Computing and Informatics, Vol. 40, 2021, 469–488, doi: 10.31577/cai_2021_3_469

CONCEPT SIMILARITY IN FORMAL CONCEPT ANALYSIS WITH MANY-VALUED CONTEXTS

Anna Formica

Istituto di Analisi dei Sistemi ed Informatica (IASI)
National Research Council
Via dei Taurini 19, I-00185, Rome, Italy
e-mail: anna.formica@iasi.cnr.it

Abstract. Formal Concept Analysis (FCA) is a mathematical framework which can also support critical activities for the development of the Semantic Web. One of them is represented by Similarity Reasoning, i.e., the identification of different concepts that are semantically close, that allows users to retrieve information on the Web more efficiently. In order to model uncertainty information, in this paper FCA with many-valued contexts is addressed, where attribute values are intervals, which is referred to as FCA with Interordinal scaling (IFCA). In particular, a method for evaluating concept similarity in IFCA is proposed, which is a problem that has not been adequately investigated, although the increasing interest in the literature in this topic.

Keywords: Formal concept analysis, similarity reasoning, many-valued contexts, FCA with interordinal scaling

1 INTRODUCTION

Formal Concept Analysis (FCA) is a formal framework based on lattice theory which is commonly used for data analysis [15, 28]. In the basic setting, FCA attributes are crisp, i.e., any object either has or does not have an attribute of a given context. This is the case of the so-called *one-valued* contexts. However, in real life most of attributes are fuzzy rather than crisp, i.e., "it is a matter of degree to which an object has a (fuzzy) attribute" [2]. In other words, an object may have different attributes with different values, and an attribute may apply to different objects with different values. This is the case of *many-valued* contexts [15]. *Fuzzy Formal Concept*

Analysis (FFCA) is a generalization of FCA which provides a formal framework for structuring, analyzing and visualizing data in the presence of uncertainty information [32]. In particular, in FFCA contexts are many-valued, and the attribute values are real numbers in the range $[0, 1]$. In this paper, this kind of FCA is referred to as OFCA, in line with the notion of FCA with *Ordinal scaling* defined in [15].

Regarding the notion of fuzzy sets [7], *Type-1 Fuzzy Sets* (T1 FSs) and *Type-2 Fuzzy Sets* (T2 FSs) were both introduced by Zadeh, the former in 1965 in the seminal paper [36], and the latter in 1975 in [37]. T2 FSs provide a way to overcome one of the early objections made about T1 FSs, i.e., that "it sounds contradictory for something that is *fuzzy* to have a perfectly defined membership function" [26]. In this paper we focus on *Interval Type-2 FSs* (IT2 FSs) [22], which represent a simplification about T2 FSs that is receiving much attention in the literature in different research areas, with different purposes, as for instance in [1, 8] just to mention a couple of examples.

Similarity Reasoning, i.e., the identification of syntactically different concepts that are semantically close, is fundamental in several research areas such as Cognitive Science, Artificial Intelligence, Software Engineering, and in the Semantic Web [4, 17]. Concept similarity in the framework of FCA with *Interordinal scaling* (IFCA), i.e., in many-valued contexts where attribute values are intervals, is a problem that has been marginally investigated in the literature, despite of an increasing interest in this topic.

In this paper, a concept similarity measure for IFCA is proposed which is novel because it combines the IT2 FS framework, with regard to concept extents, and the *information content* approach [23], with regard to concept intents. The latter has been extensively investigated and experimented in the literature, and has a higher correlation with human judgment with respect to the traditional approaches. In particular, this paper is a short and revised version of the results presented in [14]. Furthermore, in this work both the basic notions and the overall approach are presented informally, by providing simple examples, in order to reach a broad audience of readers, and not only specialists working in the area.

The paper is organized as follows. In the next section, the Related Work is given, and in Section 3 the basic notions related to FCA and IFCA are recalled. In Section 4 the notion of IFCA concept similarity is presented. In particular, first the similarity between IT2 FSs is recalled and, successively, the information content similarity is addressed. Then, they are combined in order to define the similarity between IFCA concepts. Finally, in Section 5 an evaluation of the method is given, and Section 6 concludes.

2 RELATED WORK

FCA techniques and tools have been employed in different research fields, such as, Information Retrieval, e-Learning, Expert Systems, etc., and in the development of the Semantic Web [32]. Similarity reasoning is fundamental in several research areas

such as Cognitive Science, Artificial Intelligence, Software Engineering and, recently, also in the Semantic Web [13]. In the development of the Semantic Web, similarity reasoning supports all the activities that, in general, require human interaction (which are time-consuming and error-prone), such as web service discovery, query refinement techniques for search engines, extractions of patterns and trends in web users behaviors, etc. Therefore, the similarity method proposed in this paper can be employed in all the research fields that can benefit from the combination of FCA techniques and similarity reasoning.

Some research challenges about the contribution of FCA in the Semantic Web development are illustrated in [18], and concern the automatic or semi-automatic generation of ontologies (ontology *engineering*), and the critical problem of identifying the overlapping knowledge in a common domain (also referred to as ontology *mapping*, *merging*, *integration*, or *alignment*). In particular, due to the presence in the web of large and specialized ontologies, FCA has been employed for more than one decade for reusing and combining independently developed domain ontologies, see for instance [31].

FCA concept similarity has been addressed in [10], by relying on human domain expertise, and in [11, 33], according to the information content approach, but in both cases within one-valued contexts. Many-valued contexts have been addressed in [13], but in the case of FCA with *Ordinal scaling* (OFCA). Therefore, in both the mentioned papers, IT2 FSs have not been addressed and a similarity measure has been proposed, based on T1 FSs, which has been experimented and compared with the relevant similarity measures proposed in the literature. It has been used as a basis for the definition of the similarity measure proposed in this paper for FCA with *Interordinal scaling* (IFCA), as illustrated in the next sections. Note that in [3] and [12] different problems related to OFCA have been addressed, by relying on Rough Set Theory.

With regard to IFCA, a formal framework, referred to as *L-Fuzzy concept theory*, has been defined in [5] which is probably the first research paper providing a theoretical foundation about it. Successively, some interesting works have been defined in the literature which have investigated and deepened the mathematics underlying specific aspects of IFCA, as for instance [6].

