

# DSSim Results for OAEI 2009

Miklos Nagy<sup>1</sup>, Maria Vargas-Vera<sup>1</sup>, and Piotr Stolarski<sup>2</sup>

<sup>1</sup> The Open University

Walton Hall, Milton Keynes, MK7 6AA, United Kingdom  
m.nagy@open.ac.uk; m.vargas-vera@open.ac.uk

<sup>2</sup> Poznan University of Economics

al. Niepodleglosci 10, 60-967 Poznan, Poland  
P.Stolarski@kie.ae.poznan.pl

**Abstract.** The growing importance of ontology mapping on the Semantic Web has highlighted the need to manage the uncertain nature of interpreting semantic meta data represented by heterogeneous ontologies. Considering this uncertainty one can potentially improve the ontology mapping precision, which can lead to better acceptance of systems that operate in this environment. Further the application of different techniques like computational linguistics or belief conflict resolution that can contribute the development of better mapping algorithms are required in order to process the incomplete and inconsistent information used and produced during any mapping algorithm. In this paper we introduce our system called “DSSim” and describe the improvements that we have made compared to OAEI 2006, OAEI 2007 and OAEI 2008.

## 1 Presentation of the system

### 1.1 State, purpose, general statement

Ontology mapping systems need to interpret heterogeneous data in order to simulate “machine intelligence”, which is a driving force behind the Semantic Web. This implies that computer programs can achieve a certain degree of understanding of such data and use it to reason about a user specific task like question answering or data integration. In practice there are several roadblocks[1] that hamper the development of mapping solutions that perform equally well for different domains. Additionally the different combination of these challenges needs to be addressed in order to design systems that provides good quality results. Since DSSim has been originally designed in 2005 it has progressively evolved in order to address the combination of the 5 following challenges:

- Representation and interpretation problems: Ontology designers have a wide variety of languages and language variants to choose from in order to represent their domain knowledge. From the logical representation point of view each representations are valid separately and no logical reasoner would find inconsistency in them individually. However the problem occurs once we need to compare ontologies with different representations in order to determine the similarities between classes and individuals. Consider for example one ontology where the labels are described with standard class *rdfs:label* tag and an another ontology where the same is described

as *hasNameScientific* data property. 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.

- Quality of the Semantic Web data: 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. 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.
- Efficient mapping with large scale ontologies: Ontologies can get quite complex and very large, causing difficulties in using them for any application. 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 in scale. 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. As a consequence applications need to scale well in case huge ontologies need to be processed.
- 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 aims 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 the building generic systems that perform equally well on different domains and provide acceptable results is a considerable challenge for the future research.

- 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 [2] 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 unexploitable 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 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[3] 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.

As a result from the mapping point of view ontologies will always contain inconsistencies, missing or overlapping elements and different conceptualisation of the same terms, which introduces a considerable amount of uncertainty into the mapping process. In order to represent and reason with this uncertainty authors (Vargas-Vera and Nagy) have proposed a multi agent ontology mapping framework [4], which uses the Dempster-Shafer [5] theory in the context of Question Answering. Since our first proposition[6] of such solution in 2005 we have gradually developed and investigated multiple components of such system and participated in the OAEI in order to validate the feasibility of our proposed solution. Fortunately during the recent years our original concept has received attention from other researchers [7, 8], which helps to broaden the general knowledge on this area. We have investigated different aspects of our original idea namely the feasibility of belief combination[9] and the resolution of conflicting beliefs [10] over the belief in the correctness of similarities using the fuzzy voting model. A comprehensive description of the Fuzzy voting model can be found [10]. For this contest (OAEI 2009) the benchmarks, anatomy, directory, iimb, vlcr , Eprints-Rexa-

Sweto/DBLP benchmark and conference tracks had been tested with this new version of DSSim (v0.4).

