Lower and Upper Approximations for Depleting Modules of Description Logic Ontologies

William Gatens and **Boris Konev** and **Frank Wolter**
University of Liverpool, UK

Abstract. It is known that no algorithm can extract the minimal depleting Σ-module from ontologies in expressive description logics (DLs). Thus research has focused on algorithms that approximate minimal depleting modules 'from above' by computing a depleting module that is not necessarily minimal. The first contribution of this paper is an implementation (AMEX) of such a depleting module extraction algorithm for expressive acyclic DL ontologies that uses a QBF solver for checking conservative extensions relativised to singleton interpretations. To evaluate AMEX and other module extraction algorithms we propose an algorithm approximating minimal depleting modules 'from below' (which also uses a QBF solver). We present experiments based on NCI (the National Cancer Institute Thesaurus) that indicate that our lower approximation often coincides with (or is very close to) the upper approximation computed by AMEX, thus proving for the first time that an approximation algorithm for minimal depleting modules can be almost optimal on a large ontology. We use the same technique to evaluate locality-based module extraction and a hybrid approach on NCI.

1 Introduction

Module extraction is the task of computing, given an ontology and a signature Σ of interest, a subset (called module) of the ontology such that for certain applications that use the signature Σ only, the original ontology can be equivalently replaced by the module [16, 17, 15, 13]. In most applications of module extraction it is desirable to compute a small (and, if possible, even minimal) module. In logic-based approaches to module extraction, the most robust and popular way to define modules is via model-theoretic Σ-inseparability, where two ontologies are called Σ-inseparable iff the Σ-reducts of their models coincide. Then, a Σ-module of an ontology is defined as a Σ-inseparable subset of the ontology [8, 4]. It is often helpful and necessary to refine this notion of Σ-module by considering self-contained Σ-modules (modules that are inseparable from the ontology not only w.r.t. Σ but also w.r.t. their own signature) and depleting Σ modules (modules such that the remaining axioms in the ontology are inseparable from the empty ontology w.r.t. Σ and the signature of the module) [3, 11, 9]. Note that every depleting module is self-contained and every self-contained module is a module. In all three cases it is often not possible to compute Σ-modules: by results in [9, 12], for acyclic $\mathcal{ALC}$-TBoxes and general $\mathcal{EL}$-TBoxes it is undecidable whether a given subset of a TBox is a (self-contained, depleting) Σ-module. The "maximal" description logics (DLs) for which efficient algorithms computing minimal self-contained and depleting Σ-modules have been developed are acyclic $\mathcal{ELI}$ [9] and DL-Lite [10, 11, 7].[1] For this reason, for module extraction in ontologies given in expressive DLs one has to employ approximation algorithms: instead of computing a minimal (self-contained, depleting) Σ-module, one computes some (self-contained, depleting) Σ-module and the main research problem is to minimise the size of the module (or, equivalently, to approximate minimal modules). Currently, the most popular and successful approximation algorithm is based on locality and computes the so-called $\top\bot^*$-module [18] which is a (possibly not minimal) depleting module. The size of $\top\bot^*$-modules and the performance of algorithms extracting $\top\bot^*$-modules has been analysed systematically and in great detail [18]. However, for expressive DLs neither alternative implemented depleting module extraction algorithms nor any lower approximation algorithms for depleting modules were available. So it remained open how large and significant the difference between a $\top\bot^*$-module and the minimal depleting module is and in how far it is possible to improve upon the approximation obtained by $\top\bot^*$-modules.[2]

The aim of this paper is to start to fill this gap and investigate how close one can approximate the ideal minimal depleting modules of an ontology by a sound extraction algorithm. To this end, we present algorithms and experiments that extract *two approximations* of depleting minimal modules of an ontology:

- an *upper approximation* which is a depleting module that is possibly larger than the minimal one; and
- a *lower approximation* which is contained in the minimal depleting module but which is not guaranteed to be a depleting module.

Clearly, if the upper and lower approximation coincide (or are very close to each other), then we know that the upper approximation coincides with (or comes very close to) the minimal depleting module. In detail, the contribution of this paper is as follows.

1. For the upper approximation, we have extended, optimised, and implemented the depleting module extraction algorithm introduced in [9] for acyclic $\mathcal{ALCQI}$-TBoxes. The implementation (called AMEX) covers repeated concept inclusions (as present in NCI) and uses a QBF solver as an oracle. AMEX is available from `http://www.csc.liv.ac.uk/~wgatens/software/amex.html`.

[1] For typical DL-Lite dialects, model-theoretic Σ-inseparability is decidable. Experimental evaluations of module extraction algorithms are, however, available only for language dependent notions of inseparability.

[2] An implementation of semantic locality-based $\Delta\emptyset^*$-modules and a comparison between $\top\bot^*$ and $\Delta\emptyset^*$-modules have been presented in [18]; however, the authors found no significant difference between the two approaches.

2. For the lower approximation, we introduce an algorithm that extracts the minimal *1-depleting* Σ*-module* of an arbitrary $\mathcal{ALCQI}$-TBox; that is, the minimal subset $\mathcal{M}$ of an ontology $\mathcal{T}$ such that $\mathcal{T} \setminus \mathcal{M}$ is inseparable on singleton interpretations from the empty ontology w.r.t. Σ and the signature of the module. Again, the implementation uses a QBF solver as an oracle.
3. Using NCI and minimal 1-depleting Σ-modules, we present the first experimental evaluation of how close $\top\bot^*$-modules, AMEX-modules, and hybrid modules (that result from iterating both algorithms) approximate the 'real' minimal depleting module.

Interestingly, QBF solvers have been used before in module extraction for DL-Lite [10, 11]. However, our application is completely different from their application in [10, 11]. This paper extends the workshop paper [5].

2 Preliminaries

We use standard notation from logic and description logic (DL), details can be found in [1]. In a DL, concepts are constructed from countably infinite sets $\mathsf{N_C}$ of *concept names* and $\mathsf{N_R}$ of *role names* using the concept constructors defined by the DL. For example, $\mathcal{ALCQI}$*-concepts* are built according to the rule

$$C ::= A \mid \neg C \mid \geq n\, r.C \mid \geq n\, r^-.C \mid C \sqcap D,$$

where $A \in \mathsf{N_C}$, n is a natural number, and $r \in \mathsf{N_R}$. As usual, we use the following abbreviations: $\bot$ stands for $A \sqcap \neg A$, $\top$ denotes $\neg\bot$, $\exists r.C$ denotes $\geq 1\, r.C$, $\forall r.C$ denotes $\neg\exists\, r.\neg C$, $C \sqcup D$ denotes $\neg(\neg C \sqcap \neg D)$, $\leq n\, r.C$ denotes $\neg(\geq (n+1)\, r.C)$, and $(= n\, r.C)$ stands for $((\geq n\, r.C) \sqcap (\leq n\, r.C))$.

A *general TBox* $\mathcal{T}$ is a finite set of *axioms*, where an axiom can be either a *concept inclusion (CI)* $C \sqsubseteq D$ or a *concept equality (CE)* $C \equiv D$, where C and D are concepts. A general TBox $\mathcal{T}$ is *acyclic* if all its axioms are of the form $A \sqsubseteq C$ or $A \equiv C$, where $A \in \mathsf{N_C}$, no concept name occurs more than once on the left-hand side and $A \not\prec^+_{\mathcal{T}} A$, for any $A \in \mathsf{N_C}$, where $\prec^+_{\mathcal{T}}$ is the transitive closure of the relation $\prec_{\mathcal{T}} \subseteq \mathsf{N_C} \times (\mathsf{N_C} \cup \mathsf{N_R})$ defined by setting $A \prec_{\mathcal{T}} X$ iff there exists an axiom of the form $A \sqsubseteq C$ or $A \equiv C$ in $\mathcal{T}$ such that X occurs in C.

The semantics of DLs is given by *interpretations* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where the *domain* $\Delta^{\mathcal{I}}$ is a non-empty set and $\cdot^{\mathcal{I}}$ is an *interpretation function* that maps each $A \in \mathsf{N_C}$ to a subset $A^{\mathcal{I}}$ of $\Delta^{\mathcal{I}}$ and each $r \in \mathsf{N_R}$ to a binary relation $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. The function $\cdot^{\mathcal{I}}$ is inductively expanded to complex concepts C in the standard way [1]. An interpretation $\mathcal{I}$ *satisfies* a CI $C \sqsubseteq D$ (written $\mathcal{I} \models C \sqsubseteq D$) if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$, it *satisfies* a CE $C \equiv D$ (written $\mathcal{I} \models C \equiv D$) if $C^{\mathcal{I}} = D^{\mathcal{I}}$. $\mathcal{I}$ is a *model* of $\mathcal{T}$ if it satisfies all axioms in $\mathcal{T}$.

