor, Marichal, and Roubens (1995); Yager (1998)] and WOWA operators [Torra (1998)]. However, these studies are incomplete when we have values that present dependencies among them. In these situations the use of Choquet integrals [Choquet (1953); Grabisch (1995)] can be very useful as a tool for doing a proper analysis without loss of information as we did in [Alcalde and Burusco (2017)].

The paper addresses the study of the relevance of objects and attributes using overlap indexes and proposes a method for the reduction of the size of the L -fuzzy context when $L = [0, 1]$ using Choquet integrals. Firstly, in Section 2 we will introduce the L -fuzzy concept analysis, overlap indexes and Choquet integrals that will be used in the paper. The rest of the paper is organized as follows: Section 3 applies overlap indexes to establish orderings among objects and attributes, Section 4 studies the use of Choquet integrals to reduce the size of the L -fuzzy contexts aggregating some rows or columns. Finally, in Section 5 we apply the developed methods to deal with the problem of differential diagnosis of erythemato-squamous diseases. Conclusions are detailed in last section.

2. Preliminaries

2.1. L -Fuzzy concept analysis

The R. Wille's Formal Concept Analysis [Wille (1982)] extracts information from a binary table that represents a formal context (X, Y, R) with X and Y finite sets of objects and attributes respectively and $R \subseteq X \times Y$. The hidden information consists of pairs (A, B) with $A \subseteq X$ and $B \subseteq Y$, called formal concepts, verifying $A^* = B$ and $B^* = A$, where $(\cdot)^*$ is the derivation operator that associates the attributes related to the elements of A to every object set A , and the objects related to the attributes of B to every attribute set B . These formal concepts can be interpreted as a group of objects A that shares the attributes of B .

In previous works [Burusco and Fuentes-González (1994, 1998)] we defined the L -fuzzy contexts (L, X, Y, R) , with L a complete lattice, X and Y sets of objects and attributes respectively and $R \in L^{X \times Y}$ a fuzzy relation between the objects and the attributes. This is an extension of Wille's formal contexts to the fuzzy case when we want to study the relations between the objects and the attributes with values in a complete lattice L , instead of binary ones.

In our case, to work with these L -fuzzy contexts, we defined the derivation operators

1 and 2 given by means of these expressions [Burusco and Fuentes-González (1998); Burusco and Fuentes-González (2000)]:

$$\forall A \in L^X, \forall B \in L^Y$$

$$A_1(y) = \inf_{x \in X} \{I(A(x), R(x, y))\}$$

$$B_2(x) = \inf_{y \in Y} \{I(B(y), R(x, y))\}$$

with I a fuzzy implication operator defined in (L, \leq) .

The information stored in the context is visualized by means of the L -fuzzy concepts that are pairs $(A, A_1) \in L^X \times L^Y$ with $A \in \text{fix}(\varphi)$, set of fixed points of the operator φ , being defined from the derivation operators 1 and 2 as $\varphi(A) = (A_1)_2 = A_{12}$. These pairs, whose first and second components are said to be the fuzzy extension and intension respectively, represent a group of objects that share a group of attributes.

Using the usual order relation between fuzzy sets, that is, $\forall A, C \in L^X, A \leq C \iff A(x) \leq C(x), \forall x \in X$, we define the set $\mathcal{L} = \{(A, A_1) \mid A \in \text{fix}(\varphi)\}$ with the order relation \preceq defined as: $\forall (A, A_1), (C, C_1) \in \mathcal{L}, (A, A_1) \preceq (C, C_1)$ if $A \leq C$ (or $A_1 \geq C_1$).

As φ is an order preserving operator, then the set $\text{fix}(\varphi)$ is a complete lattice and (\mathcal{L}, \preceq) is also a complete lattice that is said to be the L -fuzzy concept lattice [Burusco and Fuentes-González (1994, 1998)].

In addition, in the case of using a residuated implication ($I(a, b) = \sup\{x \mid T(a, x) \leq b\}$, with T a t-norm), given $A \in L^X$, (or $B \in L^Y$) we can obtain the associated L -fuzzy concept applying twice the derivation operators. In this case, the L -fuzzy concept associated to A is (A_{12}, A_1) (or (B_2, B_{21})). In the paper, residuated implication operators will be used.

In particular, if we represent every object $x \in X, i \in \{1, \dots, n\}$, by the L -fuzzy set \mathbf{x}_i such that $\mathbf{x}_i(x_i) = 1$ and $\mathbf{x}_i(x) = 0$, for any $x \neq x_i$, we can define the following L -fuzzy concept:

Definition 2.1. For every $x_i \in X, i \in \{1, \dots, n\}$, the pair $\mathcal{C}_{\mathbf{x}_i} = ((\mathbf{x}_i)_{12}, (\mathbf{x}_i)_1)$ is said to be the L -fuzzy concept derived from \mathbf{x}_i .

Analogously $\mathcal{C}_{\mathbf{y}_j} = ((\mathbf{y}_j)_2, (\mathbf{y}_j)_{21}), j \in \{1, \dots, m\}$ is the L -fuzzy concept derived from \mathbf{y}_j , being \mathbf{y}_j the L -fuzzy set that represents the attribute y_j .

In the rest of the paper we will denote $\mathcal{C}_{\mathbf{x}_i} = (\underline{\mathcal{C}}_{\mathbf{x}_i}, \overline{\mathcal{C}}_{\mathbf{x}_i})$ and $\mathcal{C}_{\mathbf{y}_j} = (\underline{\mathcal{C}}_{\mathbf{y}_j}, \overline{\mathcal{C}}_{\mathbf{y}_j})$. These concepts are the closest to the departure sets represented by \mathbf{x}_i or \mathbf{y}_j (study of a single object or attribute), and they give us a general idea about the behavior of the objects and the attributes in the L -fuzzy concept lattice with a low computational cost.

Our last results are related to the use of two relations in the definition of the L -fuzzy context [Alcalde and Burusco (2015a)], the study of fuzzy context sequences [Alcalde, Burusco, and Fuentes-González (2013)], the composition of L -fuzzy contexts [Alcalde and Burusco (2012)] and the study of multivalued contexts associated with criteria [Alcalde and Burusco (2018a)]. We have also developed this theory in different areas as the treatment of incomplete information [Alcalde et al. (2009); Alcalde, Burusco, and Fuentes-González (2005)] or Mathematical Morphology [Alcalde, Burusco, and Fuentes-González (2014); Alcalde et al. (2017a,b)].

Other important works that generalize the Formal Concepts Analysis using residuated implication operators are due to Bělohlávek [Bělohlávek (1999, 2002)] and S.

Pollandt [Pollandt (1997)]. Moreover, extensions of Formal Concept Analysis to the interval-valued case are in [Alcalde et al. (2009); Djouadi and Prade (2010, 2011)] and to the fuzzy property-oriented and multi-adjoint concept lattices framework in [Medina and Ojeda-Aciego (2010, 2013)].

In this paper, we are going to work with L -fuzzy contexts when $L = [0, 1]$.

2.2. Choquet integrals

In order to aggregate information, several operators have been developed in the literature. The best known are weighted means [Calvo and Mesiar (2003b,a)] and OWA operators [Yager (1998); Fodor, Marichal, and Roubens (1995)] which use weighting vectors, in the first case to weight the reliability, and to weight the values according to their ordering in the second one. When the use of both orderings is interesting, WOWA operators [Torra (1998)] can be a good option. We have applied these operators to aggregate values in different situations [Alcalde, Burusco, and Fuentes-González (2013); Alcalde et al. (2016); Alcalde and Burusco (2014, 2015b)]. Choquet integrals [Grabisch (1995)] are defined as a generalization of the previous ones when not only individual observations but also groups are relevant. We will see next the main definitions.

We will begin with the definition of generalized measure [Klir and Wang (2009)] defined in $\mathcal{P}(X)$, the set of parts of X .