## 1.2 Specific techniques used

This year within the tasks preparing the results for conference track we focused mainly on improvements and fine-tuning the algorithms for obtaining better effects in terms of both precision and recall. Moreover in order to conform to the extended terms of the track - we additionally implemented a simple enhancement for supplying subsumption correspondences as the DSSim system allowed only detection of equivalence between ontological entities. Below we will cover both types of changes more thoroughly. The first type of mentioned changes concentrates on improvements made to the compound nouns comparison method introduced in the last year's version of the system. The presented compound nouns comparison method deals with interpretation of compound nouns based on earlier works done in - among others - language understanding as well as question-answering and machine translation. The essence of the method focuses on establishing the semantic relations between items of compound nouns. During the development we reviewed some of the most interesting approaches [11] [12] [13]. Although all of them should be regarded as partial solutions, they manifest a good starting point for our experiments. Most of the cases uses either manually created rules [11] or machine learning techniques [12] in order to automatically build classification rules that will enable to rate any given compound noun phrase into one of a set of pre-selected semantic relations which best reflects the sense and nature of that phrase. We extended the initial set of simple rules by additional ones. We also made the rule engine more flexible so as it the semantic relation categories can now be assessed not only on the basis of comments or labels but also their id names. This last option is useful in some cases identified earlier in the analysis stage of the last year's results. Finally we extended also the set of semantic relation categories itself by another few categories. The compound nouns semantic relation detection algorithm is used in DSSim system as a determiner of such relations within ontology entities' identifiers, labels or comments. After the relation  $r^{1,n}$  has been classified independently for entities in the first of aligned ontologies  $O^1$  and  $r^{2,m}$  separately for entities from the other ontology  $O^2$ , the alignments may be produced between the entities from  $O^1$  and  $O^2$  on the basis of similarity between the relations  $r^{1,n}$  and  $r^{2,m}$  itself. In order to eliminate the drawbacks of this approach the algorithm is viewed as a helper rather than independent factor of alignment establishment process. Nevertheless, because of the superb, multi-criterion architecture of the DSSim [14] such approach to the algorithm fits especially well allowing easy integration. As the number of elements in the set of isolated semantic relations is usually limited only to very general ones, the probability of detecting the same or similar relations is subjectively high, therefore the method itself is rather sensitive to the size of the set. Thus this year innovations concentrated on extending the rules and supplying another important categories. Moving on to another type of changes, we called the subsumption detection facility a simple one as it in fact does not alter the DSSim system algorithms to cover other types of correspondences. On the contrast the facility in this year's shape uses the results of the algorithm itself to post-produce the possible weaker (non-equivalent) correspondences basing on the algorithm result set. In order

to achieve that we implemented a straightforward inference rules over the taxonomical trees of matched ontologies. We hope to move the function to the main algorithm in the future as the simple approach introduces a number of limitations.

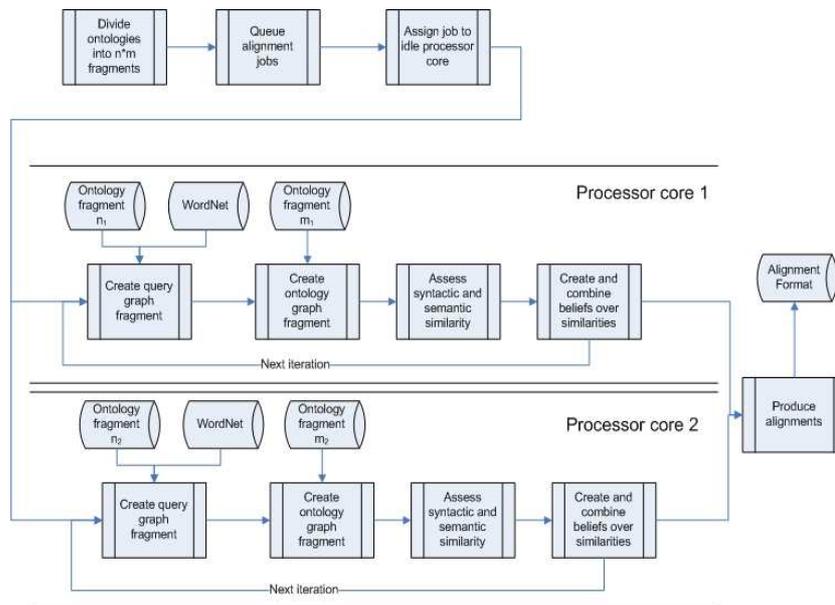
To sum up the introduced improvements, we made selected and subtle yet important alterations of the system. The modifications of last year proved to be useful and supplied promising results thus our intention is to build on the top of this achievements rather than starting completely different ideas. The changes introduced for this year's version of the system were backed up by the thorough interpretation and in-depth analysis of OAEI 2008 [14] outcomes.

