Towards Intelligent Ontology Alignment Systems for Question Answering: Challenges and Roadblocks

Maria Vargas-Vera, Miklos Nagy
Open University, Walton Hall, Milton Keynes, MK7 6AA, England, UK
Email: {M.Vargas-Vera, M.Nagy}@open.ac.uk

Abstract— **This paper introduces the main challenges and future research directions for the Ontology Alignment problem. To date a good number of ontology alignment solutions have been proposed. These solutions utilise a wide variety of techniques from machine learning to uncertain reasoning. However, none of the approaches have proved to be an integrated solution, which can be used by different communities. Since 2004, the Ontology Alignment Initiative (OAEI) has established an annual evaluation for systems that could be tested using the same datasets. This, of course, has helped to improve the work on ontology alignment, as the ontology community now has a set of common datasets to make comparisons on the performance of different algorithms for ontology alignment. In this paper we discuss the main challenges and roadblocks that need to be addressed in order to built successful mapping frameworks. Finally this paper presents DSSim and our results on the ontology evaluation 2008.**

Index Terms— **ontology mapping, uncertain reasoning**

I. Challenges on ontology alignment

The Ontology Alignment Community has identified several challenges [1], [2], which are considered as major roadblocks for successful future implementations of ontology mapping systems. In our context (ontology mapping for question answering), we have identified five critical and interrelated challenges that can be considered as roadblocks for future successful mapping frameworks. The first challenge relates to the representation, the second to the quality of the data and the third one to the efficient ontology alignment for large ontologies. The fourth and the fifth challenge goes beyond the information related concepts and tries to address the overall difficulties namely the problem of generic and intelligent systems. We will discuss each our suggested challenges in turn.

A. Representation and interpretation problems

The vision of the Semantic Web is to achieve machine-processable interoperability through the annotation of the content. This implies that computer programs can achieve a certain degree of understanding of such data and use it to reason about user specific tasks like question answering or data integration. Data on the semantic web is represented by ontologies, which typically consist of a number of classes, relations, instances and axioms. These elements are expressed using a logical language. The W3C has proposed RDF(S) [3] and OWL [4] as Web ontology language however OWL has three increasingly-expressive sublanguages (OWL Lite, OWL DL, OWL Full) with different expressiveness and language constructs. In addition to the existing Web ontology languages W3C has proposed other languages like SKOS [5], which is a standard to support the use of knowledge organization systems (KOS) such as thesauri, classification schemes, subject heading systems and taxonomies within the framework of the Semantic Web. SKOS are based on the Resource Description Framework (RDF) and it allows information to be passed between computer applications in an interoperable way. Ontology designers can choose between these language variants depending on the intended purpose of the ontologies. The problem of interpreting semantic web data however stems not only from the different language representations [6] but the fact that ontologies especially OWL Full has been designed as a general framework to represent domain knowledge, which in turn can differ from designer to designer. Consider the following excerpts Fig. 1, 2 from different FAO (Food and Agricultural Organization of the United Nations) ontologies.

Assume we need to assess similarity between classes and individuals between the two ontologies. In fragment one a class *c_8375* is modelled as named OWL individuals. In the class description only the ID is indicated therefore to determine the properties of the class one needs to extract the necessary information from the actual named individual. In Fig. 2 the classes are represented as RDF individuals where the individual properties are defined as OWL data properties. One can note the difference how the class labels are represented on Fig. 1 through *rdfs:label* and Fig. 2 through *hasNameScientific* and *hasNameLongEN* tags. From the logical representation point of view both ontologies are valid separately and no logical reasoner would find inconsistency in them individually. However the problem occurs once we need to compare them in order to determine the similarities between classes and individuals. It is easy to see that once we need to compare the two ontologies a considerable amount of uncertainty arises over the classes and its properties and in a way they can be compared. This uncertainty can be contributed to the fact that due to the different representation certain elements will be missing for the comparison e.g. we have label in fragment Fig.

doi:10.4304/jetwi.2.3.244-257

```
...
<owl:Class rdf:ID="c_8375">
  <rdfs:subClassOf>
    <owl:Class rdf:ID="c_7033"/>
  </rdfs:subClassOf>
</owl:Class>
...
< c_8375 rdf:ID="i_8375">
  <aos:hasScopeNote xml:lang="EN">Isscaap group
b-52</aos:hasScopeNote>
  <aos:hasScopeNote xml:lang="FR">Groupe b-52 de
la csitapa</aos:hasScopeNote>
...
  <rdfs:label xml:lang="en">Demospongiae</rdfs:label>
</c_8375>
...
```

Figure 1. Ontology fragment from the AGROVOC ontology

```
...
<owl:Class rdf:about="#species">
  <rdfs:subClassOf rdf:resource="#biological_entity"/>
  <owl:disjointWith rdf:resource="#family"/>
  <owl:disjointWith rdf:resource="#order"/>
  <owl:disjointWith rdf:resource="#group"/>
</owl:Class>
...
<rdf:Description rdf:about="http://www.fao.org/aims
/aos/fi/species_v1.0.owl#31005_17431">
  <j.0:hasNameLongEN>Barrel
sponge</j.0:hasNameLongEN> <j.0:hasMeta>31005
  </j.0:hasMeta> <j.0:hasNameScientific>
  DEMOSPONGIAE</j.0:hasNameScientific>
</rdf:Description> ...
```

Figure 2. Ontology fragment from the ASFA ontology

1 but is missing from fragment Fig. 2 but there is *hasNameLongEN* tag in fragment Fig. 2 but missing in fragment Fig. 1.

As a result of these representation differences ontology mapping systems will always need to consider the uncertain aspects of how the semantic web data can be interpreted.

B. Quality of the Semantic Web Data

Data quality problems [7] [8] in the context of database integration [9] have emerged long before the Semantic Web concept has been proposed. The major reason for this is the increase in interconnectivity among data producers and data consumers, mainly spurred through the development of the Internet and various Web-based technologies. For every organisation or individual the context of the data, which is published can be slightly different depending on how they want to use their data. Therefore from the exchange point of view incompleteness of a particular data is quite common. The problem is that fragmented data environments like the Semantic Web inevitably lead to data and information quality problems causing the applications that process this data deal with ill-defined, inaccurate or inconsistent information on the domain. The incomplete data can mean different things to data consumer and data producer in a given application scenario. In traditional integration scenarios resolving these data quality issues represents a vast amount of time and resources for human experts before any integration can take place. Data quality has two aspects

- Data syntax covers the way data is formatted and gets represented
- Data semantics addresses the meaning of data

Data syntax is not the main reason of concern as it can be resolved independently from the context because it can be defined what changes must occur to make the data consistent and standardized for the application e.g. defining a separation rule of compound terms like "MSc-Thesis", "MSc_Thesis". The main problem what Semantic Web applications need to solve is how to resolve semantic data quality problems i.e. what is useful and meaningful because it would require more direct input from the users or creators of the ontologies. Clearly considering any kind of designer support in the Semantic Web environment is unrealistic therefore applications itself need to have built in mechanisms to decide and reason about whether the data is accurate, usable and useful in essence, whether it will deliver good information and function well for the required purpose. Consider the following example Fig. 3 from the directory ontologies.

```
...
<owl:Class rdf:about="http://matching.com/source
  /3887.owl#Windows_Vista">
  <rdfs:label xml:lang="en"> Windows Vista Home
Edition </rdfs:label>
  <j:hasSerialNumber>
    <rdfs:label >00043-683-036-658</rdfs:label>
  </j:hasSeriamNumber>
  <rdfs:subClassOf>
    <owl:Class rdf:about="http://matching.com
    /source/3887.owl#Operating_Systems">
    </owl:Class>
  </rdfs:subClassOf>
</owl:Class>
...
```

Figure 3. Ontology fragment from the Web directories ontology

As figure Fig. 3 shows we can interpret Windows Vista as the subclass of the operating systems however the designer has indicated that it has a specific serial number therefore it can be considered as an individual as well. At any case the semantic data quality is considered as low as the information is dubious therefore the Semantic Web application has to create its own hypothesises over the meaning of this data.