Definition 2.2. A function $\mu : \mathcal{P}(X) \rightarrow [0, 1]$ is a *generalized measure* on X if and only if it satisfies the $\mu(\emptyset) = 0$.

Usually monotonicity is also required: $A \subseteq B \subseteq X$ implies $\mu(A) \leq \mu(B)$. Besides, although it is not necessary, we will assume here $\mu(X) = 1$.

Definition 2.3. A generalized measure μ is said to be *additive* if for all $A, B \subseteq X$ such that $A \cap B = \emptyset$,

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

The Choquet integral was formulated by Grabisch [Grabisch (1995)] as follows:

Definition 2.4. Given a generalized measure μ , the Choquet integral with respect to μ is defined as:

$$Ch_{\mu}(a_1 \dots a_N) = \sum_{k=1}^N a_{\sigma(k)} (\mu(A_{\sigma(k)}) - \mu(A_{\sigma(k-1)}))$$

where $\{\sigma(1), \dots, \sigma(N)\}$ is a permutation of $\{1, \dots, N\}$ such that $a_{\sigma(1)} \geq a_{\sigma(2)} \geq \dots \geq a_{\sigma(N)}$ and $A_{\sigma(k)} = \{\sigma(j) | j \leq k\}$ (therefore $A_{\sigma(r)} = \{\sigma(1), \dots, \sigma(r)\}$ when $r \geq 1$ and $A_{\sigma(0)} = \emptyset$.)

Remark 1. From this definition we can see that weighted means, OWA and WOWA operators are particular types of Choquet integral [Torra (1998)].

Remark 2. If μ is an additive generalized measure, then the Choquet integral with

respect to μ can be calculated as:

$$Ch_{\mu}(a_1 \dots a_N) = \sum_{k=1}^N a_{\sigma(k)} \mu(\{\sigma(k)\})$$

A key issue in the use of Choquet integrals is the election of the measure to calculate the weights of the elements that will be aggregated. We propose in this work measures defined from overlap indexes [Paternain et al. (2016)]. In this way the defined measure will give a higher value to those sets of elements that have a stronger relationship among them.

2.3. Measures constructed from overlap indexes

Given $L = [0, 1]$ and L^U the fuzzy sets of U , in [Paternain et al. (2016)] the authors define an overlap index as a mapping $O : L^U \times L^U \rightarrow [0, 1]$, such that $\forall A, B \in L^U$:

- i) $O(A, B) = 0$ if and only if in A and B have disjoint supports; that is, $A(i)B(i) = 0$ for every $i \in U$
- ii) $O(A, B) = O(B, A)$
- iii) If $B \subseteq C$, then $O(A, B) \leq O(A, C)$

An overlap index such that:

- iv) $O(A, B) = 1$ if there exists $i \in U$ such that $A(i) = B(i) = 1$

is called a *normal* overlap index.

In addition, if the overlap index verifies that

- v) $O(E, A \cap B) + O(E, A \cup B) = O(E, A) + O(E, B) \forall A, B \in L^U$,

O is called *E-modular*.

Examples of overlap indexes are the following ones:

- (1) Zadeh's consistency index:

$$O_Z(A, B) = \max_{1 \leq i \leq n} (\min(A(i), B(i)))$$

- (2) Let $M : [0, 1]^2 \rightarrow [0, 1]$ be a symmetric aggregation function such that $M(x, y) = 0$ if and only if $xy = 0$. We have that:

$$O_{M,Z}(A, B) = \max_{1 \leq i \leq n} (M(A(i), B(i)))$$

is a normal overlap index that generalizes the Zadeh's index.

- (3) If in the previous example, we consider a symmetric, increasing function $M : [0, 1]^2 \rightarrow [0, 1]$ such that $M(1, 1) < 1$ and $M(x, y) = 0$ if and only if $xy = 0$, then we obtain an overlap index which is not normal. For instance, when taking $M(x, y) = (xy)^p/2$ with $p > 0$, we arrive at the overlap index:

$$O(A, B) = \max_{1 \leq i \leq n} \left(\frac{(A(i), B(i))^p}{2} \right)$$

(4) The following is also an example of overlap index:

$$O_\pi(A, B) = \frac{1}{n} \sum_{i=1}^n A(i)B(i)$$

Let $E \in L^U$ be a fixed non-empty fuzzy set. Given $A \subseteq U$, we define:

$$E_A(i) = \begin{cases} E(i) & \text{if } i \in A \\ 0 & \text{otherwise} \end{cases}$$

Observe that E_A is the intersection of the fuzzy set E and the crisp set A .

Now it is possible to introduce the definition of a measure in terms of a fuzzy set and an overlap index, as shown in the following theorem [Paternain et al. (2016)].

Theorem 2.5. *If $E \in L^U$ is a fixed, non-empty fuzzy set, then the mapping $m_{O,E} : \mathcal{P}(U) \rightarrow [0, 1]$ given by $m_{O,E}(A) = \frac{O(E, E_A)}{O(E, E)}$ is a generalized measure for every overlap index O .*

Remark 3. The generalized measure $m_{O,E}$ is additive if the overlap index O is E -modular (see [Paternain et al. (2016)]).

Example 2.6. The generalized measure $m_{O_\pi, E}$ is additive, and hence a classical measure, for every fuzzy set E .

3. Ordering and elimination of objects and attributes using overlap indexes

As it has been explained in the introduction, the size of an L -fuzzy context is one of the factors that determines the size of the L -fuzzy concept lattice and its handling.

So far, in all our works regarding the reduction of the size of the L -fuzzy context we have used the method of removing rows or columns in the relation (eliminating objects or attributes). In [Burusco and Fuentes-González (2003)] we removed the objects and/or attributes of little significance, that is, that did not appear as relevant in any L -fuzzy concept. To do this, we first obtained the L -fuzzy concept lattice, a quite laborious task.

In another different field and in order to work with missing values, in [Alcalde, Burusco, and Fuentes-González (2005)] infrequently appearing objects and attributes were studied. We removed them when they did not exceed a minimum support. To do this, we defined support of an L -fuzzy set in terms of the membership values of the elements in its derived set. The aim was to eliminate some rows or columns of missing values.

We will try in this work to reduce the size of the L -fuzzy context by establishing rankings among the objects or attributes in order to eliminate those that appear in not relevant positions.

It is important to point out that by eliminating some elements of the context the information corresponding to the others does not change, as shown in the following proposition.

Proposition 3.1. *Let us consider the subset of objects $\hat{X} \subset X$ and the relation $\hat{R} \in \hat{X} \times Y$ such that $\forall(x, y) \in \hat{X} \times Y \hat{R}(x, y) = R(x, y)$. Given an object $x_0 \in \hat{X}$, let $\mathcal{C}_{\mathbf{x}_0}$ and $\hat{\mathcal{C}}_{\mathbf{x}_0}$ be the L -fuzzy derived concepts in (L, X, Y, R) and (L, \hat{X}, Y, \hat{R}) respectively. Then, $\forall(x, y) \in \hat{X} \times Y$, $\underline{\mathcal{C}}_{\mathbf{x}_0}(x) = \hat{\underline{\mathcal{C}}}_{\mathbf{x}_0}(x)$ and $\overline{\mathcal{C}}_{\mathbf{x}_0}(y) = \hat{\overline{\mathcal{C}}}_{\mathbf{x}_0}(y)$.*

Proof. Given $x_0 \in \hat{X}$, $\forall(x, y) \in \hat{X} \times Y$,

$$\begin{aligned}\overline{\mathcal{C}}_{\mathbf{x}_0}(y) &= \inf_{x \in X} \{I(\mathbf{x}_0(x), R(x, y))\} = R(x_0, y) \\ \hat{\overline{\mathcal{C}}}_{\mathbf{x}_0}(y) &= \inf_{x \in \hat{X}} \{I(\mathbf{x}_0(x), \hat{R}(x, y))\} = \hat{R}(x_0, y)\end{aligned}$$

Therefore, as $x_0 \in \hat{X}$, $\overline{\mathcal{C}}_{\mathbf{x}_0}(y) = \hat{\overline{\mathcal{C}}}_{\mathbf{x}_0}(y)$ for all $y \in Y$.