### 1.3 Adaptations made for the evaluation

Our ontology mapping system is based on a multi agent architecture where each agent built up a belief for the correctness of a particular mapping hypothesis. Their beliefs are then combined into a more coherent view in order to provide better mappings. Although for the previous OAEI contests we have re-implemented our similarity algorithm as a standalone mapping process which integrates with the alignment api, we have recognised the need for possible parallel processing for tracks which contain large ontologies e.g. very large cross-lingual resources track. This need is indeed coincide with our original idea of using distributed multi-agent architecture, which is required for scalability purposes once the size of the ontology is increasing. Our modified mapping process can utilise multi core processors by splitting up the large ontologies into smaller fragments. Both the fragment size and the number of cores that should be used for processing can be set in the "param.xml" file. Based on the previous implementation we have modified our process for the OAEI 2009 which works as follows:

1. Based on the initial parameters divide the large ontologies into  $n*m$  fragments.
2. Parse the ontology fragments and submit them into the alignment job queue.
3. Run the job scheduler as long as we have jobs in the queue and assign jobs into idle processor cores.
  - 3.1 We take a concept or property from ontology 1 and consider (refer to it from now) it as the query fragment that would normally be posed by a user. Our algorithm consults WordNet in order to augment the query concepts and properties with their hypernyms.
  - 3.2 We take syntactically similar concepts and properties to the query graph from ontology 2 and build a local ontology graph that contains both concepts and properties together with the close context of the local ontology fragments.
  - 3.3 Different similarity and semantic similarity algorithms (considered as different experts in evidence theory) are used to assess quantitative similarity values (converted into belief mass function) between the nodes of the query and ontology fragment which is considered as an uncertain and subjective assessment.
  - 3.4 Then the similarity matrixes are used to determine belief mass functions which are combined using the Dempster's rule of combination. Based on the combined evidences we select those mappings in which we calculate the highest belief function.
4. The selected mappings are added into the alignment.

The overview of the mapping process is depicted on figure 1.



**Fig. 1.** The mapping process on a dual-core processor

#### 1.4 Link to the system and parameters file

<http://kmi.open.ac.uk/people/miklos/OAEI2009/tools/DSSim.zip>

#### 1.5 Link to the set of provided alignments (in align format)

<http://kmi.open.ac.uk/people/miklos/OAEI2009/results/DSSim.zip>

## 2 Results

### 2.1 benchmark

Our algorithm has produced the same results as last year. 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 this year. As our architecture is easily expandable with adding more mapping agents it is possible to enhance our semantic mapping performance in the future. The overall conclusion is that our system produces stable quality mappings, which is good however we still see

room for improvements. Based on the 2009 results the average precision(0.97) cannot be improved significantly however considerable improvements can be made from the recall(0.66) point of view. According to the benchmarks tests our system need to be improved for cases, which contain systematic: scrambled labels + no comments + no hierarchy and systematic: scrambled labels + no comments + expanded hierarchy + no instance.

## **2.2 anatomy**

The anatomy track contains two reasonable sized real world ontologies. Both the Adult Mouse Anatomy (2.744 classes) and the NCI Thesaurus (3.304 classes) describes anatomical concepts. The classes are represented with standard owl:Class tags with proper rdfs:label tags. Our mapping algorithm has used the labels to establish syntactic similarity and has used the rdfs:subClassOf tags to establish semantic similarities between class hierarchies. We could not make use of the owl:Restriction and oboInOwl:has-RelatedSynonym tags as this would require ontology specific additions. The anatomy track represented a number of challenges for our system. Firstly the real word medical ontologies contain classes like “outer renal medulla peritubular capillary”, which cannot be easily interpreted without domain specific background knowledge. Secondly one ontology describes humans and the second describes mice. To find semantically correct mappings between them requires deep understanding of the domain. The run time per test was around 10 min, which is an improvement compared to last year. Further we have realised significant improvement both in terms of precision and recall compared to the last year’s results. Our system ranks in the middle positions out of 10 participating systems.

