DSSim - Managing Uncertainty on the Semantic Web

Miklos Nagy[1], Maria Vargas-Vera[2] and Enrico Motta[2]

[1] Knowledge Media Institute (Kmi)
The Open University
Walton Hall, Milton Keynes, MK7 6AA, United Kingdom
mn2336@student.open.ac.uk;e.motta@open.ac.uk
[2] Department of Information Systems
Poznan University of Economics
al. Niepodleglosci 10, 60-967 Poznan, Poland
maria@kie.ae.poznan.pl

Abstract. Managing uncertainty on the Semantic Web can potentially improve the ontology mapping precision which can lead to better acceptance of systems that operate in this environment. Further ontology mapping in the context of Question Answering can provide more correct results if the mapping process can deal with uncertainty effectively that is caused by the incomplete and inconsistent information used and produced by the mapping process. In this paper we introduce our algorithm called "DSSim" and describe the improvements that we have made compared to OAEI 2006.

1 Presentation of the system

1.1 State, purpose, general statement

The problem of mapping two ontologies effectively and efficiently is a necessary precondition to integrate information on the Semantic Web. In recent years different research communities have proposed[1] a wide range of methods for creating such mappings. The proposed methods usually combine syntactic and semantic measures by introducing different techniques ranging from heuristics to machine learning. While these methods perform well in certain domains the quality of the produced mappings can differ from domain to domain depending on the specific parameters defined in the methods e.g. tuning similarity threshold. Considering Question Answering systems like AQUA [2, 3] which answers queries over heterogeneous sources described by their ontologies, it is very important how its mapping algorithm performs in terms of mapping precision. Our objective is to produce a method that does not depend on any fine tuned internal parameters for a specific domain or does not assume having large amount of data samples a-priory for machine learning or Bayesian probability assessment[4]. Our hypothesis is that the correctness of different similarity mapping algorithms is always heavily dependent on the actual content and conceptual structure of these ontologies which are different even if two ontologies have been created on the same domain but with different purpose. Therefore from the mapping point of view these 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.

1.2 Specific techniques used

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 similariy measure between the i and j elements of the matrix as follows:

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

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 that 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 method we assume that each expert carries only partial knowledge of the domain and can observe it from its own perspective where available prior knowledge is generally uncertain and subjective. In order to represent subjective probabilities in our system we use the Dempster-Shafer theory of evidence [7], which provides a mechanism for modeling 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 modeled by Dempster-Shafer approach and additionally evidences from two or more sources can be combined using Dempster's rule of combination. 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 [8] is highly similarity to ”hasArticle” in O_2. Therefore after similarity assessment our variables will have the following belief mass value:

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

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 experts.

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. 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.

1.3 Adaptations made for the evaluation

Our mapping algorithm which is originally based on multi agent architecture has been re-implemented as a standalone mapping process which uses the common WordNet dictionary which is considered more general knowledge then originally we assume in or architecture. Originally our mapping process receives query fragments from the AQUA system where the query fragments contain several concept names and their properties. For the evaluation we modified our mapping process so we consider the individual concept or property names as query fragments which contain less information about the possible mapping then the query fragments that we originally receive from the AQUA system.

Creating the particular ontology mappings in the context of question answering is ideally an iterative process where the users are involved in the loop as well. In a real case scenario the users pose different questions that contain both concepts and properties of a particular domain. This information then can be used to query the different ontologies, create mapping between its concepts and properties that can be used to answer the particular query. For the OAEI 2006[6] we have implemented an iterative closed loop which creates the mapping without any human interaction. Based on this implementation we have modified our process for the OAEI 2007 which works as follows:

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.
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 ans properties together with the close context of the local ontology fragments.
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.

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.
5. The selected concepts are added into the alignment.

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

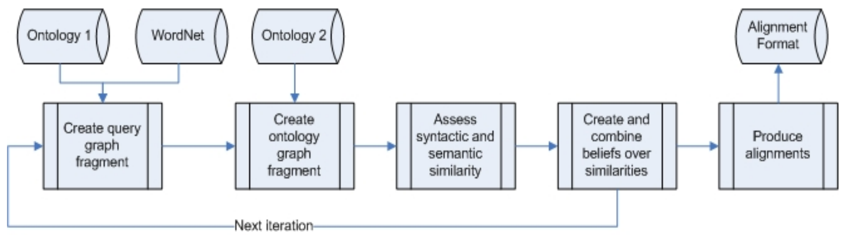

Fig. 1. The iterative mapping process

In order to avoid a complex graph of relationships in the query and the ontology fragments we need to define a reasonable limit on the number of hypernyms, which are extracted from the WordNet. To define such a limit is also desirable when we carry out the belief combination since all extracted terms represent a variable where each similarity value needs to be combined with the Dempster's rule of combination. The combination rule implies that the problem space increases exponentially with the number of variables therefore the proper definition of this limit can considerably affect the scalability of our system.