On the other hand, for all $x \in \hat{X}$,

$$\underline{\mathcal{C}}_{\mathbf{x}_0}(x) = \inf_{y \in Y} \{I(R(x_0, y), R(x, y))\} = \inf_{y \in Y} \{I(\hat{R}(x_0, y), \hat{R}(x, y))\} = \hat{\underline{\mathcal{C}}}_{\mathbf{x}_0}(x).$$

□

Similarly, we can prove that by removing a subset of attributes from the context the information corresponding to the others does not change either.

Proposition 3.2. *Given the subset of attributes $\hat{Y} \subset Y$ and the relation $\hat{R} \in X \times \hat{Y}$ such that $\forall(x, y) \in X \times \hat{Y} \hat{R}(x, y) = R(x, y)$, consider the attribute $y_0 \in \hat{Y}$, and let $\mathcal{C}_{\mathbf{y}_0}$ and $\hat{\mathcal{C}}_{\mathbf{y}_0}$ be the L -fuzzy derived concepts in (L, X, Y, R) and (L, X, \hat{Y}, \hat{R}) respectively. Then, $\forall(x, y) \in X \times \hat{Y}$, $\underline{\mathcal{C}}_{\mathbf{y}_0}(x) = \hat{\underline{\mathcal{C}}}_{\mathbf{y}_0}(x)$ and $\overline{\mathcal{C}}_{\mathbf{y}_0}(y) = \hat{\overline{\mathcal{C}}}_{\mathbf{y}_0}(y)$.*

3.1. Ranking obtained from supports

The definition of support of an L -fuzzy subset of objects or attributes given in [Alcalde, Burusco, and Fuentes-González (2005)] was extended in [Alcalde and Burusco (2018b)] in order to assign a support to each object and attribute in the context that allows establish rankings among them as follows.

Definition 3.3. Given an object $x \in X$, we define its *support* as

$$\text{supp}(x) = \frac{\sum_{y \in Y} \overline{\mathcal{C}}_{\mathbf{x}}(y)}{|Y|}$$

In the same way, the *support of an attribute* $y \in Y$ can be defined as

$$\text{supp}(y) = \frac{\sum_{x \in X} \underline{\mathcal{C}}_{\mathbf{y}}(x)}{|X|}$$

This support measures the average membership degree in the intention (extension) of the concept associated with an object (attribute) and, therefore, it can be considered

as a measure of the relevance of this object (attribute). The greater the support, that is, the greater the presence of the attributes (objects) in the associated concept, the more relevant the object (attribute) will be.

Based on the support it is possible to define the following preorder relation among the objects in the L -fuzzy context:

Definition 3.4. Given two objects x_i and $x_j \in X$,

$$x_i \geq_{\text{supp}} x_j \iff \text{supp}(x_i) \geq \text{supp}(x_j)$$

Similarly, $\forall y_i, y_j \in Y$,

$$y_i \geq_{\text{supp}} y_j \iff \text{supp}(y_i) \geq \text{supp}(y_j)$$

Note that \geq_{supp} is not an antisymmetric relation and therefore it is not an order relation.

3.2. Other rankings defined from overlap indexes

Overlap indexes can be used to define other different rankings among the objects or attributes of the context by means of the following preorder relations:

Definition 3.5. For every $x_i, x_j \in X$. Let $E_X \in L^X$ be a fuzzy set and O an overlap index.

$$x_i \geq_{O, E_X} x_j \iff O(E_X, \bar{\mathcal{C}}_{x_i}) \geq O(E_X, \bar{\mathcal{C}}_{x_j})$$

Analogously, $\forall y_i, y_j \in Y$ and $\forall E_Y \in L^Y$,

$$y_i \geq_{O, E_Y} y_j \iff O(E_Y, \underline{\mathcal{C}}_{y_i}) \geq O(E_Y, \underline{\mathcal{C}}_{y_j})$$

The obtained ranking differs if we change the considered set E and overlap index O . The election of the preorder relation will depend on the situation that we want to analyze.

Taking into account that the support of an object summarizes the information of its relevance in the context, an interesting case is the one obtained when we chose the set E_X such that $E_X(x) = \text{supp}(x) \in L^X$ for every $x \in X$. Considering this set E_X , the relation \geq_{O, E_X} orders the objects in function, not only of the support of these objects, but also of the support of all those that appear related to them in the intentions of the concepts that represent them. From the point of view of the attributes we can proceed in a similar way considering the set $E_Y(y) = \text{supp}(y) \in L^Y$ for every $y \in Y$.

4. Reducing the size of an L-fuzzy context by means of the aggregation of rows or columns

We are aware that when objects or attributes of the contexts are removed, some information is also eliminated, so we want to study a new method that minimizes the impact of the missing information.

Instead of deleting rows and/or columns from the context, we can think of aggregating these rows or columns among them or with others. That is, a new row or column of the context is added by aggregating others. Again, the aim of this action is to reduce the size of the context and its L -fuzzy concept lattice.

In this way we can aggregate objects or attributes of low relevance with others. In addition, it is possible that we also want to aggregate other objects or attributes that we consider important to the context in order to reduce its size. In this case we aggregate them with the other connected ones.

On the other hand, if we are interested specifically in a group of objects, we will do an independent study of the L -fuzzy subcontext formed by their corresponding rows before the aggregation process as we did in [Alcalde and Burusco (2017)].

Furthermore, if we change a row or column in the context (that occurs when we aggregate several rows or columns), we can see that the L -fuzzy concepts obtained from the non modified objects or attributes have the same membership degrees (see [Alcalde and Burusco (2017)]).

Proposition 4.1. *Let (L, X, Y, R) and (L, X, Y, \hat{R}) be L -fuzzy contexts such that $\hat{R}(x, y) = R(x, y), \forall x \in X \setminus \{x_0\}, \forall y \in Y$. Consider $x_l \in X \setminus \{x_0\}$ and let \mathcal{C}_{x_l} and $\hat{\mathcal{C}}_{x_l}$ be the L -fuzzy derived concepts in (L, X, Y, R) and (L, X, Y, \hat{R}) respectively. For any $x \in X \setminus \{x_0\}$ and for any $y \in Y$, the membership degrees in both L -fuzzy concepts are coincident.*

Analogously, from the point of view of the attributes.

Proposition 4.2. *Let (L, X, Y, R) and (L, X, Y, \hat{R}) be L -fuzzy contexts such that $\hat{R}(x, y) = R(x, y), \forall x \in X, \forall y \in Y \setminus \{y_0\}$. Consider $y_l \in Y \setminus \{y_0\}$ and let \mathcal{C}_{y_l} and $\hat{\mathcal{C}}_{y_l}$ be the L -fuzzy derived concepts in (L, X, Y, R) and (L, X, Y, \hat{R}) . For any $y \in Y \setminus \{y_0\}$ and for any $x \in X$, the membership degrees are coincident in both L -fuzzy concepts.*

The use of these L -fuzzy concepts derived from the objects or attributes will be an interesting tool in order to aggregate some related objects and/or attributes.

For any $x_k \in X, k \in \{1, \dots, n\}$, $\underline{\mathcal{C}}(x_k)$ denotes the membership degree of the object x_k in \mathcal{C} . Analogously, the membership degree of the attribute y_l in \mathcal{C} is represented by $\overline{\mathcal{C}}(y_l)$ for every $y_l \in Y$.