In [30] the need for IT2 fuzzy analytical systems for the development of the Semantic Web is emphasized, and a similarity measure for IFCA is proposed. It is based on the similarity measure for IT2 FSs defined in [35], the approach presented in [11], and relies on the experimental results given in [13]. In Section 5, a discussion about the evaluation of the proposed method is given.

As mentioned, the proposed *T2ConSim* combines the similarity of the concept extents and the concept intents. Concept extents are evaluated according to *IT2FSim*, which is the widely accepted crisp similarity measure for IT2 FSs defined in [35]. It is used in most applications of general T2 FSs due to the simpler underlying mathematics. Such a notion has been adopted here because it allows a relevant simplification about the definition of similarity between general T2 FSs, in line with the scope of this paper which is intended for non-specialist readers.

3 FCA WITH MANY-VALUED CONTEXTS

In this section the basic notions related to *Formal Concept Analysis* (FCA) are briefly recalled. In order to illustrate them, the context named *Sardinia Hotels* presented in [14] is used, which also allows us to introduce, in the next section, the notions underlying IFCA in an intuitive way.

In FCA [15], a *one-valued context* (*context* for short) is a triple (O, A, R), where O is a set of *objects*, A is a set of *attributes*, and R is a binary relation between O and A. In the *Sardinia Hotels* context mentioned below, the set O is defined by the following six objects representing six different hotels:

$$O = \{H1, H2, H3, H4, H5, H6\},$$

and the set A is defined by the three following attributes:

$$A = \{SwPool, Sea, Meal\}$$

where *SwPool* stands for swimming pool. Furthermore, the relation R among hotels and attributes is defined by Table 1.

A concept of the *Sardinia Hotels* context is, for instance, the pair (E, I) where E is a set of objects, referred to as concept *extent*, and I is a set of attributes, referred to as concept *intent*, defined as follows:

$$((H1, H3, H5), (Sea, Meal))$$

since the objects $H1$, $H3$, and $H5$ have both the attributes Sea and $Meal$, and vice versa, both these attributes apply to the objects $H1$, $H3$, and $H5$.

Intuitively, we can say that concepts correspond to maximal rectangles of crosses in the context, after appropriate permutations of rows and columns. It is possible to establish an *inheritance relation* ($\leq$) between concepts of a context, say (E_1, I_1), (E_2, I_2), as follows:

$$(E_1, I_1) \leq (E_2, I_2) \text{ iff } E_1 \subseteq E_2 (\textit{iff } I_2 \subseteq I_1).$$

In particular, (E_1, I_1) is called *subconcept* of (E_2, I_2) and (E_2, I_2) is called *superconcept* of (E_1, I_1). For instance, the concept $((H1, H2, H3, H5), (Meal))$ is a superconcept of the previous one, i.e.:

$$((H1, H3, H5), (Sea, Meal)) \leq ((H1, H2, H3, H5), (Meal))$$

and, vice versa, the former concept is a subconcept of the latter.

Given a context (O, A, R), consider the set of all the concepts of this context, indicated as $\mathcal{L}(O, A, R)$. Then:

$$(\mathcal{L}(O, A, R), \leq)$$

	SwPool	**Sea**	**Meal**
H1		×	×
H2	×		×
H3		×	×
H4	×	×	
H5		×	×
H6	×	×	

Table 1. The FCA *Sardinia Hotels* context

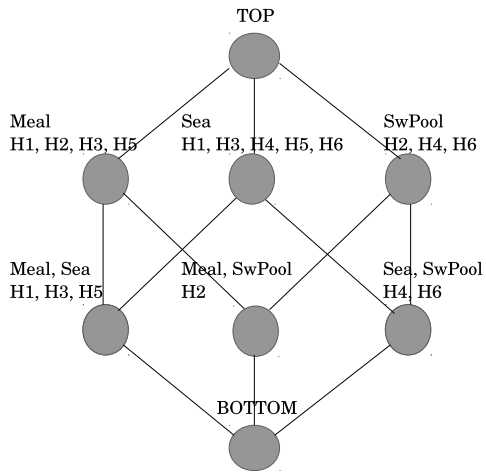

Figure 1. Concept Lattice of the *Sardinia Hotels* context [14]

is a complete lattice called *Formal Concept Lattice* (*Concept Lattice* for short), i.e., for each subset of concepts, the greatest lower bound (the greatest common subconcept) and the least upper bound (the least common superconcept) exist. For instance, the Concept Lattice constructed from the context of Table 1 is shown in Figure 1. Note that the Concept Lattice has two special nodes, the maximum and minimum nodes, grouping all the objects and the attributes of the context, respectively. The number of objects in the concept extent (the cardinality) is also referred to as the *support* the concept [19], therefore the concept corresponding to the maximum node has maximum support.

3.1 From One- to Many-Valued Contexts

In a one-valued context an attribute is a property that an object may have or may not have. For instance, according to the one-valued context *Sardinia Hotels* above, each of the attributes *SwPool*, *Sea*, and *Meal* applies or does not apply to each of the hotel objects. However, in real world, an attribute may apply to different objects with different values, i.e., it can be many-valued.

In FCA, a *many-valued context* is a quadruple (O, A, V, R), where O is a set of *objects*, A is a set of *many-valued attributes*, V is a set of *attribute values*, and R is a ternary relation among O, A, and V such that:

$$(o, a, v) \in R \text{ and } (o, a, w) \in R \Rightarrow v = w$$

where $(o, a, v) \in R$ can be read as "the attribute a has the value v for the object o". Note that (O, A, V, R) is referred to as *one-valued context* if V has one element [15].

Analogously to one-valued contexts, many-valued contexts can be represented by tables, where rows are labeled by objects and columns are labeled by attributes. Many-valued contexts can be transformed into one-valued contexts according to a *conceptual scaling* process [15]. In particular, in this process, each attribute of a many-valued context is interpreted by means of a context, referred to as *conceptual scale* (for details about the transformation process of a many-valued context into a one-valued context see [15]). Typical conceptual scales are *Nominal*, *Ordinal*, and *Interordinal* scales. Nominal scales are used for attribute values which mutually exclude each other, for instance in the case of the attribute values {*human, animal, plant*}. Ordinal scales are suitable when attribute values are ordered, and each value implies the weaker ones, e.g., {*extremely active, very active, active*}. *Interordinal scales* are used for attributes which have a range of possible values (intervals), e.g., {*fully, very much, very few, not at all*}. In the next subsection we focus on FCA with Interordinal scaling.

