CompAMa

www.compama.co.usb.ve

Bull. Comput. Appl. Math., Vol.6, No.2, 2018

ISSN 2244-8659

Integrating Fuzzy Formal Concept Analysis and Rough Set Theory for the Semantic Web

Anna Formica [1,2]

CompAMa Vol.6, No.2, pp.65-84, 2018 - Accepted November 4, 2018

Abstract

Formal Concept Analysis and *Rough Set Theory* provide two mathematical frameworks in information management which have been developed almost independently in the past. Currently, their integration is revealing very interesting in different research fields, such as knowledge discovery, data mining, information retrieval, elearning, and ontology engineering. In this paper, we show how *Rough Set Theory* can be employed in combination with a generalization of *Formal Concept Analysis* for modeling uncertainty information (*Fuzzy Formal Concept Analysis*) to perform Semantic Web search. In particular this paper presents an updated evaluation of a previous proposal of the author which has been addressed because of the increasing interest in this topic and, at the same time, the absence in the literature of significant proposals combining these two frameworks.

Keywords: Semantic Web, Fuzzy Formal Concept Analysis, Rough Set Theory.

1 Introduction

Formal Concept Analysis (FCA) [1] provides a mathematical framework which can support several activities in different research fields as, for instance, software engineering, requirements analysis, component retrieval, etc... [2]. It

[1]Istituto di Analisi dei Sistemi ed Informatica “Antonio Ruberti” (IASI-CNR), Via dei Taurini 19, I-00185, Rome, Italy. (anna.formica@iasi.cnr.it).

[2]Corresponding author.

is also a well-founded methodological approach for the construction of ontologies for the Semantic Web development [3]. In fact, FCA can serve as a guideline for ontology building because it allows the identification of concepts by factoring out their commonalities while preserving concept specialization relationships [4]. *Fuzzy Formal Concept Analysis* (FFCA) is a generalization of FCA for modeling uncertainty information [5]. FFCA can support ontology construction when some information is more relevant than other data, or Semantic Web search when the user is not sure about what he/she is looking for. However, often, in real life applications user needs cannot be described on the basis of formal concepts only [6], and "approximate concepts will become increasingly more important as Semantic Web becomes a reality" [7].

Rough Set Theory (RST) [8] is an extension of set theory for data analysis in the presence of inexact, uncertain or vague information. RST and FCA have been extensively investigated in the literature within several research areas and with different purposes [9]. In particular, their combination provides an interesting framework for Semantic Web search and development, although in this research area most of the proposals address FCA and RST separately [10].

This paper shows how FFCA and RST can be combined in order to perform Semantic Web search. In particular, the work addresses an updated evaluation of [11], which represents "one of the most thorough applications of the combination of FCA with fuzzy attributes and rough set theory" [9]. Indeed, on the basis of the current literature, there are still no significant proposals combining these two frameworks, although the increasing interest in their integration. In the mentioned proposal, in the case the required data are not modeled by any formal concept, the user can search and discover information in the Web that is closer to his/her preferences by following a two-fold approach. Thanks to the notions of lower/upper approximations, the user can select super/subsets of the data (objects) he/she is looking for. Furthermore, the notion of a fuzzy context in FFCA allows the user to choose, within the selected sets, specific objects that, on the basis of "grades of membership", allow him/her to quantify "how much" they are described by the required attributes and, therefore, "how much" these objects correspond to the user needs.

The paper is organized as follows. In Sections 2 and 3 Formal Concept Analysis and Fuzzy Formal Concept Analysis are recalled, respectively. In Section 4, the relationship between FCA and RST is given and, in Subsection

4.1, Web search based on RST and FFCA is shown. Successively in Section 5 the Related Work is presented with a discussion about the proposed method, and Section 6 concludes.

2 FCA

In FCA a *formal concept* is defined within a *formal context*.

Definition 1. [Formal Context] *A formal context (context for short) is a triple (O,A,R), where O and A are two sets of elements called objects and attributes, respectively, and R is a binary relation between O and A.*

In a context, if oRa, $o \in O$ and $a \in A$, then we say that "the object o has the attribute a" or "the attribute a applies to the object o". The definition of a formal concept follows.

Definition 2. [Formal Concept] *Given a context (O,A,R), let E, I be two sets such that $E \subseteq O$ and $I \subseteq A$. Then, consider the dual sets E' and I', i.e., the sets defined by the attributes applying to all the objects belonging to E and the objects having all the attributes belonging to I, respectively, i.e.:*

$E' = \{a \in A \mid oRa \; \forall o \in E\}$

$I' = \{o \in O \mid oRa \; \forall a \in I\}$.

A formal concept (concept for short) of the context (O,A,R) is a pair (E,I) such that:

$E \subseteq O, \; I \subseteq A$

and the following conditions hold:

$E' = I, \; I' = E.$

The sets E and I represent the concept extensional and intensional components respectively, and are referred to as the extent and the intent of the concept, respectively.