Definition 4.3. Consider $0 < \alpha \leq 1$. Let (L, X, Y, R) be an L -fuzzy context and let \mathcal{L} be its L -fuzzy concept lattice. For any L -fuzzy concept $\mathcal{C} \in \mathcal{L}$, we define the α -objects associated with \mathcal{C} as the set:

$$T^\alpha(\mathcal{C}) = \{x_k \in X \mid \underline{\mathcal{C}}(x_k) \geq \alpha\}.$$

Analogously, the α -attributes associated with \mathcal{C} are defined as:

$$Q^\alpha(\mathcal{C}) = \{y_l \in Y \mid \overline{\mathcal{C}}(y_l) \geq \alpha\}.$$

Definition 4.4. If $x_k \in T^\alpha(\mathcal{C}_{x_i}), k \neq i$, then x_i is said to be α -dependent on x_k . On the contrary, we say that x_i is α -independent on x_k .

Similarly, if $y_l \in Q^\alpha(\mathcal{C}_{y_j}), l \neq j$, then y_j is said to be α -dependent on y_l . Otherwise, we say that y_j is α -independent on y_l .

That is, x_i is α -dependent on x_k when in the L -fuzzy concept derived from x_i, x_k

stands out in a level bigger than or equal to α . In the case $\alpha = 1$, the membership degree of x_k in the L -fuzzy concept is equal to 1.

Remark 4. x_k can belong to $T^\alpha(\mathcal{C}_{x_i})$, $k \neq i$ for different objects x_i . The same can be said for the attributes.

Proposition 4.5. *The relation of α -dependence is not necessarily symmetric.*

That is, if x_k is α -dependent on x_j , x_j is not necessarily α -dependent on x_k fulfilled. The same holds for the attributes.

Definition 4.6. Given α , we say that x_i is α -independent if $T^\alpha(\mathcal{C}_{x_i}) = \{x_i\}$.

Analogously, y_j is an α -independent attribute if $Q^\alpha(\mathcal{C}_{y_j}) = \{y_j\}$.

The relation of α -dependence among the objects allows to set up a graph $G_X^\alpha = (X, \mathcal{E}_X^\alpha)$ which nodes are the objects of the context. The edge $(x_i, x_j) \in \mathcal{E}_X^\alpha$ is in the graph if $x_j \in T^\alpha(\mathcal{C}_{x_i})$.

Similarly, we can consider the graph $G_Y^\alpha = (Y, \mathcal{E}_Y^\alpha)$ on the set of attributes of the context.

Note that, although the relation of α -dependence establishes a direction in each edge of the context, our goal is to aggregate those objects or attributes that are related among them and therefore we can work with the undirected graph.

The connected components of the graph G_X^α form a partition of the set X . We will denote this partition as $X^\alpha = \{\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_k\}$, where the elements \tilde{X}_i represent the different connected components of the graph. From the elements of this obtained partition we can define a new L -fuzzy context.

Definition 4.7. Let be $0 < \alpha \leq 1$. The L -fuzzy context α -reduced by objects $(L, X^\alpha, Y, R_O^\alpha)$ is obtained aggregating the values of R corresponding to the elements that form each connected component in the graph G_X^α . That is:

$$\forall \tilde{X}_i = \{x_{i1}, x_{i2}, \dots, x_{iN_i}\} \in X^\alpha, \forall y \in Y,$$

$$R_O^\alpha(\tilde{X}_i, y) = \text{aggr}(R(x_{i1}, y), R(x_{i2}, y), \dots, R(x_{iN_i}, y))$$

The L -fuzzy context α -reduced by attributes, $(L, X, Y^\alpha, R_A^\alpha)$, can be defined in an analogous way.

Remark 5. It is obvious that given α , the L -fuzzy context α -reduced by objects $(L, X^\alpha, Y, R_O^\alpha)$ and the L -fuzzy context α -reduced by attributes $(L, X, Y^\alpha, R_A^\alpha)$ are not coincident.

In all the cases, we have to choose one of the sets (objects or attributes) in order to carry out our study. The election will depend on the particularities of the problem and the aim of study.

4.1. Choquet integral derived from objects or attributes

Choquet integrals allow to aggregate elements taking into account the existing relation among them. We can use, for example, measures that give a higher weight to those subsets of elements that have a closer relationship.

With this purpose of aggregate rows or columns of the L -fuzzy context when there exists an strong relation among them from the point of view of the study of L -fuzzy

concepts, we will adapt the idea of Choquet integral as follows:

Definition 4.8. Given the L -fuzzy context (L, X, Y, R) , consider the subset of objects $\hat{X} = \{x_1, x_2, \dots, x_N\} \subseteq X$. Let μ be a generalized measure defined on the set $\mathcal{P}(\hat{X})$. We define the \hat{X} -Choquet integral with respect to μ as a function $Ch_\mu^{\hat{X}} : L \rightarrow L$ such that for all $y \in Y$

$$Ch_\mu^{\hat{X}}(y) = \sum_{k=1}^N R(x_{\sigma(k)}, y)(\mu(A_{\sigma(k)}) - \mu(A_{\sigma(k-1)}))$$

where $\{\sigma(1), \dots, \sigma(N)\}$ is a permutation of $\{1, \dots, N\}$ such that $R(x_{\sigma(1)}, y) \geq R(x_{\sigma(2)}, y) \geq \dots \geq R(x_{\sigma(N)}, y)$, and $A_{\sigma(k)} = \{x_{\sigma(j)} | j \leq k\}$.

Analogously, from the point of view of the attributes:

Definition 4.9. Given the subset of attributes $\hat{Y} = \{y_1, y_2, \dots, y_M\} \subseteq Y$, Let ρ be a generalized measure defined on the set $\mathcal{P}(\hat{Y})$. We define the \hat{Y} -Choquet integral with respect to ρ as a function $Ch_\rho^{\hat{Y}} : L \rightarrow L$ such that for all $x \in X$

$$Ch_\rho^{\hat{Y}}(x) = \sum_{k=1}^M R(x, y_{\sigma(k)})(\rho(B_{\sigma(k)}) - \rho(B_{\sigma(k-1)}))$$

where $\{\sigma(1), \dots, \sigma(M)\}$ is a permutation of $\{1, \dots, M\}$ such that $R(x, y_{\sigma(1)}) \geq R(x, y_{\sigma(2)}) \geq \dots \geq R(x, y_{\sigma(M)})$ and $B_{\sigma(k)} = \{y_{\sigma(j)} | j \leq k\}$.

Proposition 4.10. Given the subset of objects $\hat{X} = \{x_1, x_2, \dots, x_N\} \subseteq X$, the \hat{X} -Choquet integral with respect to μ verify the following properties.

- i) If $R(x_i, y) = 1 \quad \forall i \leq N$, then $Ch_\mu^{\hat{X}}(y) = 1$.
- ii) $\min_{1 \leq i \leq N} \{R(x_i, y)\} \leq Ch_\mu^{\hat{X}}(y) \leq \max_{1 \leq i \leq N} \{R(x_i, y)\}$.

4.2. Context Reduction Algorithms

As commented above, the first election that we have to make is if we want to reduce the number of objects or attributes in the context. We present here the process in order to reduce the number of objects (see Algorithm 1).

The election of α establishes the exigency level and depends on every practical case. We will begin with a level $\alpha = 1$ and we can try with a smaller value when the size of the α -reduced context is not manageable (the lower α , the larger will be the number of aggregated objects).

An analogous process can be performed in order to reduce the number of attributes, obtaining the Algorithm 2.

The computational cost of the process is low since it only uses the L -fuzzy concepts derived from objects and/or attributes and, therefore, the entire construction of the L -fuzzy concept lattice is not necessary.

We will choose the most appropriate measure μ in each practical case.

In [Alcalde and Burusco (2017)] we presented a first approximation to the problem of context reduction using Choquet integrals. In this paper, with the aim of improving

Algorithm 1 Context Reduction by Objects Algorithm (CROA)

Input: (L, X, Y, R) : an L -fuzzy context, α : exigence level.

Output: $(L, X^\alpha, Y, R_O^\alpha)$: the L -fuzzy context α -reduced by objects.