3.1.1 FCA with Interordinal Scaling

As mentioned above, in many-valued contexts attributes do not describe objects in a uniform way, i.e., a given attribute applies to different objects in different ways. For instance, in the *Sardinia Hotels* context above, consider the attribute $Meal$. In general, when reserving an hotel, we would like to know whether the hotel provides both lunch and dinner, or half-board. Without the introduction of fuzzy information, we have no way to specify how appropriate is an attribute to a given object.

In order to deal with fuzzy contexts, we need to recall the following definitions.

A *Type-1 Fuzzy Set* (T1 FS) A (also called *fuzzy set*) in a space of points X is characterized by a *membership function* $\mu_A(x)$ which associates each point x in X with a real number in the interval $[0, 1]$ representing the *grade of membership* of x in A [36]. Note that for an ordinary set, the membership function can take only the values 1 and 0, depending on x does or does not belong to A, respectively.

For instance, the following set A:

$$A = ((H1, 1.0), (H2, 0.5), (H3, 0.5), (H5, 1.0))$$

is a T1 FS in the space of point $X = \{H1, H2, H3, H5\}$.

An *Interval Type-2 Fuzzy Sets* (IT2 FS) $\tilde{A}$ in a space of points X is characterized by two membership functions, an *upper membership function* $\bar{\mu}_{\tilde{A}}$ and a *lower membership function* $\underline{\mu}_{\tilde{A}}$ which are both T1 FSs, such that each point x in X is associated with an interval $[\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)]$ representing the *grade of membership* of x in $\tilde{A}$ [34].

For instance, the following set $\tilde{A}$:

$$\tilde{A} = ((H2, [0.6, 0.7]), (H4, [0.6, 0.8]), (H5, [0.4, 0.9]))$$

is a IT2 FS in the space of point $X = \{H2, H4, H5\}$.

In this subsection we address FCA contexts where grades of memberships are intervals, and in particular *words*. Indeed, words are closer to human judgment when we need to quantify "how much" an object is described by an attribute or, vice versa, an attribute applies to an object [5, 30]. Possible words representing grades of membership are:

$$\{Fully, Very\ Much, Very, Few, Very\ Few, Not\ at\ all\}.$$

For instance, consider the many-valued context *Sardinia Hotels* which is specified by the fuzzy relation given in Table 2 where crosses in Table 1 have been replaced by *words*, each allowing us to specify "how much" an object has, or is described by, an attribute, and vice versa an attribute applies to an object.

Consider for instance the hotel $H2$ in Table 2. It has the attribute *SwPool* with grade of membership *Fully*, which means that such an attribute fully applies to the hotel $H2$ (and vice versa the hotel $H2$ can be properly described by the attribute *SwPool*). Instead, the object $H2$ has the attribute *Meal* with a membership value *Very*, which means that such an attribute partially applies to this hotel (for instance it could provide meals just for lunch). In order to address only objects related to attributes with relevant grades of membership, a threshold is fixed such that the pairs with membership values under the threshold are ignored. For instance, assume that in the *Sardinia Hotels* context the intervals *Very Few* and *Not at all* are below the threshold. With this assumption, the pair $(H5, Sea)$ is ignored.

For instance, consider the IFCA context for the *Sardinia Hotels* shown in Table 2.

	SwPool	**Sea**	**Meal**
H1		Fully	Fully
H2	Fully		Very
H3		Very much	Very
H4	Fully	Fully	
H5		Very Few	Fully
H6	Fully	Very much	

Table 2. The IFCA *Sardinia Hotels* context, by using words

In order to elaborate such grades of membership, words are replaced by intervals (IT2 FS grades of membership). The association of words with intervals is a problem which has been extensively investigated in the literature and is still attracting a lot of attention [25], [27]. A simple association of words with intervals is shown in Table 3. Therefore, the context of Table 2 becomes the IFCA context shown in Table 4.

Not at all	$[0.0, 0.1]$
Very few	$[0.1, 0.3]$
Few	$[0.3, 0.5]$
Very	$[0.5, 0.7]$
Very much	$[0.7, 0.9]$
Fully	$[0.9, 1.0]$

Table 3. Mapping words to intervals

	SwPool	**Sea**	**Meal**
H1		$[0.9, 1.0]$	$[0.9, 1.0]$
H2	$[0.9, 1.0]$		$[0.5, 0.7]$
H3		$[0.7, 0.9]$	$[0.5, 0.7]$
H4	$[0.9, 1.0]$	$[0.9, 1.0]$	
H5		$[0.1, 0.3]$	$[0.9, 1.0]$
H6	$[0.9, 1.0]$	$[0.7, 0.9]$	

Table 4. The IFCA *Sardinia Hotels* context

In IFCA, Concept Lattices are defined similarly to FCA Concept Lattices. For instance, the IFCA Concept Lattice constructed from the context of Table 4 is shown in Figure 2. In IFCA an object of a concept is associated with an interval, standing for the related grade of membership. In the case two or more attributes apply to an object with different grades of membership (i.e., different intervals) the object is associated with the interval having, as lower bound and upper bound, the minimum between the lower bounds and the upper bounds, respectively. For instance, consider again the concept involving the attributes Sea and $Meal$, which in this case is defined as follows:

$$(((H1, [0.9, 1.0]), (H3, [0.5, 0.7])), (Sea, Meal))$$

since, as mentioned above, the pair ($(H5$,[0.1,0.3]), Sea) is not considered because under the threshold. According to the context shown in Table 4, Sea and $Meal$ apply to $H3$ with different intervals, that are $[0.7, 0.9]$ and $[0.5, 0.7]$, respectively. Since 0.5 is the minimum between their lower bounds, and 0.7 is the minimum between their upper bounds, in the concept above the object $H3$ has been associated with the interval $[0.5, 0.7]$. Indeed this interval represents the highest common grade of

membership that allows $H3$ to be described by both the attributes Sea and $Meal$ (and, vice versa, both the attributes Sea and $Meal$ to be applied to $H3$).

In the following section, the similarity between concepts in IFCA is addressed.

