



**HAL**  
open science

# Efficient Computation of General Modules for ALC Ontologies (Extended Version)

Hui Yang, Patrick Koopmann, Yue Ma, Nicole Bidoit

► **To cite this version:**

Hui Yang, Patrick Koopmann, Yue Ma, Nicole Bidoit. Efficient Computation of General Modules for ALC Ontologies (Extended Version). 2023. hal-04096870

**HAL Id: hal-04096870**

**<https://hal.science/hal-04096870v1>**

Preprint submitted on 13 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Efficient Computation of General Modules for $\mathcal{ALC}$ Ontologies (Extended Version)

Hui Yang<sup>1</sup>, Patrick Koopmann<sup>2</sup>, Yue Ma<sup>1</sup> and Nicole Bidoit<sup>1</sup>

<sup>1</sup>LISN, CNRS, Université Paris-Saclay

<sup>2</sup>KAI, Vrije Universiteit Amsterdam, Netherlands

yang@lisn.fr, p.k.koopmann@vu.nl, ma@lisn.fr, nicole.bidoit@lisn.fr

## Abstract

We present a method for extracting general modules for ontologies formulated in the description logic  $\mathcal{ALC}$ . A module for an ontology is an ideally substantially smaller ontology that preserves all entailments for a user-specified set of terms. As such, it has applications such as ontology reuse and ontology analysis. Different from classical modules, general modules may use axioms not explicitly present in the input ontology, which allows for additional conciseness. So far, general modules have only been investigated for lightweight description logics. We present the first work that considers the more expressive description logic  $\mathcal{ALC}$ . In particular, our contribution is a new method based on uniform interpolation supported by some new theoretical results. Our evaluation indicates that our general modules are often smaller than classical modules and uniform interpolants computed by the state-of-the-art, and compared with uniform interpolants, can be computed in a significantly shorter time. Moreover, our method can be used for, and in fact improves, the computation of uniform interpolants and classical modules.

## 1 Introduction

Ontologies are used to formalize terminological knowledge in many domains such as biology, medicine and the Semantic Web. Usually, they contain a set of statements (axioms) about concept and role names (unary and binary predicates). Using a formalization based on description logics (DLs) allows DL reasoners to infer implicit information from an ontology. Modern ontologies are often large and complex, which can make ontology engineering challenging. For example, as of 3 January 2023, the medical ontology SNOMED CT [Donnelly and others, 2006], used in the health-case systems of many countries, formalizes over 360,000 concepts, and the BioPortal repository of ontologies from the bio-medical domain [Noy *et al.*, 2009] currently hosts 1,043 ontologies that use over 14 million concepts. Often, one is not interested in the entire content of an ontology, but only in a fragment, for instance if one wants to reuse content from a larger ontology for a more specialized application, or to analyse what the

ontology states about a given set of names. In such a case, one may provide a *signature*  $\Sigma$  of concept and role names of interest, and would like to compute an ontology  $\mathcal{M}$  that captures all the logical entailments of the original ontology  $\mathcal{O}$  expressed using only the names in  $\Sigma$ . Our aim is to compute such an  $\mathcal{M}$  that is as simple as possible, in terms of number and size of axioms. This problem has received a lot of attention in the past years, and for an  $\mathcal{M}$  satisfying those requirements, one may use *modules*, *uniform interpolants*, or *general modules*. We discuss those approaches and their limitations, before presenting our idea.

*Modules* are subsets of an ontology that preserve entailments within a given signature. There is a variety of notions of modules and properties they can satisfy that have been investigated in the literature [Grau *et al.*, 2008; Konev *et al.*, 2009]. *Semantic modules* preserve all models of the ontology modulo the given signature  $\Sigma$ . This makes them undecidable already for light-weight DLs such as  $\mathcal{EL}$  [Konev *et al.*, 2013], which is why existing methods often compute approximations of minimal modules [Gatens *et al.*, 2014; Romero *et al.*, 2016]. A popular example are *locality-based modules*, which can be computed in a very short time [Grau *et al.*, 2008]. However, locality-based modules can be comparatively large, even if the provided signature is small [Chen *et al.*, 2014]. In contrast to semantic modules, *deductive modules* are decidable, and focus only on entailments in  $\Sigma$  that can be expressed in the DL under consideration. Practical methods to compute them are presented in [Koopmann and Chen, 2020; Yang *et al.*, 2023]. However, while the size of minimal deductive modules is often half of that of the corresponding locality-based modules, for the more expressive DL  $\mathcal{ALC}$ , those methods are time-consuming, and can also not always guarantee minimality. An approximation of modules are *extracts*, whose size is bounded by the user, but which may not preserve all entailments in the given signature [Chen *et al.*, 2017].

Since modules can only use axioms that occur in the original ontology, they may use names that are not part of the given signature. *Uniform interpolants*, on the other hand, only use names from the signature provided, which means their axioms do not necessarily occur in the input ontology. This makes them useful also for applications other than ontology reuse, such as for logical difference [Ludwig and Konev, 2014], abduction [Del-Pinto and Schmidt, 2019], in-

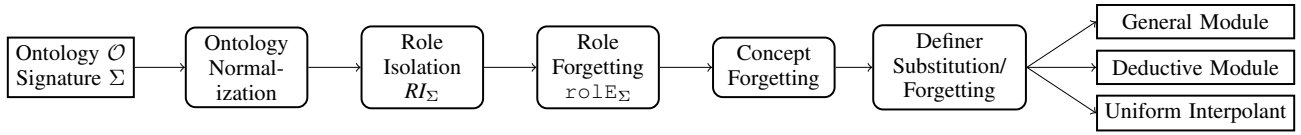


Figure 1: Overview of our unified method for computing general modules, deductive modules, and uniform interpolants

formation hiding [Grau, 2010], and proof generation [Alrabbaa *et al.*, 2020]. The strict requirement on the signature means that uniform interpolants may not always exist, and, in case of  $\mathcal{ALC}$ , can be of a size that is triple exponential in the size of the input [Lutz and Wolter, 2011]. Despite this high complexity, practical implementations for computing uniform interpolants exist [Zhao and Schmidt, 2018; Koopmann, 2020]. However, their computation times are much higher than for module extraction and can produce very complex axioms.

We present a method that is able to compute both uniform interpolants and deductive modules in a significantly shorter time than the state-of-the-art. At the center of our method is the computation of *general modules*, which generalize both uniform interpolants and deductive modules, and neither have to use only axioms from the input ontology, nor only names from the provided signature. This allows our general modules to be smaller than both modules and uniform interpolants, and thus better suited for applications such as ontology reuse. While general modules have been investigated for the lightweight DLs  $\mathcal{EL}$  and  $\mathcal{ELH}$  [Nikitina and Glimm, 2012; Alghamdi *et al.*, 2021], to our knowledge, this is the first time they are investigated for  $\mathcal{ALC}$ .

The main steps of our approach are shown in Figure 1. Our method essentially works by performing uniform interpolation on a normalized version of the input ontology (Section 3). During normalization, so-called *definer names* are introduced, which are eliminated in the final step. This idea is inspired by the method for uniform interpolation in [Koopmann, 2020]. However, different from this approach, we put fewer constraints on the normal form and do not allow the introduction of definers after normalization which changes the mechanism of uniform interpolation. As a result, our definer elimination step may reintroduce names eliminated during uniform interpolation, which is not a problem for the computation of general modules. In contrast, eliminating definers as done in [Koopmann, 2020] can cause an exponential blowup, and introduce concepts with the non-standard *greatest fixpoint* constructor [Calvanese and De Giacomo, 2003].

A particular challenge in uniform interpolation is the elimination of role names, for which existing approaches either rely on expensive calls to an external reasoner [Zhao *et al.*, 2019; Koopmann, 2020] or avoid the problem partially by introducing universal roles [Zhao and Schmidt, 2017; Koopmann and Chen, 2020], leading to results not representable in  $\mathcal{ALC}$ . We overcome this problem by introducing an intermediate step called *role isolation*, after which role names can be eliminated more efficiently (Section 4).

Our evaluation shows that all our methods, including the one for uniform interpolation, can compete with the run times of locality-based module extraction, while at the same time

resulting in substantially smaller ontologies (Section 7).

Our main contributions are: 1) the first method dedicated to general modules in  $\mathcal{ALC}$ , 2) a formal analysis of some properties of the general modules we compute, 3) new methods for module extraction and uniform interpolation that significantly improve the state-of-the-art, 4) an evaluation on real-world ontologies indicating the efficiency of our technique.

## 2 Preliminaries

We recall the DL  $\mathcal{ALC}$  [Baader *et al.*, 2017]. Let  $N_C = \{A, B, \dots\}$  and  $N_R = \{r, s, \dots\}$  be pair-wise disjoint, countably infinite sets of *concept* and *role names*, respectively. A *signature*  $\Sigma \subseteq N_C \cup N_R$  is a set of concept and role names. *Concepts*  $C$  are built according to the following grammar rules.

$$C ::= \top \mid A \mid \neg C \mid C \sqcap C \mid C \sqcup C \mid \exists r.C \mid \forall r.C \quad (1)$$

For simplicity, we identify concepts of the form  $\neg \neg C$  with  $C$ . In this paper, an *ontology*  $\mathcal{O}$  is a finite set of *axioms* of the form  $C \sqsubseteq D$ ,  $C$  and  $D$  being concepts. We denote by  $\text{sig}(\mathcal{O})/\text{sig}(C)$  the set of concept and role names occurring in  $\mathcal{O}/C$ , and we use  $\text{sig}_C(*)/\text{sig}_R(*)$  to refer to the concept/role names in  $\text{sig}(*)$ . For a signature  $\Sigma$ , a  $\Sigma$ -*axiom* is an axiom  $\alpha$  s.t.  $\text{sig}(\alpha) \subseteq \Sigma$ .

An *interpretation*  $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$  consists of a non-empty set  $\Delta^{\mathcal{I}}$  and a function  $\cdot^{\mathcal{I}}$  mapping each  $A \in N_C$  to  $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$  and each  $r \in N_R$  to  $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ . The interpretation function  $\cdot^{\mathcal{I}}$  is extended to concepts as follows:

$$\begin{aligned} \top^{\mathcal{I}} &= \Delta^{\mathcal{I}}, & (\neg C)^{\mathcal{I}} &= \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}, \\ (C \sqcap D)^{\mathcal{I}} &= C^{\mathcal{I}} \cap D^{\mathcal{I}}, & (C \sqcup D)^{\mathcal{I}} &= C^{\mathcal{I}} \cup D^{\mathcal{I}}, \\ (\exists r.C)^{\mathcal{I}} &= \{a \in \Delta^{\mathcal{I}} \mid \exists b \in \Delta^{\mathcal{I}} : (a, b) \in r^{\mathcal{I}}\}, \\ (\forall r.C)^{\mathcal{I}} &= \{a \in \Delta^{\mathcal{I}} \mid \forall b : (a, b) \in r^{\mathcal{I}} \rightarrow b \in C^{\mathcal{I}}\}. \end{aligned}$$

An axiom  $C \sqsubseteq D$  is *satisfied* by an interpretation  $\mathcal{I}$  ( $\mathcal{I} \models C \sqsubseteq D$ ) if  $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ .  $\mathcal{I}$  is a *model* of an ontology  $\mathcal{O}$  ( $\mathcal{I} \models \mathcal{O}$ ) if  $\mathcal{I}$  satisfies every axiom in  $\mathcal{O}$ .  $\mathcal{O}$  *entails* an axiom  $\alpha$  if  $\mathcal{I} \models \alpha$  for every model  $\mathcal{I}$  of  $\mathcal{O}$ . If  $\alpha$  holds in every interpretation, we write  $\models \alpha$  and call  $\alpha$  a *tautology*.

The *length*  $|\cdot|$  of *concepts* and *axioms* is defined inductively by  $|\top| = |A| = 1$ , where  $A \in N_C$ ,  $|C \sqcup D| = |C \sqcap D| = |C \sqsubseteq D| = |C| + |D|$ ,  $|\forall r.C| = |\exists r.C| = |C| + 1$ , and  $|\neg C| = |C|$ . Then, the *length of an ontology*, denoted  $\|\mathcal{O}\|$ , is defined by  $\|\mathcal{O}\| = \sum_{\alpha \in \mathcal{O}} |\alpha|$ .

A central notion for us is (deductive) inseparability [Konev *et al.*, 2009; Koopmann and Chen, 2020]. Given two ontologies  $\mathcal{O}_1$  and  $\mathcal{O}_2$  and a signature  $\Sigma$ ,  $\mathcal{O}_1$  and  $\mathcal{O}_2$  are  $\Sigma$ -*inseparable*, in symbols  $\mathcal{O}_1 \equiv_{\Sigma} \mathcal{O}_2$ , if for every  $\Sigma$ -axiom  $\alpha$ ,  $\mathcal{O}_1 \models \alpha$  iff  $\mathcal{O}_2 \models \alpha$ . In this paper, we are concerned with the computation of general modules, defined in the following.

$\mathcal{O}$ :	$A_1 \sqsubseteq \exists r. \exists s. B_1 \sqcup \exists r. B_2$	$B_1 \sqcap B_3 \sqsubseteq \perp$	$A_2 \sqsubseteq A_3 \sqcup \forall s. B_3$	$B_4 \sqsubseteq A_4$	$B_2 \sqsubseteq B_4$
$cl(\mathcal{O})$ :	$\neg A_1 \sqcup \exists r. D_1 \sqcup \exists r. D_3$	$\neg D_1 \sqcup \exists s. D_2$	$\neg D_2 \sqcup B_1$	$\neg D_3 \sqcup B_2$	$\neg A_2 \sqcup A_3 \sqcup \forall s. D_4$
	$\neg D_4 \sqcup B_3$	$\neg B_1 \sqcup \neg B_3$	$\neg B_2 \sqcup B_4$	$\neg B_4 \sqcup A_4$	
$RI_\Sigma(\mathcal{O})$ :	$\neg A_1 \sqcup \exists r. D_1 \sqcup \exists r. D_3$	$\neg D_1 \sqcup \exists s. D_2$	$\neg D_3 \sqcup B_2$	$A_2 \sqcup A_3 \sqcup \forall s. D_4$	$\neg B_1 \sqcup \neg B_3$
	$\neg B_2 \sqcup B_4$	$\neg B_4 \sqcup A_4$	$\neg D_2 \sqcup \neg D_4$		
$role_\Sigma(RI_\Sigma(\mathcal{O}))$ :	$\neg A_1 \sqcup \exists r. D_1 \sqcup \exists r. D_3$	$\neg D_3 \sqcup B_2$	$\neg B_1 \sqcup \neg B_3$	$\neg B_2 \sqcup B_4$	$\neg B_4 \sqcup A_4$
	$\neg D_2 \sqcup \neg D_4$	$\neg D_1 \sqcup \neg A_2 \sqcup A_3$			
$conE_\Sigma(role_\Sigma(RI_\Sigma(\mathcal{O})))$ :	$\neg A_1 \sqcup \exists r. D_1 \sqcup \exists r. D_3$	$\neg D_1 \sqcup \neg A_2 \sqcup A_3$	$\neg D_3 \sqcup A_4$		
$gm_\Sigma(\mathcal{O})$ :	$A_1 \sqsubseteq \exists r. \exists s. B_1 \sqcup \exists r. B_2$	$A_2 \sqcap \exists s. B_1 \sqsubseteq A_3$	$B_2 \sqsubseteq A_4$		
$gm_\Sigma^*(\mathcal{O})$ :	$A_1 \sqsubseteq \exists r. (\neg A_2 \sqcup A_3) \sqcup \exists r. A_4$				

Table 1: Ontologies generated throughout the running example.

**Definition 1** (General module). *Given an ontology  $\mathcal{O}$  and a signature  $\Sigma$ , an ontology  $\mathcal{M}$  is a general module for  $\mathcal{O}$  and  $\Sigma$  iff (i)  $\mathcal{O} \equiv_\Sigma \mathcal{M}$  and (ii)  $\mathcal{O} \models \mathcal{M}$ .*

Observe that an ontology is always trivially a general module of itself. What we are interested in in particular is general modules that are simpler than the input ontology, and more focussed on the provided signature. Two extreme cases of general modules are uniform interpolants and deductive modules.

**Definition 2** (Uniform interpolant & deductive module). *Let  $\mathcal{O}$  be an ontology,  $\Sigma$  a signature, and  $\mathcal{M}$  a general module for  $\mathcal{O}$  and  $\Sigma$ . Then, (i)  $\mathcal{M}$  is a uniform interpolant for  $\mathcal{O}$  and  $\Sigma$  if  $sig(\mathcal{M}) \subseteq \Sigma$ , and (ii)  $\mathcal{M}$  is a deductive module for  $\mathcal{O}$  and  $\Sigma$  if  $\mathcal{M} \subseteq \mathcal{O}$ .*

### 3 Ontology Normalization

Our method performs forgetting on a normalized view of the ontology, which is obtained via the introduction of fresh names as in [Koopmann, 2015]. An ontology  $\mathcal{O}$  is in *normal form* if every axiom is of the following form:

$$\top \sqsubseteq L_1 \sqcup \dots \sqcup L_n \quad L_i ::= A \mid \neg A \mid Qr.A,$$

where  $A \in \mathbb{N}_C$ , and  $Q \in \{\forall, \exists\}$ . We call the disjuncts  $L_i$  *literals*. For simplicity, we omit the “ $\top \sqsubseteq$ ” on the left-hand side of normalized axioms, which are regarded as *sets*, in order to avoid dealing with duplicated literals and order. As an example, the axiom  $A_2 \sqsubseteq A_3 \sqcup \forall s. B_3$  is equivalent to  $\neg A_2 \sqcup A_3 \sqcup \forall s. B_3$  in normal form.