C. Efficient mapping with large scale ontologies

Ontologies can get quite complex and very large, causing difficulties in using them for any application [10] [11]. This is especially true for ontology mapping where overcoming scalability issues becomes one of the decisive factors for determining the usefulness of a system. Nowadays with the rapid development of ontology applications, domain ontologies can become very large and complex. This can partly be contributed to the fact that a number of general knowledge bases or lexical databases have been and will be transformed into ontologies in order to support more applications on the Semantic Web. Consider for example WordNet. Since the project started in 1985 WordNet [1] has been used for a number of different purposes in information systems. It is popular general background knowledge for ontology mapping systems because it contains around 150.000 synsets and their semantic relations. Other efforts to represent common sense knowledge as ontology is the Cyc project [2], which consists of more than 300.000 concepts and nearly 3.000.000 assertions or the Suggested Upper Merged Ontology(SUMO)[3] with its 20.000 terms and 70.000 axioms when all domain ontologies are combined. However the far largest ontology so far (according to our knowledge) in terms of concept number is the DBPedia [4] , which contains over 2.18 million resources or "things", each tied to an article in the English language Wikipedia. Discovering correspondences between these large scale ontologies is an ongoing effort however only partial mappings have been established i.e. SUMO-Wordnet due to the vast amount of human and computational effort involved in these tasks. The Ontology Alignment Initiative 2008 [12] has also included a mapping track for very large cross lingual ontologies, which includes establishing mappings between Wordnet, DBPedia an GTAA (Dutch acronym for Common Thesaurus for Audiovisual Archives) [13], which is a domain specific thesaurus with approximately 160.000 terms. A good number of researchers might argue that the Semantic Web is not just about large ontologies created by the large organisations but more about individuals or domain experts who can create their own relatively small ontologies and publish it on the Web. Indeed might be true however from the scalability point of view it does not change anything if thousands of small ontologies or a small number of huge ontologies need to be processed. Consider that in 2007 Swoogle [14] has already indexed more than 10.000 ontologies, which were available on the Web. The large number of concepts and properties that is implied by the scale or number of these ontologies poses several scalability problems from the reasoning point of view. Any Semantic Web application not only from ontology mapping domain has to be designed to cope with these difficulties otherwise it is deemed to be a failure from the usability point of view.

D. Task specific vs. generic systems

Existing mapping systems can clearly be classified into two categories. First group includes domain specific systems, which are build around well defined domains e.g. medical, scientific etc. These systems use specific rules, heuristics or background knowledge. As a consequence domain specific systems perform well on their own domain but their performance deteriorate across different domains. As a result the practical applicability of these systems on the Semantic Web can easily be questioned. The second group includes systems that aim to perform equally well across different domains. These systems utilise generic methods e.g. uncertain reasoning, machine learning, similarity combination etc. These systems has the potential to support a wide variety of applications on the Semantic Web in the future.

Based on this classification it is clear that building generic systems that perform equally well on different domains and provide acceptable results is a considerable challenge for the future research.

E. Incorporating intelligence

To date the quality of the ontology mapping was considered to be an important factor for systems that need to produce mappings between different ontologies. However competitions organised on ontology mapping has demonstrated that even if systems use a wide variety techniques, it is difficult to push the mapping quality beyond certain limits. It has also been recognised [15] that in order to gain better user acceptance, systems need to introduce cognitive support for the users i.e. reduce the difficulty of understanding the presented mappings.

There are different aspects of this cognitive support i.e. how to present the end results, how to explain the reasoning behind the mapping, etc. Ongoing research focuses on how the end results can be represented in a way that end users can understand better the complex relations of large-scale ontologies. Consider for example a mapping representation between two ontologies with over 10.000 concepts each. The result file can contain thousands of mappings. To visualise this mapping existing interfaces will most likely present an unrecognizable web of connections between these properties. Even though this complex representation can be presented in a way that users could better understand the problem still arises once the users need to understand why actually these mappings have been selected. This aspect so far has totally been hidden from the end users and has formed an internal and unexpoitable part of mapping systems itself.

Nevertheless in order to further improve the quality of the mapping systems these intermediary details need to be exposed to the users who can actually judge if the certain reasoning process is flawed or not. This important feedback or the ability to introspect can then be exploited by the system designers or ultimately the system itself through improving the reasoning processes, which is carried out behind the scenes in order to produce the end results. This ability to introspect the internal reasoning steps is a fundamental component of how human beings reason, learn and adapt. However, many existing ontology mapping systems that use different

[1] http://wordnet.princeton.edu/
[2] http://www.cyc.com/
[3] http://www.ontologyportal.org/
[4] http://dbpedia.org/About

forms of reasoning exclude the possibility of introspection because their design does not allow a representation of their own reasoning procedures as data. Using a model of reasoning based on observable effect it is possible to test the ability of any given data structure to represent reasoning. Through such a model we present a minimal data structure [16] necessary to record a computable reasoning process and define the operations that can be performed on this representation to facilitate computer reasoning. This model facilitates the introduction and development of basic operations, which perform reasoning tasks using data recorded in this format. It is necessary that we define a formal description of the structures and operations to facilitate reasoning on the application of stored reasoning procedures. By the help of such framework provable assertions about the nature and the limits of numerical reasoning can be made.

II. Approach to Ontology alignment considering uncertainty

For ontology mapping in the context of Question Answering over heterogeneous sources we propose a multi agent system called DSSim [17] because as a particular domain becomes larger and more complex, open and distributed, a set of cooperating agents are necessary in order to address the ontology mapping task effectively. In real scenarios, ontology mapping can be carried out on domains with large number of classes and properties. Without the multi agent architecture the response time of the system can increase exponentially when the number of concepts to map increases. The main objective of DSSim architecture is to be able to use it in different domains for creating ontology mappings. These domains include Question Answering, Web services or any application that need to map database metadata e.g. Extract, Transform and Load (ETL) tools for data warehouses. Therefore DSSim is not designed to have its own user interface but to integrate with other systems through well defined interfaces. In our implementation we have used the AQUA Question Answering system, which is the user interface that creates First Order Logic(FOL) statements based on natural language queries posed by the user. As a consequence the inputs and outputs for the DSSim component are valid FOL formulas.

An overview of our system is depicted on Fig. 4 The two real word ontologies[5][6] describe BibTeX publications from the University of Maryland, Baltimore County (UMBC) and from the Massachusetts Institute of Technology (MIT) . The AQUA [18] system and the answer composition component are described just to provide the context of our work (our overall framework) but these are not our major target in this paper. The user poses a natural language query to the AQUA system, which converts it into FOL (First Order Logic) terms. The main components and its functions of the system are as follows:

[5]http://ebiquity.umbc.edu/ontology/publication.owl

[6]http://visus.mit.edu/bibtex/0.01/bibtex.owl

1) Broker agent receives FOL term, decomposes it(in case more than one concepts are in the query) and distributes the sub queries to the mapping agents.
2) Mapping agents retrieve sub query class and property hypernyms from WordNet.
3) Mapping agents retrieve ontology fragments from the external ontologies, which are candidate mappings to the received sub-queries. Mapping agents use WordNet as background knowledge in order to enhance their beliefs on the possible meaning of the concepts or properties in the particular context.
4) Mapping agents build up coherent beliefs by combining all possible beliefs over the similarities of the sub queries and ontology fragments. Mapping agents utilize both syntactic and semantic similarity algorithms build their beliefs over the correctness of the mapping.
5) Broker agent passes the possible mappings into the answer composition component for particular sub-query ontology fragment mapping in which the belief function has the highest value.
6) Answer composition component retrieves the concrete instances from the external ontologies or data sources, which is included into the answer.
7) Answer composition component creates an answer to the user's question.

The main novelty in our solution is that we propose solving the ontology mapping problem based on the principles of collective intelligence, where each mapping agent has its own individual belief over the solution. However before the final mapping is proposed the broker agent creates the result based on a consensus between the different mapping agents. This process reflects well how humans reach consensus over a difficult issue.