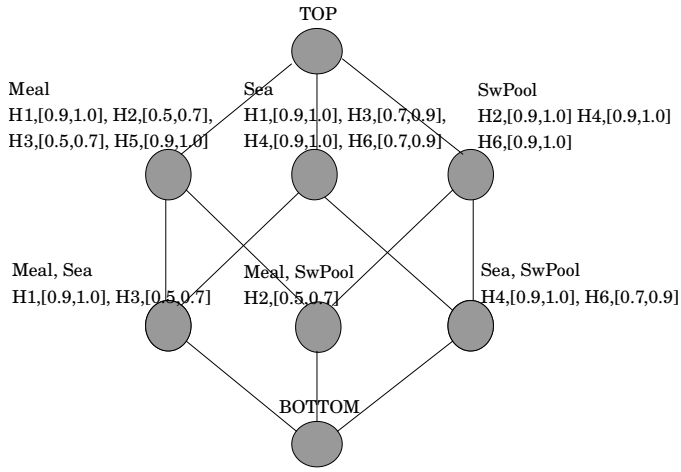

Figure 2. Concept Lattice of the IFCA *Sardinia Hotels* context [14]

4 IFCA CONCEPT SIMILARITY

In this section IFCA concept similarity is computed by combining the similarity of concept extents, i.e., the IT2 FSs of objects, and the similarity of concept intents, i.e., the sets of attributes.

4.1 Concept Extent Similarity

With regard to the similarity of concept extents, we need to recall a few basic notions about IT2 FSs. Note that in the literature, the notions of similarity between T2 FSs have been proposed by several authors, and are based on different underlying definitions (as for instance the notion of cardinality) [16, 34]. Below we focus on the definitions that are the most frequently used in the literature, which require a simpler mathematics with respect to the others.

4.1.1 Similarity Between IT2 FSs

In this subsection, the notions of cardinality and *average cardinality* of an IT2 FS are recalled [16]. To this end, we first need to remind that the *cardinality* of a T1 FS A in a space of points X, also referred to as *power* of the T1 FS A, and denoted

as $p(A)$, is given by the sum of all membership grades, i.e.:

$$p(A) = p(\mu_A(x)) = \sum_{i=1}^{N} \mu_A(x_i). \tag{1}$$

For instance, the cardinality of the set A:

$$A = ((H1, 1.0), (H2, 0.5), (H3, 0.5), (H5, 1.0))$$

is:

$$p(A) = 1.0 + 0.5 + 0.5 + 1.0 = 3.0.$$

Given an IT2 FS $\tilde{A}$, the *cardinality* of $\tilde{A}$, denoted as $P(\tilde{A})$, is an interval defined as follows:

$$P(\tilde{A}) = [p(\underline{\mu}_{\tilde{A}}(x)), p(\bar{\mu}_{\tilde{A}}(x))] \tag{2}$$

where $p(\underline{\mu}_{\tilde{A}})$, and $p(\bar{\mu}_{\tilde{A}})$ are the cardinalities of the lower and upper membership functions, respectively, which are T1 FSs. The *average cardinality* of an IT2 FS $\tilde{A}$, indicated as $AC(\tilde{A})$, is defined as the average of its minimum and maximum cardinalities, i.e.:

$$AC(\tilde{A}) = \frac{p(\underline{\mu}_{\tilde{A}}(x)) + p(\bar{\mu}_{\tilde{A}}(x))}{2}. \tag{3}$$

For instance, consider the IT2 FS $\tilde{A}$:

$$\tilde{A} = ((H2, [0.6, 0.7]), (H4, [0.6, 0.8]), (H5, [0.4, 0.9])).$$

The cardinality of $\tilde{A}$, $P(\tilde{A})$, is the interval having as lower and upper bounds the sums of the grades of the lower and upper membership functions, respectively, therefore:

$$P(\tilde{A}) = [1.6, 2.4]$$

because:

$$p(\underline{\mu}_{\tilde{A}}) = 0.6 + 0.6 + 0.4 = 1.6$$

and:

$$p(\bar{\mu}_{\tilde{A}}) = 0.7 + 0.8 + 0.9 = 2.4.$$

Then, the average cardinality $AC(\tilde{A})$ is the following:

$$AC(\tilde{A}) = (1.6 + 2.4)/2 = 2.$$

Let us now address the intersection and union of IT2 FSs. The intersection, $\tilde{A} \cap \tilde{B}$, and union, $\tilde{A} \cup \tilde{B}$, of the IT2 FSs $\tilde{A}$ and $\tilde{B}$ are both IT2 FSs. In particular, the membership grades of an element x are intervals defined, respectively, according to the lower and upper membership functions as follows:

$$\tilde{A} \cap \tilde{B}(x) = [\min(\underline{\mu}_{\tilde{A}}(x_i), (\underline{\mu}_{\tilde{B}}(x_i)), \min(\bar{\mu}_{\tilde{A}}(x_i), (\bar{\mu}_{\tilde{B}}(x_i))], \tag{4}$$

$$\tilde{A} \cup \tilde{B}(x) = [\max(\underline{\mu}_{\tilde{A}}(x_i), (\underline{\mu}_{\tilde{B}}(x_i)), \max(\bar{\mu}_{\tilde{A}}(x_i), (\bar{\mu}_{\tilde{B}}(x_i))]. \tag{5}$$

On the basis of these notions, we are now able to recall the similarity between IT2 FSs. In particular, we follow the crisp similarity measure proposed by [35], here referred to as *IT2FSim*, which is defined below:

$$IT2FSim(\tilde{A}, \tilde{B}) = \frac{AC(\tilde{A} \cap \tilde{B})}{AC(\tilde{A} \cup \tilde{B})} \tag{6}$$

and, therefore:

$$IT2FSim(\tilde{A}, \tilde{B}) = \frac{\sum_{i=1}^{N} \min(\bar{\mu}_{\tilde{A}}(x_i), (\bar{\mu}_{\tilde{B}}(x_i)) + \sum_{i=1}^{N} \min(\underline{\mu}_{\tilde{A}}(x_i), (\underline{\mu}_{\tilde{B}}(x_i))}{\sum_{i=1}^{N} \max(\bar{\mu}_{\tilde{A}}(x_i), (\bar{\mu}_{\tilde{B}}(x_i)) + \sum_{i=1}^{N} \max(\underline{\mu}_{\tilde{A}}(x_i), (\underline{\mu}_{\tilde{B}}(x_i))}. \tag{7}$$