A *signature* Σ is a finite subset of $\mathsf{N_C} \cup \mathsf{N_R}$. The signature $\mathsf{sig}(C)$ ($\mathsf{sig}(\alpha)$, $\mathsf{sig}(\mathcal{T})$) of a concept C (axiom α, TBox $\mathcal{T}$, resp.) is the set of concept and role names that occur in C ($\alpha, \mathcal{T}$, resp.). If $\mathsf{sig}(C) \subseteq \Sigma$ we call C a Σ-concept. The Σ*-reduct* $\mathcal{I}|_\Sigma$ of an interpretation $\mathcal{I}$ is obtained from $\mathcal{I}$ by setting $\Delta^{\mathcal{I}|_\Sigma} = \Delta^{\mathcal{I}}$, and $X^{\mathcal{I}|_\Sigma} = X^{\mathcal{I}}$ for all $X \in \Sigma$, and $X^{\mathcal{I}|_\Sigma} = \emptyset$ for all $X \notin \Sigma$. Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be TBoxes and Σ a signature. Then $\mathcal{T}_1$ and $\mathcal{T}_2$ are Σ*-inseparable*, in symbols $\mathcal{T}_1 \equiv_\Sigma \mathcal{T}_2$, if

$$\{\mathcal{I}|_\Sigma \mid \mathcal{I} \models \mathcal{T}_1\} = \{\mathcal{I}|_\Sigma \mid \mathcal{I} \models \mathcal{T}_2\}.$$

It is proved in [9] that TBoxes $\mathcal{T}_1$ and $\mathcal{T}_2$ are Σ-inseparable if, and only if, $\mathcal{T}_1 \models \varphi$ iff $\mathcal{T}_2 \models \varphi$ holds for any second-order sentence φ using symbols from Σ only. Thus, Σ-inseparable TBoxes are indistinguishable not only in applications using entailed CIs between Σ-concepts but also in data access applications with data given in Σ. We use Σ-inseparability to define modules.

> **Input**: $\mathcal{ALCQI}$-TBox $\mathcal{T}$, Signature Σ
> **Output**: Minimal $\mathcal{M} \subseteq \mathcal{T}$ s.t. $\mathcal{T} \setminus \mathcal{M} \equiv^1_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$.
>
> Set $\mathcal{M} := \emptyset$ and apply exhaustively the following rule
> **(1-insep)** If $\alpha \in \mathcal{T} \setminus \mathcal{M}$ is a *1-separability causing axiom* then set $\mathcal{M} := \mathcal{M} \cup \{\alpha\}$

Figure 1. Minimal 1-depleting module extraction in $\mathcal{ALCQI}$

Definition 1 *Let $\mathcal{M} \subseteq \mathcal{T}$ be TBoxes and Σ a signature. Then $\mathcal{M}$ is a* depleting Σ-module *of $\mathcal{T}$ if $\mathcal{T} \setminus \mathcal{M} \equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$.*

Every depleting module $\mathcal{M}$ of $\mathcal{T}$ is inseparable from $\mathcal{T}$ for its signature, that is, if $\mathcal{M}$ is a depleting Σ-module of $\mathcal{T}$ then $\mathcal{T} \equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \mathcal{M}$, and, in particular, $\mathcal{T} \equiv_\Sigma \mathcal{M}$. Thus, a TBox and its depleting Σ-module can be equivalently replaced by each other in applications which concern Σ only. Throughout this paper we use the fact that minimal depleting Σ-modules of a TBox are uniquely determined [11]. For further discussion of the advantages of depleting modules we refer the reader to [3, 9]. Unfortunately, checking if a subset $\mathcal{M}$ of $\mathcal{T}$ is a depleting Σ-module of $\mathcal{T}$ for some given signature Σ is undecidable already for general TBoxes formulated in $\mathcal{EL}$ and for acyclic $\mathcal{ALC}$-TBoxes [9, 12].

3 Lower Approximation: 1-Depleting Modules

We introduce a lower approximation of depleting Σ-modules and give an algorithm extracting such approximations from arbitrary $\mathcal{ALCQI}$-TBoxes. The results of this section can be easily extended to arbitrary first-order ontologies.

Assume that $\mathcal{T}_1$ and $\mathcal{T}_2$ are TBoxes and Σ a signature. Then $\mathcal{T}_1$ and $\mathcal{T}_2$ are *1-Σ-inseparable*, in symbols $\mathcal{T}_1 \equiv^1_\Sigma \mathcal{T}_2$, if

$$\{\mathcal{I}|_\Sigma \mid \sharp\Delta^{\mathcal{I}} = 1 \text{ and } \mathcal{I} \models \mathcal{T}_1\} = \{\mathcal{I}|_\Sigma \mid \sharp\Delta^{\mathcal{I}} = 1 \text{ and } \mathcal{I} \models \mathcal{T}_2\}.$$

If $\mathcal{T}_1$ and $\mathcal{T}_2$ are Σ-inseparable, then they are 1-Σ-inseparable. The following example shows that the converse does not hold.

Example 2 *Let $\mathcal{T}$ be the following subset of NCI:*

$\mathsf{Thoracic\_Cavity} \sqsubseteq \exists \mathsf{has\_Location}.\mathsf{Thorax}$
$\mathsf{Pleural\_Tissue} \sqsubseteq \forall \mathsf{has\_Location}.\mathsf{Thoracic\_Cavity}$

and let $\Sigma = \{\mathsf{Pleural\_Tissue}, \mathsf{has\_Location}\}$. Then one can show that $\mathcal{T} \equiv^1_\Sigma \emptyset$, but $\mathcal{T} \not\equiv_\Sigma \emptyset$.

In contrast to Σ-inseparability which is undecidable, 1-Σ-inseparability can be decided by reduction to the validity of quantified Boolean formulas (QBF). For simplicity, we consider 1-Σ-inseparability between the empty TBox and $\mathcal{ALCQI}$-TBoxes. Given $\mathcal{T}$ and Σ, take a propositional variable p_A for each concept name $A \in \Sigma$ and a (distinct) propositional variable q_X for each symbol $X \in \mathsf{sig}(\mathcal{T}) \setminus \Sigma$. Translate concepts D in the signature $\mathsf{sig}(\mathcal{T})$ into propositional formulas $D^\dagger$ by setting

$$\begin{aligned} A^\dagger &= p_A, \quad \text{for all } A \in \Sigma \\ A^\dagger &= q_A, \quad \text{for all } A \in \mathsf{sig}(\mathcal{T}) \setminus \Sigma \\ (D_1 \sqcap D_2)^\dagger &= D_1^\dagger \wedge D_2^\dagger \\ (\neg D)^\dagger &= \neg D^\dagger \end{aligned}$$

$$(\geq n\, r.D)^\dagger = (\geq n\, r^-.D)^\dagger = \begin{cases} p_r \wedge D^\dagger, \text{if } n = 1 \text{ and } r \in \Sigma \\ q_r \wedge D^\dagger, \text{if } n = 1 \text{ and } r \in \mathsf{sig}(\mathcal{T}) \setminus \Sigma \\ \bot, \text{ else} \end{cases}$$

(1)	$\mathsf{Renal\_Pelvis\_and\_U}$	$\sqsubseteq$	$\exists\mathsf{partOf.K\_and\_U}$
(2)	$\mathsf{K\_and\_U\_Neoplasm}$	$\equiv$	$\mathsf{U\_T\_Neoplasm} \sqcap (\forall\mathsf{hasSite.K\_and\_U})$
(3)	$\mathsf{Malignt\_U\_T\_Neoplasm}$	$\equiv$	$\mathsf{U\_T\_Neoplasm} \sqcap (\forall\mathsf{hasAbnCell.Malignt\_Cell})$
(4)	$\mathsf{Benign\_U\_T\_Neoplasm}$	$\equiv$	$\mathsf{U\_T\_Neoplasm} \sqcap (\forall\mathsf{excludesAbnCell.Malignt\_Cell})$

Figure 2. TBox for Example 6

Now let $\mathcal{T}^\dagger = \bigwedge_{C \sqsubseteq D \in \mathcal{T}} C^\dagger \rightarrow D^\dagger \wedge \bigwedge_{C \equiv D \in \mathcal{T}} C^\dagger \leftrightarrow D^\dagger$ and let $\vec{p}$ denote the sequence of variables p_A, $A \in \Sigma$, and $\vec{q}$ denote the sequence of variables q_X, $X \in \mathsf{sig}(\mathcal{T}) \setminus \Sigma$.

Lemma 3 *$\mathcal{T} \equiv^1_\Sigma \emptyset$ iff the QBF $\varphi_\mathcal{T} := \forall\vec{p}\exists\vec{q}\mathcal{T}^\dagger$ is valid.*

We define 1-depleting Σ-modules in the same way as depleting Σ-modules except that 1-Σ-inseparability replaces inseparability:

Definition 4 *Let $\mathcal{M} \subseteq \mathcal{T}$ be TBoxes and Σ a signature. Then $\mathcal{M}$ is a* 1-depleting Σ-module *of $\mathcal{T}$ if $\mathcal{T} \setminus \mathcal{M} \equiv^1_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$.*

Example 2 shows that there are acyclic $\mathcal{ALC}$-TBoxes in which not every 1-depleting Σ-module is a depleting Σ-module. Note, however, that using results from [9] one can prove that for acyclic $\mathcal{ELI}$-TBoxes 1-depleting Σ-modules coincide with depleting Σ-modules.

