Prioritized base Debugging in Description Logics

Julien Corman, Nathalie Aussenac-Gilles, Laure Vieu

Prioritized base debugging in Description Logics

Julien Corman
IRIT, University of Toulouse

Nathalie Aussenac-Gilles
CNRS
IRIT, University of Toulouse

Laure Vieu
CNRS
IRIT, University of Toulouse

Abstract

The problem investigated is the identification within an input knowledge base of axioms which should be preferably discarded (or amended) in order to restore consistency, coherence, or get rid of undesired consequences. Most existing strategies for this task in Description Logics rely on conflicts, either computing all minimal conflicts beforehand, or generating conflicts on demand, using diagnosis. The article studies how prioritized base revision can be effectively applied in the former case. The first main contribution is the observation that for each axiom appearing in a minimal conflict, two bases can be obtained for a negligible cost, representing what part of the input knowledge must be preserved if this axiom is discarded or retained respectively, and which may serve as a basis to obtain a semantically motivated preference relation over these axioms. The second main contributions is an algorithm which, assuming this preference relation is known, selects some of the maximal consistent/coherent subsets of the input knowledge base accordingly, without the need to compute all of of them.

1 Introduction

The focus of this work is the automated or semi-automated debugging of some knowledge base (KB) expressed in the Description Logic (DL) $\mathcal{SROIQ}$ (Horrocks *et al.* 2006), which underlies the OWL 2 W3C recommendation. More exactly, the task considered here consists in suggesting axioms of the input knowledge base which could be discarded in order to restore consistency, coherence, or get rid of some undesired consequence. This is of particular interest for OWL KBs within the Semantic Web framework, where aggregated knowledge issued from different sources with overlapping signatures may have unexpected consequences.

The problem has been studied in the field of knowledge base engineering, as well as belief base revision/contraction[1], from a different perspective though. In order to account for both, it will be assumed that the user wants to preserve a (consistent) part Θ of the input knowledge, but that Θ may be empty.

[1] to be distinguished from *belief set* revision/contraction, which primarily considers (deductively closed) theories

As an example, consider the following inconsistent set of statements $K \cup \Theta$,[2] and assume that Θ should be preserved :

Ex 1. $K = \{$
(1) owningCompany(*Smithsonian Networks*, *Smithsonian Institution*),
(2) publisher(*Birds of South Asia*, *Smithsonian Institution*),
(3) award(*James Dewar*, *Smithsonian Institution*),
(4) doctoralAdvisor(*Thaddeus S.C. Lowe*, *Smithsonian Institution*),
(5) $\top \sqsubseteq \forall$award.Award,
(6) $\top \sqsubseteq \forall$doctoralAdvisor.Person,
(7) $\top \sqsubseteq \forall$owningCompany.Company,
(8) Company $\sqsubseteq$ Organization $\}$
$\Theta = \{$Person $\sqsubseteq \neg$Organization,
Award $\sqsubseteq \neg$Organization $\}$

An important issue for base debugging is the number of candidate sets of axioms for removal. Even if the set of remaining axioms is required to be maximal (wrt set inclusion), which is intuitive, there are still 7 candidates here, i.e. one could alternatively discard {1}, {7}, {8}, {3,4}, {3,6}, {4,5} or {5,6} to restore the consistency of $K \cup \Theta$, while preserving Θ.

Let $\mathcal{R}_\subseteq$ designate the 7 corresponding complements in K, i.e. the 7 possible sets of remaining statements. If one refuses to choose arbitrarily one element of $\mathcal{R}_\subseteq$, a common solution consists in keeping by default the intersection $\bigcap \mathcal{R}_\subseteq$ of all of them, extended with Θ (or possibly their disjunction $\bigvee \mathcal{R}_\subseteq$ extended with Θ, depending on the application, as explained in section 3).

But the output in this case may be quite weak, i.e. result in an important information loss. In this example for instance, $\bigcap \mathcal{R}_\subseteq = \{2\}$, which is clearly not satisfying. Therefore the need for a principled way of selecting a subset $\mathcal{R}'$ of $\mathcal{R}_\subseteq$ such that $\bigcap \mathcal{R}'$ (resp. $\bigvee \mathcal{R}'$) is stronger that $\bigcap \mathcal{R}_\subseteq$ (resp. $\bigvee \mathcal{R}_\subseteq$).