For instance, consider the previous set $\tilde{A}$, and the set $\tilde{B}$ below:

$$\tilde{B} = ((H1, [0.4, 0.9]), (H2, [0.7, 0.8]), (H5, [0.3, 1.0])).$$

Then:

$$\tilde{A} \cap \tilde{B} = ((H2, [0.6, 0.7]), (H5, [0.3, 0.9])),$$
$$AC(\tilde{A} \cap \tilde{B}) = ((0.6 + 0.3) + (0.7 + 0.9))/2 = 1.25.$$

and:

$$\tilde{A} \cup \tilde{B} = ((H1, [0.4, 0.9]), (H2, [0.7, 0.8]), (H4, [0.6, 0.8]), (H5, [0.4, 1.0])),$$
$$AC(\tilde{A} \cup \tilde{B}) = ((0.4 + 0.7 + 0.6 + 0.4) + (0.9 + 0.8 + 0.8 + 1))/2 = 2.8.$$

Therefore:

$$IT2FSim(\tilde{A}, \tilde{B}) = \frac{AC(\tilde{A} \cap \tilde{B})}{AC(\tilde{A} \cup \tilde{B})} = \frac{1.25}{2.8} = 0.45,$$

which is a crisp measure of the similarity between the IT2 FSs $\tilde{A}$, and $\tilde{B}$. Such a measure is used in order to evaluate the similarity of concept extents in IT2 Fuzzy Concept Lattices.

4.2 Concept Intent Similarity

In order to address the similarity of concept intents, we need to briefly recall the notion of *information content similarity*. It is based on the well-known notion of *information content*, which has been extensively investigated in the literature [23].

4.2.1 Information Content Similarity

Let us consider a *lexical database for the English language* as, for instance, *WordNet* [9]. Besides English concept nouns, *WordNet* contains verbs, adjectives and

adverbs, each associated with the related natural language definition and *frequency*. Frequencies are estimated using noun frequencies from large text corpora, as for instance the *Brown Corpus of American English*. Concept nouns are organized according to the *ISA* and *PartOf* relationships, and for each concept noun, a set of synonyms is given. In order to deal with the *information content* approach, below we focus on (fragments of) *WordNet ISA* hierarchies and, for the sake of simplicity, without addressing sets of synonyms. The *probability* of a concept noun c, $p(c)$, is defined as:

$$p(c) = \frac{freq(c)}{M} \tag{8}$$

where $freq(c)$ is the frequency of c from a text corpus, and M is the total number of observed instances of nouns in the corpus. In this paper probabilities have been assigned according to the *SemCor* project, which labels subsections of the *Brown Corpus* to senses in the *WordNet* lexicon. In Figure 3, the simple fragment of *ISA* hierarchy presented in [13] is recalled, where each concept is associated with the related probability.

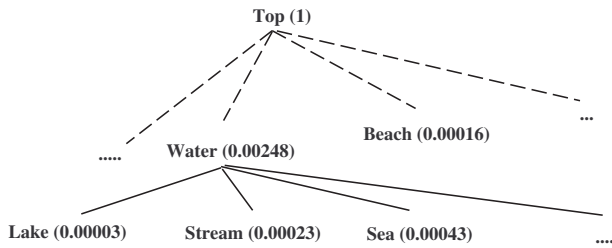

Figure 3. A fragment of *ISA* hierarchy from *WordNet* [13]

The *information content* of a concept noun c is defined as $-\log p(c)$, that is, as the probability of a concept noun increases, the informativeness decreases, therefore the more abstract a concept noun, the lower its information content. The similarity between hierarchically organized concept nouns is given by the maximum information content shared by the concepts, that is, the more information two concepts share, the more similar they are. Given a hierarchy of concept nouns organized according to a tree (also referred to as *taxonomy*), consider two concept nouns of this hierarchy, say c_1, c_2. Then, the maximum information content shared by c_1, c_2 in the taxonomy is provided by the superconcept of c_1, c_2 whose information content is maximum, i.e., the *least common superconcept* (*lcs*). In this paper we focus on concept hierarchies which are trees, therefore the *lcs* of two concept nouns always exists. Starting from these assumptions, the *information content similarity* (*ics*) of two concept nouns is defined by the maximum information content shared by the concepts divided by the information contents of the comparing concepts [23].

For instance, in the case of *Lake* and *Sea*, *Water* is their *lcs* in the hierarchy, and therefore:

$$ics(Lake, Sea) = \frac{2 \log p(Water)}{\log p(Lake) + \log p(Sea)} = \frac{2 \cdot 8.66}{14.85 + 11.18} = 0.67.$$

Below, the *ics* is used in order to compute the similarity between sets of concept nouns, i.e., between concept intents.

4.2.2 Similarity Between Sets of Attributes

In the following, since concept intents are defined by sets of attributes, we refer to attributes rather than concept nouns. The comparison between concept intents is performed according to the *Hungarian algorithm* in polynomial time [21]. Informally, given a lexical database for the English language, consider two sets of attributes, say I_1, I_2, defined in the lexical database. Let a *candidate set of pairs* be a subset of $I_1 \times I_2$ such that there are no two pairs in the set sharing an element. For instance, assume that I_1 and I_2 represent a set of boys and a set of girls, respectively, a candidate set of pairs defines a possible set of marriages (when polygamy is not allowed). Within all possible candidate sets of pairs, consider (one of) the set(s) such that the sum of the information content similarity (*ics*) of the pairs is maximal (*maximum weighted matching* problem in bipartite graphs [11]). Such a sum is indicated as $\mathcal{M}(I_1, I_2)$. Then, the similarity between the sets of attributes I_1, and I_2, $ASim(I_1, I_2)$ is defined as follows:

$$ASim(I_1, I_2) = \frac{\mathcal{M}(I_1, I_2)}{n} \tag{9}$$

where n is the greatest between the cardinalities of I_1, and I_2.