- 1: **for all** $x \in X$ **do**
- 2: Obtain its derived L -fuzzy concept \mathcal{C}_x .
- 3: Calculate the set $T^\alpha(\mathcal{C}_x)$ of α -objects associated with \mathcal{C}_x .
- 4: **end for**
- 5: Establish the graph $G_X^\alpha = (X, \mathcal{E}_X^\alpha)$.
- 6: Obtain the set of connected components X^α .
- 7: **for all** $\tilde{X} \in X^\alpha$ **do**
- 8: **for all** $y \in Y$ **do**
- 9: Calculate the relation of the L -fuzzy context α -reduced by objects:

$$R_O^\alpha(\tilde{X}, y) = Ch_\mu^{\tilde{X}}(y).$$

- 10: **end for**
 - 11: **end for**
-

Algorithm 2 Context Reduction by Attributes Algorithm (CRAA)

Input: (L, X, Y, R) : an L -fuzzy context, α : exigence level.

Output: $(L, X, Y^\alpha, R_A^\alpha)$: the L -fuzzy context α -reduced by attributes.

- 1: **for all** $y \in Y$ **do**
- 2: Obtain its derived L -fuzzy concept \mathcal{C}_y .
- 3: Calculate the set $Q^\alpha(\mathcal{C}_y)$ of α -attributes associated with \mathcal{C}_y .
- 4: **end for**
- 5: Establish the graph $G_Y^\alpha = (Y, \mathcal{E}_Y^\alpha)$.
- 6: Obtain the set of connected components Y^α .
- 7: **for all** $\tilde{Y} \in Y^\alpha$ **do**
- 8: **for all** $x \in X$ **do**
- 9: Calculate the relation of the L -fuzzy context α -reduced by attributes:

$$R_A^\alpha(x, \tilde{Y}) = Ch_\mu^{\tilde{Y}}(x).$$

- 10: **end for**
 - 11: **end for**
-

the results, we will use the measure obtained from overlap indexes (see Theorem 2.5).

In order to measure the relationship grade among the objects, for every connected component $\tilde{X}_i \in X^\alpha$, $\tilde{X}_i = \{x_{i1}, x_{i2}, \dots, x_{iN_i}\} \subset X$, we will consider the fuzzy set $E_{\tilde{X}_i} \in L^{\tilde{X}_i}$ such that $\forall x_{ik} \in \tilde{X}_i$,

$$E_{\tilde{X}_i}(x_{ik}) = \frac{\sum_{j=1}^{N_i} \mathcal{C}_{x_{ij}}(x_{ik})}{N_i}$$

Fixed an overlap index O , we will define a measure μ on $\mathcal{P}(\tilde{X}_i)$ as:

$$\forall A \subset \tilde{X}_i, \quad \mu(A) = m_{O, E_{\tilde{X}_i}}(A) = \frac{O(E_{\tilde{X}_i}, E_{\tilde{X}_i A})}{O(E_{\tilde{X}_i}, E_{\tilde{X}_i})}$$

Thus, the higher the dependency among the objects of A the greater its measure $\mu(A)$.

In terms of computational cost, if $|X| = n$ and $|Y| = m$, Algorithm 1 has a complexity of $\mathcal{O}(n^2m)$ and Algorithm 2 is $\mathcal{O}(nm^2)$.

5. Practical application: Differential diagnosis of similar diseases

An interesting application of the results developed above is the help in the differential diagnosis of similar diseases that share a set of symptoms. We will illustrate here this problem analyzing the diagnosis of six dermatological diseases. This problem was treated in [Güvenir, Demiröz, and Ilter (1998)] with a classification algorithm that provided good results. We will see here that same results can be obtained applying the developed theory with a very low computational cost.

The differential diagnosis of erythemato-squamous diseases is a difficult problem in dermatology because they share the clinical features of erythema and scaling and, at first sight, they look very much alike.

The diseases in this group are:

- x_1 : Psoriasis
- x_2 : Seboreic dermatitis
- x_3 : Lichen planus
- x_4 : Pityriasis rosea
- x_5 : Cronic dermatitis
- x_6 : Pityriasis rubra pilaris

Patients are usually evaluated clinically with the following eleven features:

- y_1 : Erythema
- y_2 : Scaling
- y_3 : Definite borders
- y_4 : Itching
- y_5 : Koebner phenomenon
- y_6 : Polygonal papules
- y_7 : Follicular papules
- y_8 : Oral mucosal involvement
- y_9 : Knee and elbow involvement
- y_{10} : Scalp involvement
- y_{11} : Family history, (0 or 1)

Some patients can be diagnosed with these clinical features only but, frequently, a biopsy is necessary in order to evaluate the following histopathological features:

- y_{12} : Melanin incontinence
- y_{13} : Eosinophils in the infiltrate
- y_{14} : PNL infiltrate
- y_{15} : Fibrosis of the papillary dermis

- y_{16} : Exocytosis
- y_{17} : Acanthosis
- y_{18} : Hyperkeratosis
- y_{19} : Parakeratosis
- y_{20} : Clubbing of the rete ridges
- y_{21} : Elongation of the rete ridges
- y_{22} : Thinning of the suprapapillary epidermis
- y_{23} : Spongiform pustule
- y_{24} : Munro microabcess
- y_{25} : Focal hypergranulosis
- y_{26} : Disappearance of the granular layer
- y_{27} : Vacuolisation and damage of basal layer
- y_{28} : Spongiosis
- y_{29} : Saw-tooth appearance of retes
- y_{30} : Follicular horn plug
- y_{31} : Perifollicular parakeratosis
- y_{32} : Inflammatory monoluclear infiltrate
- y_{33} : Band-like infiltrate

In order to define our context we have used the dataset available in the UCI Machine Learning Repository [Dua and Taniskidou (2017)]. This dataset contains 366 instances consisting in the disease suffered by each patient and the presented features.

We have constructed the L -fuzzy context (L, X, Y, R) , in which the objects are the different diseases, the set of attributes is formed by the analyzed 33 features, and the relation among them is given by the proportion of cases where the corresponding feature has been observed for each of the diseases. The obtained relation R is represented in Table 1.

The construction of the whole L -fuzzy concept lattice has a high computational cost. So, we will apply the theory developed above to reduce the size of the L -fuzzy context.

We will start by establishing a ranking of attributes in order to determine those that could be eliminated because they are the less relevant in the extraction of information. In case that with this step the context is not sufficiently reduced to allow extracting information without great difficulty, we will apply the second step and the context will be reduced by aggregation.

In this situation we have decided to consider the ranking obtained from supports of attributes which represent the presence of the symptoms in the diseases. Once the supports are calculated we will eliminate those attributes that occupy the highest positions in the ranking, since the fact that the support is high means that the corresponding symptom is present in almost all the diseases and, therefore, it does not provide relevant information for a differential diagnosis.

We eliminate those attributes which support is greater than 0.7, that is, they are present in at least 70% of the patients and they are not determining symptoms to make an appropriate diagnosis. So the eliminated attributes are (See Table 2),

$$\{y_1, y_2, y_{17}, y_{32}, y_3, y_{16}, y_{19}\}$$

We define a new L -fuzzy context (L, X, \hat{Y}, \hat{R}) without the eliminated attributes. The new relation is in Table 3.