An ontology  $\mathcal{O}$  is translated into normal form (or normalized) using standard transformations (see for example [Koopmann, 2015]). In particular, we replace concepts  $C$  occurring under role restrictions  $Qr.C$  by fresh names  $D$  taken from a set  $\mathbb{N}_D \subseteq \mathbb{N}_C \setminus sig_C(\mathcal{O})$  of *definers*. We use  $D, D', D_1, D_2, \dots$  to denote definers. For each introduced definer  $D$ , we remember the concept  $C_D$  that was replaced by it. We assume that distinct occurrences of the same concept are replaced by distinct definers. Thus, in the resulting normalization of  $\mathcal{O}$  denoted  $cl(\mathcal{O})$ , every literal  $Qr.D$  satisfies  $D \in \mathbb{N}_D$ , and for every  $D \in \mathbb{N}_D$ ,  $cl(\mathcal{O})$  contains at most one literal of the form  $Qr.D$ . Every definer  $D$  has to occur only in literals of the form  $\neg D$  or  $Qr.D$ , that is, positive literals of the form  $D$  are not allowed. Obviously, we require  $cl(\mathcal{O}) \equiv_{sig(\mathcal{O})} \mathcal{O}$ .

**Example 1.** *Let  $\mathcal{O}$  be the ontology defined in the first row of Table 1. By normalizing  $\mathcal{O}$ , we obtain the set  $cl(\mathcal{O})$  shown in*

*the second row of Table 1. The definers  $D_1, D_2$  and  $D_3$  in  $cl(\mathcal{O})$  replace the concepts  $C_{D_1} = \exists s. B_1, C_{D_2} = B_1$  and  $C_{D_3} = B_2$ , respectively.*

For a fixed normalization, we define a partial order  $\preceq_d$  over all introduced definers, which is defined as the smallest reflexive-transitive relation over  $\mathbb{N}_D$  s.t.

- $D' \preceq_d D$  if  $\neg D \sqcup C \in cl(\mathcal{O})$  and  $D' \in sig(C)$ .

Intuitively,  $D' \preceq_d D$  whenever  $C_{D'}$  is contained in  $C_D$ . In Example 1, we have  $D_2 \preceq_d D_1$ , since  $\neg D_1 \sqcup \exists s. D_2 \in cl(\mathcal{O})$ . Our normalization ensures that  $\preceq_d$  is acyclic.

In the following, we assume that the ontology  $\mathcal{O}$  and the signature  $\Sigma$  do not contain definers, unless stated otherwise.

## 4 Role Forgetting

An ontology  $\mathcal{M}$  is called a *role forgetting* for  $\mathcal{O}$  and  $\Sigma$  iff  $\mathcal{M}$  is a uniform interpolant for  $\mathcal{O}$  and  $\Sigma' = \Sigma \cup sig_C(\mathcal{O})$ . Existing methods to compute role forgetting either rely on an external reasoner [Zhao *et al.*, 2019; Koopmann, 2020] or use the *universal role*  $\nabla$  [Zhao and Schmidt, 2017; Koopmann and Chen, 2020]. The former approach can be expensive, while the latter produces axioms outside of  $\mathcal{ALC}$ . The normalization allows us to implement a more efficient solution within  $\mathcal{ALC}$ , which relies on an integrated reasoning procedure and an additional transformation step that produces so-called *role isolated ontologies*.

### 4.1 Role Isolated Ontologies

The main idea is to separate names  $A \in \mathbb{N}_C$  that occur with roles outside of the signature, using the following notations.

$$Rol(A, \mathcal{O}) = \{r \in sig(\mathcal{O}) \mid Qr.A \text{ appears in } \mathcal{O}, Q \in \{\forall, \exists\}\}$$

$$Out_\Sigma(\mathcal{O}) = \{A \in sig(\mathcal{O}) \mid Rol(A, \mathcal{O}) \not\subseteq \Sigma\}$$

**Definition 3** (Role-isolated ontology). *An ontology  $\mathcal{O}$  is role isolated for  $\Sigma$  if (i)  $\mathcal{O}$  is in normal form, and (ii) every axiom  $\alpha \in \mathcal{O}$  is of one of the following forms:*

- (c1)  $L_1 \sqcup \dots \sqcup L_n, L_i ::= \neg A$  with  $A \in Out_\Sigma(\mathcal{O})$  for all  $i$ ;
- (c2)  $L_1 \sqcup \dots \sqcup L_m, L_i ::= Qr.A \mid B \mid \neg B$  with  $r, A \in sig(\mathcal{O}), B \notin Out_\Sigma(\mathcal{O})$  for all  $i$ .

Thus, an axiom in a *role isolated* ontology falls into two disjoint categories: either (c1) it contains literals built only over concepts in  $Out_\Sigma(\mathcal{O})$  or (c2) it contains role restrictions or literals built over concepts outside  $Out_\Sigma(\mathcal{O})$ .

$$\begin{array}{l}
\text{A-Rule : } \frac{C_1 \sqcup A_1 \quad \neg A_1 \sqcup C_2}{C_1 \sqcup C_2} \\
\text{r-Rule : } \frac{C_1 \sqcup \exists r.D_1, \bigcup_{j=2}^n \{C_j \sqcup \forall r.D_j\}, K_D}{C_1 \sqcup \dots \sqcup C_n}, \\
\text{where } K_D = \neg D_1 \sqcup \dots \sqcup \neg D_n \text{ or } \neg D_2 \sqcup \dots \sqcup \neg D_n.
\end{array}$$

Figure 2: Inference rules for computing  $\mathcal{D}_\Sigma(\mathcal{O})$

**Example 2** (Example 1 cont'd). For  $\Sigma = \{r, A_1, A_2, A_3, A_4\}$ , we have  $\text{Out}_\Sigma(\text{cl}(\mathcal{O})) = \{D_2, D_4\}$ .  $\text{cl}(\mathcal{O})$  is not role isolated for  $\Sigma$  because of  $\neg D_2 \sqcup B_1$ .

Given an ontology, we compute its role isolated form using the following definition.

**Definition 4.** The role isolated form  $RI_\Sigma(\mathcal{O})$  of  $\mathcal{O}$  is defined as  $RI_\Sigma(\mathcal{O}) := \text{cl}_\Sigma(\mathcal{O}) \cup \mathcal{D}_\Sigma(\mathcal{O})$ , where

- $\text{cl}_\Sigma(\mathcal{O}) \subseteq \text{cl}(\mathcal{O})$  contains all  $\alpha \in \text{cl}(\mathcal{O})$  s.t. if  $\neg D$  is a literal of  $\alpha$ , then  $\text{Rol}(D', \text{cl}(\mathcal{O})) \subseteq \Sigma$  for all  $D' \in \mathbb{N}_D$  s.t.  $D \preceq_d D'$ .
- $\mathcal{D}_\Sigma(\mathcal{O})$  is the set of axioms  $\neg D_1 \sqcup \dots \sqcup \neg D_n$  s.t. (i)  $\text{cl}_\Sigma(\mathcal{O})$  contains axioms of the form  $C_1 \sqcup Q_1 r.D_1, C_2 \sqcup \forall r.D_2, \dots, C_n \sqcup \forall r.D_n$ , where  $r \in \mathbb{N}_R \setminus \Sigma$ ,  $Q_1 \in \{\forall, \exists\}$ , and (ii)  $\{D_1, \dots, D_n\}$  is a minimal set of definers s.t.  $\text{cl}(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$ .

Intuitively, if a definer  $D$  appears in  $\text{cl}_\Sigma(\mathcal{O})$ , then it should not depend on definers in  $\text{Out}_\Sigma(\mathcal{O})$ .

**Example 3** (Example 2 cont'd). We have:

- $\text{cl}_\Sigma(\mathcal{O}) = \text{cl}(\mathcal{O}) \setminus \{\neg D_2 \sqcup B_1, \neg D_4 \sqcup B_3\}$  because  $\text{Rol}(D_2, \text{cl}(\mathcal{O})) = \text{Rol}(D_4, \text{cl}(\mathcal{O})) = \{s\} \not\subseteq \Sigma$  and,
- $\mathcal{D}_\Sigma(\mathcal{O}) = \{\neg D_2 \sqcup \neg D_4\}$ .

**Theorem 1.**  $RI_\Sigma(\mathcal{O})$  is role isolated for  $\Sigma$  and we have  $\mathcal{O} \equiv_{\Sigma \cup \text{sig}_C(\mathcal{O})} RI_\Sigma(\mathcal{O})$ .

To compute  $\mathcal{D}_\Sigma(\mathcal{O})$ , we saturate  $\text{cl}(\mathcal{O})$  using the inference rules shown in Fig. 2, which is sufficient due to the following lemma.

**Lemma 1.** Let  $\mathcal{S}$  be the set of axioms  $\neg D_1 \sqcup \dots \sqcup \neg D_n$ , obtained by applying the rules in Fig. 2 exhaustively on  $\text{cl}(\mathcal{O})$ . Then, for all  $D_1, \dots, D_n \in \mathbb{N}_D$ , we have  $\text{cl}(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$  iff  $\neg D_{i_1} \sqcup \dots \sqcup \neg D_{i_k} \in \mathcal{S}$  for some subset  $\{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$ .

**Example 4** (Example 3 cont'd). The axiom  $\neg D_2 \sqcup \neg D_4$  in  $\mathcal{D}_\Sigma(\mathcal{O})$  is obtained by applying two A-Rule inferences:

$$\frac{\frac{\neg D_2 \sqcup B_1, \quad \neg B_1 \sqcup \neg B_3}{\neg D_2 \sqcup \neg B_3}, \quad \neg D_4 \sqcup B_3}{\neg D_2 \sqcup \neg D_4}$$

## 4.2 Role Forgetting for Role Isolated Ontologies

If  $\mathcal{O}$  is role isolated for  $\Sigma$ , a role forgetting for  $\mathcal{O}$  and  $\Sigma$  can be obtained using the *r-Rule* in Figure 2. Here, because  $\mathcal{O}$  may not be normalized and thus may contain axioms of the form  $C \sqcup Qr.A$  with  $A \notin \mathbb{N}_D$ , the concept names  $D_1, \dots, D_n$  in the *r-Rule* can include also concept names outside  $\mathbb{N}_D$ .

**Definition 5.**  $\text{role}_\Sigma(\mathcal{O})$  is the ontology obtained as follows:

1. apply the *r-Rule* exhaustively for each  $r \in \text{sig}_R(\mathcal{O}) \setminus \Sigma$ ,
2. remove all axioms containing some  $r \in \text{sig}_R(\mathcal{O}) \setminus \Sigma$ .

The second step ensures that all role names in the resulting ontology  $\text{role}_\Sigma(\mathcal{O})$  are in  $\Sigma$  and therefore, we have  $\text{sig}(\text{role}_\Sigma(\mathcal{O})) \subseteq \Sigma \cup \text{sig}_C(\mathcal{O})$ .

**Example 5** (Example 4 cont'd). For the ontology  $RI_\Sigma(\mathcal{O})$ , Table 1 (fourth row) shows  $\text{role}_\Sigma(RI_\Sigma(\mathcal{O}))$  which is obtained through the following two steps:

1. The new axiom  $\neg D_1 \sqcup \neg A_2 \sqcup A_3$  is generated by the *r-Rule* inference:

$$\frac{\neg D_1 \sqcup \exists s.D_2, \quad \neg A_2 \sqcup A_3 \sqcup \forall s.D_4, \quad \neg D_2 \sqcup \neg D_4}{\neg D_1 \sqcup \neg A_2 \sqcup A_3}$$

2. The two axioms  $\neg D_1 \sqcup \exists s.D_2, \quad \neg A_2 \sqcup A_3 \sqcup \forall s.D_4$  are removed because they contain  $s \in \text{sig}(RI_\Sigma(\mathcal{O})) \setminus \Sigma$ .

**Theorem 2.** If  $\mathcal{O}$  is role isolated for  $\Sigma$ , then  $\text{role}_\Sigma(\mathcal{O})$  is a role-forgetting for  $\mathcal{O}$  and  $\Sigma$ .

## 5 Computing General Modules via $\text{role}_\Sigma$

We compute our general module from  $\text{role}_\Sigma$  by forgetting also the concept names (obtaining a hopefully more concise representation), and eliminating all definers (ensuring the result is entailed by  $\mathcal{O}$ ).

### 5.1 Concept Forgetting

We say that an ontology  $\mathcal{M}$  is a *concept forgetting* for  $\mathcal{O}$  and  $\Sigma$  iff  $\mathcal{M}$  is a uniform interpolant for  $\mathcal{O}$  and the signature  $\Sigma' = \Sigma \cup \text{sig}_R(\mathcal{O}) \cup \mathbb{N}_D$ . A concept forgetting can be computed through the *A-Rule* in Figure 2.

**Definition 6.**  $\text{con}_\Sigma(\mathcal{O})$  is the ontology obtained as follows:

1. apply the *A-Rule* exhaustively for each  $A \in \text{sig}_C(\mathcal{O}) \setminus \Sigma$ ,
2. delete every axiom  $\alpha$  that contains  $A$  or  $\neg A$ , where  $A \in \mathbb{N}_C \setminus \Sigma$  and no axiom contains  $Qr.A$  for  $Q \in \{\forall, \exists\}$  and  $r \in \mathbb{N}_R$ .

**Example 6** (Example 5 cont'd). Table 1 (the 5th row) shows the axioms in  $\text{con}_\Sigma(\text{role}_\Sigma(RI_\Sigma(\mathcal{O})))$  obtained as follows.

1.  $\neg D_3 \sqcup B_4, \quad \neg B_2 \sqcup A_4$ , and  $\neg D_3 \sqcup A_4$  are first generated by applying the *A-Rule* on  $B_2$  and  $B_4$ .
2. Axioms containing  $B_i$  or  $\neg B_i$ ,  $i \in \{1, \dots, 4\}$ , are removed since  $B_i \notin \Sigma$ .  $\neg D_2 \sqcup \neg D_4$  is also removed because there are no literals of the form  $Qr.D_2$  or  $Qr.D_4$ .

The following is a consequence of [Zhao and Schmidt, 2017, Theorem 1].

**Theorem 3.** If  $\mathcal{O}$  is in normal form, then  $\text{con}_\Sigma(\mathcal{O})$  is a concept forgetting for  $\mathcal{O}$  and  $\Sigma$ .

Theorems 1, 2 and 3, give us the following corollary.

**Corollary 1.**  $\text{con}_\Sigma(\text{role}_\Sigma(RI_\Sigma(\mathcal{O}))) \equiv_\Sigma \mathcal{O}$ .

## 5.2 Constructing the General Module

Now, in order to obtain our general modules, we have to eliminate the definers from  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$ . To improve the results, we delete *subsumed axioms* (i.e., axioms  $\top \sqsubseteq C \sqcup D$  for which we also derived  $\top \sqsubseteq C$ ) and also simplify the axioms.

**Theorem 4.** *Let  $gm_\Sigma(\mathcal{O})$  be the ontology obtained from  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$  by*

- deleting subsumed axioms,
- replacing each definer  $D$  by  $C_D$ , and
- exhaustively applying  $C_1 \sqsubseteq \neg C_2 \sqcup C_3 \Rightarrow C_1 \sqcap C_2 \sqsubseteq C_3$  and  $C_1 \sqsubseteq \text{Qr}.\neg C_2 \sqcap C_3 \Rightarrow C_1 \sqcap \text{Qr}.C_2 \sqsubseteq C_3$ , where  $\exists = \forall$  and  $\forall = \exists$ .

Then,  $gm_\Sigma(\mathcal{O})$  is a general module for  $\mathcal{O}$  and  $\Sigma$ .

**Example 7** (Example 6 cont'd). *Table 1 (the 8th row) shows the general module  $gm_\Sigma(\mathcal{O})$ , which has been obtained using  $C_{D_1} = \exists s.B_1$  and  $C_{D_3} = B_2$ .*

Eliminating definers in this way may reintroduce previously forgotten names, which is why our general modules are in general not uniform interpolants. This has the advantage of avoiding the triple exponential blow-up caused by uniform interpolation (see Section 1). In contrast, the size of our result is at most single exponential in the size of the input.