For instance in our running example, assume $I_1 = \{SwPool, Sea\}$, and $I_2 = \{SwPool, Lake\}$. In this simple case, within the two possible sets of pairs of attributes that can be formed with I_1 and I_2 as described above, the set of pairs with maximal sum is the following:

$$\{(SwPool, SwPool), (Sea, Lake)\},$$

because, of course, $ics(SwPool, SwPool) = 1$, and $ics(Sea, Lake) = 0.67$. Therefore:

$$\mathcal{M}((SwPool, Sea), (SwPool, Lake)) = 1.67,$$

whereas the other possible set of pairs:

$$\{(SwPool, Lake), (Sea, SwPool)\}$$

leads to a null value (the *ics* of both the pairs are null because, according to the *ISA* hierarchy of Figure 3, *SwPool* does not share any information content neither

with *Lake*, nor with *Sea*). As a result:

$$ASim(I_1, I_2) = \frac{\mathcal{M}(I_1, I_2)}{2} = 0.84.$$

Now we are able to evaluate the similarity between IFCA concepts, on the basis of the similarity of concept extents and the similarity of concept intents defined above.

4.3 Similarity Between IFCA Concepts

In this section, the notion of similarity between IFCA concepts, referred to as *T2ConSim*, is presented. It is essentially given by the weighted average between the similarity of the concept extents and the similarity of concept intents above. Formally, given two concepts of an IFCA Concept Lattice, namely $C_1 = (\tilde{E}_1, I_1)$, and $C_2 = (\tilde{E}_2, I_2)$, their similarity $T2ConSim(C_1, C_2)$ is defined as follows:

$$T2ConSim(C_1, C_2) = IT2FSim(\tilde{E}_1, \tilde{E}_2) \cdot w + ASim(I_1, I_2) \cdot (1 - w) \tag{10}$$

where *IT2FSim* is the similarity between the IT2 FSs $\tilde{E}_1, \tilde{E}_2$, $ASim(I_1, I_2)$ is the similarity between the sets of attributes I_1, and I_2, and w is a weight, $0 \leq w \leq 1$, defined by domain experts depending on the characteristics of the application domain. In the case of very small values of w, concept similarity is evaluated by taking into account mainly the concept intents, i.e., the sets of attributes associated with the objects of the application domain whereas, in the opposite case, values of w very close to 1 mean that the computation of similarity is performed by focusing on the specific objects of the application domain, rather than their intensional descriptions.

For instance, in our running example assume $w = \frac{1}{2}$, and consider the concept:

$$C_1 = (((H4, [0.9, 1.0]), (H6, [0.7, 0.9])), (SwPool, Sea))$$

of the Concept Lattice of Figure 2. Furthermore, consider the concept C_2 below, defined as follows:

$$C_2 = (((H2, [0.8, 1.0]), (H4, [0.7, 0.9]), (H7, [0.6, 0.8])), (SwPool, Lake))$$

and suppose C_2 belongs to a different context (it contains the object $H7$ and the attribute *Lake* which do not belong to the context *Sardinia Hotels*). The similarity between C_1 and C_2 is computed as follows:

$$IT2FSim(\tilde{E}_1, \tilde{E}_2) = \frac{AC(\tilde{E}_1 \cap \tilde{E}_2)}{AC(\tilde{E}_1 \cup \tilde{E}_2)} = \frac{0.8}{3.35} = 0.24$$

because:

$$\begin{aligned}
\tilde{E_1} \cap \tilde{E_2} &= (H4, [0.7, 0.9]), \\
AC(\tilde{E_1} \cap \tilde{E_2}) &= (0.7 + 0.9)/2 = 0.8, \\
\tilde{E_1} \cup \tilde{E_2} &= ((H2, [0.8, 1.0])(H4, [0.9, 1.0]), (H6, [0.7, 0.9]), (H7, [0.6, 0.8])), \\
AC(\tilde{E_1} \cup \tilde{E_2}) &= ((0.8 + 0.9 + 0.7 + 0.6) + (1.0 + 1.0 + 0.9 + 0.8))/2 = 3.35.
\end{aligned}$$

We have seen that:

$$ASim(I_1, I_2) = \mathcal{M}((SwPool, Sea), (SwPool, Lake))/2 = 0.84.$$

Therefore:

$$T2ConSim(C_1, C_2) = \frac{1}{2} \cdot 0.24 + \frac{1}{2} \cdot 0.84 = 0.54.$$

5 EVALUATION AND DISCUSSION

In line with the work for non-fuzzy concepts presented in [11], the information content approach and the use of a lexical database for the English language lead to a fundamental difference with respect to other proposals. In fact, in the absence of them, the evaluation of the attribute similarity (independently of the related objects), such as *Sea* and *Lake* in the example of the previous section, requires "additional knowledge" which, in general, is provided by a panel of experts in the given application domain [10]. Furthermore, note that *T2ConSim* is not a distance-based similarity measure and, in line with the notion of *information content similarity* on which it relies, the *triangle inequality* does not hold [23].

Finally, it is important to note that setting the parameter w in *T2ConSim* is a complex problem whose definition is, in general, left to the domain expert according to the context, which plays a crucial role when measuring concept similarity [20]. Within similarity measures, this topic has been addressed by several authors in different research areas as, for instance, in [29] where the problem of determining features' relevance in the context of Geographical Information Systems has been analyzed. However, the definition of (semi-)automatic criteria to evaluate context-dependent parameters is still a challenging topic which requires human expertise (and goes beyond the scope of this paper).

With regard to concept intents, which are non-fuzzy sets, it is important to recall that their similarity can also be evaluated by following several different approaches defined in the literature, as for instance *Dice*, *Jaccard*, *Cosine* [24], etc. Here we only recall the *Jaccard* measure since it is the one on which the similarity between IT2 FSs is based (of course reformulated for crisp sets), and we show the reason why it is not indicated in order to evaluate the similarity of concept intents.

Let I_1, and I_2 be two concept intents, the *Jaccard* similarity, $Jaccard(I_1, I_2)$, is defined on the basis of the cardinalities of their intersection and union sets as follows:

$$Jaccard(I_1, I_2) = \frac{\mid I_1 \cap I_2 \mid}{\mid I_1 \cup I_2 \mid}. \tag{11}$$

For instance, in our running example, consider the intents:

$$I_1 = (SwPool, Sea),$$

$$I_2 = (SwPool, Lake),$$

then, according to their intersection and union sets:

$$Jaccard((SwPool, Sea), (SwPool, Lake)) = 1/3 = 0.33.$$

Note that *Sea* and *Lake* do not contribute to the intersection since they are evaluated as different strings, independently of their semantics. Vice versa, in Subsection 4.2.2 we have seen that, according to the information content approach, the *ics* between *Sea* and *Lake* is:

$$ics(Sea, Lake) = 0.67$$

and:

$$ASim((SwPool, Sea), (SwPool, Lake)) = 0.84,$$

which is closer to human judgment. Indeed, Lin's approach has been extensively experimented in the literature and shows a higher correlation with human judgment than other methods such as Resnik, Wu and Palmer, etc., and the traditional *edge-counting* approach [23].