Therefore, a concept is a pair of sets where the former consists of precisely those objects which have all attributes from the latter and, conversely, the latter consists of precisely those attributes that apply to all objects from the former. For instance, consider a context named *Sardinia Hotels*, suppose that the set O is defined by the following six objects representing six different hotels:

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

and that the set A is defined by six possible attributes of these objects:

$A = \{\text{Tennis, Theater, SwPool, Meal, Sea, Cinema}\}$

where $SwPool$ stands for swimming pool. Furthermore, suppose the hotels are related to the above attributes according to the binary relation R defined by the following table:

Tab. 1: The *Sardinia Hotels* context in FCA

	Te	**Th**	**Sw**	**Me**	**Se**	**Ci**
H1		×		×	×	
H2	×		×	×		×
H3		×		×	×	
H4	×		×		×	
H5				×	×	
H6	×		×		×	

where Te, Th, Sw, Me, Se, and Ci stand for $Tennis$, $Theater$, $SwPool$, $Meal$, Sea, and $Cinema$, respectively. According to Table 1, we say that, for instance, the hotel $H4$ has, or is described by, three attributes, namely $Tennis$, $SwPool$, and Sea, and vice versa, these three attributes apply to the object $H4$. A concept of the *Sardinia Hotels* context is, for instance:

$((H4, H6), (Tennis, SwPool, Sea))$

since both $H4$ and $H6$ have the attributes $Tennis$, $SwPool$, and Sea, and vice versa, all these attributes apply to both the objects $H4$,$H6$.

Given two concepts of a context, (E_1,I_1), (E_2,I_2), it is possible to establish an *inheritance relation* ($\leq$) between them according to the following condition:

$(E_1, I_1) \leq (E_2, I_2) \iff E_1 \subseteq E_2 \; (\iff I_2 \subseteq I_1)$

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

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

is a complete lattice called *Concept Lattice*, i.e., for each subset of concepts, the greatest common subconcept and the least common superconcept exist [1]. The Concept Lattice that can be constructed from the context of Table 1 is shown in Figure 1. Note that nodes are labeled with the concepts of the context, and arcs are established among the nodes whose associated concepts are in $\leq$ relation. The Concept Lattice has two special nodes, the maximum and minimum nodes, grouping all the objects and the attributes of the context, respectively.

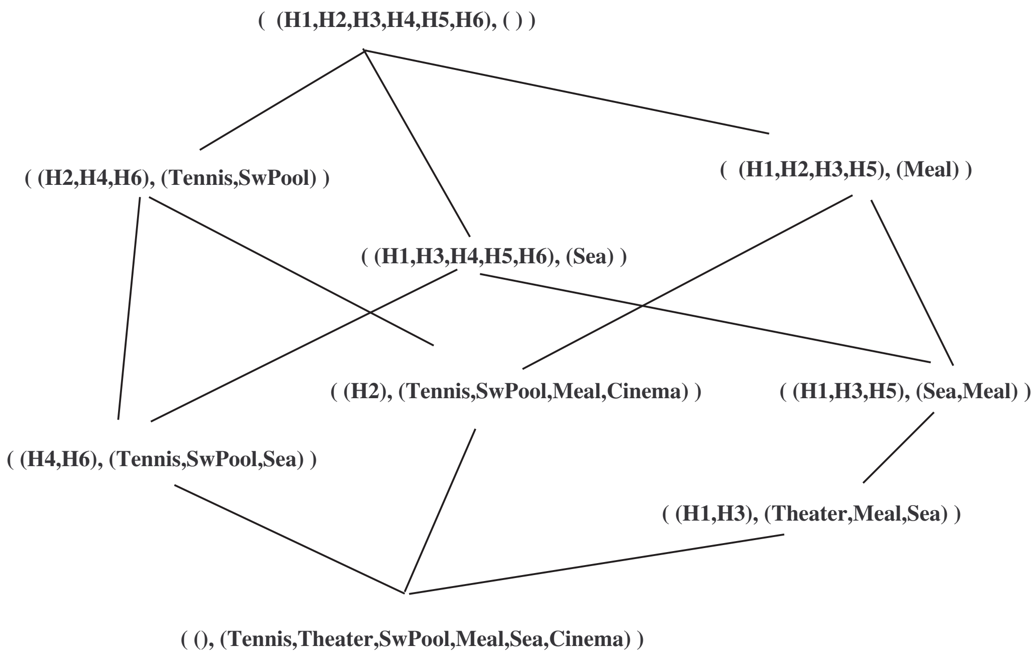

Fig. 1: Concept Lattice of the *Sardinia Hotels* context

In Figure 1, the least common superconcept of, for instance, the concepts (($H4$,$H6$), ($Tennis$,$SwPool$,Sea)) and (($H1$,$H3$),($Theater$,$Meal$,Sea)) is the concept (($H1$,$H3$, $H4$,$H5$,$H6$), (Sea)), having as set of attributes the intersection of the sets of attributes of the concepts. Whereas the greatest common subconcept of the concepts (($H2$,$H4$,$H6$), ($Tennis$,$SwPool$)) and (($H1$,$H2$,$H3$,$H5$),($Meal$)) is the concept (($H2$),($Tennis$, $SwPool$,$Meal$, $Cinema$)), having as set of objects the intersection of the sets of objects of the concepts.

Unfortunately, modeling a domain of interest with traditional FCA (i.e., with non-fuzzy sets) can be inaccurate when the attributes do not describe the objects in a uniform way or, in other words, a given attribute applies to different objects in different ways. For instance, in our example, consider the attribute Sea. One should be able to distinguish the hotels located just on the sea, from that having a walking distance seaside (reachable in, for instance, ten or twenty minutes). Analogously, regarding the attribute $Meal$, we would like to be aware about the hotels providing both lunch and dinner, rather than half-board. Without the introduction of fuzzy information, we have no way to specify how appropriate is a feature, or an attribute, to a given object, therefore describing all the objects in a uniform way.