**Proposition 1.** *For any ontology  $\mathcal{O}$  and signature  $\Sigma$ , we have  $\|gm_\Sigma(\mathcal{O})\| \leq 2^{\mathcal{O}(\|\text{cl}(\mathcal{O})\|)}$ . On the other hand, there exists a family of ontologies  $\mathcal{O}_n$  and signatures  $\Sigma_n$  s.t.  $\|\mathcal{O}_n\|$  is polynomial in  $n \geq 1$  and  $\|gm_{\Sigma_n}(\mathcal{O}_n)\| = n \cdot 2^{\mathcal{O}(\|\text{cl}(\mathcal{O}_n)\|)}$ .*

For some module extraction methods, such as for locality-based modules [Grau *et al.*, 2008], iterating the computation can lead to smaller modules. The following result shows that this is never the case for our method.

**Proposition 2.** *Let  $(\mathcal{M}_i)_{i \geq 1}$  be the sequence of ontologies defined by (i)  $\mathcal{M}_1 = gm_\Sigma(\mathcal{O})$  and (ii)  $\mathcal{M}_{i+1} = gm_\Sigma(\mathcal{M}_i)$  for  $i \geq 1$ . Then, we have  $\mathcal{M}_i \subseteq \mathcal{M}_{i+1}$  for  $i \geq 1$ . Moreover, there exists  $i_0 \geq 0$  s.t.  $\mathcal{M}^k = \mathcal{M}^{i_0}$  for all  $k \geq i_0$ .*

This property holds thanks to the substitution step of Theorem 4. This step may reintroduce in  $gm_\Sigma(\mathcal{O})$  concept and role names outside of  $\Sigma$ . As a result, the repeated application of  $\text{rolE}_\Sigma$  and  $\text{conE}_\Sigma$  on  $gm_\Sigma(\mathcal{O})$  can produce additional but unnecessary axioms. However, for ontologies in normal form, our method is *stable* in the sense that repeated applications produce the same ontology.

**Proposition 3.** *Let  $\mathcal{O}$  be an ontology in normal form and  $\mathcal{M} = gm_\Sigma(\mathcal{O})$ . Then,  $gm_\Sigma(\mathcal{M}) = \mathcal{M}$ .*

## 5.3 Optimizing the Result

The general module  $gm_\Sigma(\mathcal{O})$  may contain complex axioms since definers  $D$  can stand for complex concepts  $C_D$ . To make the result more concise, we eliminate some definers before substituting them. In particular, we use the following operations on  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$ , inspired by [Sakr and Schmidt, 2022].

*conD-Elim:*

$$\frac{C_1 \sqcup \text{Qr}.D_1, \bigcup_{j=2}^n \{C_j \sqcup \forall r.D_j\}, \neg D_1 \sqcup \dots \sqcup \neg D_n}{C_1 \sqcup \dots \sqcup C_n \sqcup \text{Qr}.\perp}$$

$$\text{D-Prop : } \frac{C_1 \sqcup \text{Qr}.D, \bigcup_{j=2}^n \{\neg D \sqcup C_j\}}{C_1 \sqcup \text{Qr}.(C_2 \sqcap \dots \sqcap C_n)}$$

where  $\bigcup_{j=2}^n \{\neg D \sqcup C_j\}$  ( $n \geq 1$ ) are all the axioms of the form  $\neg D \sqcup C$ . This rule is applicable only if no  $C_j$  contains definers.

Figure 3: Rules to eliminate definers

**Op1. Eliminating conjunctions of definers** aims to eliminate disjunctions of negative definers ( $\neg D_1 \sqcup \dots \sqcup \neg D_n$ ). This is done in two steps: (i) Applying the *conD-Elim* rule of Figure 3 on  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$ , and then (ii) deleting all axioms of the form  $\neg D_1 \sqcup \dots \sqcup \neg D_n$ .

**Op2. Eliminating single definers** aims to get rid of definers  $D$  that do not occur in axioms of the form  $\neg D \sqcup \neg D_1 \sqcup C$ . This is done in two steps: (i) applying the *D-Prop* rule of Figure 3 exhaustively and then (ii) deleting all axioms containing definers for which *D-Prop* has been applied.

**Theorem 5.** *Let  $gm_\Sigma^*(\mathcal{O})$  be the ontology obtained by:*

- successive application of *Op1* and *Op2* over  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$ , and then
- application of the steps described in Theorem 4.

Then,  $gm_\Sigma^*(\mathcal{O})$  is a general module for  $\mathcal{O}$  and  $\Sigma$ .

**Example 8.** *Assume  $\Sigma = \{r, A, A_1\}$  and*

$$\mathcal{O} = \{A \sqsubseteq \forall r.\exists s.B_1, A_1 \sqsubseteq \forall r.\forall s.B_2, B_1 \sqcap B_2 \sqsubseteq \perp\}.$$

Then,  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$  is:

$$\{\neg A \sqcup \forall r.D_1, \neg A_1 \sqcup \forall r.D_2, \neg D_1 \sqcup \neg D_2\},$$

where  $C_{D_1} = \exists s.B_1$ ,  $C_{D_2} = \forall s.B_2$ . And thus, by replacing  $D_i$  by  $C_{D_i}$ , we obtain  $gm_\Sigma(\mathcal{O}) =$

$$\{A \sqsubseteq \forall r.\exists s.B_1, A_1 \sqsubseteq \forall r.\forall s.B_2, \exists s.B_1 \sqcap \forall s.B_2 \sqsubseteq \perp\},$$

which is actually more intricate than  $\mathcal{O}$ . We can avoid this by applying the two optimizations described above.

The elimination of definer conjunctions (*Op1*) produces

$$\{\neg A \sqcup \forall r.D_1, \neg A_1 \sqcup \forall r.D_2, \neg A \sqcup \neg A_1 \sqcup \forall r.\perp\}. \quad (2)$$

(i) The first step of *Op1* applies the *conD-Elim* inference:

$$\frac{\neg A \sqcup \forall r.D_1, \neg A_1 \sqcup \forall r.D_2, \neg D_1 \sqcup \neg D_2}{\neg A \sqcup \neg A_1 \sqcup \forall r.\perp}.$$

(ii) The second step of *Op1* removes the axiom  $\neg D_1 \sqcup \neg D_2$ .

Then, the elimination of definers (*Op2*) produces

$$\{\neg A \sqcup \forall r.\top, \neg A_1 \sqcup \forall r.\top, \neg A \sqcup \neg A_1 \sqcup \forall r.\perp\} \quad (3)$$

by replacing  $D_1, D_2$  by  $\top$  as there is no axioms with negative  $D_i, i = 1, 2$  in Equation (2). Note that the first two axioms in Equation (3) are tautologies and thus can be ignored.

Finally, we have  $gm_\Sigma^*(\mathcal{O}) = \{A \sqcap A_1 \sqsubseteq \forall r.\perp\}$ .

## 6 Deductive Modules and Uniform Interpolants

**Deductive modules** Depending on the situation, users might prefer the axioms in the original ontology  $\mathcal{O}$  rather than newly introduced axioms (e.g., axioms in  $gm_{\Sigma}(\mathcal{O})$  or  $gm_{\Sigma}^*(\mathcal{O})$ ). For such situations, we can compute a deductive module for  $\mathcal{O}$  and  $\Sigma$  by tracing back the inferences performed when computing the general module  $gm_{\Sigma}^*(\mathcal{O})$ .

Let  $Res_{\Sigma}(\mathcal{O})$  be the set of all axioms generated by the computation progress of  $gm_{\Sigma}^*(\mathcal{O})$ . Clearly,  $gm_{\Sigma}^*(\mathcal{O}) \subseteq Res_{\Sigma}(\mathcal{O})$ . We iteratively construct a relation  $R$  on  $Res_{\Sigma}(\mathcal{O})$  during the computation  $gm_{\Sigma}^*(\mathcal{O})$  as follows: we start with  $R = \emptyset$ , and each time a new axiom  $\beta$  is generated from a premise set  $\{\alpha_1, \dots, \alpha_n\}$  (e.g., if  $\beta$  is obtained by applying r-Rule on  $\{\alpha_1, \dots, \alpha_n\}$ ), we add to  $R$  the relations  $\alpha_1 R \beta, \dots, \alpha_n R \beta$ . Let  $R^*$  be the smallest transitive closure of  $R$ . Then the deductive module is defined as follows.

**Theorem 6.** *Let us define  $dm_{\Sigma}(\mathcal{O})$  by*

$$dm_{\Sigma}(\mathcal{O}) = \{\alpha \in \mathcal{O} \mid \alpha R^* \beta \text{ for some } \beta \in gm_{\Sigma}^*(\mathcal{O})\}.$$

*Then,  $dm_{\Sigma}(\mathcal{O})$  is a deductive module for  $\mathcal{O}$  and  $\Sigma$ .*

**Uniform interpolants** While general modules can be a good alternative to uniform interpolants for ontology reuse, uniform interpolation has applications that require the ontology to be fully in the selected signature, as stated in the introduction. If instead of substituting definers  $D$  by  $C_D$ , we eliminate them using existing uniform interpolation tools, we can compute a uniform interpolant for the input.

When computing  $gm_{\Sigma}^*(\mathcal{O})$  for an ontology  $\mathcal{O}$  and a signature  $\Sigma$ , if all definers have been eliminated by Op1 and Op2 from Section 5.3 (as in Example 8), then  $gm_{\Sigma}^*(\mathcal{O})$  is indeed a uniform interpolant for  $\mathcal{O}$  and  $\Sigma$ . If this is not the case, we compute a uniform interpolant by forgetting the remaining definers using an existing uniform interpolation tool such as LETHE or FAME [Koopmann, 2020; Zhao and Schmidt, 2018]. As we see in Section 7, this allows us to compute uniform interpolants much faster than using the tool alone.

## 7 Evaluation

We implemented a prototype in Python (version 3.7) called GEMO. To validate our methods, we compare them with each other and with the state-of-the-art tools implementing module extraction and uniform interpolation methods for  $\mathcal{ALC}$ . As evaluation metrics, we are interested in run time, length of computed ontologies, and length of largest axiom in the result. All the experiments were run on a machine with an Intel Xeon Silver 4112 2.6GHz, 64 GiB of RAM, Ubuntu 18.04, and OpenJDK 11.

**Corpus** The ontologies used in our experiment are generated from the OWL Reasoner Evaluation (ORE) 2015 classification track [Parsia *et al.*, 2017] by the two following steps. First, we removed axioms outside of  $\mathcal{ALC}$  from each ontology in ORE 2015. Then, we kept the ontologies  $\mathcal{O}$  for which  $cl(\mathcal{O})$  contained between 100 and 100,000 names. This resulted in 222 ontologies.

$\top\perp^*$ -module	minM	LETHE	FAME	GEMO	gmLethe
100%	84.34%	85.27%	91.25%	97.34%	96.17%

Table 2: Success rate evaluation. The **first** (resp. **second**) best-performing method is highlighted in **red** (resp. **blue**).

**Signatures** For each ontology, we generated 50 signatures consisting of 100 concept and role names. As in [Koopmann and Chen, 2020], we selected each concept/role name with a probability proportional to their occurrence frequency in the ontology. In the following, a *request* is a pair consisting of an ontology and a signature.

**Methods** For each request  $(\mathcal{O}, \Sigma)$ , GEMO produced three different (general) modules  $gm_{\Sigma}(\mathcal{O})$ ,  $gm_{\Sigma}^*(\mathcal{O})$  and  $dm_{\Sigma}(\mathcal{O})$ , respectively denoted by gm, gm\*, and dm. gmLethe denotes the uniform interpolation method described in Section 6, where we used GEMO for computing gm\* and then LETHE for definer forgetting. In the implementation, for each request, we first extracted a locality based  $\top\perp^*$ -module [Grau *et al.*, 2008] to accelerate the computation. Since removing subsumed axioms as mentioned in Theorems 4 and 5 can be challenging, we set a time limit of 10s for this task.

We compared our methods with four different alternatives: (i)  $\top\perp^*$ -modules [Grau *et al.*, 2008] as implemented in the OWL API [Horridge and Bechhofer, 2011]; (ii) minM [Koopmann and Chen, 2020] that computes *minimal deductive modules* under  $\mathcal{ALCH}^{\nabla}$ -semantics; (iii) LETHE (version 0.6)<sup>1</sup>[Koopmann, 2020] and FAME<sup>2</sup> (version 1.0) [Zhao and Schmidt, 2018] that compute uniform interpolants.

**Success rate** We say a method *succeeds* on a request if it outputs the expected results within 600s. Table 2 summarizes the success rate for the methods considered. After the  $\top\perp^*$ -modules, our method GEMO had the highest success rate.

**Module length and run time** Because the axiom shape in general modules is arbitrary in  $\mathcal{ALC}$ , one can easily transform any ontology into an equivalent one with just one axiom. Thus, we chose to use length as defined in Section 2, rather than size, for our evaluation. Table 3 shows the length and run time for the requests on which all methods were successful (78.45% of all requests).

We observe that dm and gmLethe have the best overall performance: their results had a substantially smaller average length and were computed significantly faster than others. Note that the average size of results for dm was even smaller than that for minM. The reason is that minM preserves entailments over  $\mathcal{ALCH}^{\nabla}$ , while we preserve only entailments over  $\mathcal{ALC}$ . Therefore, the minM results may contain additional axioms compared to the  $\mathcal{ALC}$  deductive modules.

Comparing gm and gm\* regarding length lets us conclude that the optimization in Section 5.3 is effective. On the other hand, minM produced results of small length but at the cost of long computation times. FAME and  $\top\perp^*$ -module were quite time-efficient but less satisfactory in size, especially for FAME, whose results are often considerably larger than for

<sup>1</sup><https://lat.inf.tu-dresden.de/~koopmann/LETHE/>

<sup>2</sup><http://www.cs.man.ac.uk/~schmidt/sf-fame/>

Methods	Resulting ontology length	Time cost
minM	2355 / 392.59 / 264	595.88 / 51.82 / 8.86
$\perp^*$ -module	4,008 / 510.77 / 364	5.94 / 1.03 / 0.90
FAME	9,446,325 / 6,661.01 / 271	526.28 / 3.20 / 1.17
LETHE	131,886 / 609.30 / 196	598.20 / 49.21 / 13.57
gm	179,999 / 2,335.05 / 195	
GEMO gm*	21,891 / 466.15 / 166	17.50 / 2.44 / 1.63
dm	2,789 / 366.36 / 249	
gmLethe	21,891 / 364.10 / 162	513.15 / 3.08 / 1.68

Table 3: Comparison of different methods (max. / avg. / med.).

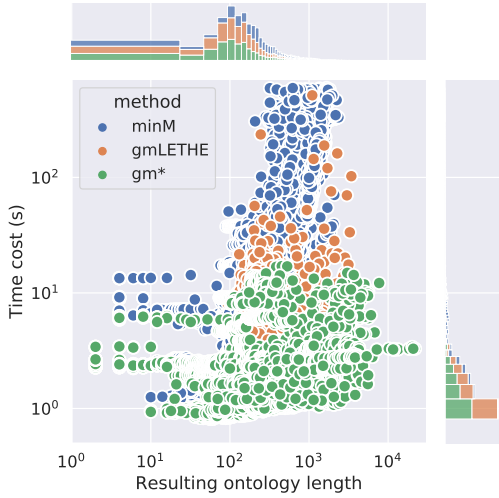


Figure 4: Comparison of minM, gmLETHE, and gm\* using the logarithmic scale.

the other methods. LETHE took more time than FAME, but produced more concise uniform interpolants on average.

For 87.87% of the requests reported in Table 3, gm\* already computed a uniform interpolant, so that gmLethe did not need to perform any additional computations.

Figure 4 provides a detailed comparison of minM, gm\* and gmLethe. gm\* was often faster but produced larger results. In contrast, gmLethe produced more concise results at the cost of longer computation time. While minM avoided large modules, it was generally much slower than our methods.

Table 4 summarizes the results concerning all requests for which GEMO (resp. gmLethe) was successful. We see that the results of dm had a small average size. However, as for gm\* and gmLethe, the median size of results was much smaller, which suggests that gm\* and gmLethe perform better over relatively simple cases.

**Uniform interpolants** For 80.23% of GEMO successful requests,  $gm^*_\Sigma(\mathcal{O})$  are uniform interpolants. In the cases where gm\* did not already produce a uniform interpolant, the success rate for gmLethe was 93.96%. In the cases where LETHE failed, the success rate for gmLethe was 36.23%.

The comparison of LETHE with gmLethe in Figure 5 shows that gmLethe was significantly faster than LETHE in most of the cases.

Methods	Resulting ontology length	Time cost
gm	17,335,040 / 35,008.2 / 310	
GEMO gm*	2,318,878 / 2,978.77 / 214	585.97 / 4.89 / 1.75
dm	18,218 / 638.74 / 309	
gmLethe	353,107 / 1,006.34 / 192	579.70 / 7.56 / 2.02

Table 4: GEMO and gmLethe: Summary of results for all their own successful experiments (max. / avg. / med.).

