A system for repairing missing is-a structure in ontologies

Patrick Lambrix, Qiang Liu, and He Tan

Department of Computer and Information Science
Linköpings universitet, 581 83 Linköping, Sweden

Abstract. Developing ontologies is not an easy task and often the resulting ontologies are not consistent or strucurally complete. Such ontologies, although often useful, also lead to problems when used in semantically-enabled applications. Wrong conclusions may be derived or valid conclusions may be missed. To deal with this problem we may want to repair the ontologies. In this demo we present a system that supports the repair of the is-a hierarchy in ontologies. We have developed a tool that, given missing is-a relations, generates and recommends relevant ways to repair the is-a structure of the ontology and that allows a domain expert to do the repair in a semi-automatic way.[1]

1 Introduction

Developing ontologies is not an easy task and often the resulting ontologies are not consistent or structurally complete. Such ontologies, although often useful, lead to problems when used in semantically-enabled applications. Wrong conclusions may be derived or valid conclusions may be missed. Defects in ontologies can take different forms. Syntactic defects are usually easy to find and to resolve. Defects regarding style include such things as unintended redundancy. More severe defects are the modeling defects which require domain knowledge to detect and resolve, and semantic defects such as unsatisfiable concepts and inconsistent ontologies. There are a number of approaches to deal with semantic defects (e.g. [11, 6, 5, 4, 9, 3]).

In this demo we present a system, RepOSE (*Rep*air of *O*ntological *S*tructure *E*nvironment), that tackles a special case of the problem of repairing modeling defects, i.e. the repairing of missing is-a relations, and to our knowledge this system is the first in its kind. In the given setting it is known that a number of intended is-a relations are not present in the source ontology, and these are given. These missing is-a relations could be generated by automated tools. For instance, in the case of task 4 in the Anatomy track in the 2008 Ontology Alignment Evaluation Initiative (OAEI) [8], two ontologies, Adult Mouse Anatomy Dictionary [1] (MA, 2744 concepts) and the NCI Thesaurus - anatomy [10] (NCI-A, 3304 concepts), and 988 mappings between the two ontologies are given. Based on the structure of the source ontologies and the given mappings, it can be derived that, after removing redundancy, 121 is-a relations in MA and 83 in NCI-A are missing. Another approach for detecting missing is-a relations is given in [2].

[1] This paper is an abbreviated and slightly revised version of [7].

The problem is then to find is-a relations (called a *structural repair*) such that when these are added to the ontology, all missing is-a relations can be derived from the extended ontology. For formal definitions we refer to [7]. Although the missing is-a relations themselves constitute a structural repair, this may not be the most interesting solution for the domain expert. For instance, in MA, we know that an is-a relation between *wrist joint* and *joint* is missing and could be added to the ontology. However, knowing that there is an is-a relation between *wrist joint* and *limb joint*, a domain expert will most likely prefer to add an is-a relation between *limb joint* and *joint* instead. This is more informative and would lead to the fact that the missing is-a relation can be derived. In this particular case, it would also lead to the repairing of 6 other missing is-a relations (e.g between *elbow joint* and *joint*).

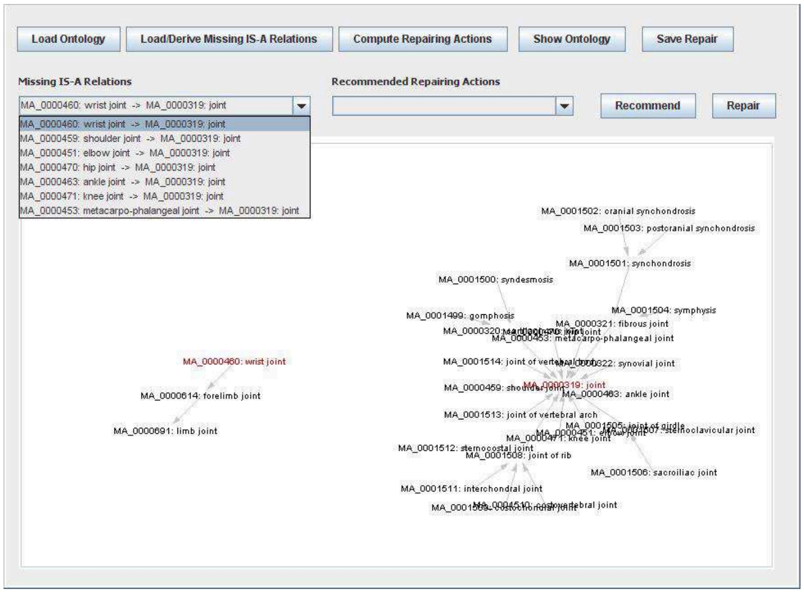
2 RepOSE

We have developed a tool that generates and recommends possible ways to repair the structure of the ontology (based on named concepts and subsumption axioms) and that allows a domain expert to repair the structure of an ontology in a semi-automatic way. As input our system takes an ontology in OWL format as well as a list of missing is-a relations. We use a framework and reasoner provided by Jena (version 2.5.7). The domain knowledge that we use is WordNet and the Unified Medical Language System. The ontology and missing is-a relations can be imported using the *Load/Derive Missing IS-A Relations* button. The user can see the list of missing is-a relations under the *Missing IS-A Relations* menu (see figure 1). In our example there are 7 missing is-a relations. Clicking on the *Compute Repairing Actions* button, results in the computation of possible repairing actions.

