Resolution of conflicts among ontology mappings: a fuzzy approach $^{(*)}$

Alfio Ferrara[1], Davide Lorusso[1],
Giorgos Stamou[2], Giorgos Stoilos[2], Vassilis Tzouvaras[2], Tassos Venetis[2]

[1] Università degli Studi di Milano,
DICo, via Comelico 39, 20135 Milano, Italy,
{ferrara, lorusso}@dico.unimi.it

[2] National Technical University of Athens,
IVML, Iroon Polytechniou 9, 15780 Athens, Greece,
{gstam,gstoil,tzouvaras,avenet}@image.ece.ntua.gr

Abstract. Interoperability is a strong requirement in open distributed systems and in the Semantic Web. The need for ontology integration is not always completely met by the available ontology matching techniques because, in most cases, the semantics of the compared ontologies is not considered, thus leading to inconsistent mappings. Probabilistic approaches has been proposed to validate mappings and solve the inconsistencies, based on a mapping confidence measure. As probabilistic approaches suffer from the lack of well-founded likelihood measures of mapping correctness, we propose a validation approach based on fuzzy interpretation of mappings, which better models the notion of degree of similarity between ontology elements. Moreover, we describe a conflict resolution method which computes the minimal sets of conflicting mappings and can be the ground of different validation strategies.

1 Introduction

In the context of the Semantic Web, the available information is organized in ontologies. Ontologies are controlled vocabularies describing objects and relations between them in a formal way, and have a grammar for using the vocabulary terms in order to express something meaningful within a specified domain of interest. However, ontologies themselves can be heterogeneous: given two ontologies describing a reference domain, the same real entity can be denoted in the two ontologies with different names or it can be defined in different ways (an entity of one ontology may be the union of two of the entities of the other ontology) whereas both ontologies may be expressed in different languages, though expressing the same knowledge. In order to achieve the goal of ontology interoperability, we need to align heterogeneous ontologies by (semi-)automatically discovering mappings between the elements in two different ontologies. Most of

$^{(*)}$ This paper has been partially funded by the BOEMIE Project, FP6-027538, 6th EU Framework Programme.

the existing matching techniques do not take into account the semantics of the compared ontologies, therefore the resulting mappings can not be interpreted as semantic relations among the ontology elements, which is a necessary condition to perform integration and, subsequently, query answering over the integrated schema. Recently, several studies have focused on mapping validation with respect to the semantics of the ontologies involved and, at the same time, by maintaining the uncertain nature of mappings. In [1] is proposed a language for representation and reasoning with uncertain mappings by combining ontology and rule languages with probabilistic reasoning. This method represents confidence values as error probabilities in order to resolve inconsistencies by using trust probabilities, and to reason about these on a numeric level. In our previous work [2] we presented a tool for mapping validation with the help of probabilistic reasoning. The idea is to assume a semantic interpretation of ontology mappings as probabilistic and hypothetical relations among ontology elements in order to build a unique distributed knowledge base from the two independent ontologies and, subsequently, check for inconsistencies.

Probabilistic approaches for mapping validation suffer of limitations due to the nature of mappings and the way the probability values are computed. Our idea is to adopt a completely different interpretation in order to be able to validate mappings even in the absence of a precise semantics and in the presence of uncertainty. Assuming that an ontology mapping states the generic similarity of two concepts, we can assert that the objects modeled by the first concept can be also modeled by the second concept to a certain degree. In other words, the individuals of the first concept belong to the second concept with a certain degree, which is exactly the semantics of fuzzy membership functions. The degree of membership is determined by the strength of the similarity relation, computed by the same matching technique which produced the mapping. By using the acquired mappings to create fuzzy individual assertions, we provide a formal interpretation of mappings. Moreover, on the grounds of the Fuzzy Description Logics theory, we are able to perform reasoning on the integrated ontologies in order to detect and solve inconsistencies by mapping refinement, which is another difference compared to [2].

2 Ontology Mappings and Fuzzy Interpretation

In this section, we provide an introduction to a fuzzy extension of Description Logics (DL) by adding degrees to DL facts; we call this extension f-DL. This extension is based on Fuzzy Sets and Fuzzy Logic [3] and on previous work on fuzzy Description Logics [4, 5].