3 FFCA

FFCA incorporates fuzzy logic into FCA in order to represent vague information. Similarly to FCA, in FFCA a concept is defined within a *fuzzy formal context*. Below, we start by recalling the notion of a *fuzzy set* [12].

Definition 3. [Fuzzy Set] *Given a domain X, a fuzzy set A in X is characterized by a membership function* $\mu_A(x)$ *which associates each point in X with a real number in the interval [0,1]:*

$$A = \{(x, \mu_A(x)) \mid x \in X\}$$

The value $\mu_A(x)$ *represents the "grade of membership" of* x *in* A*.*

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. Just to provide an example, assume X is a set of people, a fuzzy set $Young$ is defined by associating with each person in X a real number in [0,1] establishing the degree of youth of the person, such that the nearer this value to unity, the higher the grade of membership of a person in the set $Young$. The notion of a *fuzzy relation* can be obtained by generalizing the notion of a fuzzy set as follows. A fuzzy relation R in $X \times Y$ is a fuzzy set in the product space $X \times Y$.

Given a traditional set of items S (crisp set), we denote as $\phi(S)$ a *fuzzy set generated from* S, i.e., $\phi(S)$ is a fuzzy set where each item in S has a membership value in [0,1]. Analogously, given two crisp sets S, T, $\phi(S \times T)$ is a *fuzzy relation* in $S \times T$.

For instance, consider the set of objects O and the set of attributes A of the *Sardinia Hotels* context in the previous section. A fuzzy relation $\phi(O \times A)$ can be defined as follows:

$\{((H1, Te), 0), ((H2, Te), 0.6), ((H3, Te), 0), \ldots,$
$((H1, Th), 0.7), ((H2, Th), 0), ((H3, Th), 0.8), \ldots,$
$((H1, Ci), 0), ((H2, Ci), 0.9), ((H3, Ci), 0), \ldots\}$

where each pair in $(O \times A)$ is associated with a membership value in [0,1] (where Te, Th and Ci stand for $Tennis$, $Theater$ and $Cinema$, respectively).

Below, the notions of a *Fuzzy Formal Context* and a *Fuzzy Formal Concept* are given.

Definition 4. [Fuzzy Formal Context] *A fuzzy formal context (fuzzy context for short) is a triple:*
$K = (O, A, R = \phi(O \times A))$
where O is a set of objects, A is a set of attributes, and R is a fuzzy relation in $O \times A$. Each pair $(o, a) \in R$ has a membership value $\mu(o, a)$ in [0,1].

Given a pair $(o, a) \in R$ with membership value $\mu(o, a)$, we say that "the object o has the attribute a" or "the attribute a applies to the object o" with the grade of membership $\mu(o, a)$.

Definition 5. [Fuzzy Formal Concept] *Given a fuzzy context:*
$K = (O, A, R = \phi(O \times A))$
a confidence threshold T, and two sets E, I, such that $E \subseteq O$ and $I \subseteq A$, consider the dual sets E' and I', defined respectively as follows:
$E' = \{a \in A \mid \mu(o, a) \geq T \ \forall o \in E\}$
$I' = \{o \in O \mid \mu(o, a) \geq T \ \forall a \in I\}$.
A fuzzy formal concept (fuzzy concept for short) of the fuzzy context K with confidence threshold T is a pair $(\phi(E), I)$, $E \subseteq O$, $I \subseteq A$, and $E' = I$, $I' = E$. Each object $o \in E$ has a membership value μ_o defined as:

$$\mu_o = \min_{a \in I} \mu(o, a)$$

where $\mu(o, a)$ is the membership value between the object o and the attribute a. If $I = \emptyset$, $\mu_o = 1$ for every o. The sets E and I represent the concept extensional and intensional components respectively, and are referred to as the extent and the intent of the fuzzy concept, respectively.

The definition of Fuzzy Formal Concept above has been given in line with [4]. As shown by the authors in the mentioned paper, with respect to other approaches proposed in the literature, this definition allows to generate simpler Fuzzy Concept Lattices (in terms of the number of formal concepts) and is also more suited for evaluating concept similarity in the context of the Semantic Web.

Note that the above definition can also be formulated in terms of a *Pattern Concept* of a *Pattern Structure* [13]. Pattern Structures consist of objects with descriptions (patterns) that allow a semi-lattice operation on them, referred to as *similarity operation.* In essence, for an arbitrary set of objects, the similarity operation gives a description representing the similarity of the objects from the set. The connection between FFCA and Pattern Structures, and the related concepts, is an interesting topic that has been quite investigated in the literature, see for instance [14].

Consider the *Sardinia Hotels* fuzzy context specified by the fuzzy relation given in Table 2. Note that crosses in Table 1 have been replaced by grades of membership, from 0 to 1, each allowing us to quantify "how much" an object has, or is described by, an attribute and vice versa an attribute applies to an object.