The user can select which missing is-a relation to repair first. They are ranked with respect to the number of possible repairing actions. The first missing is-a relation in the list has the fewest possible repairing actions, and may therefore be a good starting point. The repairing actions are represented using Source and Target sets. A possible repairing action is then an is-a relation “A is-a B” where A is an element from the Source set and B is an element from the Target set. For instance, in figure 1 we have the Source and Target sets in the panels on the left and the right, respectively, for the missing is-a relation between *wrist joint* and *joint*. The concepts in the missing is-a relation are highlighted in red. Any pair from Source x Target would allow us, when added to the ontology, to derive the missing is-a relation. For instance, we could choose *limb joint* from Source and *joint* from Target. We have implemented two algorithms that compute possible repairing actions: a basic algorithm and a more complex algorithm that takes into account influences of other missing is-a relations that are valid for all possible choices for repairing actions for the other missing is-a relations. Both algorithms implement three heuristics. The first heuristic prefers not to use redundant or non-contributing is-a relations for repairing. The second heuristic prefers to use the most informative repairing actions. The third heuristic prefers not to change is-a relations in the original ontology into equivalence relations. For details we refer to [7].

The user can also ask for recommendations by clicking the *Recommend* button. In our case, the system recommends to add an is-a relation between *limb joint* and *joint*.

In general, the system presents a list of recommendations. By selecting an element in the list, the concepts in the repairing action are highlighted in the panels. The user can repair a missing is-a relation by selecting a concept in the left panel and a concept in the right panel and clicking on the *Repair* button. The is-a relation is then added to the ontology, and may lead to updates for other missing is-a relations. At all times during the process the user can inspect the ontology by clicking the *Show Ontology* button (see figure 2). Newly added is-a relations will be highlighted. After adding the is-a relation between *limb joint* and *joint*, not only (*wrist joint*, *joint*) is repaired, but all other missing is-a relations in our example as well, as they can be derived in the extended ontology. The list of missing is-a relations is therefore updated to be empty. After completing the repair of all missing is-a relations, the repaired ontology can be exported into an OWL file by clicking the *Save Repair* button.


Load Ontology
Load/Derive Missing IS-A Relations
Compute Repairing Actions
Show Ontology
Save Repair
Recommended Repairing Actions
Recommend
Repair
MA_0001502: cranial synchondrosis
MA_0001503: postcranial synchondrosis
MA_0001501: synchondrosis
MA_0001500: syndesmosis
MA_0001499: gomphosis
MA_0001504: symphysis
MA_0000321: fibrous joint
MA_0000320: cartilaginous joint
MA_0000470: hip joint
MA_0000453: metacarpo-phalangeal joint
MA_0000460: wrist joint
MA_0001514: joint of vertebral body
MA_0000322: synovial joint
MA_0000614: forelimb joint
MA_0000459: shoulder joint
MA_0000319: joint
MA_0000463: ankle joint
MA_0000691: limb joint
MA_0001513: joint of vertebral arch
MA_0001505: joint of girdle
MA_0000451: elbow joint
MA_0001507: sternoclavicular joint
MA_0000471: knee joint
MA_0001512: sternocostal joint
MA_0001508: joint of rib
MA_0001506: sacroiliac joint
MA_0001511: interchondral joint
MA_0001509: costochondral joint
MA_0001510: costovertebral joint

Missing IS-A Relations

MA_0000460: wrist joint -> MA_0000319: joint

MA_0000460: wrist joint -> MA_0000319: joint
MA_0000459: shoulder joint -> MA_0000319: joint
MA_0000451: elbow joint -> MA_0000319: joint
MA_0000470: hip joint -> MA_0000319: joint
MA_0000463: ankle joint -> MA_0000319: joint
MA_0000471: knee joint -> MA_0000319: joint
MA_0000453: metacarpo-phalangeal joint -> MA_0000319: joint

Fig. 1. Screenshot.