As usual fuzzy DLs are defined by an alphabet of distinct *concept names* (*class names*) $\mathbf{C}$, *role names* (*property names*) $\mathbf{R}$ and *individuals* $\mathbf{I}$. The set of roles (properties) is defined as $\mathbf{R} \cup \{R^- \mid R \in \mathbf{R}\}$, where R^- represents the *inverse* of R. Elementary descriptions are *atomic concepts* and *atomic roles*, and by using concept constructors we can define complex concept descriptions. More precisely, if $A, C, D \in \mathbf{C}$, $R, S \in \mathbf{R}$ and $p \in \mathbb{N}$, where A is an atomic concept,

C, D are complex concepts, and S is an atomic role [6], then f-$\mathcal{SHIN}$-concepts are defined inductively by the following abstract syntax:

$$C, D \longrightarrow \bot \mid \top \mid A \mid C \sqcup D \mid C \sqcap D \mid \neg C \mid \forall R.C \mid \exists R.C \mid \geq pS \mid \leq pS$$

A fuzzy DL Knowledge Base Σ is a triple $\Sigma = \langle \mathcal{T}, \mathcal{R}, \mathcal{A} \rangle$, where $\mathcal{T}$ is a *TBox*, $\mathcal{R}$ a *RBox* and $\mathcal{A}$ an *ABox*. A TBox is a set of *concept subsumption axioms* of the form, $C \sqsubseteq D$ and *concept equivalence axioms* of the form $C \equiv D$, where C, D are f-$\mathcal{SHIN}$-concepts. An RBox is a set of *transitive role axioms* of the form $\mathsf{Trans}(R)$ and *role subsumption axioms* of the form $R \sqsubseteq S$, where R, S are f-$\mathcal{SHIN}$-roles, while an ABox is a set of *fuzzy concept* and *fuzzy role assertions* of the form $(a : C)\bowtie n$ and $((a, b) : R)\bowtie n$, or individual equalities and inequalities of the form $a = b$ or $a \neq b$, where $a, b \in \mathbf{I}$, $\bowtie \in \{\geq, >, \leq, <\}$ and $n \in [0, 1]$.

The semantics of f-DL are based on *fuzzy interpretations*. A fuzzy interpretation $\mathcal{I}$ is a pair $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where the domain $\Delta^{\mathcal{I}}$ is, like the crisp case, a non-empty set of objects and $\cdot^{\mathcal{I}}$ is a fuzzy interpretation function, which maps

- an *individual name* o to an object $o^{\mathcal{I}} \in \Delta^{\mathcal{I}}$,
- a *concept name* C to a membership function $C^{\mathcal{I}} : \Delta^{\mathcal{I}} \rightarrow [0, 1]$ [1], and
- a *property name* R to a membership function $R^{\mathcal{I}} : \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \rightarrow [0, 1]$.

Complex f-$\mathcal{SHIN}$-concepts, roles and axioms are interpreted by extending fuzzy interpretation, making use of fuzzy set theoretic operators and notions, like subsethood, from the fuzzy set literature. The complete semantics are presented in Table 1, where sup is the *supremum*, inf is the *infimum*, c is a *fuzzy complement*, t is a *fuzzy conjunction* (*t-norm*), u is a *fuzzy disjunction* (*t-conorm*) and $\mathcal{J}$ is a *fuzzy implication*.

A fuzzy knowledge base Σ is satisfiable iff there exists a fuzzy interpretation $\mathcal{I}$ which satisfies all axioms in Σ. Basic inference problems in f-DL are: (i) check if a fuzzy knowledge base is *consistent* i.e. has a model, (ii) check if D *subsumes* C w.r.t. Σ, i.e. $\Sigma \models C \sqsubseteq D$, (iii) check if a is an instance of C to degree $\bowtie n$, i.e. $\Sigma \models a : C\bowtie n$, where $\bowtie \in \{\geq, >, \leq, <\}$ and (iv) determine the *greatest lower bound* of a w.r.t. Σ, denoted $glb(\Sigma, a)$, where $glb(\Sigma, a) = \sup\{n \mid \Sigma \models a \geq n\}$.

2.1 Fuzzy Interpretation of Ontology Mappings

In order to achieve ontology interoperability heterogeneous ontologies should be (semi-)automatically aligned. The problem called "Ontology Alignment" or "Ontology Matching" can be described as follows: given two ontologies each describing a set of discrete entities (which can be classes, properties, predicates, etc.), find the relationships (e.g., equivalence or subsumption) that hold between these entities. In a more formal way we could say that a mapping $\mathcal{M}$ is a set of tuples

$$m_i = \langle C_i, C'_i, n_i, R_i \rangle$$

for $i \in I$, where

[1] For instance, given an object $a \in \Delta^{\mathcal{I}}$ and a class name C, $C^{\mathcal{I}}(a)$ gives a degree of confidence (such as 0.8) that the object a belongs to the fuzzy concept C.

