

Integrating Uncertainty Into Ontology Mapping

Ying Wang

School of Electronics, Electrical Engineering and Computer Science,
Queen's University Belfast, Belfast, BT7 1NN, UK
{ywang14}@qub.ac.uk

Abstract. This paper gives an outline of my PhD thesis which describes the integration of managing uncertainty into ontology mapping. Ontology mapping is one of the most important tasks for ontology interoperability and its main aim is to find semantic relationships between entities (i.e. concept, attribute, and relation) of two ontologies. However, in the process of mapping, uncertainty and incompleteness of semantics in the syntactic representation and description of relations between entities in ontologies will lead to imprecise results. If we want to obtain better results, it becomes more significant for the ontology mapping to be able to deal with uncertainty.

1 Introduction to the Problem

Ontologies are cores in the Semantic Web because they are the carriers of the meaning contained in the Semantic Web. However in many cases, different domains define different ontologies containing the same concepts. Even in the same domain, different organizations construct different ontologies. Therefore, it is necessary to find mappings between ontologies.

In recent years, researchers have developed a number of tools for finding these mappings. In most cases, the mappings produced are imprecise. For instance, most automatic ontology mapping tools use heuristics or machine learning techniques, which are imprecise by their very nature. Even experts sometimes could not be sure about the exact match between concepts and just assign some certainty ratings to matches [1]. The reason why we get imprecise mapping results is to some extent because of uncertainty. Uncertainty means that an agent, i.e. a computer or a human, has only partial knowledge about the truth value of a given piece of information. Peter Haase thought that uncertainty can be distinguished to objective uncertainty and subjective uncertainty [2]. I think that subjective uncertainty plays the most prominent role that prevent ontology mapping from getting better results. For instance, in the process of mapping, one may use diverse range of rules to judge the mapping, but different rules imply different uncertainty and one does not know the accuracy and reliability of rules. There has been little focus on dealing with uncertainty in ontology mapping. However this problem began to attract attention of researchers in this area.

The rest of the paper is organized as follows. Section 2 presents the current state of existing solutions. Section 3 discusses the current approaches we have

achieved so far and also problems we have encountered. Section 4 concludes the paper.

2 Uncertainty, Ontology Mapping and Ontology Mapping with Uncertainty

Uncertainty There are many models to deal with uncertain information in artificial intelligence and the most popular methods are:

1. Probability Theory It attempts to quantify the notion of probable. An increasing important approach to using probability theory in computing is probability networks, also called Bayesian networks or Causal networks. Bayesian networks can be used to represent and reason more efficiently with probabilistic information.

2. Evidence Theory The Dempster-Shafer theory (Evidence Theory) is a mechanism for representing and reasoning with uncertainty, imprecise and incomplete information [3]. Evidence theory has an ability to model information flexibly and it provides a convenient and simple mechanism for combining two or more pieces of evidence when they come from distinct sources.

3. Possibility Theory It was introduced in 1978 by L.A. Zadeh [4], in connection with the fuzzy set theory, to allow a reasoning to be carried out on imprecise or vague knowledge, making it possible to deal with uncertainties on knowledge. Two typical merging operations in possibility theory are max and min which combine two possibility distribution into one.

Ontology Mapping Ontology mapping is a critical problem in many application domains and many different mapping solutions have been proposed with diverse range of mapping techniques so far. For example, an integrated ontology mapping approach [5] is proposed based on rules and the method in quick ontology mapping [6] puts attention to the runtime of program. Anchor-PROMPT is a tool for ontology merging and mapping [7, 8]. It contains a sophisticated prompt mechanism for possible mapping entities. In [9], researchers from several universities are working together to create an ontology mapping approach based on information flow. The approach of semantic enrichment for ontology mapping exploits text categorization to automatically assign documents to the concept in the ontology and use the documents to calculate the similarities between concepts in ontologies [10].

Ontology Mapping with Uncertainty In the mapping process, if only syntactic or element-level matching is performed, as in the case for name matching without the use of a thesaurus, inaccuracies can occur [11]. It affects the results of mapping, but so far only a few ontology mapping methods considered dealing with uncertainty issue.

Nagy et al [12] and Besana [13] both recognized the importance of uncertainty in ontology mapping, and both of them used Dempster-Shafer theory to assist mapping. They believed that different matchers have uncertainties associated with them, so they combine the results obtained from different matchers using

DS theory and it is possible to give a uniform interpretation, consistent with the uncertainty inherent in the problem.