Tab. 2: The *Sardinia Hotels* context in FFCA

	Te	**Th**	**Sw**	**Me**	**Se**	**Ci**
H1		0.7		1.0	1.0	
H2	0.6		1.0	0.5		0.9
H3		0.8		0.5	0.7	
H4	0.8		1.0		1.0	
H5				1.0	0.3	
H6	0.8		1.0		0.8	

For instance, consider the hotel $H2$ in Table 2. It has the attribute $SwPool$ with grade of membership 1.0, which means that such 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 0.5, which means that such an attribute partially applies to this hotel (for instance it could provide meals just for dinner). Analogously, in the case of $H3$, the value 0.7 in correspondence with the attribute Sea means that this feature better describes the hotels $H1$, $H4$ or $H6$ than $H3$, but it is more appropriate to $H3$ than $H5$ (having $H5$ a lower grade of membership with Sea, i.e., 0.3). 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 less than the threshold are ignored. For instance, consider our running example and assume that a threshold is fixed equal to 0.5. The grade of membership 0.3 between $H5$ and Sea is ignored and treated analogously to the grades of membership that in Table 2 are not specified (they are equal to zero). A fuzzy concept of the *Sardinia Hotels* fuzzy context is, for instance, the pair:

$(((H1, 0.7), (H3, 0.5)), (Theater, Meal, Sea))$.
In fact the objects $H1$,$H3$ share the attributes $Theater$, $Meal$ and Sea and, vice versa, these three attributes apply to the objects $H1$ and $H3$ with membership values which are not less than the threshold. According to the definition of fuzzy formal concept above, in the case the attributes apply to an object with different grades of membership, the minimum among them is selected. For instance, the object $H3$ has the attributes $Theater$, $Meal$ and Sea with different grades of membership, that are 0.8, 0.5 and 0.7, respectively. In the concept above, the minimum value between them has been selected because it represents the highest common grade of membership that allows $H3$ to be described by the all the three attributes $Theater$, $Meal$ and Sea.

Fuzzy Concept Lattices can be defined similarly to Concept Lattices, on the basis of fuzzy set theory. Below, the *fuzzy set intersection* and *fuzzy set union* are briefly recalled.

Definition 6. [Fuzzy Set Intersection] *The intersection of two fuzzy sets A and B, denoted as $A \cap B$, with respective membership functions $\mu_A(x)$, and $\mu_B(x)$, is a fuzzy set whose membership function is defined as:*

$$\mu_{A \cap B}(x) = min(\mu_A(x), \mu_B(x)).$$

Definition 7. [Fuzzy Set Union] *The union of two fuzzy sets A and B, denoted as $A \cup B$, with respective membership functions $\mu_A(x)$, and $\mu_B(x)$, is a fuzzy set whose membership function is defined as:*

$$\mu_{A \cup B}(x) = max(\mu_A(x), \mu_B(x)).$$

The Fuzzy Concept Lattice that can be constructed from the context of Table 2 is shown in Figure 2. Analogously to Concept Lattices, nodes are labeled with the fuzzy concepts of the context, and arcs are established among the nodes that are in inheritance relation. The Fuzzy Concept Lattice has two special nodes, the maximum and minimum nodes, grouping all the objects and the attributes of the context, respectively. In particular, the membership values associated with the objects of the maximum node are all equal to one. Also in Fuzzy Concept Lattices, for any subset of concepts, the greatest common subconcept and the least common superconcept are always defined. For instance, consider the concepts:

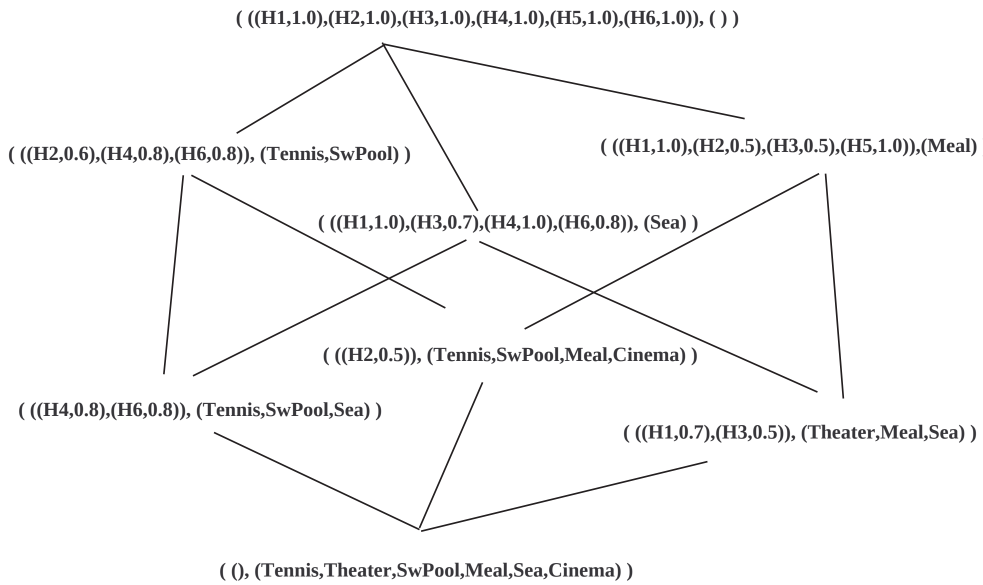

Fig. 2: Fuzzy Concept Lattice of the *Sardinia Hotels*

$(((H2, 0.6), (H4, 0.8), (H6, 0.8)), (Tennis, SwPool))$
$(((H1, 1.0), (H3, 0.7), (H4, 1.0), (H6, 0.8)), (Sea))$.
The greatest common subconcept is:
$(((H4, 0.8), (H6, 0.8)), (Tennis, SwPool, Sea))$
where, in the fuzzy set intersection, the minimum among different grades of membership associated with the same object has been selected (for instance, in the case of $H4$, 0.8).