3 Discussion

In [7] we showed how to repair the two ontologies from the 2008 Anatomy track in OAEI. For MA our basic algorithm generates for 15 missing is-a relations only 1 repairing action (the missing is-a relation itself). Therefore these could be immediately repaired. For NCI-A this number is 8. Of the remaining missing is-a relations there are 65 for MA that have only 1 element in the Source and 2 that have 1 element in the

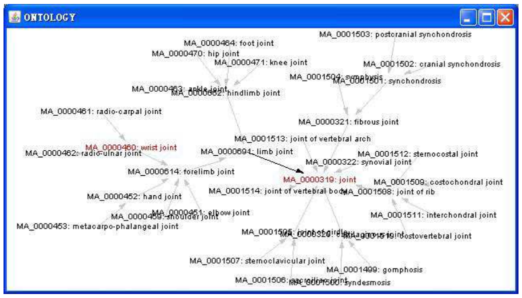

ONTOLOGY
MA_0001503: postcranial synchondrosis
MA_0000464: foot joint
MA_0000470: hip joint
MA_0000471: knee joint
MA_0001502: cranial synchondrosis
MA_0001504: symphysis
MA_0001501: synchondrosis
MA_0000463: ankle joint
MA_0000662: hindlimb joint
MA_0000461: radio-carpal joint
MA_0000321: fibrous joint
MA_0001513: joint of vertebral arch
MA_0000460: wrist joint
MA_0000462: radio-ulnar joint
MA_0000694: limb joint
MA_0001512: sternocostal joint
MA_0000322: synovial joint
MA_0000614: forelimb joint
MA_0000319: joint
MA_0001509: costochondral joint
MA_0001514: joint of vertebral body
MA_0001508: joint of rib
MA_0000452: hand joint
MA_0000451: elbow joint
MA_0000459: shoulder joint
MA_0001511: interchondral joint
MA_0000453: metacarpo-phalangeal joint
MA_0001505: joint of girdle
MA_0000320: cartilaginous joint
MA_0001510: costovertebral joint
MA_0001507: sternoclavicular joint
MA_0001499: gomphosis
MA_0001506: sacroiliac joint
MA_0001500: syndesmosis

Fig. 2. The repaired ontology.

Target set. For NCI-A these numbers are 20 and 3, respectively. These are likely to be good starting points for repairing. For most of the missing is-a relations the Source and Target sets are small and thus can be easily visualized in the panels of our system. The running time for generating recommendations for all missing is-a relations was circa 40 minutes for MA and circa 1 hour for NCI-A. In our tool, however, we do not generate recommendations for all missing is-a relations at once, but only on demand for a particular missing is-a relation. For NCI-A the system recommended repairing actions other than the missing is-a relation itself, for only 5 missing is-a relations and each of those received one recommended repairing action. For MA 22 missing is-a relations received 1 recommended repairing action, 12 received 2 and 2 received 3. The recommendation can come from small sets of repairing actions or from large sets.

References

1. AMA. Adult mouse anatomical dictionary. http://www.informatics.jax.org/searches/AMA_form.shtml.
2. M Bada and L Hunter. Identification of OBO nonalignments and its implication for OBO enrichment. *Bioinformatics*, 24(12):1448–1455, 2008.
3. Q Ji, P Haase, G Qi, P Hitzler, and S Stadtmuller. RaDON - repair and diagnosis in ontology networks. In *Demo at the 6th European Semantic Web Conference*, pages 863–867, 2009.
4. A Kalyanpur. *Debugging and Repair of OWL Ontologies*. PhD thesis, University of Maryland, College Park, 2006.
5. A Kalyanpur, B Parsia, E Sirin, and B Cuenca-Grau. Repairing unsatisfiable concepts in OWL ontologies. In *Proc 3rd European Semantic Web Conference*, pages 170–184, 2006.
6. A Kalyanpur, B Parsia, E Sirin, and J Hendler. Debugging unsatisfiable classes in OWL ontologies. *Journal of Web Semantics*, 3(4):268–293, 2006.
7. Patrick Lambrix, Qiang Liu, and He Tan. Repairing the missing is-a structure of ontologies. In *Proc 4th Asian Semantic Web Conference*, 2009.
8. C Meilicke and H Stuckenschmidt. Anatomy track at the 2008 Ontology Alignment Evaluation Initiative. Anatomy at http://oaei.ontologymatching.org/2008/.
9. C Meilicke, H Stuckenschmidt, and A Tamilin. Repairing ontology mappings. In *Proc 22nd National Conference on Artificial Intelligence - AAAI*, pages 1408–1413, 2007.
10. NCI-A. National cancer institute - anatomy. http://www.cancer.gov/cancerinfo/terminologyresources/.
11. S Schlobach. Debugging and semantic clarification by pinpointing. In *Proc 2nd European Semantic Web Conference*, pages 226–240, 2005.