1.4 Link to the system and parameters file

http://kmi.open.ac.uk/people/miklos/OAEI2007/tools/DSSim.zip

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

http://kmi.open.ac.uk/people/miklos/OAEI2007/results/DSSim.zip

2 Results

2.1 Benchmark

Based on the results of the benchmarks we have improved our algorithm compared to the OAEI 2006 results in terms of recall. The improvement was achieved by introducing instance level comparisons for the classes and properties. Nevertheless there is considerable room for improvement since we did not achieve similar results compared to the best performing systems for the tests 248-266 in terms of recall. The performance of our algorithm has also been improved considering the execution time. In general the benchmarks are excellent for improving the algorithm since we can calculate the recall and precision any time which helps a lot evaluating the impact of a particular improvement. This year we have tried to make use of the *rdfs:label* tags but it did not improve the recall or precision.

2.2 Anatomy

The anatomy test has proved quite manageable considering the execution time. During our experiments our algorithm has always created the mappings within 2 hours. We had to use the *rdfs:labels* for the comparison but we could not make use of the *oboInOwl:Synonym* tags. The usage of the labels has introduced complexity retrieving hypernyms from WordNet since it is quite challenging to split the label into terms that can be used for querying the WordNet.

2.3 Directory

The directory test as well has been managable 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" which introduces certain level of complexity for determining synomyms using any background knowledge.

2.4 Food

The food test was extremely challenging due to the large number of concepts in the ontologies. We had to split up the original files into 8 parts and carry out the mapping one by one. Additionally we have developed a SKOS parser which can create smaller OWL chunks from the SKOS and run the mapping algorithm on it. As a result of this split we could not consider *rdfs:subClassOf* relationships between classes since it cannot be guaranteed that we would find the super classes in the same ontology chunk. The run time was around 1 week even though 2 parallel processors were used to run the mapping algorithm. During the SKOS OWL conversion we did not consider *skos:ConceptScheme* elements.

2.5 Environment

The environment test was the extension of the food test therefore it represented similar complexity in terms of run time performance. Nevertheless the GEMET SKOS contain smaller number of concepts compared to the food ontologies but we had to split up the ontologies into 2 separate parts. This implies that the *rdfs:subClassOf* relations have also not been considered which might have a negative impact on the mapping precision and recall. The mapping was also carried out on 2 parallel processors and the run time was around 2 days.

2.6 Library

The library test was also large therefore we also had to split it into 2 parts. The cumulative run time was around 1 day. Additionally not all labels were available in English therefore we have used the original Dutch labels. The implication is that we could not determine hypernyms from WordNet which might impact our mapping precision negatively.

3 General comments

3.1 Comments on the results

Most of the benchmark tests proved that when different similarity assessments have to be combined handling uncertainty can lead to a high precision rate which is a definite strength of our system. Another strength of our system is that the produced mappings are not very dependent on the structure and hierarchy of the concepts and properties in the ontology (see group 2xx). The reason is our mapping algorithm takes mainly concepts (classes) and properties (object and data type) to capture the specific restrictions in the particular ontologies and converts them into directed graph fragments. As a consequence our method is not heavily dependent on subclass, sub property, disjointness or equivalency relationships among classes and properties hence on the logical constraints imposed by the ontology language itself. Additionally the query terms are extended with their inherited hypernyms from WordNet so the uncertainty can be distributed sufficiently that can lead to a large number of possibly valid choices. However since Dempser's combination rule is computationally expensive operation we need to reduce the problem space therefore the number of additional variables per query fragment. This can lead to the loss of valuable information and consequently more irrelevant mappings.

3.2 Discussions on the way to improve the proposed system

Based on the results we have identified the following improvement possibilities that can further improve our system:

1. We need to consider Natural Language descriptions where available in the ontologies. This can lead to a definite improvement of precision for the particular mapping.
2. We need to further exploit the properties of the instances or individuals in the ontologies. This can lead to a definite improvement of recall for the particular mappings.
3. The possible application of additional multi lingual background knowledge can provide added value for improving both recall and precision of the system.

3.3 Comments on the OAEI 2007 procedure

The OAEI procedure and the provided alignment api works very well out of the box for the benchmarks, anatomy and directory tracks. However for the food, environment and library track we have developed 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 chunk SKOS parser which can process SKOS file iteratively in chunks avoiding memory problems.

3.4 Comments on the OAEI 2007 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. However, our system did not perform as well as we first expected probably due to the fact that most of the classes and properties in the ontologies are organized in a rather flat hierarchy so in our system the semantic similarity component did not influence the overall mappings considerably. However, in order to make use of a large group of tests (248-266) our system had to consider individuals or instances of the classes.

3.5 Comments on the OAEI 2007 measures

For our system the precision measure was the most important of all because this gives us the possibility to draw constructive conclusions on how the uncertainty handling can influence the precision of the system. The additional measures like recall and fallout can be used effectively for identifying where do we need to make further improvements in our system.

3.6 Proposed new measures

Besides the traditional measures it would be useful as well to introduce a measure that expresses the difficulty to create the particular mapping. E.g. there is a considerable difference in the level of difficulty between creating mapping with the reference ontology itself (101 to 101) and real word ontology (101 to 304). Additionally this measure then could be used to assess the how the particular system can handle mappings that involves complex comparison operations.