Table 1. Fuzzy DL Descriptions and Axioms

Abstract Syntax	DL Syntax	Semantics
Bottom	$\bot$	$\bot^{\mathcal{I}}(a) = 0$
Top	$\top$	$\top^{\mathcal{I}}(a) = 1$
Intersection	$C \sqcap D$	$(C \sqcap D)^{\mathcal{I}}(a) = t(C^{\mathcal{I}}(a), D^{\mathcal{I}}(a))$
Union	$C \sqcup D$	$(C \sqcup D)^{\mathcal{I}}(a) = u(C^{\mathcal{I}}(a), D^{\mathcal{I}}(a))$
Complement	$\neg C$	$(\neg C)^{I}(a) = c(C^{\mathcal{I}}(a))$
Existential Restriction	$\exists R.C$	$(\exists R.C)^{\mathcal{I}}(a) = \sup_{b \in \Delta^{\mathcal{I}}} t(R^{\mathcal{I}}(a,b), C^{\mathcal{I}}(b))$
Universal Restriction	$\forall R.C$	$(\forall R.C)^{\mathcal{I}}(a) = \inf_{b \in \Delta^{\mathcal{I}}} J(R^{\mathcal{I}}(a,b), C^{\mathcal{I}}(b))$
Min Cardinality Restriction	$\geq nR$	$(\geq nR)^{\mathcal{I}}(a) = \sup\limits_{b_1,\ldots,b_p \in \Delta^{\mathcal{I}}} t(\mathop{\mathrm{t}}\limits_{i=1}^{p} R^{\mathcal{I}}(a,b_i), \mathop{\mathrm{t}}\limits_{i<j}\{b_i \neq b_j\})$
Max Cardinality Restriction	$\leq nR$	$(\leq nR)^{\mathcal{I}}(a) = \inf\limits_{b_1,\ldots,b_{p+1} \in \Delta^{\mathcal{I}}} \mathcal{J}(\mathop{\mathrm{t}}\limits_{i=1}^{p+1} R^{\mathcal{I}}(a,b_i), \mathop{\mathrm{u}}\limits_{i<j}\{b_i = b_j\})$
SubClass	$C \sqsubseteq D$	$C^{\mathcal{I}}(a) \leq D^{\mathcal{I}}(a)$
Equivalent Classes	$C \equiv D$	$C^{\mathcal{I}}(a) = D\mathcal{I}(a)$
SubRole	$R \sqsubseteq S$	$R^{\mathcal{I}}(a,b) \leq S^{\mathcal{I}}(a,b)$
Class Individual	$o : C \bowtie n$	$C^{\mathcal{I}}(o^{\mathcal{I}}) \bowtie n$
Role Individual	$(o, o') : R \bowtie n$	$R^{\mathcal{I}}(o^{\mathcal{I}}, o'^{\mathcal{I}}) \bowtie n$
Disjoint Classes	$C \sqsubseteq \neg D$	$C^{\mathcal{I}}(a) \leq 1 - D^{\mathcal{I}}(a)$
Transitive Object Property	$\mathsf{Trans}(R)$	$\sup_{b \in \Delta^{\mathcal{I}}} t(R^{\mathcal{I}}(a,b), R^{\mathcal{I}}(b,c)) \leq R^{\mathcal{I}}(a,c)$

- C_i, C_i' are the discrete entities from two ontologies, $\mathcal{O}$ and $\mathcal{O}$', between which a relation is asserted by the mapping;
- n_i is a value, which is a part of structure $\langle \mathcal{D}, \leq, 0, 1 \rangle$, where $\mathcal{D}$ is the set of degrees and $\forall d \in \mathcal{D}, 0 \leq d \leq 1$ holds, that denotes the strength of the relation R_i;
- and R_i is one of the following relations $R = \{\equiv, \sqsubseteq, \sqsupseteq\}$, that holds between the entities C_i and C_i'.

Another way to represent these relations using bridge rules, as used in distributed description logics [7], is

$$C_i \xrightarrow{\equiv} C_i' : n \qquad C_i \xrightarrow{\sqsubseteq} C_i' : n \qquad C_i \xrightarrow{\sqsupseteq} C_i' : n$$

In order to take into account the uncertain and fuzzy nature of the mappings we define a fuzzy mapping as follows.