A. Example scenario

Based on the architecture depicted on Fig. 4 we present the following simplified example, which will be used in the following sections of the paper in order to demonstrate our algorithm. We consider the following user query and its FOL representation as an input to our mapping component framework: List all papers with keywords uncertain ontology mapping?

$$(\exists x)\, paper\,(x)\, and$$
$$hasKeywords\,(x, [uncertain, ontology\ mapping])$$

- Step 1: Broker agent distributes (no decomposition is necessary in this case) the FOL query to the mapping agents.
- Step 2: Mapping agents 1 and 2 consult WordNet in order to extend the concepts and properties with their inherited hypernym in the query. These hypernyms serve as variables in the hypothesis. For the concepts "paper" e.g. we have found that "article" and "communication" or "publication" are possible concepts that can appear in any of the external ontologies.
- Step 3: Mapping agents iterate through all concepts and properties from the ontologies and create sev-

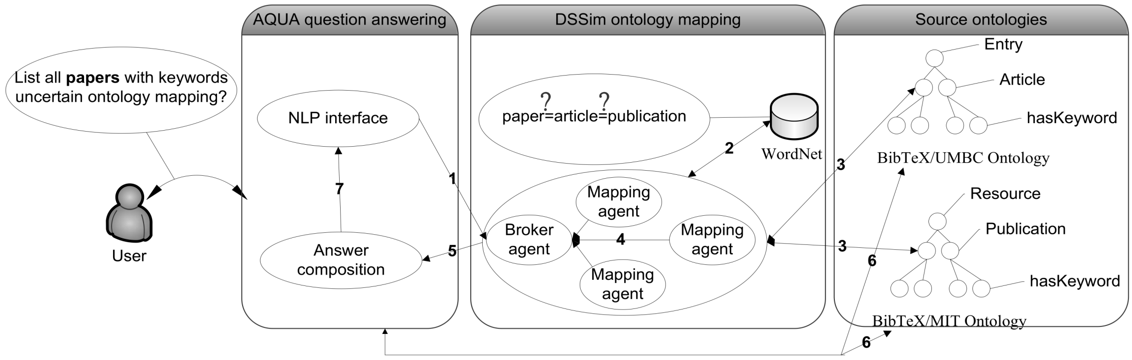

Figure 4. Overview of the mapping system

eral hypothesises that must be verified with finding evidences e.g.

$$Agent1 : H_n(\textit{mapping}) =$$
$$\textit{Query}\ \{\textit{paper,article,communation,publication}\} \Longleftrightarrow$$
$$\textit{Ontology}_{\text{MIT}}\ \{\textit{Article}\}$$
$$and$$
$$Agent2 : H_n(\textit{mapping}) =$$
$$\textit{Query}\ \{\textit{paper,article,communation,publication}\}$$
$$\Leftrightarrow\ \textit{Ontology}_{\text{UMBC}}\ \{\textit{Publication}\}$$

where Hn is the hypothesis for the mapping.
Further, we find supporting evidences for hypothesis. In this phase different syntactic and semantic similarity measures are used. These similarity measures are considered as different experts determining belief functions for the hypothesis. The last phase of this step is to combine the belief mass functions using Dempster's combination rule in order to form a coherent belief of the different experts on the hypothesises.

- Step 4: Mapping agents select the hypothesis in which they believe in most and sent it back to the broker agent. In our example the following mappings have been established:

$$Mapping_{\text{Query, MIT ontology}}(paper \leftrightarrow article)$$
$$Mapping_{\text{Query, UMBC ontology}}(paper \leftrightarrow publication)$$

- Step 5-6: The answer is composed for the user's query, which includes the relevant instances from the ontologies.

III. Uncertain reasoning and agent belief

Our proposed method works with two ontologies, which contain arbitrary number of concepts and their properties.

$$O_1 = \{C_1, .., C_n; P_1, .., P_n; I_1, .., I_n\}$$
$$O_2 = \{C_1, .., C_m; P_1, .., P_m; I_1, .., I_m\}$$

where O represents a particular ontology, C, P and I the set of concepts, properties and instances in the ontology.

In order to assess similarity we need to compare all concepts and properties from O_1 to all concepts and properties in O_2. Our similarity assessments, both syntactic and semantic produce a sparse similarity matrix where the similarity between C_n from O_1 and C_m in O_2 is represented by a particular similarity measure between the i and j elements of the matrix as follows:

$$SIM := (s_{i,j})_{n\times m}$$
$$1 \le i \le n \ and\ 1 \le j \le m$$

where SIM represents a particular similarity assessment matrix, s is a degree of similarity that has been determined by a particular similarity e.g. Jaccard or semantic similarity measure. We consider each measure as an "expert", which assess mapping precision based on its knowledge. Therefore we assume that each similarity matrix is a subjective assessment of the mapping what needs to be combined into a coherent view. If combined appropriately this combined view provides a more reliable and precise mapping than each separate mapping alone. However one similarity measure or some technique can perform particularly well for one pair of concepts or properties and particularly badly for another pair of concepts or properties, which has to be considered in any mapping algorithm.

In our ontology mapping framework each agent carries only partial knowledge of the domain and can observe it from its own perspective where available prior knowledge is generally uncertain. Our main argument is that knowledge cannot be viewed as a simple conceptualization of the world, but it has to represent some degree of interpretation. Such interpretation depends on the context of the entities involved in the process. This idea is rooted in the fact the different entities' interpretations are always subjective, since they occur according to an individual schema, which is than communicated to other individuals by a particular language. In order to represent these subjective probabilities in our system we use the Dempster-Shafer theory of evidence [19], which provides a mechanism for modelling and reasoning uncertain information in a numerical way, particularly when it is not possible to assign belief to a single element of a set of variables. Consequently the theory allows the user to represent uncertainty for knowledge representation, because the interval between support and plausibility can be easily assessed for a set of hypothesizes. Missing data (ignorance) can also be modelled by Dempster-Shafer approach and additionally evidences from two or more sources can be combined using Dempster's rule of combination. The combined support, disbelief and uncertainty can each be separately evaluated. The main advantage of the Dempster-Shafer theory is that it provides a method for combining the effect of different learned evidences to establish a new belief by using Dempster's combination rule.

The following elements have been used in our system in order to model uncertainty:

Frame of Discernment(Θ) :finite set representing the space of hypothesizes. It contains all possible mutually exclusive context events of the same kind.

$$\Theta = \{H_1, ..., H_n, ...H_N\} \tag{1}$$

In our method Θ contains all possible mappings that have been assessed by the particular expert.

Evidence:available certain fact and is usually a result of observation. Used during the reasoning process to choose the best hypothesis in Θ. We observe evidence for the mapping if the expert detects that there is a similarity between C_n from O_1 and C_m in O_2.

Belief mass function (m): is a finite amount of support assigned to the subset of Θ. It represents the strength of some evidence and

$$\sum_{A \subseteq \Theta} m_i(A) = 1 \tag{2}$$

where m_i(A) is our exact belief in a proposition represented by A that belongs to expert *i*. The similarity algorithms itself produce these assignment based on different similarity measures. As an example consider that O_1 contains the concept "paper", which needs to be mapped to a concept "hasArticle" in O_2. Based on the WordNet we identify that the concept "article" is one of the inherited hypernyms of "paper", which according to both JaroWinkler(0.91) and Jaccard(0.85) measure [20] is highly similarity to "hasArticle" in O_2. Therefore after similarity assessment our variables will have the following belief mass value:

– $m_{\text{expert1}}(O_1 \{paper, article, publication\},$
$O_2 \{hasArticle\}) = 0.85$
– $m_{\text{expert2}}(O_1 \{paper, article, publication\},$
$O_2 \{hasArticle\}) = 0.91$

In practice we assess up to 8 inherited hypernyms similarities with different algorithms (considered as experts), which can be combined based on the combination rule in order to create a more reliable mapping. Once the combined belief mass functions have been assigned the following additional measures can be derived from the available information.

Belief: amount of justified support to A that is the lower probability function of Dempster, which accounts for all evidence E_k that supports the given proposition A.

$$belief_i(A) = \sum_{E_k \subseteq A} m_i(E_k) \tag{3}$$

An important aspect of the mapping is how one can make a decision over how different similarity measures can be combined and which nodes should be retained as best possible candidates for the match. To combine the qualitative similarity measures that have been converted into belief mass functions we use the Dempster's rule of combination and we retain the node where the belief function has the highest value.

