Semi-Automatic Revision of Formalized Knowledge

Nadejda Nikitina

Abstract. As the amount of available ontologies and their size grow, ontology reuse gains in importance. However, the online available formalized knowledge in many cases need a revision which can lead to a high manual effort. In this paper, we propose an approach to support the revision of ontologies. We show that our method reduces the manual effort measured in number of decisions that have to be made by an ontology engineer by up to 83%.

1 INTRODUCTION

Constructing ontologies for real-world knowledge-intensive applications is a highly time-consuming task. The revision of ontologies is a typical part of it, since semi-automatic ontology construction, but also an extensive modification of ontologies by a human expert are error-prone tasks. One of the most practical usage scenarios for ontology revision however is ontology reuse. There are several reasons for the required manual inspection of the semantic data potentially relevant for the reuse in a new application context:

1. Ontologies often contain fragments irrelevant to the particular application scenario. Therefore, the relevant fragment has to be identified and extracted if necessary.
2. The available relevant knowledge bases tend to overlap. For instance, 25% of the available ontologies within the biomedical and chemical domain have an ontology mapping for more than the half of their concepts [2].
3. The conceptual compatibility of the selected relevant fragment and the target ontology has to be verified, since ontologies often model the same domain from different points of view and lexically similar entities can have different logical characteristics.

We propose a strategy of ontology revision support and evaluate it in an ontology reuse scenario, where the knowledge from several overlapping sources needs to be reused in a particular application context that is specified by a target ontology $T$. As we show in our experiments, a proposed reasoning-based approach can reduce the effort spent on the revision of semantic resources by up to 83%.

2 REVISION OF KNOWLEDGE BASES

When considering the reuse of knowledge from foreign ontologies, a decision needs to be taken for each of its axioms whether it complies with the requirements underlying $T$ such as requirements concerning the logical expressiveness within the ontology or the exact meaning of ontology entities. The proposed revision process allows an ontology engineer to select and reuse any part of the externally specified knowledge. During the revision, the expert reviews the axioms one by one while after each evaluation decision some axioms are evaluated automatically and disappear from the revision list. The automatic evaluation is based on the following assumptions:

- The requirements for the resulting ontology are known to the expert reviewing the ontology.
- The expert can only approve or decline axioms during the revision.
- Evaluation decisions cannot be changed during the revision.
- If an unevaluated axiom contradicts with the already approved ones, the expert would decline it.
- If an unevaluated axiom relationship is entailed by the already approved ones, the expert cannot decline it anymore.

The last two assumptions were adopted from the research by Meilicke et al. on ontology mapping revision [3].

The following succession operator underlies the automatic evaluation of axioms and incorporates the assumptions stated above.

Definition 1 (Succession operator) Let $O$ be a set of axioms that have to be reviewed to verify their compatibility with the target ontology $T$, and let $V$ be the set \{approved, declined, unevaluated\} of evaluation values. Let $\Omega$ denote the set of all possible axiom sets satisfying the SHOIQ syntax restrictions. The evaluation state function $f_T : O \cup T \rightarrow V$ with $\forall \gamma \in T \ (f_T(\gamma) = \text{approved})$ can be transformed into a more advanced evaluation state function using the succession operator $\Phi : V^{O\cup T} \rightarrow V^{O\cup T}$ as follows:

$$\Phi(f_T)(\alpha) = \begin{cases} \text{approved} & \text{if } f_T^{-1}(\text{approved}) \models \alpha \\ \text{declined} & \text{if } f_T^{-1}(\text{approved}) \cup \{\alpha\} \models \beta, \\ & \beta \in f_T^{-1}(\text{declined}) \cup \{\bot\} \\ \text{or } f_T^{-1}(\text{approved}) \cup \{\alpha\} \notin \Omega \\ \text{unevaluated} & \text{otherwise} \end{cases}$$

$\alpha \succ \beta$, iff $f_T^{-1}(\text{approved}) \cup \{\alpha\} \models f_T^{-1}(\text{approved}) \cup \{\beta\}$

Notice that due to the monotonicity of reasoning in SHOIQ, $\Phi$ preserves the values \{approved, declined\} assigned by $f_T$ to the axioms and only influences the evaluation values of axioms with $f_T(\alpha) = \text{unevaluated}$.

3 RANKING

The impact of reasoning-based support depends on the order in which axioms are evaluated. One possibility to rank the axioms is to determine a minimal, logically non-redundant subset within the total set of unevaluated axioms which can be used to deduce the remaining unevaluated axioms. If no axioms are declined, an evaluation of the minimal set would suffice. Therefore, the minimal set should be ranked higher than the remainder in order to insure that it will be evaluated first. The ranking technique MINSETRANK is an approximation of this idea. By the means of the reduction rules shown in
We evaluate the proposed methodology in the scenario where search engine results are considered for reuse in a particular context represented by a target ontology. The technical details about the revision-based ontology reuse approach that is deployed in this evaluation can be found in [4]. Each of the five ontologies from the OntoFarm dataset [5] shown in Table 1 is used as a target ontology in an experiment. We obtain the potential axioms for the reuse by the means of the ontology search engine Watson [1] using the name of each entity referenced in the considered target ontology as a keyword. The relative effort reduction shown in Table 1 is calculated as

$$\varrho(O) = 100\% \cdot \frac{\#(O) - \varepsilon(O)}{\#(O)}$$

where $\varepsilon(O)$ is the number of axioms that have to be evaluated by a human expert and $\#(O)$ the number of axioms considered in the revision. In order to measure $\varepsilon(O)$, we run a simulation of the evaluation where a virtual expert evaluates the axioms. We explicitly measure the effect of ranking within the same scenario and therefore repeat the same procedure but without ranking and sorting. As you can see in Table 1, we were able to reduce the effort of the evaluation by up to 83%. The ranking and sorting of axioms results in an average improvement of 25% over the non-ranking-based reasoning support.

**ACKNOWLEDGEMENTS**

This work is supported by the German Federal Ministry of Education and Research (BMBF) under the SAW-project NanOn.

**REFERENCES**