## **2.3 Eprints-Rexa-Sweto/DBLP benchmark**

This track has posed serious challenge for our system. SwetoDblp is a large-size ontology containing bibliography data of Computer Science publications where the main data source is DBLP. It contains around 1.5 million terms including 560.792 persons, 561.895 Articles in Proceedings. The eprints and rexa ontologies were large but manageable from our system’s perspective. Based on the preliminary results our system did not perform well in terms of precision and recall. The reasons needs to be investigated further. The run time including the SweetoDblp ontology was over 1 week. In spite of the fact that it was a new and difficult track this year we were disappointed with our overall results. The performance can be due to the fact that our system was originally conceived as mapping system that does not use extensively instances for establishing the mapping. As a result where only instances are present out system does not perform as well as in the other tracks.

## **2.4 directory**

The directory test as well has been manageable in terms of execution time. In general the large number of small-scale ontologies made it possible to verify some mappings for

some cases. The tests contain only classes without any labels but in some cases different classes have been combined into one class e.g. “News\_and\_Media” that introduces certain level of complexity for determining synonyms using any background knowledge. To address these difficulties we have used a compound noun algorithms described in section 1.2. The execution time was around 15 minutes. In this track our performance was stable compared to the results in 2008. In terms of precision our system compares well to the other participating systems however improvements can be made from the recall point of view.

## 2.5 IIMB

This track contains generated benchmarks constituted using one dataset and modifying it according to various criterias. The main directory contains 37 classes and about 200 different instances. Each class contains a modified sub directory and the corresponding mapping with the instances. The different modifications introduced to the original ontology included identical copy of the original sub classes where the instance IDs are randomly changed, value transformations, structural transformations, logical transformations and several combinations of the previous transformations. The IIMB track was well manageable in terms of run time as it took under 10 minutes to run the 37 different tests. Similarly to the the task (on instance matching) described in section 2.3 our system under performed on the IIMB track. The reason for this can be attributed to the same reasons described in the E-prints-Rexa-Sweto/DBLP section.

## 2.6 vlc

This vlc track contains 3 large ontologies. The GTAA thesaurus is a Dutch public audiovisual broadcasts archive, for indexing their documents, contains around 3.800 subject keywords, 97.000 persons, 27.000 names and 14.000 locations. The DBPedia is an extremely rich dataset. It contains 2.18 million resources or “things”, each tied to an article in the English language Wikipedia. The “things” are described by titles and abstracts in English and often also in Dutch. We have converted the original format into standard SKOS in order to use it in our system. However we have converted only the labels in English and in Dutch whenever it was available. The third resource was the WordNet 2.0 in SKOS format where the synsets are instances rather than classes. In our system the WordNet 3.0 is included into as background knowledge therefore we have converted the original noun-synsets into a standard SKOS format and used our WordNet 3.0 as background knowledge. The run time of the track was over 1 week. Fortunately this year an other system also participated in this track therefore we can establish a qualitative comparison. In terms of precision our system performs well (name-dblp, subject-wn, location-wn, name-wn) however in certain tasks like location-dblp, person-dblp our system performs slightly worst compared to the other participating system. In terms of recall our system does not perform as well as we have expected, therefore this should be improved for the following years.

## **2.7 conferences**

This test set is made up of collection of 15 real-case ontologies dealing with the domain of conference organization. Although all the ontologies are well embedded in the described field, nevertheless they are heterogeneous in their nature. This heterogeneity comes mainly from: designed ontology application type, ontology expressivity in terms of formalism, and robustness. Out of given 15 ontologies the production of alignments should result in 210 possible combinations (we treat the equivalent alignment as symmetric). However, we obtained 91 non-empty alignment files in the generation. From the performance point of view the alignments took about 1 hour 20 minutes on a dual core computer<sup>3</sup>.

## **3 General comments**

### **3.1 Discussions on the way to improve the proposed system**

This year some tracks proved really difficult to work with. The new library track contains ontologies in different languages and due to its size first or during the mapping a translation needs to be carried out. This can be a challenge itself due to the number of concepts involved. Therefore from the background knowledge point of view we have concluded that based on the latest results that the additional multi lingual and domains specific background knowledge could provide added value for improving both recall and precision of the system.

