Empirical study of logic-based modules: Cheap is cheerful.


Summary: For ontology reuse and integration, a number of approaches have been devised that aim at identifying modules, i.e., suitably small sets of “relevant” axioms from ontologies. Here we consider three logically sound notions of modules: MEX modules, only applicable to inexpressive ontologies; modules based on semantic locality, a sound approximation of the first; and modules based on syntactic locality, a sound approximation of the second (and thus the first), widely used since these modules can be extracted from OWL DL ontologies in time polynomial in the size of the ontology. In this paper we investigate the quality of both approximations over a large corpus of ontologies, using our own implementation of semantic locality, which is the first to our knowledge. In particular, we show with statistical significance that, in most cases, there is no difference between the two module notions based on locality; where they differ, the additional axioms can either be easily ruled out or their number is relatively small. We classify the axioms that explain the rare differences into four kinds of “culprits” and discuss which of those can be avoided by extending the definition of syntactic locality. Finally, we show that differences between MEX and locality-based modules occur for a minority of ontologies from our corpus and largely affect (approximations of) expressive ontologies - this conclusion relies on a much larger and more diverse sample than existing comparisons between MEX and syntactic locality-based modules.

doi:10.1007/978-3-642-41335-3_6