Matching Domain and Top-level Ontologies via OntoWordNet

Daniela Schmidt[*], Rafael Basso[*], Cassia Trojahn[†], Renata Vieira[*]

[*]Pontifical Catholic University of Rio Grande do Sul (Brazil)
daniela.schmidt, rafael.basso@acad.pucrs.br, renata.vieira@pucrs.br
[†]Université de Toulouse 2 & IRIT (France)
cassia.trojahn@irit.fr

1   Context and proposed approach

Matching domain and top-level ontologies is an important task but still an open problem in the ontology matching field. The main difficulties are particularly due to their different levels of abstraction. In this paper, we propose an approach that exploits existing alignments between WordNet and top-level ontologies, as an intermediate layer, and that relies on the notion of context of concepts [1,3,5]. Contexts are constructed from all information about an ontology entity (e.g., entity naming, annotation properties and information on the neighbors of entities) and are used for disambiguating the senses that better express the meaning of ontology entities in WordNet. After selecting an appropriated synset for a given domain ontology, we verify if there is a relation between that synset and a top-level concept, via existing alignments between WordNet and the top-level ontology. Here, we focus on DOLCE top-level ontology and OntoWordNet [2]. This choice is motivated by the fact that DOLCE is one of the most used top-level ontologies and serves as a reference for the modeling and integration of ontologies [4].

2   Experiments

In order to evaluate our approach, we run experiments involving a set of 7 domain ontologies from the OAEI Conference data set[1] regarding DOLCE-Lite-Plus and OntoWordNet [2]. We focused on the first-level of domain concepts hierarchy, what corresponds to 70 concepts. This choice is motivated by the fact that correspondences can be assigned by inheritance to the child concepts. Compounds have been pre-processed and we removed the modifier (e.g. conference_document is a document). As the domain ontologies are not equipped with descriptions of their concepts, we manually enriched the first-level concepts with such definitions. For that, we adopt the Cambridge online dictionary[2] where we chosen the definition of each concept considering the most related one to the conference domain. The experiments were executed with the original and enriched versions of the domain ontologies and DOLCE-Lite-Plus (resulting in 7 pairs). This resulted in a total of 71 correspondences (including the different correspondences

---

[1] http://oaei.ontologymatching.org/2016/conference/index.html
[2] http://dictionary.cambridge.org/us/

found in the two versions of the domain ontologies). These 71 correspondences were presented, separately, to an expert on top-level ontologies, via an online form. The form shows the pair of concepts, their hierarchy and description. The expert was instructed to select one of the options among "equivalent", "sub concept", "none" or "other". For "other", a description of the kind of relation was required.

Results and discussion Regarding the expert judgment, 36 correspondences out of 63 for the original ontology were judged as correct. For the dictionary-enriched ontology, there are also 36 pairs considered as correct, from a total of 62. For 7 concepts in the original ontologies and 8 in the enriched ontologies, no corresponding concepts in OntoWordNet were found. Assuming that all first-level concepts in the domain ontology have potentially a corresponding concept in the top-level ontology, we compute precision, recall and F-measure. We observe similar results for both ontology versions. In fact, we expected that the descriptions would improve the synset selection and therefore produce an impact on the alignments, however the improvements were not that significant between the two versions. As we adopted plain dictionary descriptions for the terms, it might be the case that these descriptions were simply too general.

3 Concluding remarks and future work

This paper presented an approach to automatically match domain and top-level ontologies. We consider that existing top-level and WordNet alignments are a valuable resource for the task, at least for certain general domains. For most of the concepts from the domain ontologies we found a correspondence with the top ontology. In addition, the precision was better than available matching systems considered in previous experiments [6]. We are aware that the experiment settings were different, but it is possibly an indication that the proposed approach might be an option for certain domains and it development should be continued and refined. As future work, we intend to improve the description of the concepts to include a more closer information about the domain, apply alternative similarity metrics for measuring the overlap between contexts, deal with logical reasoning and involve more experts in the evaluation process.

References

1. Djeddi, W.E., Khadir, M.T.: A novel approach using context-based measure for matching large scale ontologies. In: Data Warehousing and Knowl. Discovery. pp. 320–331 (2014)
2. Gangemi, A., Navigli, R., Velardi, P.: The OntoWordNet Project: Extension and Axiomatization of Conceptual Relations in WordNet, pp. 820–838. Springer Berlin Heidelberg (2003)
3. Maedche, A., Staab, S.: Measuring similarity between ontologies. In: Knowledge Engineering and Knowledge Management. pp. 251–263 (2002)
4. Oberle, D., Ankolekar, A., Hitzler, P., et al.: DOLCE ergo SUMO: On foundational and domain models in the SmartWeb Integrated Ontology (SWIntO). Web Semantics: Science, Services and Agents on the World Wide Web 5(3) (2007)
5. Schmidt, D., Trojahn, C., Vieira, R., Kamel, M.: Validating Top-level and Domain Ontology Alignments using WordNet. In: Brazilian Ontology Research Seminar. pp. 119–130 (2016)
6. Schmidt, D., Trojahn, C., Vieira, R.: Analysing Top-level and Domain Ontology Alignments from Matching Systems. In: Workshop on Ontology Matching. pp. 1–12 (2016)