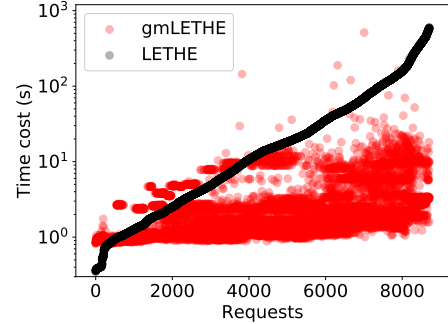


Figure 5: Run time comparison of LETHE and gmLethe using the logarithmic scale in time cost.

**Axiom size** A potential shortcoming of general modules compared to classical modules is that they could contain axioms that are more complex than those of the input, and thus be harder to handle by human end-users. For the requests reported in Table 3, the largest axiom in the output of minM had length 352, while for gm\*, it had length 5,815. In contrast, for the uniform interpolants computed by LETHE and FAME, the situation was much worse: here, the largest axiom had a length of 26,840 and 130,700, respectively, which is clearly beyond what can be understood by a human end-user. Besides these extreme cases, we can also observe differences wrt. the median values: for gm\* the longest axiom had a median length of 3, which is even lower than the corresponding value for minM (5), and, as expected, lower than for LETHE (4) and FAME (6). This indicates that, in most cases, general modules computed by gm\* are simpler.

## 8 Conclusion

This paper presents a new tool, GEMO, for computing general modules for  $\mathcal{ALC}$  ontologies, which can also be used for computing deductive modules and uniform interpolants. Our method is based on a new role isolation process that enables efficient role forgetting and an easy definer elimination. The experiments on real-world ontologies validate the efficiency of our proposal and the quality of the computed general modules.

In the future, we are interested in optimizing the concept elimination step to obtain more concise general modules. Also, we would like to investigate how to generalize our ideas to more expressive ontologies.



## Acknowledgements

Hui Yang, Yue Ma and Nicole Bidoit are funded by the BPI-France (PSPC AIDA: 2019-PSPC-09).

## References

- [Alghamdi *et al.*, 2021] Ghadah Alghamdi, Renate A. Schmidt, Warren Del-Pinto, and Yongsheng Gao. Upwardly abstracted definition-based subontologies. In Anna Lisa Gentile and Rafael Gonçalves, editors, *K-CAP '21: Knowledge Capture Conference*, pages 209–216. ACM, 2021.
- [Alrabbaa *et al.*, 2020] Christian Alrabbaa, Franz Baader, Stefan Borgwardt, Patrick Koopmann, and Alisa Kovtunova. Finding small proofs for description logic entailments: Theory and practice. In Elvira Albert and Laura Kovács, editors, *LPAR 2020: 23rd International Conference on Logic for Programming, Artificial Intelligence and Reasoning*, volume 73 of *EPiC Series in Computing*, pages 32–67. EasyChair, 2020.
- [Baader *et al.*, 2017] Franz Baader, Ian Horrocks, Carsten Lutz, and Ulrike Sattler. *An Introduction to Description Logic*. Cambridge University Press, 2017.
- [Calvanese and De Giacomo, 2003] Diego Calvanese and Giuseppe De Giacomo. Expressive description logics. In *The description logic handbook: theory, implementation, and applications*, pages 178–218. 2003.
- [Chen *et al.*, 2014] Jieying Chen, Michel Ludwig, Yue Ma, and Dirk Walther. Evaluation of extraction techniques for ontology excerpts. In *Description Logics*, volume 1193 of *CEUR Workshop Proceedings*, pages 471–482, 2014.
- [Chen *et al.*, 2017] Jieying Chen, Michel Ludwig, Yue Ma, and Dirk Walther. Zooming in on ontologies: Minimal modules and best excerpts. In *16th International Semantic Web Conference, Proceedings, Part I*, pages 173–189. Springer, 2017.
- [Del-Pinto and Schmidt, 2019] Warren Del-Pinto and Renate A. Schmidt. ABox abduction via forgetting in *ALC*. In *The Thirty-Third AAAI Conference on Artificial Intelligence*, AAAI, pages 2768–2775. AAAI Press, 2019.
- [Donnelly and others, 2006] Kevin Donnelly et al. SNOMED-CT: The advanced terminology and coding system for eHealth. *Studies in health technology and informatics*, 121:279, 2006.
- [Gatens *et al.*, 2014] William Gatens, Boris Konev, and Frank Wolter. Lower and upper approximations for deleting modules of description logic ontologies. In Torsten Schaub, Gerhard Friedrich, and Barry O’Sullivan, editors, *ECAI 2014 - 21st European Conference on Artificial Intelligence*, volume 263 of *Frontiers in Artificial Intelligence and Applications*, pages 345–350. IOS Press, 2014.
- [Grau *et al.*, 2008] B Cuenca Grau, Ian Horrocks, Yevgeny Kazakov, and Ulrike Sattler. Modular reuse of ontologies: Theory and practice. *Journal of Artificial Intelligence Research*, 31:273–318, 2008.
- [Grau, 2010] Bernardo Cuenca Grau. Privacy in ontology-based information systems: A pending matter. *Semantic Web*, 1(1-2):137–141, 2010.
- [Horridge and Bechhofer, 2011] Matthew Horridge and Sean Bechhofer. The OWL API: a java API for OWL ontologies. *Semantic Web*, 2(1):11–21, 2011.
- [Konev *et al.*, 2009] Boris Konev, Carsten Lutz, Dirk Walther, and Frank Wolter. Formal properties of modularisation. In Heiner Stuckenschmidt, Christine Parent, and Stefano Spaccapietra, editors, *Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization*, volume 5445 of *Lecture Notes in Computer Science*, pages 25–66. Springer, 2009.
- [Konev *et al.*, 2013] Boris Konev, Carsten Lutz, Dirk Walther, and Frank Wolter. Model-theoretic inseparability and modularity of description logic ontologies. *Artif. Intell.*, 203:66–103, 2013.
- [Koopmann and Chen, 2020] Patrick Koopmann and Jieying Chen. Deductive module extraction for expressive description logics. In Christian Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, pages 1636–1643, 2020.
- [Koopmann, 2015] Patrick Koopmann. *Practical uniform interpolation for expressive description logics*. PhD thesis, University of Manchester, UK, 2015.
- [Koopmann, 2020] Patrick Koopmann. LETHE: Forgetting and uniform interpolation for expressive description logics. *KI—Künstliche Intelligenz*, 34(3):381–387, 2020.
- [Ludwig and Konev, 2014] Michel Ludwig and Boris Konev. Practical uniform interpolation and forgetting for *ALC* tboxes with applications to logical difference. In Chitta Baral, Giuseppe De Giacomo, and Thomas Eiter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourteenth International Conference, KR 2014*. AAAI Press, 2014.
- [Lutz and Wolter, 2011] Carsten Lutz and Frank Wolter. Foundations for uniform interpolation and forgetting in expressive description logics. In Toby Walsh, editor, *IJ-CAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 989–995. IJ-CAI/AAAI, 2011.
- [Nikitina and Glimm, 2012] Nadeschda Nikitina and Birte Glimm. Hitting the sweetpot: Economic rewriting of knowledge bases. In Philippe Cudré-Mauroux, Jeff Heflin, Evren Sirin, Tania Tudorache, Jérôme Euzenat, Manfred Hauswirth, Josiane Xavier Parreira, Jim Hendler, Guus Schreiber, Abraham Bernstein, and Eva Blomqvist, editors, *The Semantic Web - ISWC 2012 - 11th International Semantic Web Conference, Proceedings, Part I*, volume 7649 of *Lecture Notes in Computer Science*, pages 394–409. Springer, 2012.
- [Noy *et al.*, 2009] Natalya Fridman Noy, Nigam H. Shah, Patricia L. Whetzel, Benjamin Dai, Michael Dorf, Nicholas Griffith, Clément Jonquet, Daniel L. Rubin, Margaret-Anne D. Storey, Christopher G. Chute, and

- Mark A. Musen. BioPortal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Res.*, 37(Web-Server-Issue):170–173, 2009.
- [Parsia *et al.*, 2017] Bijan Parsia, Nicolas Matentzoglou, Rafael S. Gonçalves, Birte Glimm, and Andreas Steigmiller. The OWL reasoner evaluation (ORE) 2015 competition report. *J. Autom. Reason.*, 59(4):455–482, 2017.
- [Romero *et al.*, 2016] Ana Armas Romero, Mark Kaminski, Bernardo Cuenca Grau, and Ian Horrocks. Module extraction in expressive ontology languages via datalog reasoning. *J. Artif. Intell. Res.*, 55:499–564, 2016.
- [Sakr and Schmidt, 2022] Mostafa Sakr and Renate A. Schmidt. Fine-grained forgetting for the description logic  $\mathcal{ALC}$ . In Ofer Arieli, Martin Homola, Jean Christoph Jung, and Marie-Laure Mugnier, editors, *Proceedings of the 35th International Workshop on Description Logics (DL 2022)*, volume 3263 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2022.
- [Simancik *et al.*, 2011] Frantisek Simancik, Yevgeny Kazakov, and Ian Horrocks. Consequence-based reasoning beyond Horn ontologies. In Toby Walsh, editor, *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 1093–1098. IJCAI/AAAI, 2011.
- [Yang *et al.*, 2023] Hui Yang, Yue Ma, and Nicole Bidoit. Efficient extraction of  $\mathcal{EL}$ -ontology deductive modules. In *The 37th AAAI Conference on Artificial Intelligence, AAAI*. in press, 2023.
- [Zhao and Schmidt, 2017] Yizheng Zhao and Renate A. Schmidt. Role forgetting for  $\mathcal{ALCOQH}(\nabla)$ -ontologies using an Ackermann-based approach. In Carles Sierra, editor, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017*, pages 1354–1361. ijcai.org, 2017.
- [Zhao and Schmidt, 2018] Yizheng Zhao and Renate A. Schmidt. FAME: an automated tool for semantic forgetting in expressive description logics. In *International Joint Conference on Automated Reasoning*, pages 19–27. Springer, 2018.
- [Zhao *et al.*, 2019] Yizheng Zhao, Ghadah Alghamdi, Renate A. Schmidt, Hao Feng, Giorgos Stoilos, Damir Juric, and Mohammad Khodadadi. Tracking logical difference in large-scale ontologies: A forgetting-based approach. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019*, pages 3116–3124. AAAI Press, 2019.

$$\begin{array}{l}
\mathbf{R}_A^+ \frac{}{H \sqsubseteq A} : A \in H \qquad \mathbf{R}_A^- \frac{H \sqsubseteq N \sqcup A}{H \sqsubseteq N} : \neg A \in H \\
\mathbf{R}_\sqcap^n \frac{\{H \sqsubseteq N_i \sqcup A_i\}_{i=1}^n}{H \sqsubseteq \sqcup_{i=1}^n N_i \sqcup N} : \left( \bigcap_{i=1}^n \neg A_i \right) \sqcup N \in cl_P(\mathcal{O}) \\
\mathbf{R}_\exists^+ \frac{H \sqsubseteq N \sqcup P}{H \sqsubseteq N \sqcup \exists r.D} : \neg P \sqcup \exists r.D \in cl_P(\mathcal{O}) \\
\mathbf{R}_\exists^- \frac{H \sqsubseteq N \sqcup \exists r.K, K \sqsubseteq \perp}{H \sqsubseteq N} \\
\mathbf{R}_\forall \frac{H \sqsubseteq N \sqcup \exists r.K, H \sqsubseteq N_1 \sqcup P}{H \sqsubseteq N \sqcup N_1 \sqcup \exists r.(K \sqcap D)} : \neg P \sqcup \forall r.D \in cl_P(\mathcal{O})
\end{array}$$

Table 5: Adaptation of the calculus from [Simancik *et al.*, 2011] to the syntax in  $cl_P(\mathcal{O})$ .

## 9 Proofs of Results from the Paper

The order in which we prove the results differs slightly from the order in which they appear in the main text. In particular, since the proof of Proposition 2 depends on some arguments used for Proposition 3, we show it after the proof for Proposition 3.

### 9.1 Lemma 1

Fix an ontology  $\mathcal{O}$ , and fix a set  $N_P \subseteq N_C$  of concept names such that  $N_P \cap sig(\mathcal{O}) = N_P \cap N_D = \emptyset$ . We will use the concept names  $P \in N_P$  as *placeholders* to substitute literals of the form  $Qr.D$  in  $cl(\mathcal{O})$ . In particular, we normalize the axioms in  $cl(\mathcal{O})$  further so that every axiom is of one of the following forms.

1.  $\neg P \sqcup Qr.D$ , where  $Q \in \{\forall, \exists\}$ , and  $P \in N_P$ ;
2. or  $L_1 \sqcup \dots \sqcup L_n$ , where each  $L_i$  is of the forms  $A$ , or  $\neg B$ , where  $A \in N_C \setminus N_D$ ,  $B \in N_C \setminus N_P$ .

Denote the resulting ontology by  $cl_P(\mathcal{O})$ . Similar as with the normalization introduced in Section 3, we can ensure that  $cl_P(\mathcal{O}) \equiv_{sig(cl(\mathcal{O}))} cl(\mathcal{O})$ .

In order to prove Lemma 1, we make use of the inference calculus introduced in [Simancik *et al.*, 2011], which can be seen in Table 5, adapted to our syntax of normalized axioms. Here,  $H$  and  $K$  stand for conjunctions of concept names (i.e.,  $A_1 \sqcap \dots \sqcap A_n$ ), and  $N$  for a disjunction of concept names (i.e.,  $A_1 \sqcup \dots \sqcup A_n$ ). We write

$$\mathcal{O} \vdash_P H \sqsubseteq \perp$$

if  $H \sqsubseteq \perp$  can be derived from  $\mathcal{O}$  using the rules in Table 5. We have the following theorem as the adaptation of [Simancik *et al.*, 2011, Theorem 1] to  $cl_P(\mathcal{O})$ .

**Theorem 7.** *Let  $\mathcal{O}$  be an ontology and  $H$  be any conjunction of concept names. Then,  $cl_P(\mathcal{O}) \models H \sqsubseteq \perp$  iff  $cl_P(\mathcal{O}) \vdash_P H \sqsubseteq \perp$ .*

In the following, we write

$$cl_P(\mathcal{O}) \vdash \neg D_1 \sqcup \dots \sqcup \neg D_n$$

if we can derive  $\neg D_1 \sqcup \dots \sqcup \neg D_n$  from  $cl_P(\mathcal{O})$  using applications of the A-Rule and the r-Rule. Before we show Lemma 1, we need an auxiliary lemma.

**Lemma 2.** *Let  $N = A_1 \sqcup \dots \sqcup A_m$  be an axiom s.t.  $A_j \in N_C$  for each  $1 \leq j \leq m$ , let  $D_1, \dots, D_n \in N_D$  be definers, and assume  $cl_P(\mathcal{O}) \vdash_P D_1 \sqcap \dots \sqcap D_n \sqsubseteq N$ . Then,  $D_1 \sqcap \dots \sqcap D_n \sqsubseteq N$  is a tautology, or  $cl_P(\mathcal{O}) \vdash N'$ , where  $N'$  is a disjunction over some subset of  $\{\neg D_1, \dots, \neg D_n, A_1, \dots, A_m\}$ .*

*Proof.* Let  $k_P(D_1 \sqcap \dots \sqcap D_n \sqsubseteq N)$  be the minimal number of applications of rules from Table 5 required to derive  $D_1 \sqcap \dots \sqcap D_n \sqsubseteq N$ . We prove the lemma by induction on  $k_P(D_1 \sqcap \dots \sqcap D_n \sqsubseteq N)$ .

1. If  $k_P(D_1 \sqcap \dots \sqcap D_n \sqsubseteq N) = 1$ , then we have  $N = D_i$  for some  $1 \leq i \leq n$ . In this case, the lemma holds directly.
2. Assume that the lemma holds for all  $D'_1, \dots, D'_{n'}, N'$  s.t.  $k_P(D'_1 \sqcap \dots \sqcap D'_{n'} \sqsubseteq N') < k_0$  for some  $k_0 \geq 1$ .

We show that the lemma also holds for  $D_1, \dots, D_n, N$  s.t.  $k_P(D_1 \sqcap \dots \sqcap D_n \sqsubseteq N) = k_0$ . For simplicity, let  $H_D := D_1 \sqcap \dots \sqcap D_n$ , and let

$$\rho_1, \dots, \rho_{k_0}$$

be a sequence of inferences generated by inference rules in Table 5 that derives  $H_D \sqsubseteq N$ . Then there are two different cases, depending on which was the last inference performed.

(a) The last inference  $\rho_{k_0}$  is of the form

$$\frac{\{H_D \sqsubseteq N_i \sqcup A_i\}_{i=1}^n}{H_D \sqsubseteq \bigsqcup_{i=1}^n N_i \sqcup N_0}, \left( \bigsqcup_{i=1}^n \neg A_i \right) \sqcup N_0 \in cl_P(\mathcal{O}).$$

In particular,  $N = \bigsqcup_{i=1}^n N_i \sqcup N_0$ . We show that we then also have  $cl_P(\mathcal{O}) \vdash D_1 \sqcap \dots \sqcap D_n \sqsubseteq N$ .