One can show that 1-Σ-inseparability is a monotone inseparability relation with the replacement property, as defined in [11]. Thus, it follows from [11] that for any TBox $\mathcal{T}$ and signature Σ there is a unique minimal 1-Σ-depleting module $\mathcal{M}$ of $\mathcal{T}$. Moreover, the minimal 1-depleting Σ-module of $\mathcal{T}$ is always contained in the unique minimal depleting Σ-module of $\mathcal{T}$. Thus, it is always a lower approximation of the minimal depleting Σ-module of $\mathcal{T}$.

Definition 4 can be used directly for a naïve minimal 1-depleting module extraction algorithm which goes through all subsets of $\mathcal{T}$ to identify a smallest possible $\mathcal{M}$ such that $\mathcal{T} \setminus \mathcal{M} \equiv^1_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$. Instead, in our implementation we consider an approach based on the notion of a *1-separability causing axiom*. Call an $\alpha \in \mathcal{T} \setminus \mathcal{M}$ 1-separability causing if there exists a $\mathcal{W} \subseteq \mathcal{T} \setminus \mathcal{M}$ such that

$$\alpha \in \mathcal{W}; \quad (\mathcal{W} \setminus \{\alpha\}) \equiv^1_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset; \quad \mathcal{W} \not\equiv^1_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset.$$

Clearly, if $\mathcal{T} \setminus \mathcal{M} \not\equiv^1_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$ then $\mathcal{T} \setminus \mathcal{M}$ contains a 1-separability causing axiom. Now one can show the following

Theorem 5 *Given an $\mathcal{ALCQI}$-TBox $\mathcal{T}$ and signature Σ, the algorithm in Figure 1 computes the unique minimal 1-depleting Σ-module of $\mathcal{T}$.*

The algorithm is in polynomial time with each call to a QBF solver treated as a constant time oracle call. To reduce the number of calls to the QBF solver, the search for a 1-separability causing axiom can be implemented as binary search. Hence, in the worst case one performs $\log_2(|\mathcal{T} \setminus \mathcal{M}|)$ inseparability checks to locate a 1-separability causing axiom.

Example 6 We apply the algorithm in Figure 1 to the fragment $\mathcal{T}$ of NCI given in Figure 2 with Σ defined as $\mathsf{Malignt\_U\_T\_Neoplasm}$, $\mathsf{K\_and\_U\_Neoplasm}$, and $\mathsf{Renal\_Pelvis\_and\_U}$. Here 'K', 'U', and 'T' abbreviate 'kidney', 'ureter' and 'tract', respectively.

The search for 1-separability causing axioms first establishes that $\mathcal{T} \not\equiv^1_\Sigma \emptyset$. An example showing this is $\mathcal{I}$ with $\Delta^\mathcal{I} = \{d\}$ such that $\mathsf{Renal\_Pelvis\_and\_U}^\mathcal{I} = \mathsf{Malignt\_U\_T\_Neoplasm}^\mathcal{I} = \{d\}$ and $\mathsf{K\_and\_U\_Neoplasm}^\mathcal{I} = \emptyset$. Then no $\mathcal{J}$ with $\mathcal{J}|_\Sigma = \mathcal{I}|_\Sigma$ satisfies $\mathcal{T}$.

Then the search splits $\mathcal{T}$ into two parts, $\mathcal{T}_1 = \{(1),(2)\}$ and $\mathcal{T}_2 = \{(3),(4)\}$. We have $\mathcal{T}_1 \equiv^1_\Sigma \emptyset$ and the algorithm 'grows' $\mathcal{T}_1$ with the upper part of $\mathcal{T}_2$ to construct $\mathcal{T}_3 = \{(1),(2),(3)\}$. We have $\mathcal{T}_3 \not\equiv^1_\Sigma \emptyset$ and so the algorithm identifies (3) as a 1-separability causing axiom and applies the rule (**1-insep**). It obtains $\mathcal{M} = \{(3)\}$.

The binary search proceeds with $\mathcal{T} \setminus \mathcal{M} = \{(1),(2),(4)\}$ and determines that $\mathcal{T} \setminus \mathcal{M} \not\equiv^1_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$. Then it proceeds to identify (2) as a 1-separability causing axiom and after applying the rule (**1-insep**), sets $\mathcal{M} = \{(2),(3)\}$. Finally, the algorithm proceeds with $\mathcal{T} \setminus \mathcal{M} = \{(1),(4)\}$ and identifies (1) as a 1-separability causing axiom, so $\mathcal{M}$ is set to $\{(1),(2),(3)\}$.

The rule (**1-insep**) does not apply any further and the computation finishes with the minimal 1-depleting module $\mathcal{M} = \{(1),(2),(3)\}$.

4 Upper Approximations

In this section, we provide an upper approximation algorithm for depleting Σ-modules in *acyclic* $\mathcal{ALCQI}$-TBoxes that extends and optimises the algorithm presented in [9]. The algorithm is also based on 1-Σ-inseparability but uses an additional dependency check to ensure that a depleting module is extracted. We also address the problem of extracting depleting modules from 'acyclic' TBoxes with multiple CIs for a single concept name and of combining depleting module extraction algorithms.

Let $\mathcal{T}$ be an acyclic TBox and Σ a signature. We say that $\mathcal{T}$ *has a direct Σ-dependency* if there exists $\{A, X\} \subseteq \Sigma$ with $A \prec^+_\mathcal{T} X$; otherwise we say that $\mathcal{T}$ *has no direct Σ-dependencies*. Although one can construct TBoxes $\mathcal{T}$ and depleting Σ-modules $\mathcal{M}$ of $\mathcal{T}$ such that $\mathcal{T} \setminus \mathcal{M}$ contains direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$-dependencies (see [9]), for typical depleting Σ-modules $\mathcal{M}$, the set $\mathcal{T} \setminus \mathcal{M}$ should not contain direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$-dependencies because such dependencies indicate a semantic link between two distinct symbols in $\Sigma \cup \mathsf{sig}(\mathcal{M})$. We show that Σ-inseparability reduces to 1-Σ-inseparability if one does not have direct Σ-dependencies. In detail, let for an acyclic TBox $\mathcal{T}$ and signature Σ

$$\mathsf{Lhs}_\Sigma(\mathcal{T}) = \{A \bowtie C \in \mathcal{T} \mid A \in \Sigma \text{ or } \exists X \in \Sigma \; (X \prec^+_\mathcal{T} A)\}.$$

The following lemma is proved in [9] for acyclic $\mathcal{ALCI}$-TBoxes and generalised here to acyclic $\mathcal{ALCQI}$-TBoxes.

Lemma 7 *Let $\mathcal{T}$ be an acyclic $\mathcal{ALCQI}$-TBox. If $\mathcal{T} \setminus \mathcal{M}$ has no direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$-dependencies, then the following conditions are equivalent for every $\mathcal{W} \subseteq \mathcal{T} \setminus \mathcal{M}$:*

- *$\mathcal{W} \equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$;*
- *$\mathsf{Lhs}_{\Sigma \cup \mathsf{sig}(\mathcal{M})}(\mathcal{W}) \equiv^1_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$.*

The algorithm computing a depleting Σ-module of acyclic $\mathcal{ALCQI}$-TBoxes is now given in Figure 3. In the algorithm, the extraction of depleting Σ-modules is broken into the the rule (**syn**) that checks for direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$-dependencies and the rule (**1-insep**) from the algorithm in Figure 1. It follows from Lemma 7 that if neither (**syn**) nor

Input: Acyclic $\mathcal{ALCQI}$ TBox $\mathcal{T}$, Signature Σ
Output: Minimal Module $\mathcal{M}$ s.t $\mathcal{T} \setminus \mathcal{M} \equiv_{\Sigma\cup\mathsf{sig}(\mathcal{M})} \emptyset$ and $\mathcal{T} \setminus \mathcal{M}$ has no direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$ dependencies.

Set $\mathcal{M} := \emptyset$ and apply rules (**syn**) and (**1-insep**) exhaustively, preferring rule (**syn**).

(syn) If an axiom $A \bowtie C \in \mathcal{T} \setminus \mathcal{M}$ is such that $A \in \Sigma \cup \mathsf{sig}(\mathcal{M}))$ and $A \prec^{+}_{\mathcal{T}\setminus\mathcal{M}} X$, for some $X \in (\Sigma \cup \mathsf{sig}(\mathcal{M}))$, then set $\mathcal{M} := \mathcal{M} \cup \{A \bowtie C\}$

(1-insep) If an axiom $A \bowtie C \in \mathcal{T} \setminus \mathcal{M}$ is a *1-separability causing axiom* then set $\mathcal{M} := \mathcal{M} \cup \{A \bowtie C\}$

Figure 3. Module extraction in $\mathcal{ALCQI}$

Input: Acyclic $\mathcal{ALCQI}$ TBox $\mathcal{T}$ with repeated CIs, Signature Σ
Output: A module $\mathcal{M}$ s.t $\mathcal{T} \setminus \mathcal{M} \equiv_{\Sigma\cup\mathsf{sig}(\mathcal{M})} \emptyset$.

