Towards Dialogue-Based Interactive Semantic Mediation in the Medical Domain

THESEUS
Forschungsprogramm für eine neue internetbasierte Wissensinfrastruktur

Daniel Sonntag
German Research Center for Artificial Intelligence, Saarbrücken, Germany
daniel.sonntag@dfki.de

Introduction

We think of ontology matching as a dialogue-based interactive mediation process for which we propose a three stage model. Especially, we address the challenge how to use **dialogue-based interactivity with the user** to rate partial alignments between two ontologies. Our approach allows to

- add new information sources dynamically in the context of information and knowledge retrieval applications
- benefit from dialogue applications towards implementing **interactive semantic mediation** in the context of relational database integration, ontology merging, semantic web service composition, semantic peer-to-peer networks, and semantic query answering (with increasing intensity)

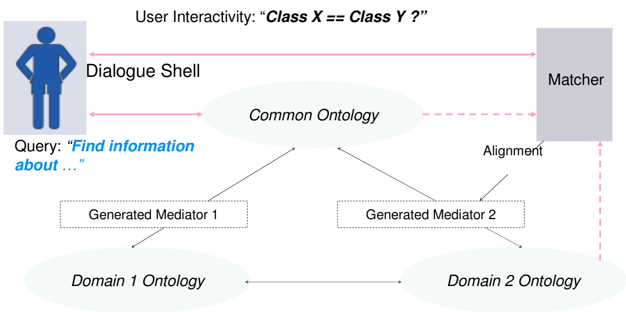
by employing a **dialogue shell**. Dialogue-based interaction systems in the sense of conversational user interfaces stand for improved efficiency and ease of use.

Idea: Consider the methods that are required for interactive ontology mapping and evaluate the impact of dialogue-based user feedback in this process. The stages are:

(1) Compute a rudimentary partial mapping by a simple string-based method;

(2) Ask the user to disambiguate some of the proposed mappings;

(3) Use the resulting alignments as input for more complex algorithms.



Evaluation

Our datasets consisted of ontologies and alignment examples (manually annotated alignments for Radlex and NCI). For the first test in the medical domain, we annotated 50 alignments, 30 perfect positives and 20 perfect negatives.

In the medical domain, the precision was 92% and recall 50% for simple string-based methods. (Corresponding concept names may differ substantially in their syntactic form.)

The best matches were taken as alignment input for similarity flooding after manually confirming their validity (which simulates positive user feedback):

- On average, the first stage of the matching execution (string-based matching) takes less than 5% of the end-to-end matching execution time.
- The input alignments (confirmed by the simulated dialogue) allow to compute a complete mapping almost 10 times faster within a 30 seconds time frame.

Dialogue for Interactive Ontology Mapping in the Medical Domain

1. U: "Show me the internal organs: lungs, liver, then spleen and colon."

2. S: Shows corresponding patient image data according to referral record.

3. U: "This lymph node here (+ pointing gesture) is enlarged; so lymphadenopathy (finding Radlex Term A). Are there any comparative cases in the hospital?"

4. S: "Sorry, there are no comparative cases, shall I continue to search for similar cases?"

5. U: "Ah okay."

6. S: "Unfortunately, external databases could not be matched appropriately. Is lymphadenopathy equivalent to lymph node enlargement (Radlex Term B) or adenopathy (NCI Term C) or a kind of hematopoietic system finding (NCI superconcept)?"

7. U: **"Yes, that could be said,"**

8. U: **"a pathophysiologic process."** (Medico now includes RadLex superconcepts for search.)

9. S: "I found comparative cases in an external knowledge base according to your feedback." Medico switches to the comparative records to help the radiologist in the differential diagnosis of the suspicious case, before the next organ (liver) is examined.

10. U: "Find similar liver lesions with the characteristics: hyper-intense and/or coarse texture ..."

11. S: Medico again displays the search results ranked by the similarity and matching of the medical terms that constrain the semantic search.

Conclusion

Use dialogue and do better than calculating the set of correspondences in a single shot.

User-confirmed perfect mappings can be used in simple name matching retrieval contexts, **but** this does not reflect the nature of real-world industrial requirements (in particular, where the user cannot be supposed to deliver a reliable judgement).

-> **Better translate formal mapping uncertainties into dialogue (questions).**

References

1. Oberle, D., Ankolekar, A., Hitzler, P., Cimiano, P., Sintek, M., Kiesel, M., Mougouie, B., Baumann, S., Vembu, S., Romanelli, M., Buitelaar, P., Engel, R., Sonntag, D., Reithinger, N., Loos, B., Zorn, H.P., Micelli, V., Porzel, R., Schmidt, C., Weiten, M., Burkhardt, F., Zhou, J.: DOLCE ergo SUMO: On foundational and domain models in the SmartWeb Integrated Ontology (SWIntO). Web Semant. 5(3) (2007) 156–174
2. Reithinger, N., Sonntag, D.: An integration framework for a mobile multimodal dialogue system accessing the Semantic Web. In: Proceedings of Interspeech'05. (2005)
3. Sonntag, D., Engel, R., Herzog, G., Pfalzgraf, A., Pfleger, N., Romanelli, M., Reithinger, N.: SmartWeb Handheld—Multimodal Interaction with Ontological Knowledge Bases and Semantic Web Services. In Huang, T.S., Nijholt, A., Pantic, M., Pentland, A., eds.: AI for Human Computing. Volume 4451 of LNCS., Springer (2007) 272–295
4. Euzenat, J., Shvaiko, P.: Ontology matching. Springer-Verlag, Heidelberg (2007)
5. Bernstein, P.A., Melnik, S., Churchill, J.E.: Incremental schema matching. In: VLDB '06: Proceedings of the 32nd international conference on Very large data bases, VLDB Endowment (2006) 1167–1170
6. Robertson, G.G., Czerwinski, M.P., Churchill, J.E.: Visualization of mappings between schemas. In: CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, ACM (2005) 431–439
7. Falconer, S.M., Noy, N., Storey, M.A.D.: Towards understanding the needs of cognitive support for ontology mapping. In Shvaiko, P., Euzenat, J., Noy, N.F., Stuckenschmidt, H., Benjamins, V.R., Uschold, M., eds.: Ontology Matching. Volume 225 of CEUR Workshop Proceedings., CEUR-WS.org (2006)
8. Melnik, S., Garcia-Molina, H., Rahm, E.: Similarity flooding: A versatile graph matching algorithm and its application to schema matching. In: ICDE '02:, Washington, DC, USA, IEEE Computer Society (2002) 117–128

This research is supported by the German Federal Ministry of Economics and Technology under the grant number 01MQ07016 (THESEUS). The potential application scenario (provided by Siemens AG in context of the THESEUS-Medico project) includes a radiologist which treats a lymphoma patient. The responsibility for this publication lies with the author.