Note that, given a context, the construction of the Concept Lattice has been extensively investigated in the literature, and also the definition of (semi-)automatic tools for the construction of fuzzy ontologies from huge amount of existing fuzzy databases [4]. The problem of reducing the large number of concepts that can be extracted from data can be addressed by using factorizations by similarity of Concept Lattices [5], factorization of Boolean matrices, extensively analyzed in [15], or conceptual clustering methods, as for instance the *Iceberg* Concept Lattices [16]. Furthermore, the assignment of fuzzy values is a critical step that is usually performed by domain experts. This problem can be addressed in line with the ontological approach proposed in [17], where a similarity measure for FCA concepts has been defined. In particular, in the mentioned approach, the similarity degrees among terms of a domain ontology are defined by a panel of experts in the given domain by means of a *Consensus System*.

In this paper, we assume that the Fuzzy Concept Lattice is given and the problems related to the construction and the reduction of the size of the Concept Lattice go beyond the scope of this work.

4 RST and FCA

RST is an extension of classical set theory with two additional operators, namely approximation operators, originally introduced in [8]. Among the various formulations that can be found in the literature, below the one given in [18] is briefly recalled.

Let U be a finite and non-empty universe of objects and E be an equivalence relation on U. E induces a partition of the universe U, indicated as U/E, and the pair apr=(U,E) is referred to as an *approximation space*. An equivalence class in U/E is referred to as an *elementary set*. Any finite union of elementary sets is called a *definable set*. Given an arbitrary set $X \subseteq U$, it may not correspond to a definable set because X may include and exclude objects that belong to different definable sets. However, X can be approximated from below and above by a pair of definable sets referred to as the *lower* and *upper approximations* of X. Intuitively, the lower approximation is the greatest definable set contained in X and the upper approximation is the least definable set containing X.

The notion of approximation can be naturally introduced into FCA. It has been extensively investigated in the literature, see for instance [6] and [19]. According to [18], a Concept Lattice can be seen as the family of *all definable concepts*. Formal concepts that do not belong to a given Concept Lattice are called *non-definable concepts*. In particular, a formal concept consists of a *definable set of objects* and a *definable set of attributes*. Given a Concept Lattice, a set of objects (attributes) that is not the extension (intension) of any formal concept can be approximated by definable sets of objects (attributes) according to RST.

In line with [19], given a Concept Lattice $\mathcal{L}$, let $Ex(\mathcal{L})$ and $In(\mathcal{L})$ be the families of all the extents and all the intents of $\mathcal{L}$, respectively. Given a set of objects Q_o that is not the extent of any concept in $\mathcal{L}$, intuitively the upper approximation $\overline{apr}(Q_o)$ is the smallest set in $Ex(\mathcal{L})$ that contains Q_o, whereas the lower approximation $\underline{apr}(Q_o)$ is the largest set in $Ex(\mathcal{L})$ that is contained in Q_o. Formally we have:

$$\overline{apr}(Q_o) = \bigcap\{X \mid X \in Ex(\mathcal{L}), Q_o \subseteq X\}$$
$$\underline{apr}(Q_o) = \{X \mid X \in Ex(\mathcal{L}), X \subseteq Q_o, \forall X' \in Ex(\mathcal{L})(X \subseteq X' \Rightarrow X' \not\subseteq Q_o\}.$$

Analogously, given the family of all the intents of $\mathcal{L}$, $In(\mathcal{L})$, consider a set of attributes Q_a. We can define the upper and lower approximations of Q_a, $\overline{apr}(Q_a)$ and $\underline{apr}(Q_a)$ respectively, as follows:

$$\overline{apr}(Q_a) = \bigcap\{X \mid X \in In(\mathcal{L}), Q_a \subseteq X\}$$
$$\underline{apr}(Q_a) = \{X \mid X \in In(\mathcal{L}), X \subseteq Q_a, \forall X' \in In(\mathcal{L})(X \subseteq X' \Rightarrow X' \not\subseteq Q_a\}.$$

It is important to observe that, since $Ex(\mathcal{L})$ ($In(\mathcal{L})$) is closed under intersection, the smallest set containing Q_o (Q_a) is unique whereas this does not hold for the largest set contained in Q_o (Q_a). Therefore, lower approximations may not be unique.

In this section we have recalled how a set of objects or a set of attributes can be approximated from above or from below by the extents or the intents, respectively, of concepts in FCA according to RST. In the next section, we will focus on FFCA and, in particular, we will address the problem of Web search supported by RST in FFCA.

4.1 Web Search based on RST and FFCA

In the following we assume we have a Fuzzy Concept Lattice and that Web queries are expressed by sets of attributes in the given formal context. Given a query, suppose there are no formal concepts having as intent the required set of attributes. Then, the goal is to find the formal concepts of the Fuzzy Concept Lattice whose intents "better approximate" the set of attributes specified by the query and, therefore, whose extents are closer to the expected answer. Furthermore, within the various approximations determined, the user can additionally select the preferred one on the basis of grades of membership of specific objects with specific attributes.