Definition 1 (Fuzzy Mapping). *Given two ontology elements C_i and C_i', a fuzzy mapping $fm_i = \langle C_i, C_i', n_i, R_i \rangle$ is a mapping m_i, whose value n_i denotes the degree that the semantic relation R_i holds between C_i and C_i', where R_i can be one of equivalence ($\equiv$) or subsumption ($\sqsubseteq$, $\sqsupseteq$).*

This way the mappings are formalized as fuzzy knowledge. The basic idea behind the formalization of mappings as fuzzy knowledge is to use the mappings so as to create fuzzy individual assertions. In order to do that we must provide semantics for the mappings and to do so we will use the Fuzzy Set Theory [3]. Let $\mathcal{I}$ be a fuzzy interpretation, while let $\mathcal{I}_c$ be a crisp interpretation. Then we have the following conditions:

$$\mathcal{I} \models C_i \xrightarrow{\equiv} C'_i : n_i \Longleftrightarrow \forall b.b \in C_i^{\mathcal{I}_c} \rightarrow C_i'^{\mathcal{I}}(b) = n_i$$
$$\mathcal{I} \models C_i \xrightarrow{\sqsubseteq} C'_i : n_i \Longleftrightarrow \forall b.b \in C_i^{\mathcal{I}_c} \rightarrow C_i'^{\mathcal{I}}(b) \geq n_i$$
$$\mathcal{I} \models C_i \xrightarrow{\sqsupseteq} C'_i : n_i \Longleftrightarrow \forall b.b \in C_i^{\mathcal{I}_c} \rightarrow C_i'^{\mathcal{I}}(b) \leq n_i$$

The above definitions imply a procedure by which we can transfer individuals from the source ontology $\mathcal{O}$ to the target ontology $\mathcal{O}$', creating a set of fuzzy assertions A_M. This procedure will be described in more detail in the following.

2.2 Fuzzy DL Reasoning with FiRE

In this section we provide a short introduction to the Fuzzy Reasoning Engine FiRE [8]. FiRE is a prototype JAVA implementation of a fuzzy algorithm for an expressive fuzzy DL language f_{KD}-$\mathcal{SHIN}$ [9]. It allows the user to create a fuzzy knowledge base, based on the description logic Knowledge Representation System Specification (KRSS) which was extended to accommodate the fuzzy elements of fuzzy assertions. The inference services that FiRE supports are: (i) checking consistency of a fuzzy knowledge base, (ii) entailment of fuzzy assertions and (iii) subsumption between two fuzzy concepts. In the following of the paper and in the evaluation procedure we will use the consistency checking inference service.

3 Mapping Validation

Our approach to mapping validation is articulated in four phases

1. *Ontology mapping acquisition.* In this phase, we acquire mappings produced by using an ontology mapping system; the matching system can rely on syntactic, structural or even semantic matching techniques.
2. *Fuzzy interpretation of mappings.* In this phase, the acquired mappings are interpreted as fuzzy assertions as presented in Section 2.1.
3. *Fuzzy reasoning over mappings.* In this phase, the ontology obtained by enriching the second ontology of the mapping with fuzzy individual assertions produced with the help of the mappings is checked for consistency by means of a fuzzy reasoning system.
4. *Mapping validation and revision.* In this phase, mappings are revised according to the reasoning results; mappings causing inconsistencies within the new ontology are refined and given a new strength.

In more detail the validation procedure, takes as input a mapping set (M) together with the respective ontologies (O_1 and O_2) and creates a new mapping set (M'), which includes refined mappings or discarded ones.

The main algorithm is described by **Algorithm-1**. Firstly, M is ordered by descending order. In this way, we first consider the stronger mappings for which similarity is higher. Then, the algorithm examines each mapping with the aforementioned order and calculates a strength. If a mapping was refined

Algorithm 1 $M' :=$ fuzzyValidation(M)

input: a mapping set M, and the mapped ontologies
output: a validated mapping set M'
while the degree of some mapping has changed **do**
  sort M w.r.t. the strength n_i of each mapping $m_i = \{C_i, C'_i, n_i, R_i\} \in M$
  $M' := \emptyset$
  for $m_i \in M$ **do**
    $newStrength_i :=$ computeStrength(m_i)
    if $newStrength_i$ is different than n_i **then**
      $m_i := \{C_i, C'_i, newStrength_i, R_i\}$
      break
    end if
    if $newStrength$ is non zero **then**
      add m_i to M'
    end if
  end for
end while
return M'

then the same method is applied again for the old set of mappings plus the new refined one, since the new degree might cause a new conflict that did not occur before. This is performed iteratively until all the mappings have been used and no inconsistencies occur. The final set of mappings is saved in M'.