If  $H_D \sqsubseteq N_{i_0} \sqcup A_{i_0}$  is a tautology for some  $1 \leq i_0 \leq n$ , then one of  $H_D \sqsubseteq N_{i_0}$ ,  $H_D \sqsubseteq A_{i_0}$  must be a tautology. There are two cases:

(i) If  $H_D \sqsubseteq N_{i_0}$  is a tautology, then  $H_D \sqsubseteq N$  is also a tautology since  $N_{i_0}$  is a sub-concept of  $N$ . Therefore, the lemma holds for this case;

(ii) If  $H_D \sqsubseteq A_{i_0}$  is a tautology, then  $A_{i_0} \in \{D_1, \dots, D_n\}$  must be a definer. By the construction of  $cl_P(\mathcal{O})$ , we have  $n = 1$  (in the formula of inference  $\rho_{k_0}$ ) and thus  $\neg A_{i_0} \sqcup N_0 \in cl_P(\mathcal{O})$ . Then  $cl_P \models \neg A_{i_0} \sqcup N_0$ . Therefore, the lemma holds also for this case.

We obtain that the lemma holds for the case where  $H_D \sqsubseteq N_{i_0} \sqcup A_{i_0}$  is a tautology for some  $1 \leq i_0 \leq n$ .

Now assume that  $H_D \sqsubseteq N_i \sqcup A_i$  is not a tautology for any  $1 \leq i \leq n$ .

Since  $k(H_D \sqsubseteq N_i \sqcup A_i) < k_0$ , by applying our inductive hypothesis on  $H_D \sqsubseteq N_i \sqcup A_i$ ,  $i \in [1, n]$ , we have

$$\begin{aligned} cl_P(\mathcal{O}) \vdash K_D^i \sqcup N_i^{sub} \\ \text{or } cl_P(\mathcal{O}) \vdash K_D^i \sqcup N_i^{sub} \sqcup A_i \end{aligned}$$

for some  $K_D^i = \bigsqcup_{j \in I_i} \neg D_j$ ,  $I_i \subseteq \{1, \dots, n\}$  and  $N_i^{sub}$  being a disjunction over some concept names occurring in  $N_i$ . We distinguish those two cases.

(i) If  $cl_P(\mathcal{O}) \vdash K_D^i \sqcup N_i^{sub}$  for some  $1 \leq i \leq n$ , then  $cl_P(\mathcal{O}) \vdash K_D^i \sqcup N_i^{sub}$  is as desired.

(ii) Otherwise,  $cl_P(\mathcal{O}) \vdash K_D^i \sqcup N_i^{sub} \sqcup A_i$  for all  $1 \leq i \leq n$ . By applying the A-Rule for all  $A_i$ ,  $1 \leq i \leq n$  on

$$\begin{aligned} K_D^i \sqcup N_i^{sub} \sqcup A_i, \quad (1 \leq i \leq n) \\ \text{and } \left( \bigsqcup_{i=1}^n \neg A_i \right) \sqcup N_0 \in cl_P(\mathcal{O}), \end{aligned}$$

we obtain the desired conclusion

$$cl_P(\mathcal{O}) \vdash K_D \sqcup N^{sub},$$

where

$$\begin{aligned} K_D &= \bigsqcup_{1 \leq i \leq n} K_D^i, \text{ and} \\ N^{sub} &= \bigsqcup_{1 \leq i \leq n} N_i^{sub} \sqcup N_0. \end{aligned}$$

Therefore, the lemma holds for this case.

(b) The last inference  $\rho_{k_0}$  is generated by Rule  $\mathbf{R}_{\exists}^{\perp}$  and is of the form

$$\frac{H_D \sqsubseteq N \sqcup \exists r.K, K \sqsubseteq \perp}{H_D \sqsubseteq N}.$$

Note that  $H_D \sqsubseteq N \sqcup \exists r.K$  must be obtained by applying

• first an  $\mathbf{R}_{\exists}^{\perp}$  inference of the form

$$\frac{H_D \sqsubseteq N_0 \sqcup P_0}{H_D \sqsubseteq N_0 \sqcup \exists r.D'_0} : \neg P_0 \sqcup \exists r.D'_0 \in cl_P(\mathcal{O});$$

• followed by  $m$   $\mathbf{R}_{\forall}$  inferences of the form

$$\frac{H_D \sqsubseteq \left( \bigsqcup_{i=0}^{j-1} N_i \right) \sqcup \exists r.K_{j-1}, H_D \sqsubseteq N_j \sqcup P_j}{H_D \sqsubseteq \left( \bigsqcup_{i=0}^j N_i \right) \sqcup \exists r.(K_{j-1} \sqcap D'_j)},$$

where  $\neg P_j \sqcup \forall r.D'_j \in cl_P(\mathcal{O})$  and  $K_j = D'_0 \sqcap \dots \sqcap D'_j$  for  $1 \leq j \leq m$ . Moreover, we have

$$N = \left( \bigsqcup_{i=0}^m N_i \right), \quad K = K_m.$$

By applying the inductive hypothesis on  $K \sqsubseteq \perp$ , we obtain

$$cl_P(\mathcal{O}) \vdash \bigsqcup_{i \in I^*} \neg D'_i, \text{ for some } I^* \subseteq \{0, \dots, m\},$$

and by applying the inductive hypothesis on  $H_D \sqsubseteq N_j \sqcup P_j, j \in [0, m]$ , we obtain

$$\begin{aligned} cl_P(\mathcal{O}) \vdash H_D^j \sqcup N_j^{sub}, \\ \text{or } cl_P(\mathcal{O}) \vdash H_D^j \sqcup N_j^{sub} \sqcup P_j, \end{aligned}$$

for some  $H_D^j = \bigsqcup_{i \in I_i} \neg D_i$  with  $I_i \subseteq \{1, \dots, n\}$  and  $N_j^{sub}$  a disjunction of concepts from  $N_j$ . We again distinguish both cases.

(i) If  $cl_P(\mathcal{O}) \vdash H_D^j \sqcup N_j^{sub}$  for some  $0 \leq j \leq m$ , then  $cl_P(\mathcal{O}) \vdash H_D^j \sqcup N_j^{sub}$  directly holds.

(ii) Otherwise,  $cl_P(\mathcal{O}) \vdash H_D^j \sqcup N_j^{sub} \sqcup P_j$  for all  $0 \leq j \leq m$ . By applying the A-Rules for all  $P_j, 1 \leq j \leq m$  on

$$\begin{aligned} H_D^0 \sqcup N_0^{sub} \sqcup P_0, \neg P_0 \sqcup \exists r.D'_0 \\ H_D^j \sqcup N_j^{sub} \sqcup P_j, \neg P_j \sqcup \forall r.D'_j, \\ (1 \leq j \leq m); \end{aligned}$$

and applying the r-Rule on

$$\begin{aligned} \bigsqcup_{i \in I^*} \neg D'_i, \\ H_D^0 \sqcup N_0^{sub} \sqcup \exists r.D'_0, \\ H_D^j \sqcup N_j^{sub} \sqcup \forall r.D'_j, \\ (j \in I^* \cap \{1, \dots, m\}), \end{aligned}$$

we obtain

$$cl_P(\mathcal{O}) \vdash H_D \sqcup N^{sub},$$

where

$$\begin{aligned} H_D &= \bigsqcup_{j \in I^* \cup \{0\}} H_D^j, \text{ and} \\ N^{sub} &= \bigsqcup_{j \in I^* \cup \{0\}} N_j^{sub}. \end{aligned}$$

We obtain that the lemma also holds in this case.  $\square$

Using Theorem 7 and Lemma 2, we can now prove Lemma 1.

**Lemma 1.** *Let  $\mathcal{S}$  be the set of axioms  $\neg D_1 \sqcup \dots \sqcup \neg D_n$ , obtained by applying the rules in Fig. 2 exhaustively on  $cl(\mathcal{O})$ . Then, for all  $D_1, \dots, D_n \in \mathbb{N}_D$ , we have  $cl(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$  iff  $\neg D_{i_1} \sqcup \dots \sqcup \neg D_{i_k} \in \mathcal{S}$  for some subset  $\{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$ .*

*Proof.* For any definers  $D_1, \dots, D_n \in \mathbb{N}_D$ , since  $cl_P(\mathcal{O}) \equiv_{sig(cl(\mathcal{O}))} cl(\mathcal{O})$ , we have  $cl(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$  iff  $cl_P(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$ .

Note that we can exchange the order of application of a A-Rule on concept name  $P \in \mathbb{N}_P$  and other rules without influencing the final result. For instance: the following two rules that produce  $C_1 \sqcup \dots \sqcup C_n \sqcup Qr.D$ .

$$\text{(r-Rule): } \frac{C_1 \sqcup P \sqcup \exists r.D_1, \bigcup_{j=2}^n \{C_j \sqcup \forall r.D_j\}, K_D}{C_1 \sqcup \dots \sqcup C_n \sqcup P}, \quad \text{(A-Rule on } P\text{): } \frac{C_1 \sqcup \dots \sqcup C_n \sqcup P, \neg P \sqcup Qr.D}{C_1 \sqcup \dots \sqcup C_n \sqcup Qr.D}$$

We also obtain  $C_1 \sqcup \dots \sqcup C_n \sqcup Qr.D$  by following two rules, where a A-Rule on  $P \in \mathbb{N}_P$  is applied first.

$$\text{(A-Rule on } P\text{): } \frac{C_1 \sqcup P \sqcup \exists r.D_1, \neg P \sqcup Qr.D}{C_1 \sqcup Qr.D \sqcup \exists r.D_1}, \quad \text{(r-Rule): } \frac{C_1 \sqcup Qr.D \sqcup \exists r.D_1, \bigcup_{j=2}^n \{C_j \sqcup \forall r.D_j\}, K_D}{C_1 \sqcup \dots \sqcup C_n \sqcup Qr.D}$$

Therefore, when applying A-Rule and r-Rule on  $cl_P(\mathcal{O})$ , we could assume that A-Rules on concept names  $P \in \mathbb{N}_P$  are applied first. Since applying A-Rules on concept names  $P \in \mathbb{N}_P$  on  $cl_P(\mathcal{O})$  produces exactly the axioms in  $cl(\mathcal{O}) \setminus cl_P(\mathcal{O})$ , we have  $cl(\mathcal{O}) \vdash \neg D_1 \sqcup \dots \sqcup \neg D_n$  iff  $cl_P(\mathcal{O}) \vdash \neg D_1 \sqcup \dots \sqcup \neg D_n$  for any definers  $D_i \in \mathbb{N}_D$ .

$$\mathbf{r}\text{-Res} : \frac{C_1 \sqcup \exists r.D_1, C_2 \sqcup \forall r.D_2, \dots, C_n \sqcup \forall r.D_n}{C_1 \sqcup \dots \sqcup C_n},$$

where  $\mathcal{M} \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp, n \geq 1$ .

Figure 6: Rule for eliminating role name  $r$ .

It is thus enough to show that for any definers  $D_i \in \mathbb{N}_D$ ,  $cl_P(\mathcal{O}) \vdash_P D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$  iff  $cl_P(\mathcal{O}) \vdash \neg D_{i_1} \sqcup \dots \sqcup \neg D_{i_k}$  for some subset  $\{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$ .

We first prove the “ $\Leftarrow$ ” direction. If  $cl(\mathcal{O}) \vdash \neg D_{i_1} \sqcup \dots \sqcup \neg D_{i_k}$ , then we have  $cl(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$ . Consequently, by Theorem 7,  $cl_P(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$ , and thus  $cl_P(\mathcal{O}) \vdash_P D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$ .

The “ $\Rightarrow$ ” direction is a direct result of Lemma 2.  $\square$

Moreover, we have the following lemma that will be used in the proof of Theorem 1 in the next section.

**Lemma 3.** *Let  $t \in \mathbb{N}_R \setminus \Sigma$  be a role name and  $D_1, \dots, D_n \in \mathbb{N}_D$  be definers s.t.  $RI_\Sigma(\mathcal{O})$  contains a literal of the form  $Qt.D_1$ , and for  $2 \leq i \leq n$ ,  $\forall t.D_i$  occurs in  $RI_\Sigma(\mathcal{O})$ . Then,  $cl(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$  iff  $RI_\Sigma(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$ .*

*Proof.* If  $RI_\Sigma(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$ , then also  $cl(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$  because  $RI_\Sigma(\mathcal{O})$  consists only of axioms from  $cl(\mathcal{O})$  or axioms that have been derived from  $cl(\mathcal{O})$ . The other direction follows directly from the definition of  $\mathcal{D}_\Sigma(\mathcal{O})$  in Definition 4.  $\square$

## 9.2 Theorem 1

In order to prove Theorem 1, we make use of a result for role forgetting from [Zhao *et al.*, 2019], which describes the method for uniform interpolation used in FAME. In [Zhao *et al.*, 2019], the rule shown in Figure 6 is used for forgetting roles, which assumes the ontology to be in normal form, and uses a fixed ontology  $\mathcal{M}$  as a side condition.

Fix a signature  $\Sigma \subseteq sig(\mathcal{O})$ . We denote by  $Res_\Sigma(cl(\mathcal{O}), \mathcal{M})$  the ontology obtained by applying the following two operations on  $cl(\mathcal{O})$ :

1. apply **r-Res** exhaustively for all role names  $r \in sig(\mathcal{O}) \setminus \Sigma$ , and
2. remove all axioms that contain a role name  $r \in sig(\mathcal{O}) \setminus \Sigma$ .

By [Zhao *et al.*, 2019, Lemma 3], we have the following result.

**Lemma 4.** *For any ontology  $\mathcal{O}$  and signature  $\Sigma \subseteq sig(\mathcal{O})$ , we have  $Res_\Sigma(cl(\mathcal{O}), cl(\mathcal{O})) \equiv_{\Sigma \cup sig_C(\mathcal{O})} cl(\mathcal{O})$ .*

Let  $Res_\Sigma(cl(\mathcal{O}), \mathcal{M})|_\Sigma$  be the sub ontology of  $Res_\Sigma(cl(\mathcal{O}), \mathcal{M})$  that contains only those axioms  $\alpha \in Res_\Sigma(cl(\mathcal{O}), \mathcal{M})$  that satisfy:

- if  $L := \neg D$  is a literal of  $\alpha$ , then for any definer  $D' \in sig(cl(\mathcal{O}))$  such that  $D \preceq_d D'$ , we have  $Rol(D', cl(\mathcal{O})) \subseteq \Sigma$ .

To prove Theorem 1, we also need the following lemma.

**Lemma 5.** *For any ontology  $\mathcal{O}$  and signature  $\Sigma \subseteq sig(\mathcal{O})$ , we have*

$$Res_\Sigma(cl(\mathcal{O}), cl(\mathcal{O})) \equiv_{\Sigma \cup sig_C(\mathcal{O})} Res_\Sigma(cl(\mathcal{O}), cl(\mathcal{O}))|_\Sigma.$$

*Proof.* Recall that we assume that each definer occurs at most once positively in  $cl(\mathcal{O})$ . In particular, for each definer  $D \in sig(cl(\mathcal{O}))$ , there is at most one occurrence of a literal of the form  $Qr.D$  in  $cl(\mathcal{O})$ .

For any definer  $D \in sig(cl(\mathcal{O}))$ , if there exists  $D' \in sig(cl(\mathcal{O}))$  such that  $D \preceq_d D'$  and  $Rol(D', cl(\mathcal{O})) = \{r_0\} \not\subseteq \Sigma$ , then we can find a sequence of axioms in  $cl(\mathcal{O})$  such as the following.

$$\begin{aligned} \alpha_0 &: C_0 \sqcup Q_0 r_0 . D_0, \\ \alpha_1 &: \neg D_0 \sqcup C_1 \sqcup Q_1 r_1 . D_1, \\ &\dots, \\ \alpha_n &: \neg D_n \sqcup C_{n+1} \sqcup Q_n r_n . D_{n+1}. \\ \alpha_{n+1} &: \neg D_{n+1} \sqcup C_{n+2}. \end{aligned}$$

where  $D_0 = D'$ ,  $D_{n+1} = D$ . Then, for every  $1 \leq i \leq n$ ,  $Q_i r_i . D_i$  is the unique literal containing  $D_i$  positively, and  $Q_i r_i . D_i$  appears only in  $\alpha_i$ .