Dempster's rule of combination:Suppose we have two mass functions $m_i(E_k)$ and $m_j(E_{k'})$ and we want to combine them into a global $m_{ij}(A)$. Following Dempster's combination rule

$$m_{ij}(A) = m_i \oplus m_j = \sum_{E_k E_{k'}} m_i(E_k) * m_j(E_{k'}) \tag{4}$$

where *i* and *j* represent two different agents.

The belief combination process is computationally very expensive and from an engineering point of view, this means that it not always convenient or possible to build systems in which the belief revision process is performed globally by a single unit. Therefore, applying multi agent architecture is an alternative and distributed approach to the single one, where the belief revision process is no longer assigned to a single agent but to a group of agents, in which each single agent is able to perform belief revision and communicate with the others. Our algorithm takes all the concepts and its properties from the different external ontologies and assesses similarity with all the concepts and properties in the query graph.

A. Voting and the best possible alternative

The idea of individual voting in order to resolve conflict and choose the best option available is not rooted in computer but political science. Democratic systems are based on voting as Condorcet jury theorem [21] [22] postulates that a group of voters using majority rule is more likely to choose the right action than an arbitrary single voter is. In these situations voters have a common goal, but do not know how to obtain this goal. Voters are informed differently about the performance of alternative ways of reaching it. If each member of a jury has only partial information, the majority decision is more likely to be correct than a decision arrived at by an individual juror. Moreover, the probability of a correct decision increases with the size of the jury. But things become more complicated when information is shared before a vote is taken. People then have to evaluate the information before making a collective decision. The same ideas apply for software agents especially if they need to reach a consensus on a particular issue. In case of ontology mapping where each agent can built up beliefs over the correctness of the mappings based on partial information we believe that voting can find the socially optimal choice. Software agents can use voting to determine the best decision for agent society but in case voters make mistakes in their judgments, then the majority alternative (if it exists) is statistically most likely to be the best choice. The application of voting for software agents is a possible way to make systems more intelligent i.e. mimic the decision making how humans reach consensus decision on a problematic issue.

B. Fuzzy voting model

In ontology mapping the conflicting results of the different beliefs in similarity can be resolved if the mapping algorithm can produce an agreed solution, even though the individual opinions about the available alternatives may vary. We propose a solution for reaching this agreement by evaluating trust between established beliefs through voting, which is a general method of reconciling differences. Voting is a mechanism where the opinions from a set of votes are evaluated in order to select the alternatives that best represent the collective preferences. Unfortunately deriving binary trust like trustful or not trustful from the difference of belief functions is not so straightforward since the different voters express their opinion as subjective probability over the similarities. For a particular mapping this always involves a certain degree of vagueness hence the threshold between the trust and distrust cannot be set definitely for all cases that can occur during the process. Additionally there is no clear transition between characterising a particular belief highly or less trustful. Therefore our argument is that the trust membership or belief difference values, which are expressed by different voters can be modeled properly by using fuzzy representation. Before each agent evaluates the trust in other agent's belief over the correctness of the mapping it calculates the difference between its own and the other agent's belief. Depending on the difference it can choose the available trust levels e.g. if the difference in beliefs is 0.2 then the available trust level can be high and medium. We model these trust levels as fuzzy membership functions. In fuzzy logic the membership function $\mu(x)$ is defined on the universe of discourse U and represents a particular input value as a member of the fuzzy set i.e. $\mu(x)$ is a curve that defines how each point in the U is mapped to a membership value (or degree of membership) between 0 and 1. Our ontology mapping system models the conflict resolution as a fuzzy system where the system components are as follows:

1) Fuzzification of input and output variables: Fuzzification is the process of decomposing a system input and/or output into one or more fuzzy sets. We have experimented different types of curves namely the triangular, trapezoidal and gauss shaped membership functions. Each fuzzy set spans a region of input (or output) value graphed with the membership. Our selected membership functions overlap to allow smooth mapping of the system. The process of fuzzification allows the system inputs and outputs to be expressed in linguistic terms so that rules can be applied in a simple manner to express a complex system.

Definition 1: Belief difference is an input variable, which represents the agents own belief over the correctness of a mapping in order to establish mappings between concepts and properties in the ontology. During conflict resolution we need to be able to determine the level of difference. We propose three values for the fuzzy membership value $\mu(x) = \{small, average, large\}$

Definition 2: belief is an input variable, which described the amount of justified support to A that is the lower probability function of Dempster, which accounts for all evidence E_k that supports the given proposition A.

$$belief_i(A) = \sum_{E_k \subseteq A} m_i(E_k) \tag{5}$$

where m Demster's belief mass function represents the strength of some evidence i.e. $m(A)$ is our exact belief in a proposition represented by A. The similarity algorithms itself produce these assignment based on different similarity measures. We propose two values for the fuzzy membership value $\nu(x) = \{weak, strong\}$

Definition 3: Similarity is an input variable and is the result of some syntactic or semantic similarity measure. We propose three values for the fuzzy membership value $\xi(x) = \{low, average, high\}$

Definition 4: Low, medium and high trusts are output variables and represent the level of trust we can assign to the combination of our input variables. We propose three values for the fuzzy membership value $\tau(x) = \{low, medium, high\}$

2) Rule set: Fuzzy sets are used to quantify the information in the rule-base, and the inference mechanism operates on fuzzy sets to produce defuzzified values. Fuzzy systems map the inputs to the outputs by a set of $condition \rightarrow action$ rules i.e. rules that can be expressed in $If - Then$ form. For our conflict resolution problem we have defined four simple rules that ensure that each combination of the input variables produce output on more than one output i.e. there is always more than one initial trust level is assigned to any input variables. As an example consider a rule for cases when the trust level is defined as low:

"IF (beliefdifference IS large OR beliefdifference IS average) AND belief IS weak AND (similarity IS low OR similarity IS average) THEN trust IS low"

The rules we have initially defined are the most general ones. In our future research we intend to investigate the impact of more fine grained rules (i.e. more rules could be defined to cover overlapping areas of our fuzzy sets) on our conflict resolution.

3) Defuzzification method: After fuzzy reasoning we have the linguistic output variables, which need to be translated into a crisp (i.e. real numbers, not fuzzy sets) value. The objective is to derive a single crisp numeric value that best represents the inferred fuzzy values of the linguistic output variable. Defuzzification is such inverse transformation, which maps the output from the fuzzy domain back into the crisp domain. In our ontology mapping system we have selected the Center-of-Area (C-o-A) defuzzification method. The C-o-A method is often referred to as the Center-of-Gravity method because it computes the centroid of the composite area representing the output fuzzy term. In our system the trust levels are proportional with the area of the membership functions therefore other defuzzification methods like Center-of-Maximum (C-o-M) or Mean-of-Maximum (M-o-M) does not correspond well to our requirements.

Definition 5: For representing trust in beliefs over similarities we have defined three membership functions, $\tau(x) = \{low, average, high\}$

in the beliefs over concept and property similarities in our ontology mapping system. Our main objective is to be able to resolve conflict between two beliefs in Dempster-Shafer theory, which can be interpreted qualitatively as one source strongly supports one hypothesis and the other strongly supports another hypothesis, where the two hypotheses are not compatible. Consider for example a situation where three agents have used WordNet as background knowledge and build their beliefs considering different concepts context, which was derived from the background knowledge e.g. agent 1 used the direct hypernyms, agent 2 the sister terms and agent 3 the inherited hypernyms. Based on string similarity measures a numerical belief value is calculated, which represent a strength of the confidence that the two terms are related to each other. The scenario is depicted in Table I.