The view adopted here, known as prioritized base revision/debugging, is that some preference relation $\preceq_a$ over the axioms of K should guide the process, trying to discard least preferred axioms first, but still not unnecessarily. Section 4 shows that no solution proposed for DLs to our knowledge satisfies both of these requirements, and section 5 provides an algorithm to do it, assuming the set $\mathcal{M}$ of all minimal

[2] K is a set of actual DBPedia statements (Mendes *et al.* 2012).

conflicts within $K \cup \Theta$ is known, as well as $\preceq_a$. Section 6 then proposes a generic approach to obtain a semantically motivated preference relation $\preceq_a$ over K, and section 7 discusses possible applications of this strategy. Before that, section 2 introduces some notation, whereas section 3 characterizes the problem, borrowing notions to the fields of KB debugging and belief base revision. Proofs are omitted due to the lack of place, but are all available online at the following url : `http://www.irit.fr/~Julien.Corman/index_en.php`.

2 Notation and conventions

2.1 Ordering

If $\preceq$ is a total preorder (transitive, reflexive, but not necessarily antisymmetric) over some finite set Δ, then :

- $\sim$ is the equivalence relation over Δ induced by $\preceq$.
- $\Delta/\sim$ denotes the quotient set of Δ by $\sim$.
- $\Delta^{\preceq} = (\Delta_1^{\preceq}, .., \Delta_n^{\preceq})$ is the list of ordered elements of $\Delta/\sim$, such that $1 \leq i < j \leq n$ iff $\forall \phi \in \Delta_i^{\preceq}, \forall \phi' \in \Delta_j^{\preceq} : \phi \not\preceq \phi'$.

If $\preceq$ is a (total or partial) preorder over Δ, then :

- $\min_{\preceq} \Delta$ (resp. $\max_{\preceq} \Delta$) is the set of minimal (resp. maximal) elements of Δ wrt $\preceq$.
- $\phi_1 \prec \phi_2$ stands for $\phi_1 \preceq \phi_2$ and $\phi_2 \not\preceq \phi_1$.

2.2 Decription logics

The reader is assumed familiar with the syntax and standard model-theoretic semantics of Description Logics (Baader *et al.* 2003). $\mathcal{L}$ will designate the DL at hand, and a DL KB is simply a finite set of DL formulas, called *axioms*.

Logical inconsistency is understood in the usual way, i.e. a set Γ of formulas is inconsistent (noted $\Gamma \vdash \bot$) iff it has no model. A slightly different notion used in the DL community is logical *incoherence*. A set Γ of formulas is said to be incoherent iff there is an atomic concept (like "Person" or "Organization") in its signature such that the interpretation of this concept is empty in every model of Γ.

3 Knowledge base debugging/belief base revision

The problem of discarding axioms of K in order to restore the consistency/coherence of $K \cup \Theta$ (or get rid of undesired consequences of it) has been studied both in the fields of belief base revision/contraction and KB debugging (usually with $\Theta = \emptyset$ in the latter case). In order to avoid possible confusions, belief base revision/contraction should be distinguished from belief set revision/contraction, whose most influential framework (the AGM framework) focuses on (deductively closed) theories, abstracting from the syntax. On the contrary, belief base revision/contraction requires the output to be a syntactic subset of $K \cup \Theta$, adapting several notions from the belief set revision/contraction literature, but with sometimes very different implications.

This syntactic requirement may be viewed as too constraining, because it results in weaker theories than what could be obtained by considering belief sets. But it is nonetheless required in many practical debugging applications where traceability matters, i.e where identifying faulty axioms is relevant from an engineering point of view, for instance in an OWL KB when axioms have been obtained from different sources. This article exclusively focuses on base revision/debugging, therefore most works in the field of belief set revision for DLs, like (Flouris 2006; Qi and Du 2009; Wang *et al.* 2010; Ribeiro 2013) fall out of its scope, and will not be reviewed.

A comprehensive series of base revision and contraction operators applicable to Description Logics have been defined in (Ribeiro and Wassermann 2009), and part of the terminology used in this article is borrowed from that work. Precisely, the operations of interest here are named partial meet and kernel base revision without negation (with weak or full success) in (Ribeiro and Wassermann 2009). Due to the lack of place though, a simplified notation will be used, and some limit and/or trivial cases will not be explicitly addressed, in order to focus on the practical problem at hand. In particular, the set of statements Θ to be preserved is supposed to be consistent.