Since  $r_0 \notin \Sigma$ , there is no axiom in  $Res_\Sigma(cl(\mathcal{O}), cl(\mathcal{O}))$  that contains a literal of the form  $Q_0 r_0 . D_0$ . Therefore,  $D_0$  does not appear positively in  $Res_\Sigma(cl(\mathcal{O}), cl(\mathcal{O}))$ . It is well known that, if a concept  $D_0$  occurs only negatively in an ontology,

we can preserve all entailments of axioms not using  $D_0$  if we replace  $D_0$  by  $\perp$ , which with our normal form means, we can delete all axioms with the literal  $\neg D_0$  without losing any consequences in  $\Sigma$  (see also [Zhao and Schmidt, 2017, Theorem 1]). Specifically, if we set  $\Sigma_{Res} = sig(Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O})))$ , and let  $Res_{\Sigma}^0(\mathcal{O})$  be the ontology obtained from  $Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O}))$  by removing all axioms that contain the literal  $\neg D_0$ , then we have

$$Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O})) \equiv_{\Sigma_{Res} \setminus \{D_0\}} Res_{\Sigma}^0(cl(\mathcal{O})).$$

Note that if a literal  $L_1 = Qr.D$  always occurs in  $cl(\mathcal{O})$  together with another literal  $L_2$  which is of the form  $A$  or  $\neg A$  (i.e., every axiom  $\alpha \in cl(\mathcal{O})$  either contains both  $L_1$  and  $L_2$  or none of them), then  $L_1$  also always appears together with  $L_2$  in  $Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O}))$ . Because the rule r-Res preserves all literals of the form  $A$  or  $\neg A$ , and since  $Q_1 r_1 . D_1$  always appears with  $\neg D_0$  in  $cl(\mathcal{O})$ ,  $D_1$  cannot appear positively in  $Res_{\Sigma}^0(cl(\mathcal{O}))$ , because we removed all occurrences of  $D_1$  along with the occurrences of  $\neg D_0$ . Then, if  $Res_{\Sigma}^1(\mathcal{O})$  is the ontology obtained by removing all axioms containing the literal  $\neg D_1$  from  $Res_{\Sigma}^0(\mathcal{O})$ , we have

$$Res_{\Sigma}^0(cl(\mathcal{O})) \equiv_{\Sigma_{Res} \setminus \{D_1\}} Res_{\Sigma}^1(cl(\mathcal{O})).$$

Consequently, we have

$$Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O})) \equiv_{\Sigma_{Res} \setminus \{D_0, D_1\}} Res_{\Sigma}^1(cl(\mathcal{O})).$$

Repeat this process for all  $D_2, \dots, D_{n+1}$ , and we have for  $0 \leq i \leq n$  that

$$Res_{\Sigma}^i(cl(\mathcal{O})) \equiv_{\Sigma_{Res} \setminus \{D_{i+1}\}} Res_{\Sigma}^{i+1}(cl(\mathcal{O})) \text{ and}$$

$$Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O})) \equiv_{\Sigma_{Res} \setminus \{D_0, \dots, D_{i+1}\}} Res_{\Sigma}^{i+1}(cl(\mathcal{O})),$$

where  $Res_{\Sigma}^{i+1}(\mathcal{O})$  is the ontology obtained from  $Res_{\Sigma}^i(\mathcal{O})$  by removing all axioms containing the literal  $\neg D_{i+1}$ . We conclude that removing all axioms that contain the literal  $\neg D_{n+1} = \neg D$  preserves all logical consequences over  $Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O}))$  in the signature that excludes  $\{D_0, \dots, D_{n+1}\}$ .

Finally, we have

$$Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O})) \equiv_{\Sigma \cup sig_c(\mathcal{O})} Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O})) \Big|_{\Sigma}$$

by repeating the process above for all  $D$  such that there exists  $D' \in sig(cl(\mathcal{O}))$  such that  $D \preceq_d D'$  and  $Rol(D', cl(\mathcal{O})) = \{r_0\} \not\subseteq \Sigma$ .  $\square$

We now have everything to prove Theorem 1.

**Theorem 1.**  $RI_{\Sigma}(\mathcal{O})$  is role isolated for  $\Sigma$  and we have  $\mathcal{O} \equiv_{\Sigma \cup sig_c(\mathcal{O})} RI_{\Sigma}(\mathcal{O})$ .

*Proof.*  $RI_{\Sigma}(\mathcal{O})$  is role isolated by definition. We show that  $\mathcal{O} \equiv_{\Sigma \cup sig_c(\mathcal{O})} RI_{\Sigma}(\mathcal{O})$ .

By Theorem 4 and Lemma 5, it is sufficient to show that

$$Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O})) \Big|_{\Sigma} = Res_{\Sigma}(RI_{\Sigma}(\mathcal{O}), RI_{\Sigma}(\mathcal{O})) \Big|_{\Sigma}. \quad (4)$$

This follows from the following observations.

1. If an axiom  $\alpha$  contains the literal  $\neg D$ , any axiom obtained from  $\alpha$  using **r-Res** also contains  $\neg D$ . Therefore, we have after Definition 4 defining  $cl_{\Sigma}(\mathcal{O})$ :

$$Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O})) \Big|_{\Sigma} = Res_{\Sigma}(cl_{\Sigma}(\mathcal{O}), cl(\mathcal{O})).$$

2. Assume that  $C_1 \sqcup \exists r.D_1, C_2 \sqcup \forall r.D_2, \dots, C_n \sqcup \forall r.D_n \in Res_{\Sigma}(cl_{\Sigma}(\mathcal{O}), cl(\mathcal{O}))$  and  $r \notin \Sigma$ . Then, by Lemma 3, we have  $cl(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$  iff  $RI_{\Sigma}(\mathcal{O}) \models D_1 \sqcap \dots \sqcap D_n \sqsubseteq \perp$ . Therefore, for a given premise set  $P = \{C_1 \sqcup \exists r.D_1, C_2 \sqcup \forall r.D_2, \dots, C_n \sqcup \forall r.D_n\} \subseteq Res_{\Sigma}(cl_{\Sigma}(\mathcal{O}), cl(\mathcal{O}))$ , an **r-Res** inference  $\rho$  is applicable on  $P$  with  $\mathcal{M} = cl(\mathcal{O})$  iff  $\rho$  is applicable on  $P$  with  $\mathcal{M} = RI_{\Sigma}(\mathcal{O})$ . Consequently, we have

$$Res_{\Sigma}(cl_{\Sigma}(\mathcal{O}), cl(\mathcal{O})) = Res_{\Sigma}(cl_{\Sigma}(\mathcal{O}), RI_{\Sigma}(\mathcal{O})).$$

3. Since  $RI_{\Sigma}(\mathcal{O}) = cl_{\Sigma}(\mathcal{O}) \cup \mathcal{D}_{\Sigma}(cl(\mathcal{O}))$  and every axiom in  $\mathcal{D}_{\Sigma}(cl(\mathcal{O}))$  has a literal  $\neg D$  with  $Rol(D, cl(\mathcal{O})) \not\subseteq \Sigma$ , we have

$$\begin{aligned} & Res_{\Sigma}(cl_{\Sigma}(\mathcal{O}), RI_{\Sigma}(\mathcal{O})) \\ &= Res_{\Sigma}(RI_{\Sigma}(\mathcal{O}), RI_{\Sigma}(\mathcal{O})) \Big|_{\Sigma}. \end{aligned} \quad \square$$

### 9.3 Theorem 2

**Theorem 2.** If  $\mathcal{O}$  is role isolated for  $\Sigma$ , then  $rolE_{\Sigma}(\mathcal{O})$  is a role-forgetting for  $\mathcal{O}$  and  $\Sigma$ .

*Proof.* Assume  $\mathcal{O}$  is role isolated for  $\Sigma$ . Recall that by Lemma 4, for any ontology  $\mathcal{O}$  and signature  $\Sigma \subseteq sig(\mathcal{O})$ , we have  $Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O})) \equiv_{\Sigma \cup sig_c(\mathcal{O})} cl(\mathcal{O})$ . We note that, since  $\mathcal{O}$  is in normal form,  $cl(\mathcal{O})$  is obtained by replacing every literal  $Qr.A$  by some  $Qr.D$ , where we also have the axioms  $\neg D \sqcup A$ . We observe furthermore that, since  $\mathcal{O}$  is role isolated, the r-Rule is applicable in  $\mathcal{O}$  exactly iff the r-Res is applicable on the corresponding normalized axioms, excluding the last premise. As a consequence, we obtain that  $Res_{\Sigma}(cl(\mathcal{O}), cl(\mathcal{O})) = cl(rolE_{\Sigma}(\mathcal{O}))$ . We obtain  $cl(rolE_{\Sigma}(\mathcal{O})) \equiv_{\Sigma \cup \mathcal{N}_c} cl(\mathcal{O})$ . Since we also have  $cl(\mathcal{O}) \equiv_{\Sigma \cup \mathcal{N}_c} \mathcal{O}$ , we obtain that  $rolE_{\Sigma}(\mathcal{O})$  is a role forgetting of  $\mathcal{O}$  for  $\Sigma$ .  $\square$

## 9.4 Proof of Theorem 4

Let  $\Sigma$  be a signature. For  $X$  being a role name or a concept, we define  $\text{copy}_\Sigma(X)$  by structural induction.

1.  $\text{copy}_\Sigma(A) = A$ , if  $A \in (\mathbf{N}_C \cap \Sigma) \cup \mathbf{N}_D$ ;
2.  $\text{copy}_\Sigma(r) = r$ , if  $r \in \mathbf{N}_R \cap \Sigma$ ;
3.  $\text{copy}_\Sigma(B) = \bar{B}$  if  $B \in \text{sig}_C(\mathcal{O}) \setminus \Sigma$ , where  $\bar{B}$  is fresh;
4.  $\text{copy}_\Sigma(r) = \bar{r}$  if  $r \in \text{sig}_R(\mathcal{O}) \setminus \Sigma$ , where  $\bar{r}$  is fresh;
5.  $\text{copy}_\Sigma(C_1 \sqcup C_2) = \text{copy}_\Sigma(C_1) \sqcup \text{copy}_\Sigma(C_2)$ ;
6.  $\text{copy}_\Sigma(C_1 \sqcap C_2) = \text{copy}_\Sigma(C_1) \sqcap \text{copy}_\Sigma(C_2)$ ;
7.  $\text{copy}_\Sigma(\mathbf{Q}r.C) = \mathbf{Q}\text{copy}_\Sigma(r).\text{copy}_\Sigma(C)$ ,
8.  $\text{copy}_\Sigma(\neg C) = \neg\text{copy}_\Sigma(C)$ .

We further define  $cl^{ex}(\mathcal{O}) = RI_\Sigma(\mathcal{O}) \cup cl(\mathcal{O}_D)$ , where

$$\mathcal{O}_D = \{D \equiv \text{copy}_\Sigma(C_D) \mid D \in \text{sig}(\text{conE}_\Sigma(\text{role}_\Sigma(RI_\Sigma(\mathcal{O}))) \cap \mathbf{N}_D)\}. \quad (5)$$

To prove Theorem 4, we need the following two lemmas.

**Lemma 6.** *We have  $\mathcal{O} \equiv_{\Sigma \cup \text{sig}_C(\mathcal{O})} cl^{ex}(\mathcal{O})$ .*

*Proof.* Since  $RI_\Sigma(\mathcal{O}) \subseteq cl^{ex}(\mathcal{O})$ , and  $\mathcal{O} \equiv_{\Sigma \cup \text{sig}_C(\mathcal{O})} RI_\Sigma(\mathcal{O})$  (by Theorem 1), we have for every axiom  $\alpha$  s.t.  $\text{sig}(\alpha) \subseteq \Sigma \cup \text{sig}_C(\mathcal{O})$  and  $\mathcal{O} \models \alpha$ , also  $cl^{ex}(\mathcal{O}) \models \alpha$ . We thus only need to show the other direction.

Let  $\alpha$  be s.t.  $\text{sig}(\alpha) \subseteq \Sigma \cup \text{sig}_C(\mathcal{O})$ , and assume  $cl^{ex}(\mathcal{O}) \models \alpha$  but  $\mathcal{O} \not\models \alpha$ . Then, also  $cl(\mathcal{O}) \not\models \alpha$ , and there is a witnessing model  $\mathcal{I}$  of  $cl(\mathcal{O})$  s.t.  $\mathcal{I} \not\models \alpha$ . Based on  $\mathcal{I}$ , we construct a model  $\mathcal{I}^{ex}$  of  $cl^{ex}(\mathcal{O})$  and also show that  $\mathcal{I}^{ex} \not\models \alpha$ , and thus  $cl^{ex}(\mathcal{O}) \not\models \alpha$ . A contradiction!

$\mathcal{I}^{ex}$  is defined as follows.

1.  $A^{\mathcal{I}^{ex}} = A^{\mathcal{I}}, r^{\mathcal{I}^{ex}} = r^{\mathcal{I}}$  for all  $A, r \in \text{sig}(\mathcal{O}) \cap \text{sig}(cl^{ex}(\mathcal{O}))$ ;
2.  $\bar{r}^{\mathcal{I}^{ex}} = r^{\mathcal{I}^{ex}}, \bar{B}^{\mathcal{I}^{ex}} = B^{\mathcal{I}^{ex}}$  for all introduced role names  $\bar{r}$  and introduced concept names  $\bar{B}$ .
3.  $D^{\mathcal{I}^{ex}} = (C_D)^{\mathcal{I}^{ex}}$  for every definer  $D \in \text{sig}(cl^{ex}(\mathcal{O})) \cap \mathbf{N}_D$ .

Since  $\mathcal{I}$  is a model of  $\mathcal{O}$ , by the item 3 above and the definition of  $C_D$ , we know  $\mathcal{I}^{ex}$  is compatible with all axioms in  $RI_\Sigma(\mathcal{O})$  and  $\mathcal{O}_D$ . Therefore,  $\mathcal{I}^{ex}$  is compatible with all axioms in  $cl(\mathcal{O}_D)$  and thus a model of  $cl^{ex}(\mathcal{O}_D)$ . Moreover, because  $A^{\mathcal{I}^{ex}} = A^{\mathcal{I}}$  and  $r^{\mathcal{I}^{ex}} = r^{\mathcal{I}}$  for all  $A, r \in \Sigma \cup \text{sig}(\mathcal{O})$ , we have  $\mathcal{I}^{ex} \models cl^{ex}(\mathcal{O})$  and  $\mathcal{I}^{ex} \not\models \alpha$ . A contradiction.  $\square$

**Lemma 7.**  $cl^{ex}(\mathcal{O}) \equiv_\Sigma gm_\Sigma(\mathcal{O})$ .

*Proof.* Assume  $\Sigma^{ex} = \Sigma \cup \text{sig}(cl(\mathcal{O}_D))$ . Then  $\text{sig}_R(cl(\mathcal{O}_D)) = \text{sig}_R(\mathcal{O}_D) \subseteq \Sigma^{ex}$ . Since  $RI_\Sigma(\mathcal{O})$  is role isolated for  $\Sigma$ , we have  $cl^{ex}(\mathcal{O}) = RI_\Sigma(\mathcal{O}) \cup cl(\mathcal{O}_D)$  is role isolated for  $\Sigma^{ex}$ . Moreover, we observe the following:

1.  $\text{role}_{\Sigma^{ex}}(cl^{ex}(\mathcal{O})) = \text{role}_{\Sigma}(RI_\Sigma(\mathcal{O})) \cup cl(\mathcal{O}_D)$ .  
This is because  $\text{sig}_R(cl(\mathcal{O}_D)) \subseteq \Sigma^{ex}$  and  $cl(\mathcal{O}_D)$  does not contain axioms of the form  $\neg D_1 \sqcup \dots \sqcup \neg D_n$ , then the r-Rule is only applicable on  $RI_\Sigma(\mathcal{O})$ .
2.  $\text{conE}_{\Sigma^{ex}}(\text{role}_{\Sigma^{ex}}(cl^{ex}(\mathcal{O}))) = \text{conE}_\Sigma(\text{role}_\Sigma(RI_\Sigma(\mathcal{O}))) \cup cl(\mathcal{O}_D)$ .  
This is because  $\text{sig}_C(cl(\mathcal{O}_D)) \subseteq \Sigma^{ex}$ , then the A-Rule is only applicable on  $\text{role}_\Sigma(RI_\Sigma(\mathcal{O}))$ .

By Theorem 2 and 3, we have

$$\begin{aligned} cl^{ex}(\mathcal{O}) &\equiv_{\Sigma^{ex}} \text{role}_{\Sigma^{ex}}(cl^{ex}(\mathcal{O})), \\ \text{role}_{\Sigma^{ex}}(cl^{ex}(\mathcal{O})) &\equiv_{\Sigma^{ex}} \text{conE}_{\Sigma^{ex}}(\text{role}_{\Sigma^{ex}}(cl^{ex}(\mathcal{O}))). \end{aligned}$$

It follows from these observations that

$$\begin{aligned} cl^{ex}(\mathcal{O}) &\equiv_{\Sigma^{ex}} \text{conE}_{\Sigma^{ex}}(\text{role}_{\Sigma^{ex}}(cl^{ex}(\mathcal{O}))) \\ &= \text{conE}_\Sigma(\text{role}_\Sigma(RI_\Sigma(\mathcal{O}))) \cup cl(\mathcal{O}_D) \\ &\equiv_\Sigma \text{conE}_\Sigma(\text{role}_\Sigma(RI_\Sigma(\mathcal{O}))) \cup \mathcal{O}_D. \end{aligned}$$

Moreover, we have

$$\text{conE}_\Sigma(\text{role}_\Sigma(RI_\Sigma(\mathcal{O}))) \cup \mathcal{O}_D \equiv_\Sigma gm_\Sigma(\mathcal{O})$$

because  $gm_\Sigma(\mathcal{O})$  can be obtained by the following three operations over  $\text{conE}_\Sigma(\text{role}_\Sigma(RI_\Sigma(\mathcal{O}))) \cup \mathcal{O}_D$ :

1. Replace all occurrences of the definers  $D$  in  $\text{conE}_\Sigma(\text{role}_\Sigma(RI_\Sigma(\mathcal{O})))$  by  $\text{copy}_\Sigma(C_D)$  and remove tautologies.