For instance, suppose we have the Fuzzy Concept Lattice related to the *Sardinia Hotels* of Figure 2 and suppose the user is looking for a hotel on the sea where he/she can eat. This query can be represented by the following set of attributes:

$Q_a = (Sea, Meal)$.

In the Fuzzy Concept Lattice of Figure 2 there are no concepts defined by the set of attributes Q_a. However we can look for the formal concepts whose intents better approximate it, i.e. the upper and lower approximations of Q_a. According to the definitions given above, the smallest intent in the Concept Lattice of Figure 2 containing Q_a is ($Theater$,$Meal$,Sea), i.e., the concept whose intent is an upper approximation of Q_a is:

$(((H1, 0.7), (H3, 0.5)), (Theater, Meal, Sea))$

The grades of membership associated with the hotels $H1$ and $H3$, 0.7 and 0.5 respectively, specify "how much" these hotels are properly described by *both* Sea and $Meal$, but *also* by the attribute $Theater$ which was not required by the user. For this reason, lower approximations of the query can be addressed. In this case, two lower approximations are identified in the Concept Lattice of Figure 2. They correspond to the following concepts whose intents are the largest sets contained in Q_a, i.e., the singletons ($Meal$) and (Sea):

$(((H1, 1.0), (H3, 0.7), (H4, 1.0), (H6, 0.8)), (Sea))$

$(((H1, 1.0), (H2, 0.5), (H3, 0.5), (H5, 1.0)), (Meal))$.

The user can therefore select as answer, on the basis of his/her needs, one of the concept extents associated with the above upper and lower approximations. In addition, within one of the above extents, he/she can choose the preferred objects also on the basis of his/her priorities according to the defined grades of membership. For instance, if Sea is the attribute with the highest priority, he/she can select the hotels $H1$ or $H4$, having both grades of membership with Sea equal to 1.0. Analogously, if $Meal$ is on the top of his/her preferences, the user can choose the hotels $H1$ or $H5$. In this case, by analyzing the lower approximations, it is reasonable to assume that the hotel $H1$ represents the closest answer to the user needs. Of course, this does not hold in general, and the user has to choose among several objects that do not perfectly match with the specified query. For instance consider the following query:

$Q_a = (Tennis,SwPool,Meal)$.

Analogously to the previous example, in Figure 2 there are no concepts whose intents correspond to the given set of attributes. Then, the upper approximation is addressed, corresponding to the intent of the following concept:

$(((H2, 0.5)), (Tennis, SwPool, Meal, Cinema))$

and two lower approximations are analyzed, whose formal concepts are:

$(((H2, 0.6), (H4, 0.8), (H6, 0.8)), (Tennis, SwPool))$

$(((H1, 1.0), (H2, 0.5), (H3, 0.5), (H5, 1.0)), (Meal))$.

The user again can choose the answer on the basis of his/her needs by analyzing first the sets of attributes that better approximate the query. In particular he/she can select the hotel $H2$ having all the three required attributes and an additional one, *Cinema*, with grade of membership at least 0.5. Otherwise the user can choose objects partially described by the required attributes, with higher grades of membership. As opposed to the previous example, in this case there are no specific objects that fully satisfy the user needs. However, he/she can identify the favorite hotels by analyzing their grades of membership with preferred attributes (e.g. if *Tennis* and *SwPool* are both preferred, the hotels $H4$ and $H6$ could provide a satisfactory answer).

5 Related Work and Discussion

The combination of FCA and RST is attracting attention within both crisp and fuzzy environments. In crisp environments, for instance, in [20], a unique framework for connecting these theories by making use of hypergraphs has been proposed, in [21], an extended view of FCA and RST has been studied based on a rich structure called *cube of oppositions*, and, in [22], a new knowledge acquisition model has been defined by introducing concept lattices in RST. With regard to the combination of FCA and RST in fuzzy environments, extended surveys can be found in [23], and [9]. Within the rich literature provided in those papers, it is worth recalling [24] and [25], although with different objectives with respect to this work. In fact, in the former, the goal is to handle very large databases efficiently. In particular, the *Infobright Community Edition* (ICE) is used, which is an open source data warehousing system. It allows the transformation of very large FCA contexts (containing up to 10^9 rows) into *rough tables*, i.e., tables with the same attributes as the original large contexts, but with combinations of objects as rows, and meta-information (*data packs*) about them as values. In the latter, the objective is the selection of relevant subsets of attributes from FCA contexts by using RST. In particular, the approach is based on a generalization of the notion of equivalence relation in RST (namely the *indiscernibility* relation), which in the mentioned paper is essentially based on a quasi-order.

Furthermore, with regard to the combination of FCA and RST, it is also worth mentioning [26], where these frameworks have been used for a comparative study of concept lattices in fuzzy contexts, [27], where the construction of fuzzy concept lattices based on generalized fuzzy rough approximation operators has been analyzed, and [28], where two new pairs of rough fuzzy set approximations within fuzzy formal contexts have been defined.