For the sake of readability still, an important reduction of scope is also made. The need for weakening K may appear when $K \cup \Theta$ is inconsistent, incoherent, or has undesired consequences. Some authors (Qi *et al.* 2006; Qi *et al.* 2008) actually distinguish revision wrt inconsistency from revision wrt to incoherence, proposing specific algorithms for each task. These distinctions will be ignored here, and the focus put by default on inconsistency, unless explicitly mentioned. But all propositions made in this article can be adapted to the cases of incoherence and undesired consequences, provided minor modifications. It is also assumed that $K \cup \Theta$ is actually inconsistent, otherwise the output of the process is trivially $K \cup \Theta$.

3.1 Remainders and selection function

An important notion for base revision/debugging is that of a (base) *remainder*, which is intuitively an admissible subset of $K \cup \Theta$ maximal wrt set inclusion. Because the focus is on inconsistency here, let $\mathcal{R}$ designate all subsets of K consistent with Θ, i.e. :

Definition 3.1. $\mathcal{R} = \{R \subseteq K \mid R \cup \Theta \nvdash \bot\}$

Then the *remainder set* $\mathcal{R}_{\subseteq}$ is defined by :[3]

Definition 3.2. $\mathcal{R}_{\subseteq} = \{R \in \mathcal{R} \mid \forall \phi \in (K \setminus R) : R \cup \{\phi\} \notin \mathcal{R}\}$

A remainder is any $R \in \mathcal{R}_{\subseteq}$. A strong intuition here is that if $R \in \mathcal{R}_{\subseteq}$, then $R \cup \Theta$ is a candidate output, and discarding any additional axiom can be viewed as an unnecessary information loss, unless there are several elements in $\mathcal{R}_{\subseteq}$, and no indication as to which one should be preferred. In this last case, according to the *fairness* principle, the output of the process may be either the in-

[3] $\mathcal{R}_{\subseteq}$ is actually defined for K, Θ and $\bot$. But given the lack of ambiguity here, it will only be noted $\mathcal{R}_{\subseteq}$, without explicit references to K and Θ.

tersection $\bigcap(\mathcal{R}_\subseteq) \cup \Theta$,[4] or, depending on the engineering constraints, the disjunction $\bigvee_{R\in\mathcal{R}_\subseteq}(R \cup \Theta)$ (Meyer *et al.* 2005). This last construction is not natively representable in DLs, but can be simulated as a multibase, by requiring that $\mathrm{Cn}(\bigvee_{R\in\mathcal{R}_\subseteq}(R\cup\Theta)) = \bigcap_{R\in\mathcal{R}_\subseteq}\mathrm{Cn}(R\cup\Theta)$.

Additionally, in order to give up less information, it may be desirable to select some remainders only (or equivalently some elements of $\mathcal{R}_\subseteq$), and yield as an output the intersection (or disjunction) of these selected remainders, or even submit them to the user if their number is small enough.

The notion of a *selection function*, once again adapted from the belief set revision literature, formalizes this idea. In the specific case of an inconsistent $K \cup \Theta$ considered here, a selection function must select a nonempty subset of the remainder set.[5] Aside from its intuitive appeal, a good argument for the notion of remainder in the context of base revision is the representation theorem given by (Ribeiro and Wassermann 2009). It states that for any selection of a nonempty subset of the remainder set, the operation that takes K and Θ as an input and yields the intersection of these selected remainders extended with Θ satisfies a set of very intuitive rationality postulates for base revision (namely inclusion, weak consistency, strong success, pre-expansion and relevance), and conversely.

3.2 Prioritized base revision/debugging

Prioritized base revision formalizes the simple idea that all axioms within K are not equal, or in other words, that some preference relation $\preceq_a$ (total preorder, i.e. intuitively a ranking) over the axioms of K is available, such that all other things being equal, if $\phi_1 \prec \phi_2$, then ϕ_1 should be preferably discarded when trying to restore the consistency of $K \cup \Theta$. The way $\preceq_a$ may be obtained is discussed in sections 6 and 7.

Assuming $\preceq_a$ is known, prioritized base revision has been characterized by (Nebel 1992) with a total order $\preceq_r$ over $\mathcal{R}$, defined by $R \prec_r R'$ iff there is $1 \leq i \leq n$ such that :

$R_i^{\preceq_a} \subset R_i'^{\preceq_a}$ and $\forall j\ _{1\leq j<i} : R_j^{\preceq_a} = R_j'^{\preceq_a}$.