In [14] a Bayesian Networks based approach was designed and a system called BayesOWL was proposed. In this approach, the source and target ontologies are first translated into Bayesian networks (BN); the concept mapping between the two ontologies are treated as evidential reasoning between the two translated BNs. Probabilities, that are required for constructing conditional probability tables (CPT) during translation and for measuring semantic similarity during mapping, are learned using text classification techniques where each concept in an ontology is associated with a set of semantically relevant text documents, which are obtained by ontology guided web mining.

Holi and Hyvönen [15] thought that in the real world, concepts are not always subsumed by each other, and cannot always be organized in a crisp subsumption hierarchies. Many concepts only partly overlap with each other, so they present a new probabilistic method to model conceptual overlap in taxonomies, and an algorithm to compute the overlap between a selected concept and other concepts of a taxonomy by using Bayesian networks.

Zhao et al [16] proposed a novel similarity measure method based on rough set theory and formal concept analysis (RFCA) to realize ontology mapping tasks. The authors combined rough set theory into the similarity computation formula of formal concept analysis (FCA). Although the authors did not consider uncertainty in the process of mapping explicitly, they applied the rough set theory to measure the similarity of concepts of ontologies. So, in some case, they considered the uncertainty problem.

3 My Approaches

We have proposed a new ontology mapping approach called ACAOM [17]. It first uses WordNet to calculate similarities between words for the node names comparison in ontologies. WordNet is organized as a classified structure in which synset is the basic unit. Every set containing many words expresses a single meaning and we can utilize jwnl to access WordNet, thereby to compute similarities between names of nodes. This is a name-based mapping strategy. Then in the instance-based strategy we follow the following assumption: the more semantics is explicitly specified about the ontologies, the more feasible their comparison becomes. Here, the instances are documents assigned to the concept nodes. According to the definition of ontology mapping given above, given a node in one ontology, a mapping function searches the node in another ontology which has the most similar semantics to it. We employ approaches in information retrieval to assist ontology mapping. In this way, we can denote the documents that have been bound with nodes by using vector space models, and then the numerical degrees of similarity give a way to rank the matches.

Following that we are developing a new mapping method which uses three different and independent matchers: *Edit distance-based matcher*, *Linguistic-based matcher* and *Structure-based matcher*. In these three matchers, a new matcher

is the *structure-based matcher* which utilizes the similarity measures between two words (w_1 and w_2), a father node of w_1 with w_2 and all the child nodes of w_1 with w_2 . This matcher takes both the semantics and the structure of an ontology into account. We then discuss how the mapping results from different matchers can be combined. We consider both the Dempster Shafer theory of evidence (DS theory) and Possibility Theory and apply them to combine the outcomes obtained by three different and independent matchers.

In our evaluation, we choose **Test 101**, **Test 103**, **Test 104**, **Test 205**, **Test 223** and **Test 302** of OAEI 2006 benchmark tests and take **Test 101** as the reference ontology. All of other ontologies are compared with **Test 101**. We use the combination mechanisms in both DS theory and Possibility Theory to combine the matching results from our three matchers. We now compare the outputs from the two combination rules to the results obtained from *falcon*, *ola* and *ctxMatch2-1* algorithms which were used in the EON 2005 Ontology Alignment Contest ¹, and the details are given in Table 1. In Table 1, p for precision, r for recall, f for f-measure, DS for Dempster’s combination rule, and PT for the minimum merging operator in Possibility Theory.

Table 1. Comparison of Experiment Results

Datasets	DS			PT			falcon			ola			ctxMatch2-1		
	p	r	f	p	r	f	p	r	f	p	r	f	p	r	f
101-103	100	98.97	99.48	100	98.97	99.48	100	100	100	100	100	100	87	34	48
101-104	100	98.97	99.48	100	98.97	99.48	100	100	100	100	100	100	87	34	48.89
101-205	46.88	46.39	46.63	30.29	29.90	30.09	88	87	87.5	43	42	42.5	36	4	7.2
101-223	100	98.97	99.48	100	98.97	99.48	100	100	100	100	100	100	83	31	45.14
101-302	45.83	45.83	45.83	43.75	43.75	43.75	97	67	79.26	37	33	34.89	0	0	0