With regard to the combination of FCA and RST in the Semantic Web research area, which is the focus of this paper, a detailed related work has been given in [11]. In particular, a classification of a set of selected proposals has been provided, showing that in this research area these two frameworks have been employed separately. Here we just recall [29], and [16], where FCA, independently of RST, has been used in the former as a knowledge acquisition framework within an e-learning community in an American University, and in the latter as a framework to support ontology building, mapping and alignment. Vice versa, regarding Semantic Web search supported by RST, independently of FCA, we recall [7], where a formal framework for defining and automatically generating approximate concepts and ontologies from traditional crisp ontologies has been presented. Note that in [17], [30], [31], the problem of defining a similarity measure for FCA concepts has been analyzed (in particular in [31], for FFCA concepts), without relying on RST.

In [11], an evaluation of the proposed method has been provided which, on the basis of the current literature, still holds. In particular, in the mentioned paper, the difficulties encountered in making a comparison of this proposal with the existing literature have been discussed, within an experiment performed in the tourism domain. The main problem in the experiment was the impossibility of comparing this work with the underlying knowledge representation and query models of the other proposals because, in none of them, Semantic Web search is supported by both RST and FFCA (or FCA). The combination of FFCA and RST is fundamental for evaluating the contribution of this paper and the adoption of only one of these two frameworks makes any comparison biased in favour of this proposal. Just to provide an example, consider [29], where FCA is employed as a knowledge acquisition framework in the absence of RST and fuzzy values. In the mentioned work, the user query is matched against the intents of the Concept Lattice without using approximation operators and without having the possibility of selecting the objects that better satisfy the user needs according to fuzzy values. In contrast, the strength of this proposal is the possibility of performing Semantic Web search leaving maximum flexibility to the user in selecting the preferred answers along two directions, i.e., by employing the approximation operators of RST from one hand, and fuzzy values of FFCA from the other hand.

6 Conclusion

In this paper RST has been employed in combination with FFCA to perform Semantic Web search and discovery of information in the Web. In the case the required data are not modeled by any formal concept, the user can search and discover the information that are closer to his/her preferences by following a two-fold approach. He/she can select (i) super/subsets of the answer that are associated with lower/upper approximations of the query and, within the proposed answers, (ii) the data that are "better" described by the required attributes, on the basis of fuzzy values. To our knowledge, in the literature there are no proposals which can be really compared with this approach, although the increasing interest in the integration of FFCA and RST.

References

[1] R. Wille. Restructuring lattice theory: an approach based on hierarchies of concepts. *In: Rival I. (eds) Ordered Sets. NATO Advanced Study Institutes Series (Series C - Mathematical and Physical Sciences)*, 83:445–470, 1982. doi: 10.1007/978-94-009-7798-3_15.

[2] T. Tilley, R. Cole, P. Becker, and P. Eklund. A survey of formal concept analysis support for software engineering activities. *In: Stumme G., Wille R. (eds) Formal Concept Analysis, Proc. of the First Int. Conference on Formal Concept Analysis (ICFCA), Lecture Notes in Computer Science*, 3626:250–271, 2003. doi: 10.1007/11528784_13.

[3] T. Lukasiewicz and U. Straccia. Managing uncertainty and vagueness in description logics for the semantic web. *J. Web Semantics: Science, Services and Agents on the World Wide Web*, 6(4):291–308, 2008. doi: 10.1016/j.websem.2008.04.001.

[4] Q.T. Tho, S.C. Hui, A. Cheuk, M. Fong, and T.H. Cao. Automatic fuzzy ontology generation for semantic web. *IEEE Transactions on Knowledge and Data Engineering*, 18(6):842–856, 2006. doi: 10.1109/TKDE.2006.87.

[5] R. Belohlávek, J. Outrata, and V. Vychodil. Fast factorization by similarity of fuzzy concept lattices with hedges. *Int. J.*

of Foundations of Computer Science, 19(2):255–269, 2008. doi: 10.1142/S012905410800567X.

[6] J. Saquer and J.S. Deogun. Concept approximations based on rough sets and similarity measures. *Int. J. Appl. Math. Comput. Sci*, 11(3):655–674, 2001. http://pldml.icm.edu.pl/pldml/element/bwmeta1.element.bwnjournal-article-amcv11i3p655bwm.

[7] P. Doherty, M. Grabowski, W. Lukaszewicz, and A. Szalas. Towards a framework for approximate ontologies. *Fundamenta Informaticae*, 57(2-4):147–165, 2003. https://content.iospress.com/articles/fundamenta-informaticae/fi57-2-4-04.

[8] Z. Pawlak. Rough sets. *International Journal of Computer and Information Sciences*, 11(5):341–356, 1982. doi: 10.1007/BF01001956.

[9] J. Poelmans, D.I. Ignatov, S.O. Kuznetsov, and G. Dedene. Fuzzy and rough formal concept analysis: a survey. *International Journal of General Systems*, 43(2):105–134, 2014. doi: 10.1080/03081079.2013.862377.

[10] Y. Zhao, W Halang, and X Wang. Rough ontology mapping in e-business integration. *In: Lu J., Zhang G., Ruan D. (eds) E-Service Intelligence. Studies in Computational Intelligence*, 37:75–93, 2007. doi: 10.1007/978-3-540-37017-8_3.

[11] A. Formica. Semantic web search based on rough sets and fuzzy formal concept analysis. *Knowledge-Based Systems*, 26:40–47, 2012. doi: 10.1016/j.knosys.2011.06.018.