Let $\mathcal{T}' := \{B \bowtie D \in \mathcal{T} \mid B$ is not repeated$\}$.
Let $\mathcal{T}^A := \{A'_i \sqsubseteq C_i \mid A \sqsubseteq C_i \in \mathcal{T}\} \cup \{A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n\}$, for A repeated, $A'_1, \ldots, A'_n$ fresh.
Let $\mathcal{M}'$ be depleting Σ-module of $\mathcal{T}' \cup \bigcup_{A \text{ is repeated}} \mathcal{T}^A$.
Let $\mathcal{M}$ be obtained from $\mathcal{M}'$ by dropping the added axioms of the form $A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n$ and replacing every occurrence of the introduced symbols $A'_1, \ldots, A'_n$ with A.

Figure 4. Module extraction in $\mathcal{ALCQI}$ with repeated CIs

(**1-insep**) is applicable then $\mathcal{T} \setminus \mathcal{M} \equiv_{\Sigma\cup\mathsf{sig}(\mathcal{M})} \emptyset$ and so the output of the algorithm in Figure 3 is a depleting Σ-module. More precisely, one can show the following characterisation of the extracted module:

Theorem 8 *Given an acyclic $\mathcal{ALCQI}$ TBox $\mathcal{T}$ and signature Σ the algorithm in Figure 3 computes the unique minimal depleting Σ-module s.t. $\mathcal{T} \setminus \mathcal{M}$ contains no direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$-dependencies.*

The algorithm is again in polynomial time with each call to the QBF solver being treated as a constant time oracle call. Note that the minimality condition in the theorem means that for any $\mathcal{M}' \subseteq \mathcal{T}$ such that $\mathcal{T} \setminus \mathcal{M}'$ has no direct $\Sigma \cup \mathsf{sig}(\mathcal{M}')$-dependencies and $\mathcal{T} \setminus \mathcal{M}' \equiv_{\Sigma\cup\mathsf{sig}(\mathcal{M}')} \emptyset$ we have $\mathcal{M} \subseteq \mathcal{M}'$. It is, however, still possible that there exists a $\mathcal{M}'' \subseteq \mathcal{T}$ with $\mathcal{T} \setminus \mathcal{M}'' \equiv_{\Sigma\cup\mathsf{sig}(\mathcal{M}'')} \emptyset$, $\mathcal{M} \not\subseteq \mathcal{M}''$ and such that $\mathcal{T} \setminus \mathcal{M}''$ has some direct $\Sigma \cup \mathsf{sig}(\mathcal{M}'')$-dependencies.

Example 9 (Example 6 continued) We apply the algorithm in Figure 3 to the same TBox and signature as in Example 6. The rule (**syn**) is not applicable. Therefore, as in Example 6, an application of the rule (**1-insep**) sets $\mathcal{M} = \{(3)\}$ and the rule (**syn**) immediately adds axioms (1) and (2) to $\mathcal{M}$. Neither (**syn**) nor (**1-insep**) apply to $\mathcal{T} \setminus \mathcal{M} = \{(4)\}$ and the computation concludes with $\mathcal{M} = \{(1), (2), (3)\}$. Thus, the computed depleting Σ-module of $\mathcal{T}$ coincides with the minimal 1-depleting Σ-module of $\mathcal{T}$ and our approximation is optimal.

Notice that axiom (4) is neither Δ- nor $\emptyset$-local for $\Sigma \cup \mathsf{sig}(\mathcal{M})$ and so the $\top\bot^*$-module of $\mathcal{T}$ w.r.t. Σ coincides with $\mathcal{T}$ (see below and [3] for definitions).

It is often the case (e.g., for the NCI Thesaurus) that a real-world ontology satisfies all conditions for acyclic TBoxes with the exception that it contains multiple CIs of the form $A \sqsubseteq C_1, \ldots, A \sqsubseteq C_n$. We call such TBoxes *acyclic with repeated CIs* and say that A is a *repeated concept name*. Clearly, one can convert such a TBox into an equivalent acyclic TBox by replacing the repeated CIs with $A \sqsubseteq C_1 \sqcap \ldots \sqcap C_n$. However, such an explicit conversion is an unattractive solution for module extraction because if such an axiom is added to a Σ-module the signature of the module now contains every symbol in the definition of every repeated name increasing the size of the resulting module considerably. The approach we take to handle acyclic TBoxes with repeated CIs is to introduce fresh concept names for different repeated occurrences of a concept name in the left-hand side of concept CIs, extract modules from the resulting acyclic TBox and then substitute away the added concept names.

Theorem 10 *Let $\mathcal{T}$ be an acyclic TBox with repeated CIs and Σ a signature. Then $\mathcal{M}$ computed by the algorithm in Figure 4 is a depleting Σ-module of $\mathcal{T}$.*

We close this section with a result about nested depleting modules which is used in the next section to guarantee that combinations of depleting module extraction algorithms extract a depleting module.

Theorem 11 *Let $\mathcal{M} \subseteq \mathcal{M}' \subseteq \mathcal{T}$ be TBoxes and Σ a signature such that $\mathcal{M}'$ is a depleting Σ-module of $\mathcal{T}$ and $\mathcal{M}$ is a depleting Σ-module of $\mathcal{M}'$. Then $\mathcal{M}$ is a depleting Σ-module of $\mathcal{T}$.*

5 Experiments and Evaluation

To evaluate how close depleting module extraction algorithms can approximate minimal depleting modules we consider the following three implementations of upper approximation algorithms for minimal depleting module extraction:

- Our system AMEX that implements the algorithm in Figure 3 and the refinement for acyclic TBoxes with repeated CIs. AMEX is written in Java aided by the OWL-API library [6] for ontology manipulation. The inseparability check was implemented using the reduction to the validity of QBF (Lemma 3) and uses the QBF solver sKizzo [2].
- $\top\bot^*$ locality-based module extraction [3, 14] as implemented in the OWL-API library version 3.2.4.1806 (called STAR-modules for ease of pronunciation).
- A hybrid approach in which one iterates AMEX and STAR-module extraction. This results in a depleting module contained in both the AMEX and the STAR-module (Theorem 11).

For the lower approximation, we implemented the algorithm in Figure 1 to compute the minimal 1-depleting module. The inseparability check was again implemented using the reduction to the validity of QBF and uses sKizzo. In our experiments we use the NCI Thesaurus version 08.09d taken from the Bioportal [19] containing 98 752 TBox axioms among which 87 934 are CIs of the form $A \sqsubseteq C$ and 10 366 are CEs of the form $A \equiv C$. In what follows, NCI$^\star(\sqsubseteq)$ denotes the TBox consisting of all such CIs, NCI$^\star(\equiv)$ denotes the TBox consisting of all such CEs, and NCI* denotes the union of both. All three TBoxes are acyclic (with repeated CIs), so AMEX can be applied to them. NCI* together with the rest of the ontology (452 axioms) is called NCI and contains, in addition, role inclusions, domain and range restrictions, and disjointness axioms.

Role%	0%					50%					100%				
$\|\Sigma\|$	Star (S)	AMEX (A)	Hybrid (H)	1-dep (D)	Diff/200	Star (S)	AMEX (A)	Hybrid (H)	1-dep (D)	Diff/200	Star (S)	AMEX (A)	Hybrid (H)	1-dep (D)	Diff/200
NCI*															
100	3834.21	722.21	710.65	671.68	10	3887.17	972.68	960.44	960.39	3	3915.18	1013.23	1000.79	1000.70	4
250	5310.96	1721.28	1705.71	1705.61	4	5452.52	1882.65	1870.87	1870.83	4	5539.39	1924.77	1912.95	1912.89	5
500	6977.33	2725.74	2700.00	2699.96	2	7186.09	2933.90	2919.23	2919.15	3	7237.22	2987.75	2977.62	2977.58	2
750	8235.36	3573.97	3542.57	3542.49	2	8437.07	3801.24	3786.05	3786.01	2	8579.98	3902.12	3892.36	3892.26	4
1000	9273.62	4341.25	4305.41	4305.38	1	9525.81	4570.55	4553.91	4553.81	4	9542.00	4621.42	4612.19	4606.46	3
NCI* ($\sqsubseteq$)															
100	58.74	69.53	58.74	58.74	0	291.91	326.68	291.91	291.89	2	345.01	357.58	345.01	344.89	5
250	330.79	386.45	330.79	330.78	1	652.09	716.64	652.09	652.09	0	775.00	808.03	775.00	775.00	0
500	852.14	1007.20	852.14	852.14	0	1173.34	1274.27	1173.34	1173.34	0	1387.67	1444.68	1387.67	1387.67	0
750	1352.47	1571.46	1352.47	1352.47	0	1681.12	1816.79	1681.12	1681.12	0	1935.47	2009.62	1935.47	1935.47	0
100	1788.02	2046.62	1788.02	1788.02	0	2152.83	2315.19	2152.83	2152.83	0	2434.06	2519.63	2434.06	2434.06	0
NCI* ($\equiv$)															
100	2760.96	310.25	310.25	309.21	122	2759.11	319.08	319.11	318.23	114	2782.54	318.79	318.79	317.73	130
250	3989.74	622.65	622.63	621.89	110	4000.93	623.38	623.25	622.50	104	3973.78	624.51	624.23	623.47	102
500	4994.77	1003.76	1003.75	1002.95	108	4983.10	1002.14	1002.04	1001.32	101	4986.77	999.87	999.87	999.08	101
750	5539.78	1310.33	1310.31	1309.38	124	5531.60	1313.51	1311.54	1310.67	90	5525.28	1307.71	1307.71	1306.85	106
1000	5886.91	1573.06	1573.14	1572.11	122	5901.34	1577.34	1572.14	1571.10	102	5903.37	1576.95	1571.18	1570.08	103

Figure 5. Modules of NCI* and its fragments

Most of NCI* (all but 4 588 axioms) are $\mathcal{EL}$-inclusions. The non-$\mathcal{EL}$ inclusions contain 7 806 occurrences of value restrictions. The signature of NCI* contains 68 862 concept and 88 role names.