The impact about the use of the information content approach within OFCA has been experimented in [13]. In the mentioned paper, the experimental results show that the correlation with human judgment has an average increment of about 0.3, with respect to the compared proposals. Besides the use of the information content approach, this significant increment is due to the combination of the concept extent and the concept intent similarities.

This strong imbalance in favor of the measure proposed within OFCA, on which this approach is based, makes us optimistic for future possible experimentations and comparisons with forthcoming proposals within IFCA.

6 CONCLUSION AND FUTURE WORK

In this paper a similarity measure for IFCA concepts has been proposed. It essentially combines the similarity of concept extents, that are IT2 FSs, and the similarity of concept intents, that are sets of concept nouns. In particular, concept extents are compared according to the *IT2FSim*, that is the widely accepted crisp similarity measure for IT2 FSs, that allows a relevant simplification about the definition of similarity between general T2 FSs. Concept intents are evaluated according to the

information content approach, which has been extensively experimented in the literature and has a higher correlation with human judgment. This combination makes us confident about future comparisons with forthcoming proposals. In addition, in this paper both IT2 FS theory and IFCA have been recalled, by providing simple examples which allow to reach a broad audience of non-specialist readers.

As future work, we are arranging a wide experiment involving the students of our Institute in order to quantify the correlation of the proposed measure with human judgment that, unfortunately, is not an easy job due to the complexity of Concept Lattices.

REFERENCES

[1] Ashrafi, M.—Prasad, D. K.—Quek, C.: IT2-GSETSK: An Evolving Interval Type-II TSK Fuzzy Neural System for Online Modeling of Noisy Data. Neurocomputing, Vol. 407, 2020, pp. 1–11, doi: 10.1016/j.neucom.2020.03.065.

[2] Belohlávek, R.: What is a Fuzzy Concept Lattice? II. In: Kuznetsov, S. O. et al. (Eds.): Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (RSFDGrC 2011). Springer, Berlin, Heidelberg, Lecture Notes in Computer Science, Vol. 6743, 2011, pp. 19–26, doi: 10.1007/978-3-642-21881-1_4.

[3] Benítez-Caballero, M. J.—Medina, J.—Ramírez-Poussa, E.—Ślęzak, D.: Rough-Set-Driven Approach for Attribute Reduction in Fuzzy Formal Concept Analysis. Fuzzy Sets and Systems, Vol. 391, 2020, pp. 117–138, doi: 10.1016/j.fss.2019.11.009.

[4] Berners-Lee, T.—Hendler, J.—Lassila, O.: The Semantic Web. Scientific American, Feature Article, May 2001.

[5] Burusco, A.—Fuentes-González, R.: The Study of the Interval-Valued Contexts. Fuzzy Sets and Systems, Vol. 121, 2001, No. 3, pp. 439–452, doi: 10.1016/S0165-0114(00)00059-2.

[6] Djouadi, Y.—Prade, H.: Interval-Valued Fuzzy Formal Concept Analysis. In: Rauch, J. et al. (Eds.): Foundations of Intelligent Systems, International Symposium on Methodologies for Intelligent Systems (ISMIS 2009). Springer, Berlin, Heidelberg, Lecture Notes in Computer Science, Vol. 5722, 2009, pp. 592–601, doi: 10.1007/978-3-642-04125-9_62.

[7] Dubois, D.—Prade, H.: Fundamentals of Fuzzy Sets. Springer Science and Business Media, 2012.

[8] Fateminia, S. H.—Sumati, V.—Fayek, A. R.: An Interval Type-2 Fuzzy Risk Analysis Model (IT2FRAM) for Determining Construction Project Contingency Reserve. Algorithms, Vol. 13, 2020, No. 7, Art. No. 163, pp. 1–22, doi: 10.3390/a13070163.

[9] Fellbaum, C.: WordNet: An Electronic Lexical Database. The MIT Press, Cambridge, MA, 1998.

[10] Formica, A.: Ontology-Based Concept Similarity in Formal Concept Analysis. Information Sciences, Vol. 176, 2006, No. 18, pp. 2624–2641, doi: 10.1016/j.ins.2005.11.014.

[11] Formica, A.: Concept Similarity in Formal Concept Analysis: An Information Content Approach. Knowledge-Based Systems, Vol. 21, 2008, No. 1, pp. 80–87, doi: 10.1016/j.knosys.2007.02.001.

[12] Formica, A.: Semantic Web Search Based on Rough Sets and Fuzzy Formal Concept Analysis. Knowledge-Based Systems, Vol. 26, 2012, pp. 40–47, doi: 10.1016/j.knosys.2011.06.018.

[13] Formica, A.: Similarity Reasoning for the Semantic Web Based on Fuzzy Concept Lattices: An Informal Approach. Information Systems Frontiers, Vol. 15, 2013, No. 3, pp. 511–520, doi: 10.1007/s10796-011-9340-y.

[14] Formica, A.: Similarity Reasoning in Formal Concept Analysis: From One- to Many-Valued Contexts. Knowledge and Information Systems, Vol. 60, 2019, No. 2, pp. 715–739, doi: 10.1007/s10115-018-1252-4.

[15] Ganter, B.—Wille, R.: Formal Concept Analysis – Mathematical Foundations. Springer, 1999.

[16] Hao, M.—Mendel, J. M.: Similarity Measures for General Type-2 Fuzzy Sets Based on the α-Plane Representation. Information Sciences, Vol. 277, 2014, pp. 197–215, doi: 10.1016/j.ins.2014.01.050.

[17] Hitzler, P.—Krötzsch, M.—Rudolph, S.: Foundations of Semantic Web Technologies. Chapman and Hall/CRC, Taylor and Francis Group, 2009, doi: 10.1201/9781420090512.

