Debugging is-a structure in and mappings between taxonomies

Patrick Lambrix, Valentina Ivanova

Department of Computer and Information Science, and Swedish e-Science Research Centre, Linköping University, Linköping, Sweden

Introduction

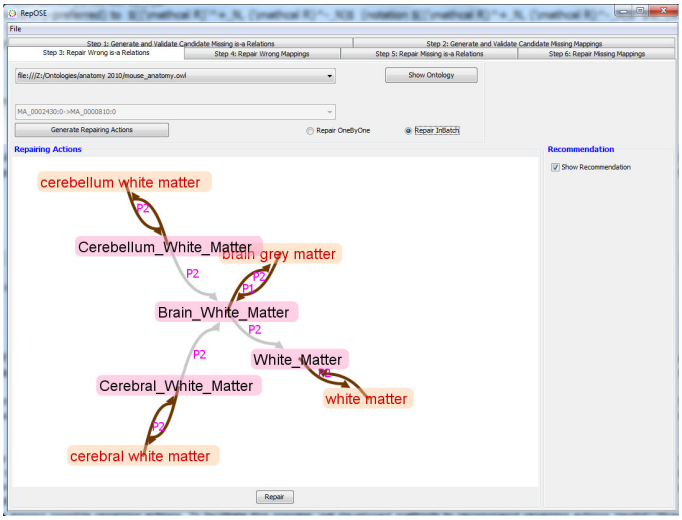
In recent years many ontologies have been developed. The benefits of using ontologies include reuse, sharing and portability of knowledge across platforms, and improved documentation, maintenance, and reliability. Ontologies lead to a better understanding of a field and to more effective and efficient handling of information in that field. Often we also want to be able to use multiple ontologies. In these cases it is important to know the relationships between the terms in the different ontologies and much research has recently been done on ontology alignment, i.e., finding mappings between terms in different ontologies.

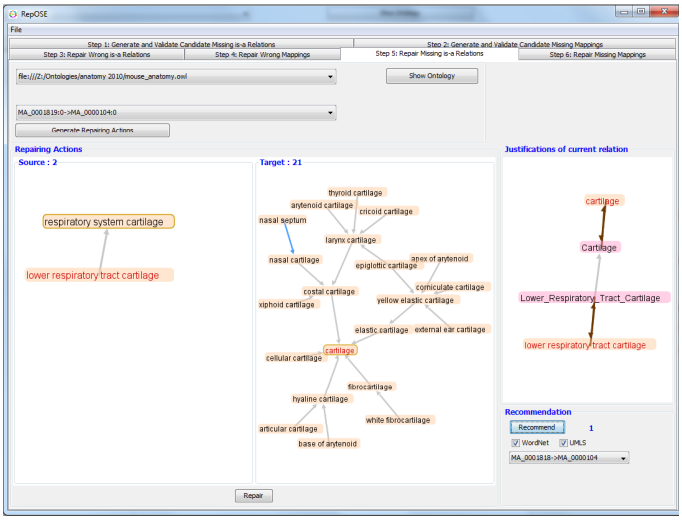
Neither developing ontologies nor aligning ontologies are easy tasks and often the resulting ontologies and mappings are not consistent or complete. With the increased use of ontologies and ontology mappings in semantically-enabled applications such as ontology-based search and data integration, the issue of detecting and repairing defects in ontologies and ontology mappings has become increasingly important. These defects can lead to wrong or incomplete results for the applications.

Results

We have developed a theoretical framework for debugging the is-a structure of and mappings between taxonomies, the most used kind of ontologies. In the poster we show an implemented system, RepOSE, that supports a domain expert to detect and repair missing and wrong is-a relations and mappings [1,2].

Using this system a domain expert has debugged the anatomy ontologies and their mappings that are used in the 2010 Ontology Alignment Evaluation Initiative. The Adult Mouse Anatomy Dictionary (AMA) contains 2744 concepts and 1807 asserted is-a relations, while the NCI Thesaurus Anatomy (NCI-A) contains 3304 concepts and 3761 asserted is-a relations. There are 986 equivalence mappings and 1 subsumption mapping between these taxonomies. Using RepOSE, we detected 102 missing is-a relations and 21 wrong is-a relations in AMA. For NCI-A we detected 61 missing is-a relations and 19 wrong is-a relations. To repair these defects 85 is-a relations were added to AMA and 57 to NCI-A, 13 is-a relations were removed from AMA and 12 from NCI-A. Further, 12 mappings were removed. In 22 cases in AMA and 8 cases in NCI-A a missing is-a relation was repaired using is-a relations that could not by derived from the original taxonomies and mappings.





References

[1] Ivanova V, Lambrix P, A System for Debugging Taxonomies and their Alignments, *First International Workshop on Debugging Ontologies and Ontology Mappings,* 2012.

[2] Lambrix P, Ivanova V, A unified approach for debugging is-a structure and mappings in networked taxonomies, *submitted.*

Acknowledgments: We acknowledge the Swedish National Graduate School in Computer Science, the Swedish Research Council and the Swedish e-Science Research Centre.

Correspondence: Patrick.Lambrix@liu.se