The following observation (proven in section **??**) states that any subbase obtained by performing prioritized revision is actually a remainder, i.e.:

Proposition 3.1. $\mathcal{R}_{\preceq_r} \subseteq \mathcal{R}_\subseteq$

In other words, each $R \in \mathcal{R}_{\preceq_r}$ is a base remainder for An immediate consequence of proposition 3.1 is that no axiom of $K \setminus \bigcup \mathcal{V}$ needs to be discarded, and therefore it is sufficient for the preference relation $\preceq_a$ introduced in section 3.2 to be defined over $\bigcup \mathcal{J}$ (or equivalently, for some authors, over $K \cup \Theta$, but with $\max_{\preceq_a}(K \cup \Theta) = (K \cup \Theta) \setminus \bigcup \mathcal{J}$).

[4] This operation should be distinguished from the *full meet contraction* operation used in the AGM framework, which is the intersection of all maximal subsets of $\mathrm{Cn}(\mathcal{R}_\subseteq)$ consitent with Θ. In particular, it is not the case in general that $\mathrm{Cn}(\bigcap(\mathcal{R}_\subseteq)\cup\Theta) = \mathrm{Cn}(\Theta)$.

[5] Two additional requirements are made when considering (deductively closed) belief sets, namely that γ selects strictly more than one and less than all remainders, but this is generally not relevant for bases.

3.3 Conflicts

To our knowledge, most practical attempts in DLs to compute the remainder set, or simply part of it, are based on so-called *justifications*, also known as MIPS or minimal conflicts, depending on the authors, and on whether $K \cup \Theta$ is inconsistent, incoherent, or has undesired consequences.[6] A justification will be defined here as an inconsistent subset of $K \cup \Theta$ which is minimal wrt set inclusion. The set $\mathcal{M}$ of all such justifications is sometimes called the *kernel* for $K \cup \Theta$ and $\bot$ (Ribeiro and Wassermann 2009), defined by :

Definition 3.3. $\mathcal{M} = \{M \subseteq (K \cup \Theta) \mid M \vdash \bot$ and $\forall M' \subset M : M' \nvdash \bot\}$

Then the family $\mathcal{J}$ will designate these justifications reduced to their respective intersections with K, and which are minimal wrt set inclusion, i.e. :

Definition 3.4. $\mathcal{J} = \min_\subseteq\{M \cap K \mid M \in \mathcal{M}\}$

As an illustration, in example 1, $\mathcal{J} = \{\{1, 7, 8, 3, 5\}, \{1, 7, 8, 4, 6\}\}$.

Computing $\mathcal{J}$ and $\mathcal{R}_\subseteq$ can be viewed as two sides of the same problem. Intuitively, discarding one axiom from each $J \in \mathcal{J}$ is necessary and sufficient to obtain a subset of K consistent with Θ. Additionally, one would like the set of discarded axioms to be minimal wrt set inclusion among those satisfying this property. More formally, let $\mathrm{hs} : 2^{2^\mathcal{L}} \mapsto 2^{2^\mathcal{L}}$ be the function which returns all *hitting sets* for an input family of sets, i.e. :

Definition 3.5. *hitting sets*
$\mathrm{hs}(\mathcal{X}) = \{\Delta \subseteq \bigcup \mathcal{X} \mid \forall X \in \mathcal{X} : X \cap \Delta \neq \emptyset\}$

A hitting set for $\mathcal{J}$ (i.e. an element of $\mathrm{hs}(\mathcal{J})$) will be called an *incision*, and a *minimal incision* iff it is minimal wrt set inclusion among all incisions. Then the following theorem, adapted (among others) from (Qi *et al.* 2008), defines $\mathcal{R}_\subseteq$ wrt $\mathcal{J}$, by stating that each element of $\mathcal{R}_\subseteq$ is the complement in K of a minimal incision, and conversely.

Theorem 3.1. $R \in \mathcal{R}_\subseteq$ iff there is a $D \in \min_\subseteq(\mathrm{hs}(\mathcal{J}))$ such that $R = K \setminus D$

For instance, in example 1, $D = \{3, 4\} \in \min_\subseteq(\mathrm{hs}(\mathcal{J}))$, so $R = \{1, 2, 5, 6, 7\} \in \mathcal{R}_\subseteq$. An immediate consequence is that no axiom of $K \setminus \bigcup \mathcal{J}$ needs to be discarded, and therefore it is sufficient for the preference relation $\preceq_a$ introduced in section 3.2 to be defined over $\bigcup \mathcal{J}$ (or equivalently, for some authors, over $K \cup \Theta$, but with $\max_{\preceq_a}(K \cup \Theta) = (K \cup \Theta) \setminus \bigcup \mathcal{J}$).