Experiments with NCI* and its Fragments. The results given in Table 5 show the average sizes of the modules extracted by our four algorithms from the TBoxes NCI*, NCI$^\star(\sqsubseteq)$, and NCI$^\star(\equiv)$ over 200 random signatures for each signature size combination of 100 to 1000 concept names and 0%, 50%, and 100% of role names. In addition, in each case we give the number of signatures (out of 200) in which there is a difference between the hybrid module and the minimal 1-depleting module. It can be seen that

- in NCI* and NCI$^\star(\sqsubseteq)$ the hybrid module almost always coincides with the minimal 1-depleting module. Thus the hybrid module almost always coincides with the minimal depleting module.
- in NCI$^\star(\equiv)$, the hybrid module coincides with the minimal 1-depleting module (and therefore the minimal depleting module) for approximately 50% of all signatures. Moreover, on average the minimal 1-depleting module (and therefore the minimal depleting module) is less than 0.3% smaller than the hybrid module.
- in all three TBoxes, hybrid modules are only slightly smaller than AMEX-modules.
- in NCI$^\star(\equiv)$, AMEX-modules are significantly smaller than STAR-modules.
- in NCI$^\star(\sqsubseteq)$, STAR-modules are slightly smaller than AMEX modules.
- in NCI*, AMEX-modules are still significantly smaller than STAR-modules, but less so than in NCI$^\star(\equiv)$.

The very different behaviour of AMEX-modules and STAR-modules in NCI$^\star(\sqsubseteq)$ and NCI$^\star(\equiv)$ can be explained as follows: it is shown in [9] that for acyclic $\mathcal{EL}$-TBoxes without CEs, AMEX-modules and STAR-modules coincide. This is not the case for acyclic $\mathcal{ALCQI}$-TBoxes (there can be axioms in STAR-modules that are not AMEX-modules and vice versa), but since the vast majority of axioms in NCI$^\star(\sqsubseteq)$ are $\mathcal{EL}$-inclusions one should not expect any significant difference between the two types of modules. Thus, it is exactly those acyclic TBoxes that contain many CEs for which AMEX-modules are significantly smaller than STAR-modules (see Example 6 for an illustration).

It would be interesting to know how often (and for which reason) the minimal 1-depleting module is a depleting module (and thus coincides with the minimal depleting module). This is the case if the hybrid module coincides with the minimal 1-depleting module. For the remaining cases, we can currently only check this "by hand" (the problem is undecidable). Among the modules extracted in the experiments we found examples in which the minimal 1-depleting module is not a depleting module (Example 2 is based on such a module) and the hybrid module is identical to the minimal depleting module and we found examples in which the minimal depleting module is identical to the minimal 1-depleting module and smaller than the hybrid module. However, we were not able to determine a general pattern.

The experiments were carried out on a PC with an Intel i5 CPU @ 3.30GHz with 2GB of Java heap space available to the program. For AMEX the average time taken per extraction was just under 3s and the maximum time taken was 15s. In 97% of all experiments the QBF solver was called just once. In those cases the AMEX-modules were computed purely syntactically and the QBF solver simply provided an assurance that the extracted axioms indeed constituted a depleting module. In the remaining 3% of all AMEX extractions the maximal number of 1-separability causing axioms recorded was 4 and the maximal number of QBF solver calls themselves was 73.

In the minimal 1-depleting module extraction case, we first extracted the hybrid module and then applied the algorithm in Figure 1. The time for minimal 1-depleting module extraction varied considerably over the input TBoxes: for NCI$^\star(\sqsubseteq)$ a single 1-depleting module extraction took no more than 2 minutes, for NCI* and NCI$^\star(\equiv)$ a single 1-depleting module extraction took up to 30 minutes. This can be attributed to the number of QBF checks: for NCI$^\star(\sqsubseteq)$ the maximum number of QBF checks needed in a single extraction was 5,052, for NCI* 193,993, and for NCI$^\star(\equiv)$ 433,564 checks were required.

Experiments with full NCI. We present preliminary experiments for module extraction from full NCI. Note that AMEX cannot be directly applied to full NCI since it is not acyclic. To tackle this problem and apply AMEX to cyclic TBoxes one can split a general TBox $\mathcal{T}$ into

Role%	0%				50%				100%			
$\lvert\Sigma\rvert$	Star(S)	Hybrid (H)	1-depl (D)	Diff/200	Star(S)	Hybrid (H)	1-depl (D)	Diff/200	Star(S)	Hybrid (H)	1-depl (D)	Diff/200
100	5274.72	1905.86	1905.86	0	5409.98	2010.51	2010.51	0	5452.89	2079.85	2079.85	0
250	7306.66	3269.68	3269.68	0	7329.96	3329.10	3329.10	0	7360.41	3365.47	3365.47	0
500	9477.80	4833.99	4833.99	0	9575.15	4880.09	4880.09	0	9572.79	4920.82	4920.82	0
750	11044.98	6050.58	6050.58	0	11132.78	6105.77	6105.77	0	11121.40	6133.92	6133.92	0
1000	12393.10	7117.74	7117.74	0	12440.65	7165.53	7165.53	0	12455.42	7215.32	7215.32	0

Figure 6. Modules of NCI

two parts $\mathcal{T}_1$ and $\mathcal{T}_2$, where $\mathcal{T}_1$ is an acyclic TBox (and as large as possible) and $\mathcal{T}_2 := \mathcal{T} \setminus \mathcal{T}_1$. Then for any signature Σ it follows from the robustness properties [8] of the inseparability relation that if $\mathcal{M}$ is a depleting $\Sigma \cup \mathsf{sig}(\mathcal{T}_2)$-module of $\mathcal{T}_1$ (note that $\mathcal{M}$ can be computed by AMEX), then $\mathcal{M} \cup \mathcal{T}_2$ is a depleting Σ-module of $\mathcal{T}$ as well. Such a direct application of AMEX to general TBoxes is unlikely to compute small modules when $\mathcal{T}_2$ is large. However, the results below indicate that this approach is beneficial when iterated with STAR-module extraction. By Theorem 11 and since both AMEX and STAR extract depleting Σ-modules, given a signature Σ and TBox $\mathcal{T}$ one can extract an AMEX module from the STAR module (and vice versa) and it is guaranteed that the resulting module is still a depleting Σ-module of $\mathcal{T}$. In this way, one can repeatedly extract from the output of one extraction approach again a module using the other approach until the sequence of modules becomes stable.

Preliminary experiments shown in Figure 6 are based on 200 signatures for each concept signature size/role percentage combination and compare the average size of modules extracted using STAR-extraction, hybrid extraction, and 1-depleting module extraction.

The results are very similar to the results for $\mathsf{NCI}^\star$. Hybrid modules are on average significantly smaller than STAR modules and are often identical to the minimal 1-depleting module (and so the minimal depleting module). In fact, in this case we found no hybrid module that does not coincide with the corresponding minimal 1-depleting module.

6 Conclusion

Using a new system, AMEX, for upper approximation and the first algorithm for lower approximation of minimal depleting modules we have shown that for the NCI Thesaurus one can compute efficiently depleting modules that are consistently very close to the minimal depleting module and often coincide with the latter. The experiments also show that for TBoxes with many axioms of the form $A \equiv C$, AMEX-modules can be significantly smaller than STAR-modules and that a hybrid approach can lead to significantly smaller modules than 'pure' STAR-modules.

This paper is only the first step towards a novel systematic evaluation of the quality of upper approximations of modules using lower approximations. It would be of great interest to compute lower approximations for a more comprehensive set of cyclic ontologies and compare them with the upper approximations given by STAR-modules and by the hybrid approach. We conjecture that for many cyclic ontologies 1-depleting modules will still be a good lower approximation and can therefore provide a suitable tool to estimate the difference between STAR/hybrid-modules and minimal depleting modules. For some ontologies, however, it will be necessary to move to n-depleting modules (based on inseparability for interpretations of size at most n) with $n > 1$. These modules can still be extracted by using QBF solvers and exactly the same algorithm; the cost is much higher, though, since the length of the encoding into a QBF is exponential in n.

REFERENCES

[1] F. Baader, D. Calvanes, D. McGuiness, D. Nardi, and P. Patel-Schneider, *The Description Logic Handbook: Theory, implementation and applications*, Cambridge University Press, Cambridge, UK, 2003.