The method that refines the degree of a mapping is described by **Algorithm-2** and proceeds as follows: A new ontology $O' = \langle T', R', A' \rangle$ is created, where $T' = T_2$, $R' = R_2$. The ABox of the new ontology is gradually constructed from the ABox of O_2 and by using the current mapping in order to transfer individuals from ontology O_1. More formally, $A' = A_2 \cup A_M$, where A_M is defined as follows:

$$\begin{aligned} A_M \quad = \quad & \{a : C'_i \geq n \mid \langle C_i, C'_i, n, \sqsubseteq \rangle \in M, O_1 \models C_i(a)\} \cup \\ & \{a : C'_i = n \mid \langle C_i, C'_i, n, = \rangle \in M, O_1 \models C_i(a)\} \cup \\ & \{a : C'_i \leq n \mid \langle C_i, C'_i, n, \sqsupseteq \rangle \in M, O_1 \models C_i(a)\}. \end{aligned}$$

As it can be noted by the above definition, both the explicit as well as inferred assertions are taken into consideration ($O_1 \models C_i(a)$). To do so we make use of a classic DL reasoner and more precisely in the current setting we have used Pellet [10]. For example, if $m_i = \langle C_i, C'_i, 0.8, \equiv \rangle$ and $O_1 \models C_i(a)$ then $A_M = A_M \cup \{a : C'_i = 0.8\}$. After, a new fuzzy individual assertion has been added in O' we call FiRE, in order to check for inconsistencies. If an inconsistency occurs the strength of the mapping is refined, while if an inconsistency does not occur the old degree is retained. The procedure that refines the strength of the mapping is `refineStrength`. This procedure takes as input low level information from the fuzzy reasoner about what conditions created the inconsistency, and according to it proceeds with the refinement of the strength of the mapping so as to restore the consistency in the ontology. For example, a pair of assertions of the form $a : C \geq 0.8$ and $a : C \leq 0.7$ obviously denotes a contradiction.

Algorithm 2 s := computeStrength(m_i)

input: $m_i := \{C_i, C'_i, n_i, R_i\}$
output: the new strength of the mapping
for every individual of C_i ($C_i(a)$) **do**
  add a to $C'_i \longrightarrow C'_i(a)$
  check consistency of O_2
  if O_2 is not consistent **then**
    remove all individuals of C_i added to C'_i
    s := refineStrength(inconsistencyInfo)
  else
    $s := n_i$
  end if
end for
return s

Example. Consider two simple ontologies, O_1 and O_2, defined as follows:

$$O_1 : MobilePhone \sqsubseteq MobileDevice$$

$$\begin{aligned} O_2 : Phone &\sqsubseteq ElectronicDevice \\ CablePhone &\sqsubseteq Phone \\ CellularPhone &\sqsubseteq Phone \\ CablePhone &\sqsubseteq \neg CellularPhone \end{aligned}$$

The two ontologies have been compared by adopting the linguistic component of HMatch 2.0 [11], which is based on a combination of terminological and syntactic techniques. The result of the matching process is the following set of mappings:

1. map($MobileDevice$, $ElectronicDevice$, 0.7)
2. map($MobilePhone$, $Phone$, 0.6)
3. map($MobilePhone$, $CablePhone$, 0.4)
4. map($MobilePhone$, $CellularPhone$, 1.0)

Since the validation process works by translating mappings into fuzzy individual assertions, suppose that each concept of the two ontologies has at least one representative individual. In particular, we assume that mp_1 is an instance of the concept $MobilePhone$ and md_1 is an instance of the concept $MobileDevice$. Sorted by the strength, one by one mappings are inserted into the second ontology as fuzzy individual assertions.

Following the example, the first mapping to be added to O_2 is mapping 4, which is translated into the assertion ($CellularPhone(mp_1)$, 1.0). Since the first mapping does not cause an inconsistency, the procedure moves to the subsequent mapping (1), which is converted into ($ElectronicDevice(md_1)$, 0.7) and ($ElectronicDevice(mp_1)$, 0.7). The latter assertion violates the fuzzy DLs interpretation of subsumption ($C \sqsubseteq D \Longleftrightarrow C^{\mathcal{I}}(a) \leq D^{\mathcal{I}}(a)$), therefore making the resulting ontology inconsistent. In this case, the solution is to increase the strength of $ElectronicDevice(mp_1)$ and $ElectronicDevice(md_1)$ to 1 in order to satisfy the semantic constraint $ElectronicDevice^{\mathcal{I}}(mp_1^{\mathcal{I}}) \geq CellularPhone^{\mathcal{I}}(mp_1^{\mathcal{I}})$. The same situation occurs when the assertions corresponding to mapping 2 are added into O_2 and the same refinement is applied to restore consistency. At last, the assertion determined by mapping 3, i.e. ($CablePhone(mp_1)$, 0.4), causes an inconsistency because it does not satisfy the semantic constraint $CablePhone^{\mathcal{I}}(mp_1^{\mathcal{I}}) \leq 1 - CellularPhone^{\mathcal{I}}(mp_1^{\mathcal{I}})$. Giving priority to the stronger mapping, the latest assertion has to be refined. Since the resulting strength would be equal to 0, the assertion corresponding to mapping 3 is definitely dropped, and the mapping is removed as well. The result of the validation process is the following mapping set:

1. map($MobileDevice$, $ElectronicDevice$, 1.0)
2. map($MobilePhone$, $Phone$, 1.0)
3. map($MobilePhone$, $CellularPhone$, 1.0)

4 Conflict Resolution

The validation process described in the previous section enforces the inconsistency detection and resolution by refining the strength of the mappings. When a conflict arises, two or more mappings are involved and, to achieve the consistency, at least one of them must be refined or removed. Generally the choice among the conflicting mappings is not trivial because it should be driven by the semantics of the mapped elements. The decision is even a harder task when is performed automatically, therefore requiring effective heuristics. Moreover, even when the choice is made by a human expert, there can be different correct decisions according to different criteria that can be adopted.

The proposed validation technique adopts a naive strategy which gives priority to the strongest mapping and forces the last added mapping to be refined or deleted. This solution has the advantage of being efficient in terms of performances but does not always lead to the expected results. In fact, for instance, one may prefer to preserve the highest number of mappings instead of the strongest ones. The limitation is more evident if we consider mapping deletion as the only possible way to solve inconsistencies. For instance, consider the two ontologies defined in the example of the previous section and assume to have the same mapping set but with the following strength values:

1. map($MobileDevice$, $ElectronicDevice$, 0.5)
2. map($MobilePhone$, $Phone$, 0.7)
3. map($MobilePhone$, $CablePhone$, 0.8)
4. map($MobilePhone$, $CellularPhone$, 0.6)

The conflicting subsets of mappings in this configuration are (1,2,3) and (3,4), due to the violation of the fuzzy DLs interpretation of subsumption and negation, respectively. If we apply a restricted version of the validation procedure of Section 3 that allows only the deletion of inconsistent mappings, the inconsistency would be solved by deleting all the mappings except for mapping 3, which is the strongest one. In this case, it is clear that giving priority to the mapping with the highest value could be not always the expected choice.

To provide a better support for the resolution of mapping inconsistencies, we propose a different approach, namely the *conflict resolution method*, based on the complete analysis of the conflicts. The underlying idea is to compute a *degree of inconsistency* of each mapping, i.e. a measure that reflects the number of times in which a mapping is involved in a conflict. To evaluate this degree, we consider the inconsistencies in all the possible mapping configurations, that are the set $\overline{\mathcal{P}(M)}$ of all the subsets of the given mapping set, except the empty set and the singleton sets. More formally, given a set of mappings M and the set $\overline{\mathcal{P}(M)} \equiv \mathcal{P}(M) \setminus \{x \in \mathcal{P}(M) \mid x = \emptyset \vee |x| = 1\}$, we define the conflicting set $\mathcal{C}(M) \subseteq \overline{\mathcal{P}(M)}$ as

$$\mathcal{C}(M) = \{c \in \overline{\mathcal{P}(M)} \mid \exists\, m, m' \in c \text{ such that } m \text{ and } m' \text{ cause an inconsistency}\}$$

$\mathcal{C}(M)$ is built by validating each subset $s_i \in \overline{\mathcal{P}(M)}$ through the validation procedure of Section 3. If the resulting set s_i' is equal to s_i then s_i does not contain any conflict and it is not included into $\mathcal{C}(M)$. Otherwise, if $s_i' \subset s_i$ then a mapping has been removed to solve an inconsistency, therefore s_i is added into $\mathcal{C}(M)$.

We define the minimal conflicting set $\mathcal{MC}(M)$ of M as the collection of all minimal subset of mappings which contains a conflict:

$$\mathcal{MC}(M) = \{mc \in \mathcal{C}(M) \mid \nexists\, mc' \in \mathcal{C}(M) \text{ such that } mc' \subseteq mc\}$$

The degree of inconsistency i_m of a mapping $m \in M$ is defined as follows:

$$i_m = |\{mc \in \mathcal{MC}(M) \mid m \in mc\}|$$

The assumption is that the higher is the degree of inconsistency of a mapping, the more benefit we will get by removing it from the mapping set. Therefore, the strategy behind this conflict resolution method is to preserve as much as possible the mappings by detecting and deleting those which participate in the highest number of conflicts. After computing the degree of inconsistency, all the mappings are added into the second ontology as fuzzy individual assertions and the resulting ontology is checked for consistency. If an inconsistency is detected, the mapping with the highest degree of inconsistency is removed and the resulting ontology is again checked for consistency. The step is repeated until consistency is achieved.