[12] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965. doi: 10.1016/S0019-9958(65)90241-X.

[13] B. Ganter and S.O. Kuznetsov. Pattern structures and their projections. *In: Delugach H.S., Stumme G. (eds) Conceptual Structures: Broadening the Base. International Conference on Conceptual Structures, ICCS. Lecture Notes in Computer Science*, 2120:129–142, 2001. doi: 10.1007/3-540-44583-8_10.

[14] V.V. Pankratieva and S.O. Kuznetsov. Relations between proto-fuzzy concepts, crisply generated fuzzy concepts, and interval pattern structures. *Fundamenta Informaticae*, 115(4):265–277, 2012. doi: 10.3233/FI-2012-655.

[15] R. Belohlávek, J. Outrata, and M. Trnecka. Toward quality assessment of boolean matrix factorizations. *Information Sciences*, 459:71–85, 2018. doi: 10.1016/j.ins.2018.05.016.

[16] G. Stumme, R. Taouil, Y. Bastide, N. Pasquier, and L. Lakhal. Computing iceberg concept lattices with titanic. *Data and Knowledge Engineering*, 42(2):189–222, 2002. doi: 10.1016/S0169-023X(02)00057-5.

[17] A. Formica. Ontology-based concept similarity in formal concept analysis. *Information Sciences*, 176(18):2624–2641, 2006. doi: 10.1016/j.ins.2005.11.014.

[18] Y.Y. Yao. Two views of the theory of rough sets in finite universes. *International Journal of Approximate Reasoning*, 15(4):291–317, 1996. doi: 10.1016/S0888-613X(96)00071-0.

[19] F. Xu, Y. Yao, and D. Miao. Rough set approximations in formal concept analysis and knowledge spaces. *In: An A., Matwin S., Ras Z.W., Slezak D. (eds) Foundations of Intelligent Systems. International Symposium on Methodologies for Intelligent Systems (ISMIS), Lecture Notes in Computer Science*, 4994:319–328, 2008. doi: 10.1007/978-3-540-68123-6_35.

[20] G. Cattaneo, G. Chiaselotti, D. Ciucci, and T. Gentile. On the connection of hypergraph theory with formal concept analysis and rough set theory. *Information Sciences*, 330:342–357, 2016. doi: 10.1016/j.ins.2015.09.054.

[21] D. Ciucci, D. Dubois, and H. Prade. The structure of oppositions in rough set theory and formal concept analysis - toward a new bridge between the two settings. *In: Beierle C., Meghini C. (eds) Int. Symp. on Foundations of Information and Knowledge Systems (FoIKS), Lecture Notes in Computer Science*, 8367:154–173, 2014. doi: 10.1007/978-3-319-04939-7_7.

[22] X. Kang and D. Miao. A rough set model based on formal concept analysis in complex information systems. *In: 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 1:206–214, 2015. doi: 10.1109/FSKD.2015.7381941.

[23] J. Atif, I. Bloch, and C. Hudelot. Some relationships between fuzzy sets, mathematical morphology, rough sets, f-transforms, and formal concept analysis. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 24(Supplement-2):1–32, 2016. doi: 10.1142/S0218488516400080.

[24] B. Ganter and S.O. Kuznetsov. Scale coarsening as feature selection. *In: Medina R., Obiedkov S. (eds) International Conference on Formal Concept Analysis, ICFCA, Lecture Notes in Computer Science*, 4933:217–228, 2008. doi: 10.1007/978-3-540-78137-0_16.

[25] B. Ganter and C. Meschke. A formal concept analysis approach to rough data tables. *In: Transactions on Rough Sets*, XIV:31–67, 2011. ISBN 978-3-642-21562-9.

[26] H. Lai and D. Zhang. Concept lattices of fuzzy contexts: Formal concept analysis vs. rough set theory. *International Journal of Approximate Reasoning*, 50(5):695–707, 2009. doi: 10.1016/j.ijar.2008.12.002.

[27] Y.Q. Yao and J.S. Mi. Fuzzy concept lattices determined by (θ, σ)-fuzzy rough approximation operators. *In: Wen P., Li Y., Polkowski L., Yao Y., Tsumoto S., Wang G. (eds) Rough Sets and Knowledge Technology (RSKT), Lecture Notes in Computer Science*, 5589:601–609, 2009. doi: 10.1007/978-3-642-02962-2_76.

[28] M.W. Shao, M. Liu, and W.X. Zhang. Set approximations in fuzzy formal concept analysis. *Fuzzy Sets and Systems*, 158(23):2627–2640, 2007. doi: 10.1016/j.fss.2007.05.002.

[29] G. Beydoun. Formal concept analysis for an e-learning semantic web. *Expert Systems with Applications: An International Journal*, 36(8):10952–10961, 2009. doi: 10.1016/j.eswa.2009.02.023.

[30] A. Formica. Concept similarity in formal concept analysis: an information content approach. *Knowledge-Based Systems*, 21(1):80–87, 2008. doi: 10.1016/j.knosys.2007.02.001.

[31] A. Formica. Concept similarity in fuzzy formal concept analysis for semantic web. *Int. J. of Uncertainty, Fuzziness And Knowledge-Based Systems*, 18(2):153–167, 2010. doi: 10.1142/S0218488510006465.