[2] M. Benedetti, 'sKizzo: a QBF decision procedure based on propositional skolemization and symbolic reasoning', Technical Report 04-11-03, ITC-irst, (2004).

[3] B. Cuenca Grau, I. Horrocks, Y. Kazakov, and U. Sattler, 'Modular reuse of ontologies: theory and practice', *Journal of Artificial Intelligence Research (JAIR)*, **31**, 273–318, (2008).

[4] B. Cuenca Grau, I. Horrocks, Y. Kazakov, and U. Sattler, 'Extracting modules from ontologies: A logic-based approach', in *Modular Ontologies*, 159–186, Springer, (2009).

[5] W. Gatens, B. Konev, and F. Wolter, 'Module extraction for acyclic ontologies', in *WoMO*, (2013).

[6] M. Horridge and S. Bechhofer, 'The OWL API: A Java API for OWL ontologies', *Semantic Web*, **2**(1), 11–21, (2011).

[7] B. Konev, R. Kontchakov, M. Ludwig, T. Schneider, F. Wolter, and M. Zakharyaschev, 'Conjunctive query inseparability of OWL 2 QL TBoxes', in *AAAI*, pp. 221–226. AAAI Press, (2011).

[8] B. Konev, C. Lutz, D. Walther, and F. Wolter, 'Formal properties of modularisation', in *Modular Ontologies*, 25–66, Springer, (2009).

[9] B. Konev, C. Lutz, D. Walther, and F. Wolter, 'Model-theoretic inseparability and modularity of description logic ontologies', *Artificial Intelligence*, **203**, 66–103, (2013).

[10] R. Kontchakov, L. Pulina, U. Sattler, T. Schneider, P. Selmer, F. Wolter, and M. Zakharyaschev, 'Minimal module extraction from DL-Lite ontologies using QBF solvers', in *IJCAI*, pp. 836–841, (2009).

[11] R. Kontchakov, F. Wolter, and M. Zakharyaschev, 'Logic-based ontology comparison and module extraction, with an application to DL-Lite', *Artificial Intelligence*, **174**(15), 1093–1141, (2010).

[12] C. Lutz and F. Wolter, 'Deciding inseparability and conservative extensions in the description logic $\mathcal{EL}$', *Journal of Symbolic Computing*, **45**(2), 194–228, (2010).

[13] R. Nortje, K. Britz, and T. Meyer, 'Reachability modules for the description logic SRIQ', in *LPAR*, pp. 636–652, (2013).

[14] U. Sattler, T. Schneider, and M. Zakharyaschev, 'Which kind of module should I extract?', in *DL*. CEUR-WS.org, (2009).

[15] J. Seidenberg, 'Web ontology segmentation: Extraction, transformation, evaluation', in *Modular Ontologies*, 211–243, (2009).

[16] *Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization*, eds., H. Stuckenschmidt, C. Parent, and S. Spaccapietra, volume 5445 of *LNCS*, Springer, 2009.

[17] B. Suntisrivaraporn, 'Module extraction and incremental classification: A pragmatic approach for ontologies', in *ESWC*, pp. 230–244, (2008).

[18] C. Vescovo, P. Klinov, B. Parsia, U. Sattler, T. Schneider, and D. Tsarkov, 'Empirical study of logic-based modules: Cheap is cheerful', in *ISWC*, pp. 84–100, (2013).

[19] P. Whetzel, N. Fridam Noy, N. Shah, P. Alexander, C. Nyulas, T. Tudorache, and M. Musen, 'BioPortal', *Nucleic Acids Research*, (Web-Server-Issue), 541–545, (2011).

A Proofs

In this section we provide proofs for results which are not a direct consequence of previously published results in [9].

Lemma 3 *$\mathcal{T} \equiv^1_\Sigma \emptyset$ iff the QBF $\varphi_\mathcal{T} := \forall\vec{p}\exists\vec{q}\mathcal{T}^\dagger$ is valid.*

Proof Recall the definition of the translation $\cdot^\dagger$:

$$\begin{aligned} A^\dagger &= p_A, && \text{for all } A \in \Sigma \\ A^\dagger &= q_A, && \text{for all } A \in \mathsf{sig}(\mathcal{T}) \setminus \Sigma \\ (D_1 \sqcap D_2)^\dagger &= D_1^\dagger \wedge D_2^\dagger \\ (\neg D)^\dagger &= \neg D^\dagger \end{aligned}$$

$$(\geq n\, r.D)^\dagger = (\geq n\, r^-.D)^\dagger = \begin{cases} p_r \wedge D^\dagger, \text{if } n = 1 \text{ and } r \in \Sigma \\ q_r \wedge D^\dagger, \text{if } n = 1 \text{ and } r \in \mathsf{sig}(\mathcal{T}) \setminus \Sigma \\ \bot, \text{ else} \end{cases}$$

Suppose that $\mathcal{T} \equiv^1_\Sigma \emptyset$. We show that $\phi_\mathcal{T}$ is valid. Consider an arbitrary assignment I of truth values to propositions in $\vec{p}$. We have to show that there exists an assignment J of truth values to the propositions in $\vec{q}$ such that the propositional formula $\mathcal{T}^\dagger$ is true under the assignment $I \cup J$. Define a singleton interpretation $\mathcal{I}$ as follows:

- $\Delta^\mathcal{I} = \{d\}$,
- $A^\mathcal{I} = \{d\}$ if I assigns *true* to p_A and $A^\mathcal{I} = \emptyset$ if I assigns *false* to p_A, for $A \in \Sigma$.
- $r^\mathcal{I} = \{(d,d)\}$ if I assigns *true* to p_r and $r^\mathcal{I} = \emptyset$ if I assigns *false* to p_r, for $r \in \Sigma$.

As $\mathcal{T} \equiv^1_\Sigma \emptyset$ there exists an interpretation $\mathcal{J}$ such that $\mathcal{J}|_\Sigma = \mathcal{I}|_\Sigma$ and $\mathcal{J} \models \mathcal{T}$. We define J by setting

- J assigns *true* to q_A if $d \in A^\mathcal{J}$ and J assigns *false* to q_A if $d \notin A^\mathcal{J}$, for $A \in \mathsf{sig}(\mathcal{T}) \setminus \Sigma$.
- J assigns *true* to q_r if $(d,d) \in r^\mathcal{J}$ and J assigns *false* to q_r if $(d,d) \notin r^\mathcal{J}$, for $r \in \mathsf{sig}(\mathcal{T}) \setminus \Sigma$.

It follows from $\mathcal{J} \models \mathcal{T}$ that $\mathcal{T}^\dagger$ is true under the assignment $I \cup J$.

Conversely, assume that $\phi_\mathcal{T}$ is valid. We show that $\mathcal{T} \equiv^1_\Sigma \emptyset$. Consider an arbitrary interpretation $\mathcal{I}$ with domain $\Delta^\mathcal{I} = \{d\}$. We have to show that there exists a model $\mathcal{J}$ of $\mathcal{T}$ such that $\mathcal{J}|_\Sigma = \mathcal{I}|_\Sigma$. Define a truth assignment I for the propositions in $\vec{p}$ as follows:

- for $A \in \Sigma$, I assigns *true* to p_A if $d \in A^\mathcal{I}$ and I assigns *false* to p_A if $d \notin A^\mathcal{J}$.
- for $r \in \Sigma$, I assigns *true* to p_r if $(d,d) \in r^\mathcal{I}$ and I assigns *false* to p_r if $(d,d) \notin r^\mathcal{I}$.

Since $\phi_\mathcal{T}$ is valid, there exists a truth assignment J for the propositions in $\vec{q}$ such that $\mathcal{T}^\dagger$ is true under $I \cup J$. Define an extension $\mathcal{J}$ of $\mathcal{I}$ as follows:

- for $A \in \mathsf{sig}(\mathcal{T}) \setminus \Sigma$, $A^\mathcal{J} = \{d\}$ if J assigns *true* to q_A and $A^\mathcal{J} = \emptyset$ if J assigns *false* to q_A.
- for $A \in \mathsf{sig}(\mathcal{T}) \setminus \Sigma$, $r^\mathcal{J} = \{(d,d)\}$ if J assigns *true* to q_r and $r^\mathcal{J} = \emptyset$ if J assigns *false* to q_r.

It can be shown that $\mathcal{J}$ is a model of $\mathcal{T}$ such that $\mathcal{J}|_\Sigma = \mathcal{I}|_\Sigma$. ∎

We now analyse the algorithm given in Figure 3. The algorithm in Figure 1 can be analysed similarly.

Theorem 8 *Given an acyclic $\mathcal{ALCQI}$ TBox $\mathcal{T}$ and signature Σ the algorithm in Figure 3 computes the unique minimal depleting Σ-module s.t. $\mathcal{T} \setminus \mathcal{M}$ contains no direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$-dependencies. The algorithm is in polynomial time with each call to the QBF solver being treated as a constant time oracle call.*

Input: TBox $\mathcal{T}$ and Σ.
Output: Minimal Module $\mathcal{M}$ s.t $\mathcal{T} \setminus \mathcal{M} \equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$ and $\mathcal{T} \setminus \mathcal{M}$ has no direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$ dependencies.

