Ontology Modularization for Knowledge Selection: Experiments and Evaluations

Mathieu d’Aquin¹, Anne Schlicht², Heiner Stuckenschmidt², and Marta Sabou¹

¹ Knowledge Media Institute (KMi), The Open University, Milton Keynes, UK
{m.daquin, r.m.sabou}@open.ac.uk
² University of Mannheim, Germany
{anne, heiner}@informatik.uni-mannheim.de

Abstract. Problems with large monolithical ontologies in terms of reusability, scalability and maintenance have led to an increasing interest in modularization techniques for ontologies. Currently, existing work suffers from the fact that the notion of modularization is not as well understood in the context of ontologies as it is in software engineering. In this paper, we experiment on applying state-of-the-art tools for ontology modularization in the context of a concrete application: the automatic selection of knowledge components to be used for Web page annotation and semantic browsing. We conclude that, in a broader context, an evaluation framework is required to guide the choice of a modularization tool, in accordance with the requirements of the considered application.

Keywords: Ontology modularization, partitioning, module extraction.

1 Introduction

Modularization is a crucial task to allow ontology reuse and exploitation on the Semantic Web. The notion of modularization comes from Software Engineering where it refers to a way of designing software in a clear, well structured way that supports maintenance and reusability. From an ontology engineering perspective, modularization should be considered as a way to structure ontologies, meaning that the construction of a large ontology should be based on the combination of self-contained, independent and reusable knowledge components. In reality, even if they implicitly relate several sub-domains, most of the ontologies are not structured in a modular way. Therefore, in order to facilitate the management and the exploitation of such ontologies, ontology modularization techniques are required to identify and extract significant modules in existing ontologies.

While there is a clear need for modularization, there are no well-defined and broadly accepted definitions of modularity for ontologies. Several approaches have been recently proposed to extract modules from ontologies, each of them implementing its own intuition about what a module should contain and what

* This work is partially funded by the Open Knowledge (IST-FF6-027253) and NeOn projects (IST-FF6-027595), and partially supported by the German Science Foundation under contract STU 266/1 as part of the Emmy-Noether Program.
should be its qualities, generally without making this intuition explicit. This lack of consensus and of clarity hinders the application of these techniques in concrete scenarios, leading to difficulties in choosing the appropriate one. Moreover, to our knowledge, no other study has focused on the evaluation and comparison of ontology modularization techniques.

Our hypothesis is that there is no universal way to modularize an ontology and that the choice of a particular technique should be guided by the requirements of the considered application. We believe that modularization criteria should be defined in terms of the applications for which the modules are catered. For this reason, we detail in this paper some experiments conducted with several ontology modularization tools on a particular application: the selection of relevant knowledge components from online available ontologies. The goal is to characterize the requirements of this particular application using criteria from the literature on ontology modularization, and thus, to analyze the results of existing ontology modularization techniques regarding these requirements. In this way, we aim at better understanding the fundamental assumptions underlying the current modularization techniques. This work can be seen as a first step towards a broader framework, guiding application developers in choosing the appropriate technique and the designers of techniques in future developments.

The paper is structured as follows. Section 2 briefly describes the concrete scenario in which we apply modularization techniques. Section 3 and Section 4 respectively overview ontology modularization techniques and evaluation criteria that have been proposed in the literature. In Section 5 we evaluate, using the considered criteria, the results of the application of modularization techniques on our case-study. We conclude in Section 6 on the need for a comprehensive evaluation framework for ontology modularizations.

2 A Case-Study for Modularization: The Knowledge Selection Scenario

Knowledge selection has been described in [1] as the process of selecting the relevant knowledge components from online available ontologies and has been in particular applied to the Magpie application. Magpie [2] is a Semantic Web browser, available as a browser plugin, in which instances of ontology classes are identified in the current Web page and highlighted with the color associated to each class. In our current work we are extending Magpie towards open semantic browsing in which the employed ontologies are automatically selected and combined from online ontologies. As such, the user is relieved from manually choosing a suitable ontology every time he wishes to browse new content. Such an extension relies on mechanisms that not only dynamically select appropriate ontologies from the Web, but also extract from these ontologies the relevant and useful parts to describe classes in the current Web page.

Our previous work and experiences in ontology selection [3] made it clear that modularization may play a crucial role in complementing the current selection techniques. Indeed, selection often returns large ontologies that are virtually