TABLE I.
BELIEF CONFLICT DETECTION

Conflict detection	Belief 1	Belief 2	Belief 3
Obvious	0.85	0.80	0.1
Difficult	0.85	0.65	0.45

The values given in Table I are demonstrative numbers just for the purpose of providing an example. In our ontology mapping framework DSSim, the similarities are considered as subjective beliefs, which is represented by belief mass functions that can be combined using the Dempster's combination rule. This subjective belief is the outcome of a similarity algorithm, which is applied by a software agent for creating mapping between two concepts in different ontologies. In our ontology mapping framework different agents assess similarities and their beliefs in the similarities need to be combined into a more coherent result. However these individual beliefs in practice are often conflicting. In this scenario applying Dempster's combination rule to conflicting beliefs can lead to an almost impossible choice because the combination rule strongly emphasizes the agreement between multiple sources and ignores all the conflicting evidence through a normalization factor. The counter-intuitive results that can occur with Dempsters rule of combination are well known and have generated a great deal of debate within the uncertainty reasoning community. Different variants of the combination rule [23] have been proposed to achieve more realistic combined belief. Instead of proposing an additional combination rule we turned our attention to the root cause of the conflict itself namely how the uncertain information was produced in our model.

The fuzzy voting model was developed by Baldwin [24] and has been used in fuzzy logic applications. However, to our knowledge it has not been introduced in the context of trust management on the Semantic Web. In this section, we will briefly introduce the fuzzy voting model theory using a simple example of 10 voters voting against or in favour of the trustfulness of an another agent's belief over the correctness of mapping. In our ontology mapping framework each mapping agent can request a number of voting agents to help assessing how trustful the other mapping agent's belief is.

According to Baldwin [24] a linguistic variable is a quintuple $(L, T(L), U, G, \mu)$ in which L is the name of the variable, $T(L)$ is the term set of labels or words (i.e. the linguistic values), U is a universe of discourse, G is a syntactic rule and μ is a semantic rule or membership function. We also assume for this work that G corresponds to a null syntactic rule so that $T(L)$ consists of a finite set of words. A formalization of the fuzzy voting model can be found in [25].

Consider the set of words { Low_trust (L_t), Medium_trust (M_t) and High_trust (H_t) } as labels of a linguistic variable trust with values in $U = [0, 1]$. Given a set "m" of voters where each voter is asked to provide the subset of words from the finite set $T(L)$, which are appropriate as labels for the value u. The membership value $\chi_{\mu_{(w)(u)}}$ is taking the proportion of voters who include u in their set of labels which is represented by w.

The main objective when resolving conflict is to have sufficient number of independent opinions that can be consolidated. To achieve our objective we need to introduce more opinions into the system i.e. we need to add the opinion of the other agents in order to vote for the best possible outcome. Therefore we assume for the purpose of our example that we have 10 voters (agents). Formally, let us define

$$V = \{A1, A2, A3, A4, A5, A6, A7, A8, A9, A10\} \quad (6)$$

$$T(L) = \{L_t, M_t, H_t\}$$

The number of voters can differ however assuming 10 voters can ensure that

1) The overlap between the membership functions can proportionally be distributed on the possible scale of the belief difference [0..1]
2) The work load of the voters does not slow the mapping process down

Let us start illustrating the previous ideas with a small example - By definition consider three linguistic output variables L representing trust levels and T(L) the set of linguistic values as $T(L) = \{Low\_trust, Medium\_trust, High\_trust\}$. The universe of discourse is U, which is defined as $U = [0, 1]$. Then, we define the fuzzy sets per output variables $\mu(Low\_trust), \mu(Medium\_trust)$ and $\mu(High\_trust)$ for the voters where each voter has different overlapping trapezoidal, triangular or gauss

membership functions. The difference in the membership functions represented by the different vertices of the membership functions, which ensures that voters can introduce different opinions as they pick the possible trust levels for the same difference in belief.
The possible set of trust levels $L = TRUST$ is defined by the Table II. Note that in the table we use a short notation L_t means Low_trust, M_t means Medium_trust and H_t means High_trust. Once the input fuzzy sets (membership functions) have been defined the system is ready to assess the output trust memberships for the input values. Both input and output variables are real numbers on the range between [0..1]. Based on the difference of beliefs represented by a real number, own belief and similarity of the different voters the system evaluates the scenario. The evaluation includes the fuzzication, which converts the crisp inputs to fuzzy sets, the inference mechanism, which uses the fuzzy rules in the rule-base to produce fuzzy conclusions (e.g. the implied fuzzy sets), and the defuzzication block, which converts these fuzzy conclusions into the crisp outputs. Therefore each input (belief difference, belief and similarity) produces a possible defuzzified output (low, medium or high trust) for the possible output variables. Each defuzzified value can be interpreted as a possible trust level where the linguistic variable with the highest defuzzified value is retained in case more than one output variable is selected. As an example consider a case where the defuzzified output for belief difference between agent 1 and agent 2 with a value 0.67 has resulted in the situation described in Table II. Note that each voter has its own membership function where the level of overlap is different for each voter. Based on a concrete input the first voting agent could map the defuzzified variables into high, medium and low trust whereas tenth voting agent to only low trust.

TABLE II.
POSSIBLE VALUES FOR THE VOTING

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
L_t	L_t	L_t	L_t	L_t	L_t	L_t	L_t	L_t	L_t
M_t	M_t	M_t	M_t	M_t	M_t				
H_t	H_t	H_t							

Note that behind each trust level there is a real number, which represents the defuzzified value. These values are used to reduce the number of possible linguistic variables in order to obtain the vote for each voting agent. Each agent retains the lingiustic variable that represents the highest value and is depicted in Table III.

TABLE III.
VOTING

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
H_t	M_t	L_t	L_t	M_t	M_t	L_t	L_t	L_t	L_t

Taken as a function of x these probabilities form probability functions. They should therefore satisfy:

$$\sum_{w \in T(L)} Pr(L = w|x) = 1 \tag{7}$$

which gives a probability distribution on words:

$$\sum Pr(L = Low\_trust|x) = 0.6 \tag{8}$$

$$\sum Pr(L = Medium\_trust|x) = 0.3 \tag{9}$$

$$\sum Pr(L = High\_trust|x) = 0.1 \tag{10}$$

As a result of voting we can conclude that given the particular difference in beliefs (represented by a real number 0.67 in this example) the combination should not consider this belief of agent 2. This is because based on its difference compared to belief of agent 1 it turns out to be a distrustful assessment. The before mentioned process of taking the ”probability distributions on words” is then repeated as many times as needed. In fact, the process is repeated as many different beliefs we have for the similarity (i.e. as many as different similarity measures exist in the ontology mapping system).

C. Possible membership functions for conflict resolution

Membership functions in fuzzy systems are subjectively specied in an ad hoc (heuristic) manner from experience or intuition. This might be possible for a real time control system, however in our case it is difficult to find intuitive choice for the membership function or the combination of the membership functions. For our conflict resolution problem we have carried out experiments in order to select the best possible membership function combination that fit well to our problem.

We have chosen the trapezodial, triangular and gauss membership function and their combinations to represent our input and output variables. For each test have generated 300 scenarios, which contain random input variables (belief difference, belief and similarity) that maps to a single trust level i.e. output variable(high, medium or low trust). In addition we have defined nine combination of membership functions that describes our input and output variables. We repeated our experiment 1000 times regenerating the 300 scenarios in each iteration.

D. Results on the use of different memberships functions

Our experiments have shown that the the fuzzy conflict resolution is really sensitive on the input membership function. The best results can be achieved using triangular membership functions. In each experiments the average wrong answers are 121 and the minimum wrong answers are 109 whereas the maximum are 134 when choosing triangular input functions. The results are promising as we are able to resolve conflict in nearly 2/3 of the cases. In practice the real improvements in the ontology mapping quality can be foreseen where the number of conflict for the candidate mapping set is high. These situations of course likely to occur where both source and target ontologies contain large number (up to 10.000) of concepts and properties. The selection of the output function does not influence the end result of the conflict resolution.