4 Conclusion

Based on the experiments of both the OAEI 2006 and 2007 we had a possibility to realise a measurable evolution in our ontology mapping algorithm and test it with 6 different mapping tracks. Our main objective to improve the mapping precision with managing the inherent uncertainty of any mapping process and information in the different ontologies on the Semantic Web can only be achieved if different mapping algorithms can be qualitatively compared and evaluated. Therefore 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.

Acknowledgements: This research project has been supported by a Marie Curie Transfer of Knowledge Fellowship of the European Community's Sixth Framework Programme under the contract number MTKD-CT-2004-509766 (enIRaF).

References

1. Choi N.,Song I-Y., Han H.(2006). A survey on ontology mapping. SIGMOD Record 35(3): 34-41.
2. Vargas-Vera M., and Motta E. (2004). An Ontology-driven Similarity Algorithm. *KMI-TR-151*, Knowledge Media Institute, The Open University, UK.
3. Vargas-Vera M.,Motta E., and Domingue J. (2003). AQUA: An Ontology-Driven Question Answering System. *AAAI Spring Symposium, New Directions in Question Answering*, Stanford, USA.: AAAI Press.
4. P. Rong,Ding Z.,Yu Y., and Peng Y. (2005). A Bayesian Network Approach to Ontology Mapping. In *Proceedings of the Fourth International Semantic Web Conference*,Galway, Ireland.
5. Nagy M.,Vargas-Vera M., and Motta E. (2005). Multi-agent Ontology Mapping Framework in the AQUA Question Answering System. the Fourth International Mexican Conference on Artificial Intelligence (MICAI-2005), Lecture Notes in Artificial Intelligence LNAI 3789, Gelbukh, A de Albornoz and H. Terashima (Eds), pp. 70-79, Monterrey Mexico.
6. Nagy M.,Vargas-Vera M., and Motta E. (2005). DSSim-ontology Mapping with Uncertainty In *Proceedings of the 1st International Workshop on Ontology Matching (OM-2006)* , Athens, Georgia, USA.
7. Shafer G. (1976). A Mathematical Theory of Evidence.: Princeton University Press.
8. Cohen W., Ravikumar P., and Fienberg S. (2003). A Comparison of String Distance Metrics for Name-Matching Tasks, In *Proceedings of Information Integration on the Web (IIWeb 2003)*, Accapulco, Mexico.

Appendix: Raw results

Matrix of results

#	Name	Prec.	Rec.	Time
101	Reference alignment	1	1	00.00.10.37
102	Irrelevant ontology	0	NaN	00.00.03.57
103	Language generalization	1	1	00.00.03.76
104	Language restriction	1	1	00.00.03.09
201	No names	1	0.16	00.00.03.16
202	No names, no comments	1	0.16	00.00.05.12
203	No comments	1	1	00.00.04.92
204	Naming conventions	0.96	0.91	00.00.03.23
205	Synonyms	0.94	0.33	00.00.03.44
206	Translation	0.97	0.39	00.00.04.38
207		0.97	0.39	00.00.04.36
208		0.95	0.9	00.00.04.18
209		0.91	0.32	00.00.03.19
210		0.97	0.39	00.00.04.26
221	No specialisation	1	1	00.00.04.15
222	Flatenned hierarchy	1	1	00.00.02.88
223	Expanded hierarchy	1	1	00.00.02.98
224	No instance	1	1	00.00.03.06
225	No restrictions	1	1	00.00.02.85
228	No properties	1	1	00.00.02.98
230	Flattened classes	0.97	1	00.00.01.59
231	Expanded classes	1	1	00.00.03.17
232		1	1	00.00.02.95
233		1	1	00.00.02.85
236		1	1	00.00.01.60
237		1	1	00.00.01.60
238		1	1	00.00.02.88
239		0.97	1	00.00.03.01
240		0.97	1	00.00.01.70
241		1	1	00.00.01.75
246		0.97	1	00.00.01.62
247		0.97	1	00.00.01.68
248		1	0.16	00.00.01.75
249		1	0.16	00.00.04.50
250		1	0.27	00.00.04.39
251		1	0.17	00.00.01.75
252		1	0.16	00.00.04.50
253		1	0.16	00.00.04.74
254		1	0.27	00.00.04.41
257		1	0.27	00.00.01.70
258		1	0.17	00.00.01.68
259		1	0.16	00.00.04.37
260		0.9	0.31	00.00.04.63
261		1	0.27	00.00.01.92
262		0.9	0.31	00.00.01.70
265		0.8	0.24	00.00.01.70
266		0.82	0.3	00.00.01.92
301	Real: BibTeX/MIT	0.85	0.6	00.00.03.44
302	Real: BibTeX/UMBC	0.85	0.8	00.00.02.69
303	Real: Karlsruhe	0.96	0.92	00.00.04.12
304	Real: INRIA	0.98	0.64	00.00.03.03