$\mathcal{M} := \emptyset$;
$\mathcal{W} := \emptyset$;
1. **while** $(\mathcal{T} \setminus \mathcal{M}) \neq \mathcal{W}$ **do**
2. *choose* $\alpha \in (\mathcal{T} \setminus \mathcal{M}) \setminus \mathcal{W}$
3. $\mathcal{W} := \mathcal{W} \cup \{\alpha\}$
4. **if** $\mathcal{W}$ contains a direct $(\Sigma \cup \mathsf{sig}(\mathcal{M}))$-dependency or
5. $\mathcal{W} \not\equiv_{(\Sigma \cup \mathsf{sig}(\mathcal{M}))} \emptyset$
6. **then**
7. $\mathcal{M} := \mathcal{M} \cup \{\alpha\}$
8. $\mathcal{W} := \emptyset$
9. **endif**
10. **end while**
11. **return** $\mathcal{M}$

Figure 7. Depleting $\mathcal{ALC}(\mathcal{Q})\mathcal{I}$ module extraction algorithm from [9].

Proof It has been proved in [9] that given an acyclic $\mathcal{ALCI}$ TBox $\mathcal{T}$ and signature Σ the algorithm given in Figure 7 computes the unique minimal depleting Σ-module of $\mathcal{T}$ such that $\mathcal{T} \setminus \mathcal{M}$ contains no direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$-dependencies. The proof carries over to the case of $\mathcal{ALCQI}$ in a straightforward way, that is, given an acyclic $\mathcal{ALCQI}$ TBox $\mathcal{T}$ and signature Σ the algorithm given in Figure 7 computes the unique minimal depleting Σ-module of $\mathcal{T}$ such that $\mathcal{T} \setminus \mathcal{M}$ contains no direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$-dependencies.

Since the output of the algorithm is uniquely determined, the output does not depend on the selection strategy in line 2 of the algorithm in Figure 7. One can use this fact to prove that the algorithm in Figure 7 and Figure 3 return the same value.

As for the running time bound, we assume w.l.o.g. that $|\Sigma| \leq |\mathcal{T}|$. Then the total number of rule applications does not exceed $|\mathcal{T}|$. Checking if **syn** is applicable can be performed by simple reachability analysis in $O(|\mathcal{T}|)$ time and checking if the **1-insep** rule is applicable can also be done in $O(|\mathcal{T}|)$ as follows: starting with $\mathcal{W} = \emptyset$ add axioms from $\mathcal{T} \setminus \mathcal{M}$ to $\mathcal{W}$ one by one while checking that $\mathcal{W} \equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$. The axiom for which the condition is violated is separability causing. Thus, the algorithm runs in $O(|\mathcal{T}|^2)$ in the worst case (assuming that each call to the QBF solver is treated as a constant time oracle call).

The search for separability causing axioms can be optimised to run in $O(\log_2 |\mathcal{T}|)$ time as shown in Figure 8. We first consider $\mathcal{T} \setminus \mathcal{M}$ itself as $\mathcal{W}$. If $\mathcal{T} \setminus \mathcal{M} \equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$ then $\mathcal{T} \setminus \mathcal{M}$ contains no separability causing axioms. Otherwise, we consider $\mathcal{W}$ to be equal to the top half of $\mathcal{T} \setminus \mathcal{M}$ (we treat $\mathcal{T} \setminus \mathcal{M}$ as an ordered set). We then check if $\mathcal{W} \equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$ and, if this is the case, we grow $\mathcal{W}$ from the bottom and if not, we half it again, etc. ∎

Theorem 10 *Let $\mathcal{T}$ be an acyclic TBox with repeated CIs and Σ a signature. Then $\mathcal{M}$ computed by the algorithm in Figure 4 is a depleting Σ-module of $\mathcal{T}$.*

Proof Let $\mathcal{T}'$ be the set of non-repeated CIs in $\mathcal{T}$. For a repeated A let $A \sqsubseteq C_1, \ldots, A \sqsubseteq C_n$, for $n > 1$, be all CIs in $\mathcal{T}$ with A on the left hand side. Then

$$\mathcal{T}^A = \{A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n\} \cup \bigcup_{i=1}^{n} \{A'_i \sqsubseteq C_i\},$$

where $A'_1, \ldots, A'_n$ are fresh concept names (not occurring in $\mathcal{T}$). For

$$\mathcal{T}^* = \mathcal{T}' \cup \bigcup_{A \text{ is repeated}} \mathcal{T}^A$$

let $\mathcal{M}'$ be the minimal depleting Σ-module of $\mathcal{T}^*$ such that $\mathcal{T}^* \setminus \mathcal{M}'$ has no direct $\Sigma \cup \mathsf{sig}(\mathcal{M}')$ dependencies. We require the following observation.

***Claim*:** If $A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n \notin \mathcal{M}'$, then for every $A'_i \sqsubseteq C_i \in \mathcal{T}^A$ we have $A'_i \sqsubseteq C_i \notin \mathcal{M}'$.

For a proof by contradiction, assume that the claim does not hold. Then there is some $i \geq 1$ such that $A'_i \sqsubseteq C_i \in \mathcal{M}'$. Consider $\mathcal{M}'' = \mathcal{M}' \setminus \{A'_i \sqsubseteq C_i\}$. Then neither $A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n$ nor $A'_i \sqsubseteq C_i$ belong to $\mathcal{M}''$, so $A'_i \notin \Sigma \cup \mathsf{sig}(\mathcal{M}'')$.

$\mathcal{T}^* \setminus \mathcal{M}''$ either has a direct $\Sigma \cup \mathsf{sig}(\mathcal{M}'')$ dependency or $\mathcal{M}''$ is not a depleting Σ-module of $\mathcal{T}^*$. Assume first that $\mathcal{T}^* \setminus \mathcal{M}''$ has a direct $\Sigma \cup \mathsf{sig}(\mathcal{M}'')$ dependency. Then for some $X \in \Sigma \cup \mathsf{sig}(\mathcal{M}'')$ either $X \prec^+_{\mathcal{T}^* \setminus \mathcal{M}'} A'_i$ or $A'_i \prec^+_{\mathcal{T}^* \setminus \mathcal{M}'} X$. In either case $\mathcal{T}^* \setminus \mathcal{M}'$ has a direct $\Sigma \cup \mathsf{sig}(\mathcal{M}')$-dependency and we have derived a contradiction.

Now assume that $\mathcal{M}''$ is not a depleting Σ-module of $\mathcal{T}^*$. Let $\mathcal{I}$ be an interpretation such that there does not exist an model $\mathcal{J}$ of $\mathcal{T}^* \setminus \mathcal{M}''$ with $\mathcal{J}|_{\Sigma \cup \mathsf{sig}(\mathcal{M}'')} = \mathcal{I}|_{\Sigma \cup \mathsf{sig}(\mathcal{M}'')}$. Consider the interpretation $\mathcal{I}'$ defined as follows:

- $\Delta^{\mathcal{I}'} = \Delta^{\mathcal{I}}$;
- $X^{\mathcal{I}'} = X^{\mathcal{I}}$, for every symbol $X \neq A'_i$;
- ${A'_i}^{\mathcal{I}'} = \emptyset$.

Since $\mathcal{M}'$ is a depleting Σ-module of $\mathcal{T}^*$ there exists a model $\mathcal{J}$ of $\mathcal{T}^* \setminus \mathcal{M}'$ such that $\mathcal{I}'|_{\Sigma \cup \mathsf{sig}(\mathcal{M}')} = \mathcal{J}|_{\Sigma \cup \mathsf{sig}(\mathcal{M}')}$. Since $A'_i \in \mathsf{sig}(\mathcal{M}')$ we have ${A'_i}^{\mathcal{J}} = \emptyset$. But then $\mathcal{J}$ is a model of $\mathcal{T}^* \setminus \mathcal{M}''$ and since $A'_i \notin \Sigma \cup \mathsf{sig}(\mathcal{M}'')$, we have $\mathcal{J}|_{\Sigma \cup \mathsf{sig}(\mathcal{M}'')} = \mathcal{I}|_{\Sigma \cup \mathsf{sig}(\mathcal{M}'')}$, and we have derived a contradiction. This finishes the proof of the claim.

Let $\mathcal{M}$ be obtained from $\mathcal{M}'$ by dropping the added axioms of the form $A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n$ and replacing every occurrence of the introduced symbols $A'_1, \ldots, A'_n$ with A. To show that $\mathcal{T} \setminus \mathcal{M} \equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$, let $\mathcal{I}$ be an arbitrary interpretation. We need to demonstrate that there exists a model $\mathcal{J}$ of $\mathcal{T} \setminus \mathcal{M}$ such that $\mathcal{I}|_{\Sigma \cup \mathsf{sig}(\mathcal{M})} = \mathcal{J}|_{\Sigma \cup \mathsf{sig}(\mathcal{M})}$.