IV. CASE STUDY

Experimental comparison of ontology mapping systems is not a straightforward task as each system is usually designed to address a particular need from a specific domain. Authors have the freedom to hand pick some specific set of ontologies and demonstrate the strengths and weaknesses of their system carrying out some experiments with these ontologies. The problem is however that it is difficult to run the same experiments with another system and compare the two results. This problem has been acknowledged by the Ontology Mapping community and as a response to this need the Ontology Alignment Evaluation Initiative [7] has been set up in 2004. The evaluation was measured with recall, precision and F-Measure, which are useful measures that have a fixed range and meaningful from the mapping point of view. Recall is 100% when every relevant entity is retrieved. However it is possible to achieve 100% by simply returning every entity in the collection for every query. Therefore, recall by itself is not a good measure of the quality of a search engine. Precision is a measure of how well the engine performs in not returning non relevant documents. Precision is 100% when every entity returned to the user is relevant to the query. There is no easy way to achieve 100% precision other than in the trivial case where no document is ever returned for any query. Both precision and recall has a fixed range: 0.0 to 1.0 (or 0% to 100%). A good mapping algorithm must have a high recall to be acceptable for most applications. The most important factor in building better mapping algorithms is to increase precision without worsening the recall. In order to compare our system with other solutions we have participated in the OAEI competitions since 2006. Each year we have been involved in more tracks than the previous year. This gave us the possibility to test our mapping system on different domains including medical, agriculture, scientific publications, web directories, food and agricultural products and multimedia descriptions. The experiments were carried out to assess the efficiency of the mapping algorithms themselves. The experiments of the question answering (AQUA) using our mappings algorithms are out of the scope of this paper. Our main objective was to compare our system and algorithms to existing approaches on the same basis and to allow drawing constructive conclusions.

A. Benchmarks

The OAEI benchmark contains tests, which were systematically generated starting from some reference ontology and discarding a number of information in order to evaluate how the algorithm behave when this information is lacking. The bibliographic reference ontology (different classifications of publications) contained 33 named classes, 24 object properties, 40 data properties. Further each generated ontology was aligned with the reference ontology. The benchmark tests were created and grouped by the following criteria:

- Group 1xx: simple tests such as comparing the reference ontology with itself, with another not related (food domain) ontology or the same ontology in its restriction to OWL-Lite
- Group 2xx: systematic tests that were obtained by discarding some features from some reference ontology e.g. name of entities replaced by random strings, synonyms, name with different conventions, strings in another language than English, comments that can be suppressed or translated in another language, hierarchy that can be suppressed, expanded or flattened.
- Group 3xx: four real-life ontologies of bibliographic references that were found on the web e.g. BibTeX/MIT, BibTeX/UMBC

Figure 5 shows the 6 best performing systems out of 13 participants. We have ordered the systems based on the their the F-Value of the H-means because the H-mean unifies all results for the test and F-Value represents both precision and recall.

In the benchmark test we have performed in the upper mid range compared to other systems. Depending on the group of tests our system compares differently to other solutions:

- Group 1xx: Our results are nearly identical to the other systems.
- Group 2xx: For the tests where syntactic similarity can determine the mapping outcome our system is comparable to other systems. However where semantic similarity is the only way to provide mappings our systems provides less mappings compared to the other systems in the best six.
- Group 3xx: Considering the F-value for this group only 3 systems SAMBO, RIMOM and Lily are ahead.

The weakness of our system to provide good mappings when only semantic similarity can be exploited is the direct consequence of our mapping architecture. At the moment we are using four mapping agents where 3 carries our syntactic similarity comparisons and only 1 is specialised in semantics. However it is worth to note that our approach seems to be stable compared to our last years performance, as our precision recall values were similar in spite of the fact that more and more difficult tests have been introduced in 2008. As our architecture is easily expandable with adding more mapping agents it is possible to enhance our semantic mapping performance in the future.

B. Directory

The purpose of this track was to evaluate performance of existing alignment tools in real world taxonomy inte-

[7] http://oaei.ontologymatching.org/

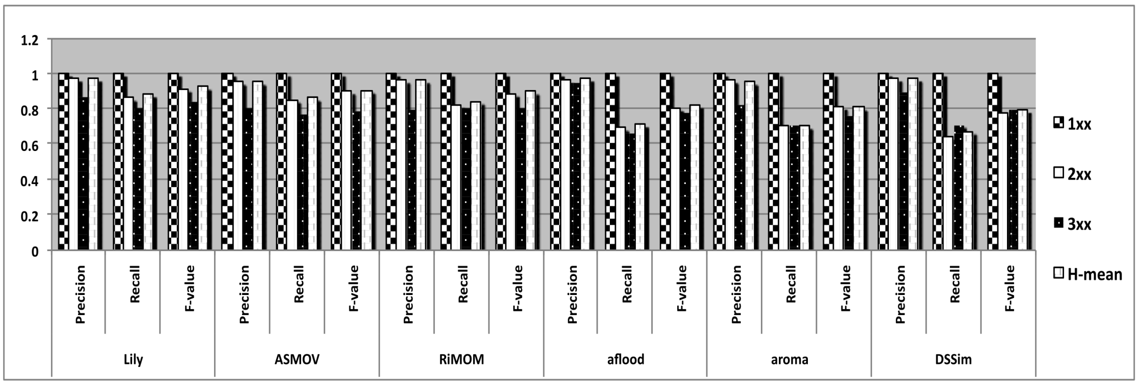

Figure 5. Best performing systems in the benchmarks track based on H-mean and F-value

gration scenario. Our aim is to show whether ontology alignment tools can effectively be applied to integration of “shallow ontologies”. The evaluation dataset was extracted from Google, Yahoo and Looksmart web directories. The specific characteristics of the dataset are:

- More than 4500 of node matching tasks, where each node matching task is composed from the paths to root of the nodes in the web directories. Expert mappings for all the matching tasks.
- Simple relationships: Basically web directories contain only one type of relationship so called ”classification relation”.
- Vague terminology and modelling principles: The matching tasks incorporate the typical ”real world” modelling and terminological errors.

These node matching tasks were represented by pairs of OWL ontologies, where classification relation is modelled as OWL *subClassOf* construct. Therefore all OWL ontologies are taxonomies (i.e. they contain only classes (without Object and Data properties) connected with subclass relation.

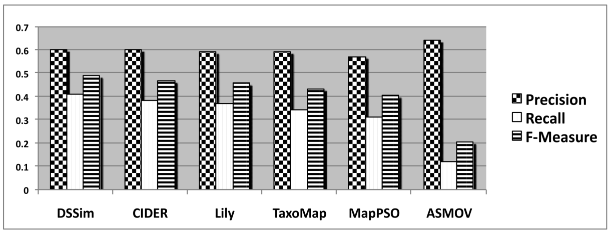

Figure 6. All participating systems in the directories track ordered by F-value

In the library track only 6 systems have participated in 2008. In terms of F-value DSSim has performed the best however the difference is marginal compared to the CIDER [26] or Lily systems. The concepts in the directory ontologies mostly can mostly be characterised as compound nouns e.g. ”News_and_Media” and we need to process(split) them properly before consulting background knowledge in order to provide better mappings in the future.

C. Library

The objective of this track was to align two Dutch thesauri used to index books from two collections held by the National Library of the Netherlands. Each collection is described according to its own indexing system and conceptual vocabulary. On the one hand, the Scientific Collection is described using the GTT, a huge vocabulary containing 35.000 general concepts ranging from “Wolkenkrabbers (Sky-scrapers)” to “Verzorging (Care)”. On the other hand, the books contained in the Deposit Collection are mainly indexed against the Brinkman thesaurus, containing a large set of headings (more than 5.000) that are expected to serve as global subjects of books. Both thesauri have similar coverage (there are more than 2.000 concepts having exactly the same label) but differ in granularity. For each concept, the thesauri provide the usual lexical and semantic information: preferred labels, synonyms and notes, broader and related concepts, etc. The language of both thesauri is Dutch, but a quite substantial part of Brinkman concepts (around 60%) come with English labels. For the purpose of the alignment, the two thesauri have been represented according to the SKOS model, which provides with all these features.

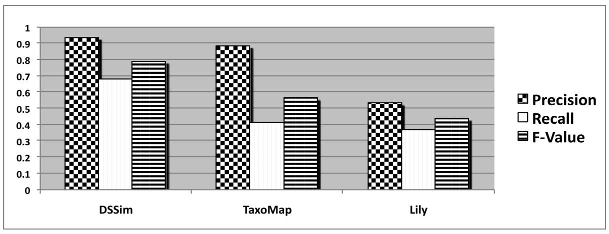

Figure 7. All participating systems in the library track ordered by F-value

In the library track DSSim has performed the best out of the 3 participating systems. The track is difficult partly because of its relative large size and because of its multilingual representation. However these ontologies contain related and broader terms therefore the mapping can be carried out without consulting multi lingual background knowledge. This year the organisers have provided instances as separate ontology as well however we did not make use of it for creating our final mappings. For further improvements in recall and precision we will need to consider these additional instances in the future.