Since the new context still has a large number of attributes, we apply the *Context*

Table 1. Relation of the L -fuzzy context

R	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}
x_1	0.99	1	1	0.51	0.44	0	0.03	0	0.79	0.79	0.29
x_2	1	1	0.66	0.85	0.02	0	0.02	0	0.07	0.08	0.05
x_3	0.99	0.97	0.99	0.97	0.72	0.96	0	0.93	0.01	0.03	0.01
x_4	1	1	0.82	0.33	0.82	0	0	0	0	0	0
x_5	0.96	0.88	0.56	0.85	0	0	0.17	0	0.04	0	0
x_6	1	1	0.75	0.45	0	0	1	0	0.95	0.3	0.5

R	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}	y_{21}	y_{22}
x_1	0	0.03	0.7	0	0.17	0.99	0.53	0.99	0.97	1	0.96
x_2	0	0.36	0.72	0	0.98	0.93	0.18	0.61	0	0.11	0.02
x_3	0.97	0.14	0	0.03	0.99	0.99	0.25	0.79	0	0	0
x_4	0	0.06	0.12	0	0.98	0.92	0.24	0.65	0	0	0
x_5	0	0.08	0	1	0.62	1	0.48	0.48	0.06	0.9	0.02
x_6	0	0	0.15	0	0.9	1	0.7	0.9	0.1	0.1	0

R	y_{23}	y_{24}	y_{25}	y_{26}	y_{27}	y_{28}	y_{29}	y_{30}	y_{31}	y_{32}	y_{33}
x_1	0.54	0.7	0	0.59	0.01	0	0	0	0	0.96	0.02
x_2	0.13	0	0	0	0	0.93	0	0.02	0.02	0.93	0.02
x_3	0	0.01	0.97	0.13	0.99	0.51	0.99	0.01	0	0.99	1
x_4	0	0.02	0	0.37	0	0.96	0.02	0	0	0.98	0
x_5	0.02	0	0	0	0	0.23	0	0.02	0	0.98	0.02
x_6	0.05	0	0.05	0	0	0.7	0	0.95	1	0.95	0.05

Table 2. Support of the attributes ordered from highest to lowest

	y_1	y_2	y_{17}	y_{32}	y_3	y_{16}	y_{19}	y_4	y_{28}	y_{18}	y_{21}
supp	0.99	0.98	0.97	0.96	0.79	0.77	0.74	0.66	0.56	0.4	0.35

	y_5	y_9	y_{14}	y_7	y_{10}	y_{20}	y_{33}	y_{26}	y_{15}	y_{25}	y_{31}
supp	0.33	0.31	0.28	0.2	0.2	0.19	0.18	0.18	0.17	0.17	0.17

	y_{29}	y_{22}	y_{30}	y_{27}	y_{12}	y_6	y_8	y_{11}	y_{23}	y_{24}	y_{13}
supp	0.17	0.17	0.17	0.17	0.16	0.16	0.16	0.14	1.12	0.12	0.11

Reduction by Attributes Algorithm (CRAA) in order to reduce them.

First of all, we calculate the L -fuzzy concept derived for each of the attributes

Table 3. Relation of the L -fuzzy context

\hat{R}	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}
x_1	0.51	0.44	0	0.03	0	0.79	0.79	0.29	0	0.03	0.7	0
x_2	0.85	0.02	0	0.02	0	0.07	0.08	0.05	0	0.36	0.72	0
x_3	0.97	0.72	0.96	0	0.93	0.01	0.03	0.01	0.97	0.14	0	0.03
x_4	0.33	0.82	0	0	0	0	0	0	0	0.06	0.12	0
x_5	0.85	0	0	0.17	0	0.04	0	0	0	0.08	0	1
x_6	0.45	0	0	1	0	0.95	0.3	0.5	0	0	0.15	0

\hat{R}	y_{18}	y_{20}	y_{21}	y_{22}	y_{23}	y_{24}	y_{25}	y_{26}	y_{27}	y_{28}	y_{29}	y_{30}
x_1	0.53	0.97	1	0.96	0.54	0.7	0	0.59	0.01	0	0	0
x_2	0.18	0	0.11	0.02	0.13	0	0	0	0	0.93	0	0.02
x_3	0.25	0	0	0	0	0.01	0.97	0.13	0.99	0.51	0.99	0.01
x_4	0.24	0	0	0	0	0.02	0	0.37	0	0.96	0.02	0
x_5	0.48	0.06	0.9	0.02	0.02	0	0	0	0	0.23	0	0.02
x_6	0.7	0.1	0.1	0	0.05	0	0.05	0	0	0.7	0	0.95

\hat{R}	y_{31}	y_{33}
x_1	0	0.02
x_2	0.02	0.02
x_3	0	1
x_4	0	0
x_5	0	0.02
x_6	1	0.05

\mathcal{C}_{y_j} and, for $\alpha = 1$ we obtain the sets of α -attributes associated with these L -fuzzy concepts:

$$\begin{aligned}
Q^1(\mathcal{C}_{y_6}) &= \{y_4, y_6, y_{12}, y_{25}, y_{27}, y_{29}, y_{33}\} \\
Q^1(\mathcal{C}_{y_8}) &= \{y_4, y_6, y_8, y_{12}, y_{25}, y_{27}, y_{29}, y_{33}\} \\
Q^1(\mathcal{C}_{y_{11}}) &= \{y_9, y_{11}, y_{18}\} \\
Q^1(\mathcal{C}_{y_{12}}) &= \{y_4, y_{12}, y_{25}, y_{27}, y_{29}, y_{33}\} \\
Q^1(\mathcal{C}_{y_{13}}) &= \{y_4, y_{13}\} \\
Q^1(\mathcal{C}_{y_{20}}) &= \{y_{20}, y_{21}\} \\
Q^1(\mathcal{C}_{y_{22}}) &= \{y_{21}, y_{22}\} \\
Q^1(\mathcal{C}_{y_{25}}) &= \{y_4, y_{25}, y_{33}\} \\
Q^1(\mathcal{C}_{y_{27}}) &= \{y_{27}, y_{33}\} \\
Q^1(\mathcal{C}_{y_{30}}) &= \{y_9, y_{30}\} \\
Q^1(\mathcal{C}_{y_{31}}) &= \{y_7, y_{31}\} \\
Q^1(\mathcal{C}_{y_j}) &= \{y_j\} \text{ in other case}
\end{aligned}$$

From these sets of attributes we construct the graph $G_Y^1(Y, \mathcal{E}_Y^1)$ represented in Figure 1 such that the edge $(y_j, y_k) \in \mathcal{E}_Y^1$ if $y_k \in Q^1(\mathcal{C}_{y_j})$.

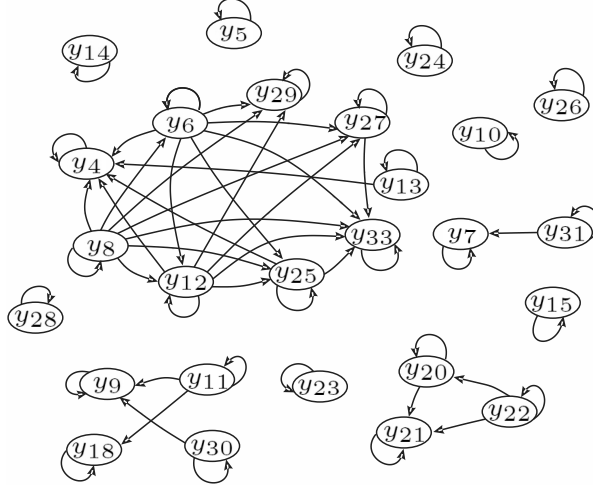


Figure 1. Graph obtained from the sets of α -attributes $Q^1(\mathcal{C}_{y_j})$

There are twelve connected components in the undirected graph (see Figure 2):

$$\begin{aligned}\tilde{Y}_1 &= \{y_4, y_6, y_8, y_{12}, y_{13}, y_{25}, y_{27}, y_{29}, y_{33}\} \\ \tilde{Y}_2 &= \{y_9, y_{11}, y_{18}, y_{30}\} \\ \tilde{Y}_3 &= \{y_{20}, y_{21}, y_{22}\} \\ \tilde{Y}_4 &= \{y_7, y_{31}\} \\ \tilde{Y}_5 &= \{y_5\} \\ \tilde{Y}_6 &= \{y_{10}\} \\ \tilde{Y}_7 &= \{y_{14}\} \\ \tilde{Y}_8 &= \{y_{15}\} \\ \tilde{Y}_9 &= \{y_{23}\} \\ \tilde{Y}_{10} &= \{y_{24}\} \\ \tilde{Y}_{11} &= \{y_{26}\} \\ \tilde{Y}_{12} &= \{y_{28}\}\end{aligned}$$

that form the set of attributes of the new L -fuzzy context α -reduced by attributes (L, X, Y^1, R_A^1) .