2. Replace every new introduced concept  $\bar{B}, \bar{r}$  by  $B, r$ , respectively. Note that  $\bar{B}, B, \bar{r}, r \notin \Sigma$  by the definition of  $\text{copy}_\Sigma$ .
3. Apply exhaustively the translations  $C_1 \sqsubseteq \neg C_2 \sqcup C_3 \Rightarrow C_1 \sqcap C_2 \sqsubseteq C_3$  and  $C_1 \sqsubseteq \text{Q}r.\neg C_2 \sqcap C_3 \Rightarrow C_1 \sqcap \bar{\text{Q}}r.C_2 \sqsubseteq C_3$ , where  $\bar{\exists} = \forall$  and  $\bar{\forall} = \exists$ .

These operations produce a new ontology that is  $\Sigma$ -inseparable to the input ontology.

In conclusion, since  $\Sigma \subseteq \Sigma^{ex}$ , we have  $cl^{ex}(\mathcal{O}) \equiv_\Sigma \text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O}))) \cup \mathcal{O}_D \equiv_\Sigma gm_\Sigma(\mathcal{O})$ . This completes the proof.  $\square$

**Theorem 4.** Let  $gm_\Sigma(\mathcal{O})$  be the ontology obtained from  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$  by

- deleting subsumed axioms,
- replacing each definer  $D$  by  $C_D$ , and
- exhaustively applying  $C_1 \sqsubseteq \neg C_2 \sqcup C_3 \Rightarrow C_1 \sqcap C_2 \sqsubseteq C_3$  and  $C_1 \sqsubseteq \text{Q}r.\neg C_2 \sqcap C_3 \Rightarrow C_1 \sqcap \bar{\text{Q}}r.C_2 \sqsubseteq C_3$ , where  $\bar{\exists} = \forall$  and  $\bar{\forall} = \exists$ .

Then,  $gm_\Sigma(\mathcal{O})$  is a general module for  $\mathcal{O}$  and  $\Sigma$ .

*Proof.* By Lemma 6 and 7, we have

$$\mathcal{O} \equiv_\Sigma cl^{ex}(\mathcal{O}) \equiv_\Sigma gm_\Sigma(\mathcal{O}),$$

which proves the theorem.  $\square$

## 9.5 Proof of Proposition 1

In the following, we use  $|\mathcal{O}|$  to denote the number of axioms in  $\mathcal{O}$ . We then have  $|\mathcal{O}| \leq \|\mathcal{O}\|$ .

**Proposition 1.** For any ontology  $\mathcal{O}$  and signature  $\Sigma$ , we have  $\|gm_\Sigma(\mathcal{O})\| \leq 2^{O(\|cl(\mathcal{O})\|)}$ . On the other hand, there exists a family of ontologies  $\mathcal{O}_n$  and signatures  $\Sigma_n$  s.t.  $\|\mathcal{O}_n\|$  is polynomial in  $n \geq 1$  and  $\|gm_{\Sigma_n}(\mathcal{O}_n)\| = n \cdot 2^{O(\|cl(\mathcal{O}_n)\|)}$ .

*Proof.* We first show the upper bound. Our construction ensures that for every axiom  $\alpha \in \text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O}))$ , we can find a sequence of axioms  $\beta_1, \dots, \beta_n \in cl_\Sigma(\mathcal{O})$  such that  $\alpha$  is obtained from  $\bigsqcup_{1 \leq i \leq n} \beta_i$  by removing all literals that contain a role  $r \notin \Sigma$ . Because there are at most exponentially many subsets of  $cl_\Sigma(\mathcal{O})$ , this limits the number of possible inferred axioms to exponentially many. We obtain

$$\begin{aligned} |\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O}))| &\leq 2^{|cl_\Sigma(\mathcal{O})|} \\ &\leq 2^{\|cl(\mathcal{O})\|} \\ \|\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O}))\| &\leq |\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O}))| \cdot \|cl(\mathcal{O})\| \\ &\leq 2^{\|cl(\mathcal{O})\|} \cdot \|cl(\mathcal{O})\|. \end{aligned}$$

Similarly, for every axiom  $\gamma = \text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$ , we can find a sequence of axioms  $\alpha_1, \dots, \alpha_n \in \text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O}))$  such that  $\gamma$  is obtained from  $\bigsqcup_{1 \leq i \leq n} \alpha_i$  by removing all literals of the form  $A$  or  $\neg A$  with  $A \notin \Sigma$ . As shown above, each  $\alpha_k$  is obtained from  $\bigsqcup_{1 \leq i \leq n_k} \beta_i^k$ , for some  $\beta_1^k, \dots, \beta_{n_k}^k \in cl_\Sigma(\mathcal{O})$ , by removing all literals that contain a role name  $r \notin \Sigma$ . We obtain that  $\gamma$  is obtained from  $\bigsqcup_{1 \leq k \leq n, 1 \leq i \leq n_k} \beta_i^k$  by removing all literals  $L$  such that (i)  $L$  contains a role name  $r \notin \Sigma$ , or (ii)  $L$  is of the form  $A$  or  $\neg A$  with  $A \notin \Sigma$ . We obtain

$$\begin{aligned} |\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))| &\leq 2^{|cl_\Sigma(\mathcal{O})|} \\ &\leq 2^{\|cl(\mathcal{O})\|} \\ \|\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))\| &\leq |\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))| \cdot \|cl(\mathcal{O})\| \\ &\leq 2^{\|cl(\mathcal{O})\|} \cdot \|cl(\mathcal{O})\|. \end{aligned}$$

For every definer  $D \in \text{sig}(cl(\mathcal{O}))$ , we have  $|C_D| < \|cl(\mathcal{O})\|$ . Taking that the lengths of axioms in  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$  are bound by  $\|cl(\mathcal{O})\|$ , we obtain that for every axiom  $\alpha \in gm_\Sigma(\mathcal{O})$ , we have  $|\alpha| \leq \|cl(\mathcal{O})\|^2$ . Consequently, we have

$$\begin{aligned} \|gm_\Sigma(\mathcal{O})\| &\leq |\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))| \cdot \|cl(\mathcal{O})\|^2 \\ &\leq 2^{\|cl(\mathcal{O})\|} \cdot \|cl(\mathcal{O})\|^2 \\ &< 2^{3 \cdot \|cl(\mathcal{O})\|}. \end{aligned}$$

Note that for the last step, we use that  $n^2 < 2^{2n}$  for all integers  $n \geq 0$ . This establishes the upper bound.

We continue to show the lower bound. For any integer  $n \geq 0$ , we define the ontology  $\mathcal{O}_n$  to contain the following axioms:

1.  $Z_1 \sqcap Z_2 \sqcap \dots \sqcap Z_n \sqsubseteq \perp$  (1 axiom of length  $n + 1$ ),
2.  $X_i \sqcup Y_i \sqsubseteq Z_i$  for all  $1 \leq i \leq n$  ( $n$  axioms of length 3),
3.  $\top \sqsubseteq A_1 \sqcup \exists s.X_1$ ,  $\top \sqsubseteq \overline{A_1} \sqcup \exists s.Y_1$  (2 axioms of length 4), and
4.  $\top \sqsubseteq A_j \sqcup \forall s.X_j$ ,  $\top \sqsubseteq \overline{A_j} \sqcup \forall s.Y_j$  for  $2 \leq j \leq n$  ( $2n - 2$  axioms of length 4).

As a signature, we define  $\Sigma_n = \{A_j\}_{j=1}^n \cup \{\overline{A_j}\}_{j=1}^n$ .

Normalizing  $\mathcal{O}_n$  introduces the definers  $D_i^X$  and  $D_i^Y$  with  $C_{D_i^X} = X_i$  and  $C_{D_i^Y} = Y_i$ . In particular, this gives  $2n$  additional axioms of length 2 each, so that we obtain  $|cl(\mathcal{O}_n)| = 5n + 1$  and

$$\begin{aligned} \|cl(\mathcal{O}_n)\| &= n + 1 + 3n + 2 \cdot 4 + (2n - 2) \cdot 4 + 2n \cdot 2 \\ &= 16n + 1 \end{aligned}$$

We continue to compute the sizes of the other axiom sets computed.

- $\mathcal{D}_{\Sigma_n}(cl(\mathcal{O}_n))$  consists of axioms of the form:

$$\neg D_1^* \sqcup \dots \sqcup \neg D_n^*, \quad \text{where for } 1 \leq i \leq n, \quad D_i^* \in \{D_i^X, D_i^Y\}.$$

We have  $|\mathcal{D}_{\Sigma_n}(cl(\mathcal{O}_n))| = 2^n$  and  $\|\mathcal{D}_{\Sigma_n}(cl(\mathcal{O}_n))\| = n \cdot 2^n$ .

- $\text{rolE}_{\Sigma_n}(RI_{\Sigma_n}(\mathcal{O}_n))$  consists of the axioms in  $\mathcal{D}_{\Sigma_n}(cl(\mathcal{O}_n))$  and axioms of the form:

$$\neg A_1^* \sqcup \dots \sqcup \neg A_n^*, \quad \text{where for } 1 \leq i \leq n, \quad A_i^* \in \{A_i, \overline{A_i}\}.$$

We obtain that  $|\text{rolE}_{\Sigma_n}(RI_{\Sigma_n}(\mathcal{O}_n))| = 2^{n+1}$  and  $\|\text{rolE}_{\Sigma_n}(RI_{\Sigma_n}(\mathcal{O}_n))\| = n \cdot 2^{n+1}$ .

- Finally, since no definer occurs positively anymore, the axioms in  $\mathcal{D}_{\Sigma_n}(cl(\mathcal{O}_n))$  are removed, so that we obtain  $|\text{conE}_{\Sigma_n}(\text{rolE}_{\Sigma_n}(RI_{\Sigma_n}(\mathcal{O}_n)))| = 2^n$  and  $\|\text{conE}_{\Sigma_n}(\text{rolE}_{\Sigma_n}(RI_{\Sigma_n}(\mathcal{O}_n)))\| = n \cdot 2^{n+1}$ .

As a final result, we obtain  $|gm_{\Sigma_n}(\mathcal{O}_n)| = 2^n$  and  $\|gm_{\Sigma_n}(\mathcal{O}_n)\| = n \cdot 2^{n+1}$ .

To summarize, we defined a sequence of ontologies  $\mathcal{O}_n$  with signatures  $\Sigma_n$  s.t.  $\|cl(\mathcal{O}_n)\| = 16n + 1$  and  $\|gm_{\Sigma_n}(\mathcal{O}_n)\| = n \cdot 2^{n+1}$ . This establishes the second claim of the proposition.  $\square$

## 9.6 Proof of Propositions 2 and 3

To simplify the proofs of Propositions 2 and 3, we define the operator  $\text{defE}$ , which applies the definer substitution explicitly.

**Definition 7.** Let  $\mathcal{O}$  be an ontology that contains definers  $D$ , for which  $C_D$  is defined. Then, the definer substitution on  $\mathcal{O}$  is the ontology  $\text{defE}(\mathcal{O})$  that is obtained from  $\mathcal{O}$  by replacing each definer  $D$  by the corresponding concept  $C_D$ .

Since Proposition 3 relies on a simpler situation than Proposition 2, namely where the input is normalized, it is more convenient to start with it, before the more complex situation of Proposition 2.

**Proposition 3.** Let  $\mathcal{O}$  be an ontology in normal form and  $\mathcal{M} = gm_{\Sigma}(\mathcal{O})$ . Then,  $gm_{\Sigma}(\mathcal{M}) = \mathcal{M}$ .

*Proof.* We first make a general observation on the effect the definers have when computing  $gm_{\Sigma}(\mathcal{O})$  for normalized ontologies  $\mathcal{O}$ . First, since  $\mathcal{O}$  is normalized, only concept names occur under role restrictions, which means that for every definer  $D$  introduced, we have  $C_D \in \mathcal{N}_C$ , and the only negative occurrence of  $D$  is in an axiom of the form  $\neg D \sqcup C_D$ . In case  $C_D \notin \Sigma$ , this means that previously eliminated concept names get reintroduced by the definer substitution when computing  $gm_{\Sigma}(\mathcal{O})$ , but they can only occur in two ways: 1) as negative literals in axioms of the form  $\neg C_D \sqcup C$ , or 2) in literals of the form  $Qr.C_D$ . This also means that  $gm_{\Sigma}(\mathcal{O})$  remains normalized. This means in particular that  $\mathcal{M}$  contains no role name  $r \notin \Sigma$ , and all concept names  $A \in \text{sig}_C(\mathcal{M}) \setminus \Sigma$  occur either negatively or under role restrictions.

Let  $RI_{\Sigma}(\mathcal{M})$  be the role isolated form for  $\Sigma$  and  $\mathcal{M}$ . We observe that  $\text{sig}_R(\mathcal{M}) \subseteq \Sigma$ , since no role name outside of  $\Sigma$  is introduced to  $\mathcal{M} = gm_{\Sigma}(\mathcal{O})$  when substituting definers  $D$  with their corresponding concepts  $C_D$ . We obtain that  $RI_{\Sigma}(\mathcal{M}) = cl(\mathcal{M})$ .

By the definition of  $gm_{\Sigma}$ , we thus have

$$gm_{\Sigma}(\mathcal{M}) = \text{defE}(\text{conE}_{\Sigma}(\text{rolE}_{\Sigma}(cl(\mathcal{M}))))).$$

We show that  $gm_{\Sigma}(\mathcal{M}) = \mathcal{M}$  using the following two results:

1.  $\text{rolE}_{\Sigma}(cl(\mathcal{M})) = cl(\mathcal{M})$ .

Since  $\text{sig}_R(\mathcal{M}) \subseteq \Sigma$ , there is no role name to be eliminated by the operator  $\text{rolE}$ . This means that  $\text{rolE}_{\Sigma}(cl(\mathcal{M})) = cl(\mathcal{M})$ .

2.  $\text{defE}(\text{conE}_\Sigma(\text{cl}(\mathcal{M}))) = \mathcal{M}$ .

First, we show that  $\mathcal{M} \subseteq \text{defE}(\text{conE}_\Sigma(\text{cl}(\mathcal{M})))$ . Let  $c \in \mathcal{M}$ . Then,  $c$  is of the form

$$\neg B_1 \sqcup \dots \sqcup \neg B_k \sqcup Q_1 r_1 . A_1 \sqcup \dots \sqcup Q_n r_n . A_n \sqcup C_1 \quad (6)$$

where (i)  $B_i \in \text{sig}_C(\mathcal{O}) \setminus \Sigma$  for  $1 \leq i \leq k$ , (ii)  $A_i \in \text{sig}_C(\mathcal{O}) \setminus \Sigma$  for  $1 \leq i \leq n$ , and (iii)  $\text{sig}(C_1) \subseteq \Sigma$ .

We show that  $c \in \text{defE}(\text{conE}_\Sigma(\text{cl}(\mathcal{M})))$ . By the definition of  $\mathcal{M} = \text{gm}_\Sigma(\mathcal{O})$ , we have the following results.

- $c \in \mathcal{M}$  must be obtained from an axiom  $c_1 \in \text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$  by replacing every definer  $D$  by the concept  $C_D$ . Then  $c_1$  is of the form

$$c_1 = \neg D_1 \sqcup \dots \sqcup \neg D_k \sqcup Q_1 r_1 . D'_1 \sqcup \dots \sqcup Q_n r_n . D'_n \sqcup C_1,$$

where  $C_{D_i} = B_i$  for  $1 \leq i \leq k$ , and  $C_{D_j} = A_j$  for  $1 \leq j \leq n$ .

By our construction,  $D_1, \dots, D_k$  must also occur positively in  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$ , since otherwise  $c_1$  is deleted through  $\text{conE}_\Sigma$ . Consequently, there are  $k$  axioms  $c_2, \dots, c_{k+1} \in \text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$  that are of the forms

$$c_2 = C_2 \sqcup Q'_1 r'_1 . D_1,$$

⋮

$$c_{k+1} = C_{k+1} \sqcup Q'_k r'_k . D_k.$$

- $c_1$  must be obtained by applying A-Rules on  $k+1$  axioms  $c'_1, \dots, c'_{k+1} \in \text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O}))$  that are of the forms

$$c'_1 = \neg B_1 \sqcup \dots \sqcup \neg B_k \sqcup Q_1 r_1 . D'_1 \sqcup \dots \sqcup Q_n r_n . D'_n \sqcup C_1, \quad (7)$$

$$c'_2 = \neg D_1 \sqcup B_1 \quad (8)$$

⋮

$$c'_{k+1} = \neg D_k \sqcup B_k. \quad (10)$$

Modulo renaming of definers,<sup>3</sup> we have (i)  $c'_1 \in \text{cl}(\mathcal{M})$  by normalizing  $c$ ; (ii)  $c'_2, \dots, c'_{k+1} \in \text{cl}(\mathcal{M})$  by normalizing the axioms in  $\text{defE}(\{c_2, \dots, c_{k+1}\}) \subseteq \mathcal{M}$ . Therefore, we can assume  $\{c'_1, c'_2, \dots, c'_{k+1}\} \subseteq \text{cl}(\mathcal{M})$ . We obtain  $c \in \text{gm}_\Sigma(\mathcal{M})$  by repeating the process of generating  $c \in \text{gm}_\Sigma(\mathcal{O})$  from  $c'_1, c'_2, \dots, c'_{k+1}$ . As a result, we obtain that  $\mathcal{M} \subseteq \text{defE}(\text{conE}_\Sigma(\text{cl}(\mathcal{M})))$ .