[18] Hitzler, P.: What's Happening in Semantic Web ... and What FCA Could Have to Do with It. In: Valtchev, P., Jäschke, R. (Eds.): Formal Concept Analysis (ICFCA 2011). Springer, Berlin, Heidelberg, Lecture Notes in Computer Science, Vol. 6628, 2011, pp. 18–23, doi: 10.1007/978-3-642-20514-9_2.

[19] Jay, N.—Nuemi, G.—Gadreau, M.—Quantin, C.: A Data Mining Approach for Grouping and Analyzing Trajectories of Care Using Claim Data: The Example of Breast Cancer. BMC Medical Informatics and Decision Making, Vol. 13, 2013, Art. No. 130, pp. 1–9, doi: 10.1186/1472-6947-13-130.

[20] Kessler, C.: Similarity Measurement in Context. In: Kokinov, B. et al. (Eds.): Modeling and Using Context (CONTEXT 2007). Springer, Berlin, Heidelberg, Lecture Notes in Computer Science, Vol. 4635, 2007, pp. 277–290, doi: 10.1007/978-3-540-74255-5_21.

[21] Kuhn, H. W.: The Hungarian Method for the Assignment Problem. Naval Research Logistics Quarterly, Vol. 2, 1955, No. 1-2, pp. 83–97, doi: 10.1002/nav.3800020109.

[22] Liang, Q.—Mendel, J. M.: Interval Type-2 Fuzzy Logic Systems: Theory and Design. IEEE Transactions on Fuzzy Systems, Vol. 8, 2000, No. 5, pp. 535–550, doi: 10.1109/91.873577.

[23] Lin, D.: An Information-Theoretic Definition of Similarity. Proceedings of the Fifteenth International Conference on Machine Learning (ICML '98), Madison, Wisconsin, USA, Morgan Kaufmann, 1998, pp. 296–304.

[24] Maarek, Y. S.—Berry, D. M.—Kaiser, G. E.: An Information Retrieval Approach for Automatically Constructing Software Libraries. IEEE Transactions on Software Engineering, Vol. 17, 1991, No. 8, pp. 800–813, doi: 10.1109/32.83915.

[25] Mendel, J. M.: Computing with Words and Its Relationship with Fuzzistics. Information Sciences, Vol. 177, 2007, No. 4, pp. 988–1006, doi: 10.1016/j.ins.2006.06.008.

[26] Mendel, J. M.: Type-2 Fuzzy Sets and Systems: A Retrospective. Informatik-Spektrum, Vol. 38, 2015, No. 6, pp. 523–532, doi: 10.1007/s00287-015-0927-4.

[27] Mendel, J. M.—Wu, D.: Perceptual Reasoning for Perceptual Computing. IEEE Transactions on Fuzzy Systems, Vol. 16, 2008, No. 6, 1550–1564, doi: 10.1109/TFUZZ.2008.2005691.

[28] Rocco, C. M.—Hernandez-Perdomo, E.—Mun, J.: Introduction to Formal Concept Analysis and Its Applications in Reliability Engineering. Reliability Engineering and System Safety, Vol. 202, 2020, Art. No. 107002, doi: 10.1016/j.ress.2020.107002.

[29] Rodríguez, A.—Egenhofer, M. J.: Comparing Geospatial Entity Classes: An Asymmetric and Context-Dependent Similarity Measure. International Journal of Geographical Information Science, Vol. 18, 2004, No. 3, pp. 229–256, doi: 10.1080/13658810310001629592.

[30] Safaeipour, H.—Fazel Zarandi, M. H.—Turksen, I. B.: Developing Type-2 Fuzzy FCA for Similarity Reasoning in the Semantic Web. Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), IEEE, 2013, pp. 1477–1482, doi: 10.1109/IFSA-NAFIPS.2013.6608620.

[31] Stumme, G.—Maedche, A.: FCA-MERGE: Bottom-Up Merging of Ontologies. Proceedings of 17th International Joint Conference on Artificial Intelligence (IJCAI '01), Seattle, Washington, USA, Vol. 1, 2001, pp. 225–230. ISBN: 1-55860-777-3.

[32] Tho, Q. T.—Hui, S. C.—Fong, A. C. M.—Cao, T. H.: Automatic Fuzzy Ontology Generation for Semantic Web. IEEE Transactions on Knowledge and Data Engineering, Vol. 18, 2006, No. 6, pp. 842–856, doi: 10.1109/TKDE.2006.87.

[33] Wang, F.—Wang, N.—Cai, S.—Zhang, W.: A Similarity Measure in Formal Concept Analysis Containing General Semantic Information and Domain Information. IEEE Access, Vol. 8, 2020, pp. 75303–75312, doi: 10.1109/ACCESS.2020.2988689.

[34] Wu, D.—Mendel, J. M.: Uncertainty Measures for Interval Type-2 Fuzzy Sets. Information Sciences, Vol. 177, 2007, No. 23, pp. 5378–5393, doi: 10.1016/j.ins.2007.07.012.

[35] Wu, D.—Mendel, J. M.: A Comparative Study of Ranking Methods, Similarity Measures and Uncertainty Measures for Interval Type-2 Fuzzy Sets. Information Sciences, Vol. 179, 2009, No. 8, pp. 1169–1192, doi: 10.1016/j.ins.2008.12.010.

[36] Zadeh, L. A.: Fuzzy Sets. Information and Control, Vol. 8, 1965, No. 3, pp. 338–353, doi: 10.1016/S0019-9958(65)90241-X.

[37] Zadeh, L. A.: The Concept of a Linguistic Variable and Its Application to Approximate Reasoning – I. Information Sciences, Vol. 8, 1975, No. 3, pp. 199–249, doi: 10.1016/0020-0255(75)90036-5.

Anna Formica received her degree (Hons.) in mathematics from the University of Rome "La Sapienza", in 1989. She is currently Senior Researcher with the "Istituto di Analisi dei Sistemi ed Informatica" (IASI) "Antonio Ruberti", Italian National Research Council (Consiglio Nazionale delle Ricerche – CNR), Rome, where she manages the "Software and Knowledge-Based Systems" (SaKS) Group. She tooks part in various research projects of the European framework programs and bilateral projects with international institutions. Her current research interests include semantic web, similarity reasoning, formal specification and validation of domain ontologies, fuzzy formal concept analysis, geographical information systems, and e-learning. She serves as a referee for several international journals and conferences.