4 Conclusion

During my PhD study, my research will mainly focus on dealing ontology mapping with uncertainty, so I think I can begin from two aspects: ontology mapping and uncertainty theory. For ontology mapping, I need to develop more scientific matchers to find out mapping relations between ontologies. These matchers should put attention on how to find different kinds of mappings and how to improve the accuracy of mappings. For uncertainty theory, firstly, I will study the different methods of representing uncertainty. Secondly, I will try to utilize these different representing methods to represent ontology mappings. Thirdly, I will study how to use different combination rules of uncertainty in different situations. I hope that I can use different methods to handle different uncertainty problems in the process of ontology mapping successfully in the future.

References

1. Choi, N., Song, I.Y., Han, H.: A survey on ontology mapping. SIGMOD Record **35**(3) (2006) 34–41

¹ <http://oaei.ontologymatching.org/2005/results/>

2. Haase, P., Völker, J.: Ontology learning and reasoning - dealing with uncertainty and inconsistency. In the International Workshop on Uncertainty Reasoning for the Semantic Web (URSW'05), collocated with the 4th International Semantic Web Conference (ISWC'05). (2005) 45–55
3. Liu, W.: Propositional, Probabilistic and Evidential Reasoning: Integrating Numerical and Symbolic Approaches. (2001)
4. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. In *Journal of Fuzzy Sets and Systems*. **100** (1999) 9–34
5. Ehrig, M., Sure, Y.: Ontology mapping - an integrated approach. In *Proceedings of the 1st European Semantic Web Symposium (ESWS'04)*. (2004) 76–91
6. Ehrig, M., Staab, S.: Qom - quick ontology mapping. In *Proceedings of the 3rd International Semantic Web Conference (ISWC'04)*. (2004) 683–697
7. Noy, N.F., Musen, M.A.: Prompt: Algorithm and tool for automated ontology merging and alignment. In *Proceedings of the 17th National Conference on Artificial Intelligence and 12th Conference on Innovative Applications of Artificial Intelligence (AAAI/IAAI'00)*. (2000) 450–455
8. Noy, N.F., Musen, M.A.: Anchor-prompt: Using non- local context for semantic matching. In *Workshop on Ontologies and Information Sharing at the 17th International Joint Conference on Artificial Intelligence (IJCAI'01)*. (2001)
9. Kalfoglou, Y., Schorlemmer, W.M.: Information-flow-based ontology mapping. In *Proceedings of the International Federated Conferences (CoopIS/DOA/ODBASE'02)*. (2002) 1132–1151
10. Su, X., Gulla, J.A.: Semantic enrichment for ontology mapping. In *Proceedings of the 9th International Conference on Applications of Natural Language to Information Systems (NLDB'04)*. (2004) 217–228
11. Cross, V.: Uncertainty in the automation of ontology matching. In *Proceedings of the 4th International Symposium on Uncertainty Modeling and Analysis (ISUMA'03)*. (2003) 135–140
12. Nagy, M., Vargas-Vera, M., Motta, E.: Dssim-ontology mapping with uncertainty. In the International Workshop on Ontology Matching (OM'06), collocated with the 5th International Semantic Web Conference (ISWC'06). (2006)
13. Besana, P.: A framework for combining ontology and schema matchers with Dempster-Shafer. In the International Workshop on Ontology Matching (OM'06), collocated with the 5th International Semantic Web Conference (ISWC'06). (2006) 196–200
14. Pan, R., Ding, Z., Yu, Y., Peng, Y.: A Bayesian network approach to ontology mapping. In *Proceedings of the 4th International Semantic Web Conference (ISWC'05)*. (2005) 563–577
15. Markus Holi, E.H.: Modeling degrees of conceptual overlap in semantic web ontologies. In the International Workshop on Uncertainty Reasoning for the Semantic Web (URSW'05), collocated with the the 4th International Semantic Web Conference (ISWC'05). (2005) 98–99
16. Zhao, Y., Wang, X., Halang, W.A.: Ontology mapping based on rough formal concept analysis. In *Proceedings of the Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services (AICT/ICIW'06)*. (2006) 180
17. Wang, Y., Gong, J., Wang, Z., Zhou, C.: A composite approach for ontology mapping. In *Proceedings of the 23rd British National Conference on Databases (BNCOD'23)*. (2006) 282–285