Furthermore, we have  $\text{defE}(\text{conE}_\Sigma(\text{cl}(\mathcal{M}))) \subseteq \mathcal{M}$ , since all the axioms in  $\text{cl}(\mathcal{M})$  are of the forms (7) – (10), and we cannot obtain axioms other than  $c$  in (6) after applying the operators  $\text{conE}$  and  $\text{defE}$ . Consequently, we have  $\mathcal{M} = \text{defE}(\text{conE}_\Sigma(\text{cl}(\mathcal{M})))$ .

From 1 and 2, it follows that  $\mathcal{M} = \text{gm}_\Sigma(\mathcal{M})$ . □

**Proposition 2.** *Let  $(\mathcal{M}_i)_{i \geq 1}$  be the sequence of ontologies defined by (i)  $\mathcal{M}_1 = \text{gm}_\Sigma(\mathcal{O})$  and (ii)  $\mathcal{M}_{i+1} = \text{gm}_\Sigma(\mathcal{M}_i)$  for  $i \geq 1$ . Then, we have  $\mathcal{M}_i \subseteq \mathcal{M}_{i+1}$  for  $i \geq 1$ . Moreover, there exists  $i_0 \geq 0$  s.t.  $\mathcal{M}^k = \mathcal{M}^{i_0}$  for all  $k \geq i_0$ .*

*Proof.* Assume  $\mathcal{M}_0 = \mathcal{O}$ . We first show that every axiom  $c \in \mathcal{M}_i = \text{gm}_\Sigma(\mathcal{M}_{i-1})$  is also in  $\mathcal{M}_{i+1}$  for all  $i \geq 1$ .

Any axiom  $c \in \text{gm}_\Sigma(\mathcal{M}_{i-1})$  is obtained from some axiom  $c_d \in \text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{M}_{i-1})))$  by replacing every definer  $D$  by the corresponding concept  $C_D$ . Then,  $\text{sig}(c_d) \subseteq \Sigma \cup \text{N}_D$  by the definitions of  $\text{rolE}_\Sigma$  and  $\text{conE}_\Sigma$ . There are two different cases.

1.  $c_d$  does not contain negative definers. Then,  $c_d$  and  $c$  are of the forms

$$c_d = C_1 \sqcup Q_1 r_1 . D_1 \sqcup \dots \sqcup Q_n r_n . D_n \quad (11)$$

$$c = C_1 \sqcup Q_1 r_1 . C_{D_1} \sqcup \dots \sqcup Q_n r_n . C_{D_n}. \quad (12)$$

where,  $\text{sig}(C_1) \subseteq \Sigma$ ,  $r_j \in \Sigma$  and  $Q_j \in \{\exists, \forall\}$  for  $1 \leq j \leq n$ .

Then,  $c \in \text{gm}_\Sigma(\mathcal{M}_i)$  because of the following observations, which hold modulo renaming of definers.

- $c_d \in \text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{M}_i))$  because  $c_d \in \text{RI}_\Sigma(\mathcal{M}_i)$  and  $\text{sig}(c_d) \subseteq \Sigma \cup \text{N}_D$ ,
- $c_d \in \text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{M}_i)))$  because  $c_d \in \text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{M}_i))$ ,  $\text{sig}(c_d) \subseteq \Sigma \cup \text{N}_D$  and  $c_d$  does not contain negative definers, and

<sup>3</sup>Note that it is in principle possible that  $\text{cl}(\mathcal{M})$  contains more definers than  $\text{cl}(\mathcal{O})$ , since occurrences of role restrictions can be multiplied. However, this does not affect the following argument, since all definers get replaced by the same concept names again.

- $\text{defE}(\{c_d\}) = \{c\}$  by definition.

2.  $c_d$  contains negative definers. Then,  $c_d$  and  $c$  are respectively of the forms

$$c_d = \neg D_1 \sqcup \dots \sqcup \neg D_n \sqcup C'_d \quad (13)$$

$$c = \neg C_{D_1} \sqcup \dots \sqcup \neg C_{D_n} \sqcup C', \quad (14)$$

where  $C'_d, C'$  do not contain negative definers, and  $C'_d$  and  $C'$  are of the forms as in (11) and (12), respectively.

In this case, normalizing  $c$  in  $\mathcal{M}_i$  produces axioms in  $cl(\mathcal{M}_i)$  that are different from  $c_d$ . However, we can still show that  $c \in \mathcal{M}_{i+1} = gm_\Sigma(\mathcal{M}_i)$ . We only consider the case where  $n = 1$ , that is, there is only one negative definer in  $c_d$ . The case for  $n > 1$  is shown by repeating the argument step-wise for each definer. We distinguish 4 possible cases based on the syntactical shape of  $C_{D_1}$ .

- $C_{D_1}$  is of the form  $A$  or  $\neg A$ . This case is proved similarly as for Proposition 3, where we considered the case of normalized ontologies for which  $C_{D_1}$  is always of the form  $A \in \mathbf{N}_C$ .
- $C_{D_1} = Qr.C_1$ . We consider only the case where  $Q = \exists$ . The other direction is shown in a similar way by just switching the quantifiers. With  $Q = \exists$ , we have

$$c_d = \neg D_1 \sqcup C'_d, \quad c = \neg(\exists r.C_1) \sqcup C'.$$

Normalizing  $c$ , we obtain

$$c'_1 = \forall r.D'_2 \sqcup C'_d \in cl(\mathcal{M}_i),$$

where  $C_{D'_2} = \neg C_1$ . If  $r \in \Sigma$ , then we have  $c \in gm_\Sigma(\mathcal{M}_i)$  as in Case 1. Assume  $r \notin \Sigma$ . We then make the following observations:

- There exists an axiom

$$c_1 = \neg D_1 \sqcup \exists r.D'_3 \in RI_\Sigma(\mathcal{M}_{i-1})$$

with  $C_{D'_3} = C_1$ . Here,  $c_1$  is introduced when normalizing the literal  $C_{D_1} = \exists r.C_1$ .

- There exists an axiom

$$c_2 = C \sqcup Qr.D_1 \in \text{conE}_\Sigma(\text{rolE}_\Sigma(RI_\Sigma(\mathcal{M}_{i-1})))$$

for some  $C, Q, r$  because  $D_1$  must also occur positively in  $\text{conE}_\Sigma(\text{rolE}_\Sigma(RI_\Sigma(\mathcal{M}_{i-1})))$ . (Otherwise,  $c_d$  will be deleted by  $\text{conE}_\Sigma$ .)

We have  $c_1 \in cl(\mathcal{M}_i)$  (modulo renaming of definers) by normalizing the axiom in  $\text{defE}(\{c_2\}) \subseteq \mathcal{M}_i$ . Furthermore, we have

$$cl(\mathcal{M}_i) \models D'_2 \sqcap D'_3 \sqsubseteq \perp.$$

This allows us to make the following further observations.

- $c_d \in \text{rolE}_\Sigma(RI_\Sigma(\mathcal{M}_i))$  due to  $c_d \in \text{conE}_\Sigma(\text{rolE}_\Sigma(RI_\Sigma(\mathcal{M}_{i-1})))$  and the following inference with the r-Rule.

$$\frac{\forall r.D'_2 \sqcup C'_d, \neg D_1 \sqcup \exists r.D'_3, \neg D'_2 \sqcup \neg D'_3}{\neg D_1 \sqcup C'_d}.$$

- $c_d \in \text{conE}_\Sigma(\text{rolE}_\Sigma(RI_\Sigma(\mathcal{M}_i)))$  since  $c_d \in \text{rolE}_\Sigma(RI_\Sigma(\mathcal{M}_i))$  and  $\text{sig}(c_d) \subseteq \Sigma \cup \mathbf{N}_D$ ;

- $\{c\} = \text{defE}(\{c_d\})$  as follows directly from (13) and (14).

We obtain that  $c \in gm_\Sigma(\mathcal{M}_i, \Sigma) = \mathcal{M}_{i+1}$ .

- $C_{D_1} = L_1 \sqcap \dots \sqcap L_n \sqcap Q_1 r_1.C_1 \sqcap \dots \sqcap Q_m r_m.C_m$ , where  $L_i = A_i$  or  $L_i = \neg A_i$  for some  $A_i \in \mathbf{N}_C$  and all  $1 \leq i \leq n$ . Then, normalizing  $c$  produces following axiom in  $cl(\mathcal{M}_i)$ :

$$\neg L_1 \sqcup \dots \sqcup \neg L_n \sqcup Q_1^* r_1.D'_1 \sqcup \dots \sqcup Q_m^* r_m.D'_m \sqcup C',$$

where  $C_{D'_i} = \neg C_i$ ,  $\{Q_i, Q_i^*\} = \{\forall, \exists\}$ ,  $1 \leq i \leq m$ . As in Case (b), we have

- $\neg D_1 \sqcup L_1, \dots, \neg D_1 \sqcup L_n, \neg D_1 \sqcup Q_1 r_1.D'_1, \dots, \neg D_1 \sqcup Q_m r_m.D'_m \in RI_\Sigma(\mathcal{M}_i)$ ;
- $C \sqcup Qr.D_1 \in \text{conE}_\Sigma(\text{rolE}_\Sigma(RI_\Sigma(\mathcal{M}_{i-1})))$  for some  $C, Q, r$  because otherwise  $c_d$  is removed in Step 3 using  $\text{conE}$ .

We obtain that  $c \in gm_\Sigma(\mathcal{M}_i)$  using the argument from Case (a) for every  $L_i$  and from Case (b) for every  $D_i$ .

- For the general case, we have  $C_{D_1} = \overline{C}_1 \sqcup \dots \sqcup \overline{C}_n$ , where each  $\overline{C}_i$  is as  $C_D$  in Case (c). In this case, we rewrite  $c$  as  $n$  different axioms.

$$\overline{c}_1 = \neg \overline{C}_1 \sqcup C', \quad \dots, \quad \overline{c}_n = \neg \overline{C}_n \sqcup C'.$$

For each  $1 \leq i \leq n$ , we then have  $\overline{c}_i \in gm_\Sigma(\mathcal{M}_i)$  as in Case (c).

We obtain in each case that  $c \in \mathcal{M}_{i+1}$ . As a consequence, we have  $\mathcal{M}_i \subseteq \mathcal{M}_{i+1}$ , for each  $i \geq 1$ .

It remains to show that there exists some  $i_0 \geq 0$  such that  $\mathcal{M}_{i_0} = \mathcal{M}_{i_0+1}$ , since all axioms in  $\mathcal{M}_i$  are axioms consisting of literals of the form

$$A, \neg A, \neg C_D, Qr.C_D,$$

where  $C_D$  is a sub-concept of a concept in  $\mathcal{O}$ . There exist only finitely many such literals and thus only finitely many such axioms. Consequently, the chain  $\mathcal{M}_0 \subseteq \mathcal{M}_1 \subseteq \dots$  must reach a fixpoint after finitely many steps.  $\square$

**Role Propagation (RP):**

$$\frac{\bigcup_{j=1}^m \{P_j \sqcup C_j\}, \quad E_0 \sqcup \text{Qr}.D_0, \quad \bigcup_{i=1}^k \{E_i \sqcup \forall r.D_i\}}{(\bigsqcup_{i=0}^n E_i) \sqcup \text{Qr}.(\prod_{j=0}^m C_j)},$$

where  $P_0 = \bigsqcup_{i=0}^n \neg D_i$ , for  $j > 0$ ,  $P_j$  is a sub-concept of  $P_0$ ,  $Q \in \{\forall, \exists\}$ , and  $C_0$  and  $C_j$  do not contain a definer.

**Reduction (Red):**

$$\frac{\mathcal{O} \cup \{\neg D_1 \sqcup \dots \sqcup \neg D_n \sqcup C\}}{\mathcal{O}},$$

where  $C$  is a general concept expression that does not contain a negative definer and  $D_1, \dots, D_n$  are definer symbols. The RP rule applies before this rule if  $\neg D_1 \sqcup \dots \sqcup \neg D_n$  takes the form of  $P_0$  in the RP rule.

Figure 7: Rules RP and Red.

## 9.7 Proof of Theorem 5

Recall the operators Op1, Op2 introduced in Section 5.3. For simplicity,

- let  $\mathcal{M}_1$  be the ontology obtained by applying the operator Op1, and
- let  $\mathcal{M}_2$  be the ontology obtained by applying the operator Op2 on  $\mathcal{M}_1$

We prove the correctness of the two operations one after the other. For the first operation, we use a result from [Sakr and Schmidt, 2022], which inspired our optimization, and which uses the inference rules shown in Figure 7.

**Lemma 8.**  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O}))) \equiv_\Sigma \mathcal{M}_1$ .

*Proof.* Let  $\mathcal{O}_{red}, \mathcal{M}_{red}$  be the ontologies obtained by applying the rules in Figure 7 on  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$  and  $\mathcal{M}_1$  respectively. By [Sakr and Schmidt, 2022, Theorem 5], we have

$$\begin{aligned} \text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O}))) &\equiv_\Sigma \mathcal{O}_{red}, \\ \mathcal{M}_1 &\equiv_\Sigma \mathcal{M}_{red}. \end{aligned}$$

Because the rule conD-Elim used for computing  $\mathcal{M}_1$  is a special case of the RP rule, we have  $\mathcal{O}_{red} = \mathcal{M}_{red}$ . As a consequence, we obtain  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O}))) \equiv_\Sigma \mathcal{M}_1$ .  $\square$

**Lemma 9.**  $\mathcal{M}_1 \equiv_\Sigma \mathcal{M}_2$ .

*Proof.* It is shown in [Zhao and Schmidt, 2017, Theorem 1], as an easy consequence of Ackermann's lemma, that if the set of all negative occurrences of a definer  $D$  in  $\mathcal{O}$  is of the form  $\{\neg D \sqcup C_j \mid 1 \leq j \leq n\}$ , then we can replace all positive occurrences of  $D$  in  $\mathcal{O}$  by  $\prod_{j=1}^n C_j$  without losing logical consequences that do not involve  $D$ . This is exactly what Op2 does.  $\square$

**Theorem 5.** Let  $gm_\Sigma^*(\mathcal{O})$  be the ontology obtained by:

- successive application of Op1 and Op2 over  $\text{conE}_\Sigma(\text{rolE}_\Sigma(\text{RI}_\Sigma(\mathcal{O})))$ , and then
- application of the steps described in Theorem 4.

Then,  $gm_\Sigma^*(\mathcal{O})$  is a general module for  $\mathcal{O}$  and  $\Sigma$ .

*Proof.* This can be shown almost in the same way as in the proof for Theorem 4. The only difference is that we change the definition of  $\mathcal{O}_D$  in Equation (5) to

$$\mathcal{O}_D = \{D \equiv \text{copy}_\Sigma(C_D) \mid D \in \text{sig}(\mathcal{M}_2) \cap \mathbb{N}_D\}.$$

The reason is that we do not need to consider definers eliminated by the rules con-Elim and D-prop anymore.  $\square$

## 9.8 Proof of Theorem 6

**Theorem 6.** Let us define  $dm_\Sigma(\mathcal{O})$  by

$$dm_\Sigma(\mathcal{O}) = \{\alpha \in \mathcal{O} \mid \alpha R^* \beta \text{ for some } \beta \in gm_\Sigma^*(\mathcal{O})\}.$$

Then,  $dm_\Sigma(\mathcal{O})$  is a deductive module for  $\mathcal{O}$  and  $\Sigma$ .

*Proof.* Set  $\mathcal{M} = dm_\Sigma(\mathcal{O})$ . We note that  $gm_\Sigma(\mathcal{M}) = gm_\Sigma(\mathcal{O})$ , since  $\mathcal{M}$  contains exactly the set of axioms that are used to compute  $gm_\Sigma(\mathcal{O})$ . By Theorem 4, we have  $\mathcal{M} \equiv_\Sigma gm_\Sigma(\mathcal{M})$  and  $\mathcal{O} \equiv_\Sigma gm_\Sigma(\mathcal{O})$ . Putting these observations together, we obtain  $\mathcal{M} \equiv_\Sigma \mathcal{O}$ . Since  $\mathcal{M} \subseteq \mathcal{O}$ ,  $\mathcal{M}$  is a deductive module of  $\mathcal{O}$  for  $\Sigma$ .  $\square$