The relation of the reduced context (L, X, Y^1, R_A^1) is calculated aggregating the values corresponding to the elements that form the connected components by the Y -Choquet integral as:

$$\forall(x_i, \tilde{Y}_j) \in X \times Y^1, R_A^1(x_i, \tilde{Y}_j) = Ch_{\rho}^{\tilde{Y}_j}(x_i)$$

where, in order to use an additive measure, ρ has been defined from the overlap index O_{π} . The obtained relation can be seen in Table 4.

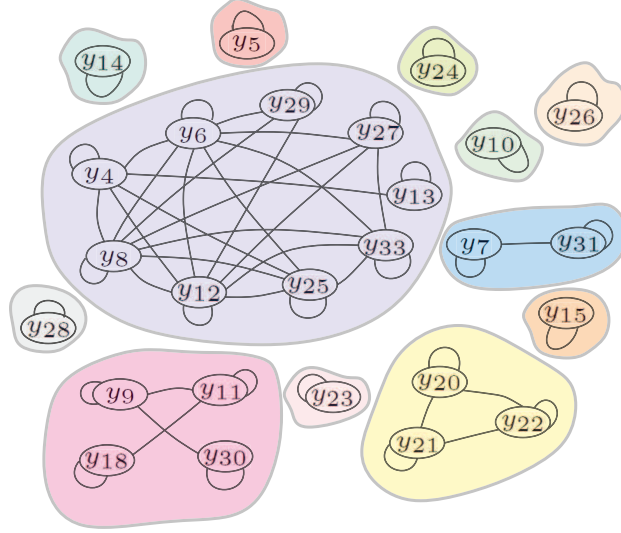


Figure 2. Connected components of the graph

Table 4. Relation of the L -fuzzy context α -reduced by attributes

R_A^1	\tilde{Y}_1	\tilde{Y}_2	\tilde{Y}_3	\tilde{Y}_4	\tilde{Y}_5	\tilde{Y}_6	\tilde{Y}_7	\tilde{Y}_8	\tilde{Y}_9	\tilde{Y}_{10}	\tilde{Y}_{11}	\tilde{Y}_{12}
x_1	0.01	0.33	0.97	0.01	0.44	0.79	0.7	0	0.54	0.7	0.59	0
x_2	0.04	0.07	0.02	0.02	0.02	0.08	0.72	0	0.13	0	0	0.93
x_3	0.9	0.06	0	0	0.72	0.03	0	0.03	0	0.01	0.13	0.51
x_4	0.01	0.04	0	0	0.82	0	0.12	0	0	0.02	0.37	0.96
x_5	0.02	0.1	0.11	0.08	0	0	0	1	0.02	0	0	0.23
x_6	0.02	0.74	0.05	1	0	0.3	0.15	0	0.05	0	0	0.7

Using this new context we can obtain diagnoses calculating the L -fuzzy concepts derived from the sets of attributes that represent the situations reported by the patients.

For example, let us suppose that the patient reports the symptoms $y_1, y_2, y_3, y_4, y_5, y_9, y_{10}, y_{17}, y_{18}, y_{19}, y_{20}, y_{21}, y_{22}, y_{23}, y_{24}$ and y_{32} . We represent this situation by the L -fuzzy set $B \in L^Y$ such that $B(y) = 1$ if the symptom y is reported by the patient and $B(y) = 0$ otherwise.

After the corresponding aggregations we obtain the subset of attributes:

$$\tilde{B} = \{\tilde{Y}_1/0.01, \tilde{Y}_2/0.33, \tilde{Y}_3/1, \tilde{Y}_4/0, \tilde{Y}_5/1, \tilde{Y}_6/1, \tilde{Y}_7/0, \tilde{Y}_8/0, \tilde{Y}_9/1, \tilde{Y}_{10}/1, \tilde{Y}_{11}/0, \tilde{Y}_{12}/0\}$$

The extension of the L -fuzzy concept derived from the set \tilde{B} is:

$$\tilde{B}_2 = \{x_1/0.44, x_2/0, x_3/0, x_4/0, x_5/0, x_6/0\}$$

The diagnosis is provided by the element that has the highest membership value. Hence, in this situation we will diagnose the disease x_1 .

If, for instance, the reported symptoms are $y_1, y_2, y_3, y_4, y_{14}, y_{16}, y_{17}, y_{28}$ and y_{32} ,

the subset of attributes given by this situation is:

$$\tilde{B} = \{\tilde{Y}_1/0.01, \tilde{Y}_2/0, \tilde{Y}_3/0, \tilde{Y}_4/0, \tilde{Y}_5/0, \tilde{Y}_6/0, \tilde{Y}_7/1, \tilde{Y}_8/0, \tilde{Y}_9/0, \tilde{Y}_{10}/0, \tilde{Y}_{11}/0, \tilde{Y}_{12}/1\}$$

and the extension of the derived L -fuzzy concept:

$$\tilde{B}_2 = \{x_1/0, x_2/0.72, x_3/0, x_4/0.12, x_5/0, x_6/0.15\}$$

Therefore, in this case, the diagnosed disease should be x_2 .

As can be seen, the diagnoses obtained in the examples are the same as those that were obtained in [Güvenir, Demiröz, and Ilter (1998)] starting from the same situations. The advantage in this case is the lower computational cost (the time complexity is $\mathcal{O}(n|Y^\alpha|) \subseteq \mathcal{O}(nm)$). It is noteworthy that, once the reduced L -fuzzy context is calculated, obtaining a new diagnosis simply consists in calculating the L -fuzzy concept derived from the set of attributes that represents the symptoms related by patient.

6. Conclusions

In this work, we have seen that Choquet integrals can be useful tools for the treatment of L -fuzzy contexts with large size. Specifically, these operators are suitable for the aggregation of rows or columns of the L -fuzzy context. In addition, the use of measures obtained from overlap indexes allows to consider the existing relationship between objects or attributes in the aggregation process. Finally, we have shown the efficiency of the developed theory to solve the problem of finding differential diagnoses of diseases that share a large number of symptoms.

Acknowledgments

This paper is partially supported by the Research Group “Intelligent Systems and Energy (SI+E)” of the University of the Basque Country - UPV/EHU, under grant GIU 16/54, and by the Research Group “Artificial Intelligence and Approximate Reasoning” of the Public University of Navarra, under TIN2016-77356-P.

References

- Alcalde, C., and A. Burusco. 2012. “Some results on the composition of L -fuzzy contexts.” In *Advances in Computational Intelligence*, Vol. 298 of *Communications in Computer and Information Science*, 305–314. Springer.
- Alcalde, C., and A. Burusco. 2014. “ L -fuzzy context sequences on complete lattices.” In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Vol. 444 of *Communications in Computer and Information Science*, 31–40. Springer.
- Alcalde, C., and A. Burusco. 2015a. “The use of two relations in L -fuzzy contexts.” *Information Sciences* 301 (1): 1–12.
- Alcalde, C., and A. Burusco. 2015b. “WOWA operators in fuzzy context sequences.” In *2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology (IFSA-EUSFLAT-15)*, 357–362.