V. Strengths and weaknesses of our solution

Based on the OAEI experiments, we can conclude that our solution compares and scales well to other well established ontology mapping systems. Nevertheless it is clear (OAEI seems to share our opinion) that it is not possible to clearly define a "winner" on these yearly competitions. Each system has its strengths and weaknesses and they tend to perform differently on different domains. However we can define some criteria to determine where we perform well and on which areas do we need to make further progress.

1) Domain independence: This is a definite strength of our system. Our solution does not rely on predefined thresholds or parameters that needs to be changed from domain to domain. Several mapping systems utilise machine learning in order to determine these parameters however these solutions are likely to be dependent on the training set. DSSim uses WordNet as the background knowledge. This ensures that we can provide equivalent mappings on different domains. Nevertheless domain specific background knowledge can influence the results positively. The anatomy track has proved that systems that use domain specific background knowledge are far superior compared to the systems with general background knowledge. Nevertheless the drawback of these systems is that they cannot produce equally good results once the domain is changing. For example the AOAS system [27] performed the best on the anatomy track on the OAEI 2007 but they did not produce result in any other track as their system was fine tuned for the medical domain.
2) Conflict management: This area needs to be improved in our system. DSSim do manage conflicting beliefs over a particular mapping, which can occur when different agents have built up conflicting beliefs for the correctness of a mapping candidate. The problem occurs when we have already selected a mapping candidate and later on in the mapping process we add an another mapping that contradicts the previous one. Systems e.g. ASMOV, which try to detect conflicting mappings in the result-set can provide better overall results compared to our solution.
3) Mapping quality: DSSim does not produce always the best precision and recall for each track however our mapping quality is stable throughout different domains. We consider this as a strength of our system because we foresee different application domains where our solution can be used. In this context it is more important that we can produce equally good enough mappings.
4) Mapping system scalability: Due to our multi-agent architecture our solution scales well with medium and large domains alike. For example in the OAEI 2008 the largest ontologies were in the Very Large Cross-Lingual Resources track. DSSim was the only system that has participated in this track. Our solution can scale well for large domains because as the domain increases we can distribute the problem space between an increasing number of agents. Additionally our solution fits well to current hardware development trends, which predicts an increasing number of processor core in order to increase the computing power.
5) Traceability of the reasoning: Unfortunately this is a weakness of our system as we cannot guarantee that running the algorithm twice on the same domain we will always get exactly the same results. The reason is that our belief conflict resolution approach [28] uses fuzzy voting for resolving belief conflicts which can vary from case to case. Additionally beliefs are based on similarities between a set of source and target variables. The set of variables are deducted from the background knowledge, which can differ depending on the actual context of our query. Therefore it is not feasible to trace exactly why a particular mapping has been selected as good mapping compared to another candidate mappings.

VI. Related work

Several ontology mapping systems have been proposed to address the semantic data integration problem of different domains independently. In this paper we consider only those systems, which have participated in the OAEI (Ontology Alignment Evaluation Initiative) competitions and has been participated more than two tracks. There are other proposed systems as well however as the experimental comparison cannot be achieved we do not include them in the scope of our analysis. Lily [29] is an ontology mapping system with different purpose ranging from generic ontology matching to mapping debugging. It uses different syntactic and semantic similarity measures and combines them with the experiential weights. Further it applies similarity propagation matcher with strong propagation condition and the matching algorithm utilises the results of literal matching to produce more alignments. In order to assess when to use similarity propagation Lily uses different strategies, which prevents the algorithm from producing more incorrect alignments. ASMOV [30] has been proposed as a domain specific mapping tool in order to facilitate the integration of heterogeneous systems, using their data source ontologies. It uses different matchers and generates similarity matrices between concepts, properties, and individuals, including mappings from object properties to datatype properties. It does not combine the similarities but uses the best values to create a pre-alignment, which are then being semantically re-validated by the system. Mappings, which pass the semantic validation will be added to the final alignment. ASMOV can use different background knowledge e.g. Wordnet or UMLS Metathesaurus(medical background knowledge) for the assessment of the similarity measures. RiMOM [31] is an automatic ontology mapping system, which models the ontology mapping problem as making decisions over entities with minimal risk. It uses the Bayesian theory to model decision making under uncertainty where observations are all entities in the two ontologies. Further it implements different matching strategies where each defined strategy is based on one kind of ontological information. RiMOM includes different methods for choosing appropriate strategies (or strategy combination) according to the available information in the ontologies. The strategy combination is conducted by a linear-interpolation method. In addition to the different strategies RiMOM uses similarity propagation process to refine the existing alignments and to find new alignments that cannot be found using other strategies. RiMOM is the only system other than DSSim in the OAEI contest that considers the uncertain nature of the mapping process however it models uncertainty differently from DSSim. RiMOM appeared for first time in the OAEI-2007 whilst DSSim appeared in the OAEI-2006. MapPSO [32] is a research prototype, which has been designed to address the need for highly scalable, massively parallel tool for both large scale and numerous ontology alignments. MapPSO method models the ontology alignment problem as an optimisation problem. It employs a population based optimisation paradigm based on social interaction between swarming animals, which provides the best answer being available at that time. Therefore it is especially suitable for providing answers under time constraint like the ontology mapping. MapPSO employs different syntactic and semantic similarity measures and combines the available base distances by applying the Ordered Weighted Average(OWA) [33] aggregation of the base distances. It aggregates the components by ordering the base distances and applying a fixed weight vector. The motivation of the MapPSO system is identical with one of the motivations of the DSSim namely to address the need of scalable mapping solutions for large scale ontologies. Surprisingly MapPSO did not participate in the Very Large Cross Lingual Resources track (especially designed for large scale thesauri) therefore experimental comparison cannot be achieved from this point of view. TaxoMap [34] is an alignment tool, which aims is to discover rich correspondences between concepts with performing oriented alignment from a source to a target ontology taking into account labels and sub-class descriptions. It uses a part-of-speech [35] and lemma information, which enables to take into account the language, lemma and an use word categories in an efficient way. TaxoMap performs a linguistic similarity measure between labels and description of concepts and it has been designed to process large scale ontologies by using partitioning techniques. TaxoMap however does not process instances, which can be a drawback in several situations. SAMBO and SAMBOdtf [36] is a general framework for ontology matching. The methods and techniques used in the framework are general and applicable to different areas nevertheless SAMBO has been designed to align biomedical ontologies. Their algorithm includes one or several matchers, which calculate similarity values between the terms from the different source ontologies. These similarities are then filtered and combined as a weighted sum of the similarity values computed by different matchers.

VII. Conclusions

This paper presented the main challenges for an alignment system in the context of question-answering. The challenges related to the data or information representation, quality and volume are addressed with introducing uncertain reasoning and representation when the available information is interpreted by our system. Our approach tries to establish an interpretation of the available information and avoids the usage of heuristics or any domain specific rules. To achieve this interpretation we have utilised Dempster-Shafer theory for managing the reasoning with vague information and have introduced fuzzy voting model for resolving conflicts during the interpretation of the Semantic Web data. Concerning the challenges related to the nature of the systems from the generic and intelligence point of view our proposed architecture is conceived to be able to exhibit a kind of "machine intelligence" through the multi-agent architecture, which is a form of collective intelligence that can emerge from the collaboration and competition of many software agents. Further we have also introduced DSSim and our performance in the benchmarks, directory and library tracks of the OAEI-2008 evaluation. The performance of our DSSim was the best among participants in the library track in 2008. Our system is conceived to be a generic mapping tool and such the performance still varies slightly across different domains. These variations and comparisons with other mappings systems are accessible from the OAEI workshop proceedings. Our participation in the Ontology Alignment Evaluation Initiative was an excellent opportunity to test and compare our system with other solutions and helped a great deal in identifying the future possibilities that needs to be investigated further.