Let us describe this method with the aforementioned set of mappings that are not correctly validated by the strength-based ordering approach. The computation of the degrees of inconsistency produces the following results:

$$\begin{aligned}
\overline{\mathcal{P}(M)} &= \{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4),(1,2,3),(1,2,4),\ldots\} \\
\mathcal{C}(M) &= \{(1,2),(1,3),(3,4),(1,2,3),(1,2,4),(2,3,4),(1,2,4),(1,2,3,4)\} \\
\mathcal{MC}(M) &= \{(1,2),(1,3)(3,4)\}
\end{aligned}$$

$$i_1 = |\{(1,2),(1,3)\}| = 2, \quad i_2 = |\{(1,2)\}| = 1, \quad i_3 = 2, \quad i_4 = 1$$

All mappings are added into the resulting ontology and, subsequently, mappings 1 and 3 are removed before consistency is restored. Compared with the results of the strength-based ordering approach, this method detected the actual incorrect mapping (3) and produced the configuration with the largest number of mappings. Moreover, in the context of semi-automatic validation tools, the analysis performed with this method can report to the user the actual minimal sets of conflicting mappings, in order to better support the decision process.

5 Related Work

Recent work [12] have focused on mapping validation as a post-processing task over mappings produced by other matchmaking tools. Grounded on the theories of the Distributed Description Logics, the process consists in translating mappings into bridge rules (i.e. inter-ontology semantic relations) and check for the consistency of the resulting distributed knowledge base. The approach does not handle the inherent uncertainty of mapping caused by the possible inaccuracy of the heuristics adopted by the matching techniques.

As a possible solution to cope with the uncertainty of automatically discovered mappings, probabilistic techniques have been developed. The approach presented in [13] translates the mapped ontologies into bayesian networks and treats concept mapping between the two ontologies as evidential reasoning between the two translated BN. In our foregoing work on mapping validation [2], starting from the crisp approach in [12], we refined the validation process by attaching to mappings a probability measure determined by the confidence value of the mapping. The probability value is interpreted as the likelihood of the mapping being correct. The resulting relations are interpreted according to the probabilistic description logics, which provides consistency check and inference services in order to perform validation. A similar approach has been presented in [1], where the combination of a rule-based framework and Probabilistic Description Logic Programs is exploited to validate and merge mappings produced by different techniques and tools. As in [2], the confidence value is interpreted as a probability measure of the mapping correctness.

To be effective, probabilistic approaches should be fed with values which actually state the confidence of the relation, therefore computed on the basis of well-founded statistical techniques or measures. This turns out to be a relevant limitation because most of the matchmaking tools do not provide such a measure but only a value representing the degree of similarity between the mapped elements. The alternative we propose is to exploit the fuzzy interpretation to handle the uncertainty of mappings but without relying on the confidence values. In the ontology matching literature, fuzzy theories have been exploited mainly with the aim of dealing with uncertainty during the process of mapping discovery and not for validation. For instance, the method described in [14] formulates the ontology mapping problem as a rule application problem in the fuzzy conceptual graph model. In our approach, based on the fuzzy description logics, the numeric value attached to a mapping is intended as a degree of truth of the relation. According to the way the numeric value is computed in most of the matching techniques, the fuzzy interpretation is more suitable compared to the probability value, especially when mappings represent a generic similarity relation between the concepts.

Regarding conflict resolution strategies, relevant work have been presented in the field of ontology repairing in order to provide debugging functionalities for logically erroneous knowledge bases. In [15], minimal incoherence-preserving sub-TBoxes (MIPS) are defined as the smallest subsets of an original TBox preserving unsatisfiability of at least one atomic concept. MIPS are detected and solved through a tableaux-like technique. Our definition of the degree of inconsistency adopts the same principle but applied to the mapping conflict resolution problem.

Other work in dealing with ontology mapping in the fuzzy context has been presented in [16] where Li et al. have introduced E-Connections integrated into extended fuzzy description Logics (EFDLs) that couple both fuzzy and distributed features within description logics and in [17], where Lu et al. propose a discrete tableau algorithm to achieve reasoning within the logical system of EFDLs. Unfortunately, not practical implementation of the algorithm is known, in order to be used in a practical setting for reasoning over such fuzzy mappings.