- Alcalde, C., and A. Burusco. 2017. "On the use of Choquet integrals in the reduction of the size of L -fuzzy contexts." In *2017 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2017*, 1–6.
- Alcalde, C., and A. Burusco. 2018a. "Multivalued contexts associated with criteria." *International Journal of General Systems* 47 (2): 118–136.
- Alcalde, C., and A. Burusco. 2018b. "Study of the relevance of objects and attributes of L -fuzzy contexts Using Overlap Indexes." In *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations - 17th International Conference, IPMU 2018, Proceedings, Part I*, 537–548.
- Alcalde, C., A. Burusco, H. Bustince, R. Fuentes-González, and M. Sesma-Sara. 2017a. "Linking mathematical morphology and L -fuzzy concepts." *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems* 25 (suppl. 1): 73–98.
- Alcalde, C., A. Burusco, H. Bustince, A. Jurio, and J.A. Sanz. 2016. "Evolution in time of the L -fuzzy context sequences." *Information Sciences* 326: 202–214.
- Alcalde, C., A. Burusco, J.C. Díaz-Moreno, and J. Medina. 2017b. "Fuzzy concept lattices and fuzzy relation equations in the retrieval processing of images and signals." *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems* 25 (suppl. 1): 99–120.
- Alcalde, C., A. Burusco, and R. Fuentes-González. 2005. "Treatment of the incomplete information in L -Fuzzy contexts." In *Proceedings of the Joint 4th Conference of the European Society for Fuzzy Logic and Technology and the 11th Rencontres Francophones sur la Logique Floue et ses Applications, Barcelona, Spain, September 7-9, 2005*, 518–523.
- Alcalde, C., A. Burusco, and R. Fuentes-González. 2013. "Application of OWA operators in the L -fuzzy concept analysis." In *Aggregation Functions in Theory and in Practise - Proceedings of the 7th International Summer School on Aggregation Operators, AGOP 2013*, Vol. 228 of *Advances in Intelligent Systems and Computing*, 129–140.
- Alcalde, C., A. Burusco, and R. Fuentes-González. 2013. "The study of fuzzy context sequences." *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems* 6 (3): 518–529.
- Alcalde, C., A. Burusco, and R. Fuentes-González. 2014. "Application of the L -fuzzy concept analysis in the morphological image and signal processing." *Annals of Mathematics and Artificial Intelligence* 72 (1-2): 115–128.
- Alcalde, C., A. Burusco, R. Fuentes-González, and I. Zubia. 2009. "Treatment of L -fuzzy contexts with absent values." *Information Sciences* 179 (1-2): 1–15.
- Belohlávek, R. 1999. "Fuzzy Galois Connections." *Mathematical Logic Quarterly* 45 (4): 497–504.
- Belohlávek, R. 2000. "Similarity relations in concept lattices." *Journal of Logic and Computation* 10 (6): 823–845.
- Belohlávek, R. 2002. *Fuzzy Relational Systems: Foundations and Principles*. Kluwer Academic Publishers.
- Belohlávek, R., and V. Vychodil. 2005. "Reducing the size of fuzzy concept lattices by hedges." In *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2005*, 663–668.
- Belohlávek, R., and V. Vychodil. 2006. "Reducing the size of fuzzy concept lattices by fuzzy closure operators." In *SCIS & ISIS 2006, Int. Conf. Soft Computing and Intelligent Systems & Int. Symposium on Intelligent Systems, Sep 20-24, 2006, Tokyo, Japan.*, 309–314.
- Belohlávek, R., and V. Vychodil. 2012. "Formal concept analysis and linguistic hedges." *International Journal of General Systems* 41 (5): 503–532.
- Belohlávek, R., and V. Vychodil. 2013. "Closure based constraints in formal concept analysis." *Discrete Applied Mathematics* 161 (13-14): 1894–1911.
- Burusco, A., and R. Fuentes-González. 1994. "The study of the L -fuzzy concept lattice." *Mathware & Soft Computing* 1 (3): 209–218.
- Burusco, A., and R. Fuentes-González. 1998. "Construction of the L -fuzzy concept lattice." *Fuzzy Sets and Systems* 97 (1): 109–114.
- Burusco, A., and R. Fuentes-González. 2000. "Concept lattices defined from implication operators." *Fuzzy Sets and Systems* 114: 431–436.

- Burusco, A., and R. Fuentes-González. 2003. “Relevant information extraction in L-Fuzzy contexts.” *Revista Internacional de Información Tecnológica* 14 (4): 65–70.
- Calvo, T., and R. Mesiar. 2003a. “Aggregation operators: ordering and bounds.” *Fuzzy Sets and Systems* 139 (3): 685–697.
- Calvo, T., and R. Mesiar. 2003b. “Weighted triangular norms-based aggregation operators.” *Fuzzy Sets and Systems* 137 (1): 3–10.
- Choquet, G. 1953. “Theory of capacities.” *Annales de L’Institut Fourier* 5: 131–295.
- Djouadi, Y., and H. Prade. 2010. “Interval-valued fuzzy Galois connections: Algebraic requirements and concept lattice construction.” *Fundamenta Informaticae* 99 (2): 169–186.
- Djouadi, Y., and H. Prade. 2011. “Possibility-theoretic extension of derivation operators in formal concept analysis over fuzzy lattices.” *FO & DM* 10 (4): 287–309.
- Dua, D., and E. Karra Taniskidou. 2017. “UCI Machine Learning Repository.” <http://archive.ics.uci.edu/ml>.
- Fodor, J.C., J.L. Marichal, and M. Roubens. 1995. “Characterization of the ordered weighted averaging operators.” *IEEE Trans. Fuzzy Systems* 3 (2): 236–240.
- Grabisch, M. 1995. “Fuzzy integral in multicriteria decision making.” *Fuzzy Sets and Systems* 69: 279–298.
- Güvenir, H. A., G. Demiröz, and N. Ilter. 1998. “Learning differential diagnosis of erythematous diseases using voting feature intervals.” *Artificial Intelligence in Medicine* 13 (3): 147–165.
- Klir, G.J., and Z. Wang. 2009. *Generalized Measure Theory*. Springer.
- Konecny, J., and M. Krupka. 2011. “Block Relations in Fuzzy Setting.” In *Proceedings of The Eighth International Conference on Concept Lattices and Their Applications, 2011*, 115–130.
- Li, L., and J. Zhang. 2013. “Attribute reduction in fuzzy concept lattices based on the T implication.” *Knowledge-Based Systems* 23 (6): 497–50.
- Medina, J. 2012. “Relating attribute reduction in formal, object-oriented and property-oriented concept lattices.” *Computers & Mathematics with Applications* 64 (6): 1992–2002.
- Medina, J., and M. Ojeda-Aciego. 2010. “Multi-adjoint t-concept lattices.” *Information Sciences* 180 (5): 712–725.
- Medina, J., and M. Ojeda-Aciego. 2013. “Dual multi-adjoint concept lattices.” *Information Sciences* 225: 47–54.
- Paternain, D., H. Bustince, M. Pagola, P. Sussner, A. Kolesárová, and R. Mesiar. 2016. “Capacities and overlap indexes with an application in fuzzy rule-based classification systems.” *Fuzzy Sets and Systems* 305: 70–94.
- Pollandt, S. 1997. *Fuzzy Begriffe*. Springer.
- Torra, V. 1998. “On some relationships between the WOWA operator and the Choquet integral.” In *Proceedings of the Seventh International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU’98)*, 818–824.
- Wei, L., and J.J. Qi. 2010. “Relation between concept lattice reduction and rough set reduction.” *Knowledge-Based Systems* 23 (8): 934–938.
- Wille, R. 1982. “Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts.” In *Ordered Sets*, edited by I. Rival, Vol. 83 of *NATO Advanced Study Institutes Series*, 445–470. Springer Netherlands.
- Yager, R. 1998. “On ordered weighted averaging aggregation operators in multi-criteria decision making.” *IEEE Transactions on Systems, Man, and Cybernetics* 18: 183–190.

Notes on authors

Cristina Alcalde received her PhD degree in Mathematics from the University of the Basque Country, Spain, in 2007. She is an Associate Professor in the Department of Applied Mathematics at the University of the Basque Country. Her current research interests are artificial intelligence, fuzzy sets theory, knowledge acquisition, L-fuzzy

concept analysis and mathematical morphology.

Ana Burusco is an Associate Professor in the Department of Statistics, Computer Science and Mathematics at the Public University of Navarra, Spain. She received her PhD degree in Mathematics from the Public University of Navarra. She is member of the research group Artificial Intelligence and Approximate Reasoning and of the research Institute of Smart Cities since its foundation. Her research interests are in fuzzy sets theory, artificial intelligence, knowledge acquisition and L-fuzzy concept analysis.