### **3.2 Comments on the OAEI 2009 procedure**

The OAEI procedure and the provided alignment api works very well out of the box for the benchmarks, IIMB, anatomy, directory and conference tracks. However for the Eprints-Rexa-Sweto/DBLP benchmark and vlc and track we had to develop an SKOS parser, which can be integrated into the alignment api. Our SKOS parser convert SKOS file to OWL, which is then processed using the alignment api. Additionally we have developed a multi threaded chunk SKOS parser which can process SKOS file iteratively in chunks avoiding memory problems. For both Eprints-Rexa-Sweto/DBLP benchmark and vlc tracks we had to develop several conversion and merging utility as the original file formats were not easily processable.

### **3.3 Comments on the OAEI 2009 test cases**

We have found that most of the benchmark tests can be used effectively to test various aspects of an ontology mapping system since it provides both real word and generated/modified ontologies. The ontologies in the benchmark are conceived in a way that allows anyone to clearly identify system strengths and weaknesses which is an important advantage when future improvements have to be identified. The anatomy, library tests are perfect to verify the additional domain specific or multi-lingual domain knowledge. Unfortunately this year we could not integrate our system with such background knowledge so the results are not as good as we expected.

---

<sup>3</sup> Intel dual Core 3,0GHz, 512MB

## 4 Conclusion

Based on the experience gained during OAEI 2006, 2007, 2008 and 2009 we had a possibility to realise a measurable evolution in our ontology mapping algorithm and test it with 7 different mapping tracks. Our main objective is to improve the mapping precision with managing the inherent uncertainty of any mapping process and information in the different ontologies. The different formalisms of the ontologies suggest that on the Semantic Web there is a need to qualitatively compare and evaluate the different mapping algorithms. Participating in the Ontology Alignment Evaluation Initiative is an excellent opportunity to test and compare our system with other solutions and helped a great deal identifying the future possibilities that needs to be investigated further.

## References

1. Shvaiko, P., Euzenat, J.: Ten challenges for ontology matching. Technical Report DISI-08-042, University of Trento (2008)
2. Falconer, S.M., Storey, M.A.D.: A cognitive support framework for ontology mapping. In: Proceedings of 6th International Semantic Web Conference (ISWC2007). (2007) 114–127
3. Nagy, M., Vargas-Vera, M.: Reasoning representation and visualisation framework for ontology mapping using 3d modelling. In: In Proceedings of the 4th edition of the Interdisciplinary in Engineering International Conference (Inter-Eng 2009). (2009)
4. Nagy, M., Vargas-Vera, M., Motta, E.: Dssim - managing uncertainty on the semantic web. In: Proceedings of the 2nd International Workshop on Ontology Matching. (2007)
5. Shafer, G.: A Mathematical Theory of Evidence. (1976)
6. Nagy, M., Vargas-Vera, M., Motta, E.: Multi agent ontology mapping framework in the aqua question answering system. In: International Mexican Conference on Artificial Intelligence (MICAI-2005). (2005)
7. Besana, P.: A framework for combining ontology and schema matchers with dempster-shafer (poster). In: Proceedings of the International Workshop on Ontology Matching. (2006)
8. Yaghlane, B.B., Laamari, N.: Owl-cm: Owl combining matcher based on belief functions theory. In: Proceedings of the 2nd International Workshop on Ontology Matching. (2007)
9. Nagy, M., Vargas-Vera, M., Motta, E.: Feasible uncertain reasoning for multi agent ontology mapping. In: IADIS International Conference-Informatics 2008. (2008)
10. Nagy, M., Vargas-Vera, M., Motta, E.: Managing conflicting beliefs with fuzzy trust on the semantic web. In: The 7th Mexican International Conference on Artificial Intelligence (MICAI 2008). (2008)
11. Turney, P.D.: Similarity of semantic relations. *Computational Linguistics* **32**(3) (2006) 379–416
12. Kim, S.N., Baldwin, T.: Interpreting semantic relations in noun compounds via verb semantics. In: Proceedings of the COLING/ACL on Main conference poster sessions. (2006) 491–498
13. Banerjee, S., Pedersen, T.: Extended gloss overlaps as a measure of semantic relatedness. In: Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence. (2003) 805–810
14. Nagy, M., Vargas-Vera, M., Stolarski, P.: Dssim results for oaei 2008. In: Proceedings of the 3rd International Workshop on Ontology Matching. (2008)