Reachability Modules for the Description Logic $SRIQ$

Riku Nortje, Katarina Britz, and Thomas Meyer

Center for Artificial Intelligence Research, University of KwaZulu-Natal and CSIR
Meraka Institute, South Africa
{rnortje,abritz,tmeyer}@csir.co.za

Abstract. In this paper we investigate module extraction for the Description Logic $SRIQ$. We formulate modules in terms of the reachability problem for directed hypergraphs. Using inseparability relations, we investigate the module-theoretic properties of reachability modules and show by means of an empirical evaluation that these modules have the potential of being substantially smaller than syntactic locality modules.

1 Introduction

Description Logics (DLs) are widely used in ontological modeling. They form a family of knowledge representation languages that are mostly decidable fragments of first-order logic. Their formal semantics not only allow for the exchange of DL ontologies but provide support for reasoning — the computation of additional logical inferences from the facts stated explicitly in an ontology.

There are many different DLs, each differing in the expressivity of the language and the complexity of reasoning. In general, the more expressive a DL the more complex the reasoning associated with it. This allows the ontology modeller to choose, for the intended application, the best balance between language expressivity on the one hand and reasoning complexity on the other. The DL $SRIQ$ is an expressive language and is a subset $SROIQ$, the W3C OWL DL Web Ontology language.

Modularization plays an important part in the design and maintenance of large scale ontologies. Modules are loosely defined as subsets of ontologies that cover some topic of interest, where the topic of interest is defined by a set of symbols. Extracting minimal modules is computationally expensive and even undecidable for expressive DLs [4,5]. Therefore, the use of approximation techniques and heuristics play an important role in the efficient design of algorithms.

Syntactic locality [4,5], because of its excellent model-theoretic properties, has become an ideal heuristic and is widely used in a diverse set of algorithms [19,30]. Suntisrivaraporn [19] showed that, for the DL $\mathcal{EL}^+$, $\bot$-locality module extraction is equivalent to the reachability problem in directed hypergraphs. Nortje et al. [14,15] extended the reachability problem to include $\top$-$\bot$-locality and introduced bidirectional reachability modules as a subset of $\bot\top^{\ast}$-locality modules. This work was further extended to the DL $SROIQ$ by Nortje et al. [16]
who showed that extracting $\bot \top^*$-reachability modules is equivalent to extracting frontier graphs in hypergraphs. Reachability modules are not only of importance in hypergraph-based reasoning support for CBoxes [16], but are potentially smaller than syntactic locality modules.

In this paper we investigate the module-theoretic properties of reachability modules for the DL $\mathcal{SRIQ}$. We show that these modules are not self-contained or depleting but they are robust under vocabulary restrictions, vocabulary extensions, replacement and joins. By showing that reachability modules preserve all justifications for entailments, we show that depleting modules are sufficient for preserving all justifications but not necessary. This paper is an extended version of the paper presented at DL2013 [17].

In Section 2 we give a brief introduction to the DL $\mathcal{SRIQ}$, hypergraphs and modularization as defined by inseparability relations. Section 3 introduces a normal form for $\mathcal{SRIQ}$ CBoxes as well as the rules necessary to transform any such CBox to normal form. In Section 4 we introduce both $\bot$- and $\top$-reachability modules and investigate all their module theoretic properties in terms of inseparability relations. In Section 5 we show the results of an empirical evaluation of these modules. Lastly in Section 6 we conclude this paper with a short summary of the results.

2 Background

In Section 2.1 we give a brief introduction to DLs and modularization with specific focus on the DL $\mathcal{SRIQ}$ [9]. In Section 2.2 we give a brief introduction to modules and module theoretic properties.

2.1 The DL $\mathcal{SRIQ}$

The syntax and semantics of $\mathcal{SRIQ}$ is listed in Table 1. $N_C$ and $N_R$ denote disjoint sets of atomic concept names and role names. The set $N_R$ includes the universal role whilst $N_C$ excludes the $\top$ and $\bot$ concepts. For a complete definition of $\mathcal{SRIQ}$, refer to Horrocks et al. [9], and for Description Logics refer to [2].

In order to ensure decidability in $\mathcal{SRIQ}$ there are some restrictions on the use of roles. $R_1 \circ \ldots \circ R_n \sqsubseteq R$, where $n \geq 1$ and $R_i, R \in N_R$, is a role inclusion axiom (RIA). A role hierarchy is a finite set of RIAs. Here $R_1 \circ \ldots \circ R_n$ denotes a composition of roles where $R, R_i$ may also be an inverse role $R^-$. A role $R$ is simple if (i) it does not appear on the right-hand side of a RIA, or (ii) is the inverse of a simple role, or (iii) appears on the right-hand side of a RIA only if the left-hand side is a simple role. $\text{Ref}(R)$, $\text{Irr}(R)$ and $\text{Dis}(R, S)$, where $R, S$ are roles other than $U$, are role assertions. A set of role assertions is simple w.r.t. a role-hierarchy $H$ if each assertion $\text{Irr}(R)$ and $\text{Dis}(R, S)$ uses only simple roles w.r.t. $H$.

A strict partial order $\prec$ on $N_R$ is a regular order if, and only if, for all roles $R$ and $S$: $S \prec R$ iff $S^- \prec R$. Let $\prec$ be a regular order on roles. A RIA $w \sqsubseteq R$ is