Define a new interpretation $\mathcal{I}'$ by setting

- $\Delta^{\mathcal{I}'} = \Delta^{\mathcal{I}}$;
- $X^{\mathcal{I}'} = X^{\mathcal{I}}$, for every symbol $X \in \Sigma \cup \mathsf{sig}(\mathcal{T})$;
- ${A'_i}^{\mathcal{I}'} = A^{\mathcal{I}}$ for every A'_i in an added axiom of the form $A \sqsubseteq A'_1 \sqcap \ldots A'_n$.

As $\mathcal{M}'$ is a depleting Σ-module of $\mathcal{T}^*$, there exists an interpretation $\mathcal{J}$ such that $\mathcal{J}|_{\Sigma \cup \mathsf{sig}(\mathcal{M}')} = \mathcal{I}'|_{\Sigma \cup \mathsf{sig}(\mathcal{M}')}$ and $\mathcal{J} \models \mathcal{T}^* \setminus \mathcal{M}'$. We show that $\mathcal{J} \models \mathcal{T} \setminus \mathcal{M}$ also holds.

As $\mathcal{T}'$ is the set of non-repeated CIs, we have $\mathcal{T}' \setminus \mathcal{M} = \mathcal{T}' \setminus \mathcal{M}'$ and so $\mathcal{J} \models \mathcal{T}' \setminus \mathcal{M}$. Suppose for some repeated $A \sqsubseteq C_i \in \mathcal{T} \setminus \mathcal{M}$ we have $\mathcal{J} \not\models A \sqsubseteq C_i$. By definition of $\mathcal{M}$ we have $A'_i \sqsubseteq C_i \in \mathcal{T}^* \setminus \mathcal{M}'$.

Input: TBox $\mathcal{T}$, subset $\mathcal{M} \in \mathcal{T}$ and signature Σ such that $\mathcal{T} \setminus \mathcal{M}$ contains no direct $\Sigma \cup \mathsf{sig}(\mathcal{M})$-dependencies and $\mathcal{T} \setminus \mathcal{M} \not\equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$
Output: Separability causing axiom α

$\mathcal{W} = \mathsf{lastAdded} = \mathit{topHalf}(\mathsf{Lhs}_{\Sigma \cup \mathsf{sig}(\mathcal{M})}(\mathcal{T} \setminus \mathcal{M}))$
$\mathsf{lastRemoved} = \mathit{bottomHalf}(\mathsf{Lhs}_{\Sigma \cup \mathsf{sig}(\mathcal{M})}(\mathcal{T} \setminus \mathcal{M}))$
while $\mathsf{lastAdded} \neq \emptyset$ **do**
  if $\mathcal{W} \equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$ **then**
    $\mathsf{lastAdded} = \mathit{topHalf}(\mathsf{lastRemoved})$
    $\mathcal{W} = \mathcal{W} \cup \mathsf{lastAdded}$
    $\mathsf{lastRemoved} = \mathsf{lastRemoved} \setminus \mathsf{lastAdded}$
  else
    $\mathsf{lastRemoved} = \mathit{bottomHalf}(\mathsf{lastAdded})$
    $\mathcal{W} = \mathcal{W} \setminus \mathsf{lastRemoved}$
    $\mathsf{lastAdded} = \mathsf{lastAdded} \setminus \mathsf{lastRemoved}$
  end if
end while
return the last axiom of $\mathcal{W}$

Figure 8. Finding 1-separability causing axiom

We make a case distinction on whether or not $A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n \in \mathcal{M}'$. Suppose that $A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n \in \mathcal{M}'$. Then $\{A, A'_1, \ldots, A'_n\} \subseteq \mathsf{sig}(\mathcal{M}')$ so ${A'_i}^{\mathcal{J}} = {A'_i}^{\mathcal{I}'} = A^{\mathcal{I}'} = A^{\mathcal{J}}$. Thus $\mathcal{J} \not\models A'_i \sqsubseteq C_i$ contradicting the assumption that $\mathcal{J}$ is a model of $\mathcal{T}^* \setminus \mathcal{M}'$.

Suppose that $A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n \notin \mathcal{M}'$. Then both $A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n \in \mathcal{T}^* \setminus \mathcal{M}'$ and $A'_i \sqsubseteq C_i \in \mathcal{T}^* \setminus \mathcal{M}'$ but $\mathcal{J} \not\models A \sqsubseteq A'_1 \sqcap \ldots \sqcap A'_n$ or $\mathcal{J} \not\models A'_i \sqsubseteq C_i$ since $\mathcal{J} \not\models A \sqsubseteq C_i$. We have derived a contradiction to the assumption that $\mathcal{J}$ is a model of $\mathcal{T}^* \setminus \mathcal{M}'$.

We have shown that $\mathcal{J} \models \mathcal{T} \setminus \mathcal{M}$. By the claim above we have $\mathsf{sig}(\mathcal{M}) \subseteq \mathsf{sig}(\mathcal{M}')$ and so $\mathcal{J}|_{\Sigma \cup \mathsf{sig}(\mathcal{M})} = \mathcal{I}|_{\Sigma \cup \mathsf{sig}(\mathcal{M})}$, as required. ∎

We illustrate Theorem 10 with an example.

Example 12 Consider an acyclic $\mathcal{EL}$ TBox with repeated CIs

$$\mathcal{T} = \{A \sqsubseteq B, A \sqsubseteq C, B \sqsubseteq C\}$$

and signature $\Sigma = \{A, B\}$. First, notice that if we (logically equivalently) rewrite $\mathcal{T}$ as an acyclic $\mathcal{EL}$ TBox

$$\mathcal{T}_{acyc} = \{A \sqsubseteq B \sqcap C, B \sqsubseteq C\}$$

then the minimal depleting Σ-module of $\mathcal{T}_{acyc}$ is $\mathcal{T}_{acyc}$ itself

On the other hand, consider $\mathcal{T}' = \{B \sqsubseteq C\}$ and

$$\mathcal{T}^A = \{A \sqsubseteq A'_1 \sqcap A'_2\} \cup \{A'_1 \sqsubseteq B\} \cup \{A'_2 \sqsubseteq C\}.$$

Then the minimal depleting Σ-module of $\mathcal{T}' \cup \mathcal{T}^A$ is

$$\mathcal{M}' = \{A \sqsubseteq A'_1 \sqcap A'_2, A'_1 \sqsubseteq B\}$$

and, by Theorem 10, $\mathcal{M} = \{A \sqsubseteq B\}$ is a depleting Σ-module of $\mathcal{T}$. In fact, it is straightforward to check that $\{A \sqsubseteq C, B \sqsubseteq C\} \equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset$ and that $\mathcal{M}$ is the minimal depleting Σ-module of $\mathcal{T}$.

Theorem 11 *Let $\mathcal{M} \subseteq \mathcal{M}' \subseteq \mathcal{T}$ be TBoxes and Σ a signature such that $\mathcal{M}'$ is a depleting Σ-module of $\mathcal{T}$ and $\mathcal{M}$ is a depleting Σ-module of $\mathcal{M}'$. Then $\mathcal{M}$ is a depleting Σ-module of $\mathcal{T}$.*

Proof Assume that $\mathcal{M} \subseteq \mathcal{M}' \subseteq \mathcal{T}$ and Σ is a signature such that $\mathcal{M}'$ is a depleting Σ-module of $\mathcal{T}$ and $\mathcal{M}$ is a depleting Σ-module of $\mathcal{M}'$ To prove that $\mathcal{M}$ is a depleting Σ-module of $\mathcal{T}$, consider an interpretation $\mathcal{I}$. We have to show that there exists a model $\mathcal{J}$ of $\mathcal{T} \setminus \mathcal{M}$ such that $\mathcal{J}|_{\Sigma\cup\mathsf{sig}(\mathcal{M})} = \mathcal{I}|_{\Sigma\cup\mathsf{sig}(\mathcal{M})}$. As $\mathcal{M}$ is a depleting Σ-module of $\mathcal{M}'$, there exists an interpretation $\mathcal{J}'$ such that $\mathcal{J}'|_{\Sigma\cup\mathsf{sig}(\mathcal{M})} = \mathcal{I}|_{\Sigma\cup\mathsf{sig}(\mathcal{M})}$ and $\mathcal{J}' \models (\mathcal{M}' \setminus \mathcal{M})$. Similarly, as $\mathcal{M}'$ is a depleting Σ-module of $\mathcal{T}$, there exists an interpretation $\mathcal{J}$ such that $\mathcal{J}|_{\Sigma\cup\mathsf{sig}(\mathcal{M}')} = \mathcal{J}'|_{\Sigma\cup\mathsf{sig}(\mathcal{M}')}$ and $\mathcal{J} \models (\mathcal{T} \setminus \mathcal{M}')$. As $\mathsf{sig}(\mathcal{M}) \subseteq \mathsf{sig}(\mathcal{M}') \subseteq \mathsf{sig}(\mathcal{T})$ we have $\mathcal{J}|_{\Sigma\cup\mathsf{sig}(\mathcal{M})} = \mathcal{I}|_{\Sigma\cup\mathsf{sig}(\mathcal{M})}$ and $\mathcal{J} \models (\mathcal{M}' \setminus \mathcal{M})$. But then $\mathcal{J} \models ((\mathcal{T} \setminus \mathcal{M})$ and so $\mathcal{J}$ is as required.

∎