References

[1] P. Shvaiko and J. Euzenat, "Ten challenges for ontology matching," University of Trento, Tech. Rep. DISI-08-042, 2008.

[2] J. Euzenat and P. Shvaiko, *Ontology matching*. Heidelberg (DE): Springer-Verlag, 2007.

[3] D. Beckett, "Rdf/xml syntax specification." [Online]. Available: http://www.w3.org/RDF/

[4] D. L. McGuinness and F. van Harmelen, "Owl web ontology language." [Online]. Available: http://www.w3.org/TR/owl-features/

[5] A. Miles and S. Bechhofer, "Skos simple knowledge organization system." [Online]. Available: http://www.w3.org/TR/skos-reference/

[6] M. Lenzerini, D. Milano, and A. Poggi, "Ontology representation & reasoning," Universit di Roma La Sapienza, Roma, Italy, Tech. Rep. NoE InterOp (IST-508011), 2004.

[7] R. Y. Wang, H. B. Kon, and S. E. Madnick, "Data quality requirements analysis and modeling," in *Proceedings of the Ninth International Conference on Data Engineering*, 1993, pp. 670–677.

[8] Y. Wand and R. Y. Wang, "Anchoring data quality dimensions in ontological foundations," *Communications of the ACM*, pp. 86–95, 1996.

[9] C. Batini, M. Lenzerini, and S. B. Navathe, "A comparative analysis of methodologies for database schema integration," *ACM Computing Surveys*, vol. 18, no. 4, pp. 323–364, 1986.

[10] W. Carlo, D. Tharam, R. Wenny, and C. Elizabeth, "Large scale ontology visualisation using ontology extraction," *International Journal of Web and Grid Services*, vol. 1, no. 1, pp. 113–135, 2005.

[11] A. Flahive, B. O. Apduhan, J. W. Rahayu, and D. Taniar, "Large scale ontology tailoring and simulation in the semantic grid environment," *International Journal of Metadata, Semantics and Ontologies*, vol. 1, no. 4, pp. 265–281, 2006.

[12] C. Caracciolo, J. Euzenat, L. Hollink, R. Ichise, A. Isaac, V. Malaisé, C. Meilicke, J. Pane, P. Shvaiko, H. Stuckenschmidt, O. Šváb-Zamazal, and V. Svátek, "First results of the ontology alignment evaluation initiative 2008," in *The 3rd International Workshop on Ontology Matching*, 2008.

[13] H. Brugman, V. Malaisé, and L. Gazendam, "A web based general thesaurus browser to support indexing of television and radio programs," in *the 5th international conference on Language Resources and Evaluation (LREC 2006)*, 2006.

[14] L. Ding, T. Finin, A. Joshi, R. Pan, R. S. Cost, Y. Peng, P. Reddivari, V. C. Doshi, and J. Sachs, "Swoogle: A search and metadata engine for the semantic web," in *Proceedings of the Thirteenth ACM Conference on Information and Knowledge Management*, 2004.

[15] S. M. Falconer and M.-A. D. Storey, "A cognitive support framework for ontology mapping," in *Proceedings of 6th International Semantic Web Conference (ISWC2007)*, 2007, pp. 114–127.

[16] M. Nagy and M. Vargas-Vera, "Reasoning representation and visualisation framework for ontology mapping using 3d modeling," in *proceedings of the 4th International Conference on Interdisciplinarity in Engineering (INTER-ENG 2009)*, 2009.

[17] M. Nagy, M. Vargas-Vera, and E. Motta, "Multi agent ontology mapping framework in the aqua question answering system," in *MICAI 2005: Advances in Artificial Intelligence, 4th Mexican International Conference on Artificial Intelligence*, 2005, pp. 70–79.

[18] M. Vargas-Vera and E. Motta, "Aqua - ontology-based question answering system," in *Third International Mexican Conference on Artificial Intelligence (MICAI-2004)*, 2004.

[19] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[20] W. W. Cohen, P. Ravikumar, and S. E. Fienberg, "A comparison of string distance metrics for name-matching tasks," in *Proceedings of Information Integration on the Web (IIWeb 2003)*, 2003.

[21] D. Austen-Smith and J. S. Banks, "Information aggregation, rationality, and the condorcet jury theorem," *The American Political Science Review*, vol. 90, no. 1, pp. 34–45, 1996.

[22] H. P. Young, "Condorcet's theory of voting," *The American Political Science Review*, vol. 82, no. 4, pp. 1231–1244, 1988.

[23] K. Sentz and S. Ferson, "Combination of evidence in dempster-shafer theory," Systems Science and Industrial Engineering Department, Binghamton University, Tech. Rep., 2002.

[24] J. F. Baldwin, *Mass assignment Fundamentals for computing with words*, ser. Selected and Invited Papers from the Workshop on Fuzzy Logic in Artificial Intelligence, Lecture Notes In Computer Science. Springer-Verlag, 1999, vol. 1566, pp. 22–44.

[25] J. Lawry, "A voting mechanism for fuzzy logic," *International Journal of Approximate Reasoning*, vol. 19, pp. 315–333, 1998.

[26] J. Gracia and E. Mena, "Ontology matching with cider:evaluation report for the oaei 2008," in *The 3rd International Workshop on Ontology Matching*, 2008.

[27] S. Zhang and O. Bodenreide, "Hybrid alignment strategy for anatomical ontologies results of the 2007 ontology alignment contest," in *The Second International Workshop on Ontology Matching*, 2007.

[28] M. Nagy, M. Vargas-Vera, and E. Motta, "Introducing fuzzy trust for managing belief conflict over semantic web data," in *Fourth International Workshop on Uncertain Reasoning on the Semantic Web*, 2008.

[29] P. Wang and B. Xu, "Lily: Ontology alignment results for oaei 2008," in *The 3rd International Workshop on Ontology Matching*, 2008.

[30] Y. R. Jean-Mary and M. R. Kabuka, "Asmov: Ontology alignment with semantic validation," in *Joint SWDB-ODBIS Workshop*, 2007.

[31] J. Tang, J. Li, B. Liang, X. Huang, Y. Li, and K. Wang, "Using bayesian decision for ontology mapping," *Web Semantics*, 2006.

[32] J. Bock and J. Hettenhausen, "Mappso results for oaei 2008," in *The 3rd International Workshop on Ontology Matching*, 2008.

[33] Q. Ji, P. Haase, and G. Qi, "Combination of similarity measures in ontology matching by owa operator," in *The 12th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'08)*, 2008.

[34] F. Hamdi, H. Zargayouna, B. Safar, and C. Reynaud, "Taxomap in the oaei 2008 alignment contest," in *The 3rd International Workshop on Ontology Matching*, 2008.

[35] H. Schmid, "Probabilistic part-of-speech tagging using decision trees," in *International Conference on New Methods in Language Processing*, 1994.

[36] P. Lambrix and H. Tan, "Sambo - a system for aligning and merging biomedical ontologies," *Journal of Web Semantics, Special issue on Semantic Web for the Life Sciences*, 2006.

Maria Vargas-Vera received her PhD from the Artificial Intelligence Department at Edinburgh University in 1995 and she was awarded Fellow of the British Computer Society (FBCS) from November 2009. Dr. Vargas-Vera is a Lecturer in Computing at the Open University, England UK. Her current research focuses on Automatic Construction of Ontologies from Text, Ontology Mapping and E-Learning Applications using Semantic Web Technologies. Dr Vargas-Vera has published many research papers in international conferences and journals and she is also a member of program committees of international conferences and workshops. Dr Vargas-Vera is an Associated Editor of the Revista Iberoamericana Computacion y Sistemas, the International Journal of Knowledge and Learning (IJKL) and the Journal of Emerging Technologies in Web Intelligence (JETWI).

Miklos Nagy is a Ph.D. candidate at the Open University's Knowledge Media Institute. His research interests are Uncertain Reasoning, Ontology Mapping, Multi-agent systems and information integration using Semantic Web technologies. His current research focuses on the development of intelligent multi-agent systems that can exploit the emerging Semantic Web's large-scale data. Miklos Nagy received his MSc in Information Engineering from the University of Miskolc, Hungary.