6 Concluding Remarks

In this paper we have discussed the application of the fuzzy DLs theories to the problem of mapping validation as a different way of handling mapping uncertainty with respect to probabilistic approaches. As a result, we described a mapping validation algorithm based on fuzzy interpretation of mappings in order to detect inconsistencies. Similarly to previous work on mapping validation, the strategy to solve inconsistencies is a simple strength-based heuristics, i.e. the conflicting mapping with the highest strength value is preserved. Although being a fast solution, this naive approach does not lead always to the expected configuration. To cope with possible different strategies, we proposed a conflict resolution approach which performs a thorough analysis of all possible inconsistencies and computes the minimal sets of conflicting mappings.

The preliminary results show that the conflict resolution method is effective and can potentially be applied to any validation semantics (e.g. probabilistic, fuzzy). Furthermore, other validation strategies can be built on top of it, for instance a strategy to maximize the number of preserved mappings. Regarding the complexity, it is obviously dependent on the number of mappings involved and, without further optimizations, the method is applicable only on relatively small alignments. Future work will be devoted to the development of optimization techniques, in particular the goal is to reduce the number of mapping subsets to be validated during the search for the minimal conflicting sets. A possible way of reducing the search space, and thus the combinatorial space, is to make some approximations, like the one proposed in [18]. Moreover, the proposed validation procedure supports only subsumption and equivalence, therefore further investigation are needed to include other kind of correspondences between aligned ontology elements.

References

1. Calì, A., Lukasiewicz, T., Predoiu, L., Stuckenschmidt, H.: A framework for representing ontology mappings under probabilities and inconsistency. In: URSW. (2007)
2. Castano, S., Ferrara, A., Lorusso, D., Nth, T.H., Moeller, R.: Mapping validation by probabilistic reasoning. In: Proceedings of the 5th European Semantic Web Conference. LNCS, Springer Verlag (2008)
3. Zadeh, L.A.: Fuzzy sets. Information and Control **8** (1965) 338–353
4. Straccia, U.: Towards a fuzzy description logic for the semantic web. In: Proceedings of the 2nd European Semantic Web Conference. (2005)
5. Stoilos, G., Stamou, G., Pan, J.: Handling imprecise knowledge with fuzzy description logics. In: Proceedings of the International Workshop on Description Logics (DL 2006), Lake District, UK. (2006)
6. Horrocks, I., Patel-Schneider, P.F., van Harmelen, F.: From $\mathcal{SHIQ}$ and RDF to OWL: The making of a web ontology language. Web Semantics **1** (2003) 7–26
7. Borgida, A., Serafini, L.: Distributed description logics: Assimilating information from peer sources. In Spaccapietra, S., March, S.T., Aberer, K., eds.: J. Data Semantics I. Volume 1 of Lecture Notes in Computer Science., Springer (2003) 153–184
8. Simou, N.: Fuzzy Reasoning Engine FiRE, (http://www.image.ece.ntua.gr/ nsimou/FiRE/)
9. Stoilos, G., Stamou, G., Tzouvaras, V., Pan, J.Z., Horrocks, I.: Reasoning with very expressive fuzzy description logics. Journal of Artificial Intelligence Research **30** (2007) 273–320
10. Sirin, E., Parsia, B., Grau, B.C., Kalyanpur, A., Katz, Y.: Pellet: A practical OWL-DL reasoner. Journal of Web Semantics **5** (2007) 51–53
11. Castano, S., Ferrara, A., Montanelli, S.: Matching Ontologies in Open Networked Systems: Techniques and Applications. Journal on Data Semantics (JoDS) **V** (2006)
12. Meilicke, C., Stuckenschmidt, H., Tamilin, A.: Repairing ontology mappings. In: AAAI. (2007) 1408–1413
13. Pan, R., Ding, Z., Yu, Y., Peng, Y.: A Bayesian Network Approach to Ontology Mapping. In: Proceedings of the Fourth International Semantic Web Conference. (2005)
14. Buche, P., Dibie-Barthélemy, J., Ibanescu, L.: Ontology mapping using fuzzy conceptual graphs and rules. In: ICCS Supplement. (2008) 17–24
15. Schlobach, S., Cornet, R.: Non-standard reasoning services for the debugging of description logic terminologies. In: IJCAI, Morgan Kaufmann (2003) 355–362
16. Li, Y., Xu, B., Lu, J., Kang, D.: A distributed and fuzzy extension of description logics. In: KES (1). (2006) 655–662
17. Lu, J., Li, Y., Zhou, B., Kang, D., Zhang, Y.: Distributed reasoning with fuzzy description logics. In: International Conference on Computational Science (1). (2007) 196–203
18. Straccia, U., Troncy, R.: oMAP: Combining classifiers for aligning automatically OWL ontologies. In: 6th International Conference on Web Information Systems Engineering (WISE-05). Number 3806, Springer Verlag (2005) 133–147