

Pattern-based Ontology Construction

Eva Blomqvist

Jönköping University, Jönköping, Sweden
blev@jth.hj.se

Abstract. Large and complex enterprise systems face the same kind of information processing problems that exist on the web in general, and constructing an ontology is a crucial part of many solutions. Construction of enterprise ontologies needs to be at least semi-automatic in order to reduce the effort required, and another important issue is to introduce further knowledge reuse in the process. In order to realise these ideas the proposed research focuses on semi-automatic ontology construction, based on the methodology of case-based reasoning.

1 Introduction

When developing semantic applications for enterprises, constructing the enterprise application ontologies is a crucial part. Manual ontology engineering is a tedious and complex task. Another issue is knowledge reuse, common practises of the business world should be exploited as well as drawing on best practises in ontology engineering. By combining patterns with a case-based reasoning view, we aim at developing a novel semi-automatic ontology construction approach.

2 Background and Related Work

Our research focuses on application ontologies within enterprises, mainly for structuring and retrieval of information. We view an ontology design pattern as an ontology template, which is self-contained, comprised of a set of consistent ontology primitives, and intended to construct a part of some ontology. Related work on ontology patterns focus mainly on templates for manual use (like in [1]).

Recent developments in ontology engineering involve ontology learning (OL) as in [2], [3] and [4]. A major problem is that much of the information in a company is not explicitly stated, this is one issue where patterns can be of assistance. Case-based reasoning (CBR) is trying to mimic human behaviour, using previous experience to solve new problems. A case is a problem situation, previously solved cases are stored in the case base for reuse. The CBR process is viewed as a cycle of four phases: retrieval, reuse, revision and retaining cases.

3 Research Hypotheses

In our research some specific research questions have been derived, that can then be reformulated as the following hypotheses:

- CBR gives a framework for further automation of the ontology construction process, compared to related semi-automatic approaches that exist today.
- Using the CBR methodology (with patterns) can improve the quality of the generated ontologies compared to existing semi-automatic approaches.
- Automation reduces the total construction effort.
- Domain knowledge and engineering experience can be reused through patterns.

To verify the hypotheses the proposed method must be evaluated and compared to manual approaches as well as the related OL approaches stated earlier. The result produced by the method must be evaluated and shown to be of better quality compared to the result of related semi-automatic approaches.

4 Proposed Approach

The basis of a CBR approach is the case base and its content. In our approach the case base corresponds to a pattern catalogue (pattern base). The design patterns are represented as small ontologies and the architecture patterns are sets of constraints on the combination of design patterns, and may also include connections to specific design patterns.

The retrieval phase constitutes the process of analysing the input text corpus and deriving its representation, then matching this to the pattern base and selecting appropriate patterns. The reuse phase concerns the reuse and adaptation of the patterns, combining them into a first ontology. The revision phase includes extending the ontology, based on evaluation results. Retaining patterns includes the discovery of new patterns as well as improving existing patterns.

In our approach there is uncertainty inherent in all the described steps. For example each ontology primitive of the input representation have a certain degree of confidence associated, and the patterns are in themselves associated with a certain level of confidence. The levels of confidence are transferred onto the constructed initial ontology and can be used when evaluating it.

The main contributions of this approach is envisioned as both further automation of the ontology construction process, but in addition an increased quality of the produced ontology, as compared to other existing OL approaches. This increased quality will mainly be due to the use of patterns and the presence of an evaluation and revision phase in the method.

References

1. Gangemi, A.: Ontology Design Patterns for Semantic Web Content. In: Proceedings of ISWC 2005. Volume 3729 of LNCS., Springer (2005) 262–276
2. Cimiano, P.: Ontology Learning and Population from Text: Algorithms, Evaluation and Applications. Springer Science (2006)
3. Fortuna, B., Grobelnik, M., Mladenic, D.: Semi-automatic Data-driven Ontology Construction System. In: Proc. of IS-2006, Ljubljana, Slovenia (2006)
4. Iria, J., Brewster, C., Ciravegna, F., Wilks, Y.: An Incremental Tri-partite Approach to Ontology Learning. In: Proc. of LREC2006, Genoa (2006)