An equivalent way of viewing this correspondence between $\mathcal{J}$ and $\mathcal{R}_\subseteq$ consists in representing $\mathcal{J}$ as a boolean formula $\tau_\mathcal{J}$, expressing the fact that at least one formula in each $J \in \mathcal{J}$ must be selected for removal. For instance, if $\mathcal{J} = \{\{\phi_1, \phi_2\}, \{\phi_1, \phi_3\}\}$, $\tau_\mathcal{J} = (\phi_1 \vee \phi_2) \wedge (\phi_1 \vee \phi_3)$. Then the minimal incisions are exactly the prime implicants of $\tau_\mathcal{J}$. The converse holds as well, i.e. if $\mathcal{D}$ is the family of all minimal incisions, and if $\rho_\mathcal{D} = \bigvee_{D\in\mathcal{D}}(\bigwedge D)$, then $\mathcal{J}$ is the set of all prime implicates of $\rho_\mathcal{D}$.

[6] Once again, for the sake of readability, only the case of inconsistency is considered here

So in theory, any of $|\mathcal{J}|$ and $|\mathcal{R}_{\subseteq}|$ could be at most exponential in the size of each other (but it can be shown that if $n = |\bigcup \mathcal{J}|$, both are bounded by $\binom{n}{\lceil \frac{n}{2} \rceil}$).

Most attempts in DLs to obtain (part of) the remainder set rely on justifications, either by computing the whole kernel beforehand (Qi *et al.* 2008; Cóbe *et al.* 2013), or by generating conflicts on demand, using Reiter's hitting set algorithm (Schlobach 2005; Friedrich and Shchekotykhin 2005; Kalyanpur *et al.* 2006). It should be noted that in the latter case, conflicts do not need to be minimal, but both (Schlobach 2005) and (Friedrich and Shchekotykhin 2005) observed that the computation of minimal conflicts (i.e. justifications), or at least small conflicts, was indeed a more efficient strategy.

A relatively optimistic assumption will be made here, which is also made by several authors (Qi *et al.* 2008; Ribeiro and Wassermann 2008; Cóbe *et al.* 2013), namely that the kernel $\mathcal{M}$ (and therefore $\mathcal{J}$ as well) for K and Θ can actually be computed in a realistic amount of time, for instance using a saturated tableau, like in (Schlobach and Cornet 2003).

But as explained above, obtaining the whole remainder set $\mathcal{R}_{\subseteq}$ out of $\mathcal{J}$ remains intractable. So two problems will be discussed in sections 5 and 6, assuming that $\mathcal{J}$ is known for K and Θ, but not necessarily $\mathcal{R}_{\subseteq}$: how to compute $\mathcal{R}_{\preceq_r}$ given a preference relation $\preceq_a$ over $\bigcup \mathcal{J}$, and how to obtain such a preference relation on a semantically grounded basis. But before that, the following section reviews different solutions proposed in the literature in order to deal with the automated selection of axioms of K to be discarded.

4 State of the art

A first straightforward strategy to select a subset of $\mathcal{R}_{\subseteq}$ consists in selecting only the set $\mathcal{R}_{\leq}$ of elements of $\mathcal{R}$ with maximal cardinality. Obviously, $\mathcal{R}_{\leq} \subseteq \mathcal{R}_{\subseteq}$, so for each $R \in \mathcal{R}_{\leq}$, $R \cup \Theta$ is indeed a remainder. For DLs, the computation of (some elements of) $\mathcal{R}_{\leq}$ has been investigated by (Friedrich and Shchekotykhin 2005) in the case neither $\mathcal{J}$ nor $\mathcal{R}_{\subseteq}$ is known, using Reiter's hitting set algorithm, and by (Qi *et al.* 2008) in the case $\mathcal{J}$ only is known. But the rationale behind this selection function remains questionable, especially for relatively large datasets : for instance, if $|K| > 1000$, one may arguably wonder why discarding 12 axioms instead of 13 (or even 20) is necessarily preferable. For some real datasets too, (Corman *et al.* 2015b) observed that this heuristic often failed to discard the expected axioms, or was biased towards the removal of TBox axioms rather than ABox axioms. This is illustrated by example 1, where the best element of $\mathcal{R}_{\subseteq}$ is intuitively $K \setminus \{3, 4\}$, but it is not retained in $\mathcal{R}_{\leq}$, because there are 4 subbases in $R_{\subseteq}$ with strictly higher cardinality (or equivalently, 4 disagnoses in $\mathcal{D}$ with strictly lower cardinality).

As an alternative, the notion of *core* has been proposed in (Schlobach and Cornet 2003). They define a core of arity n as as set of axioms which appear in n elements of $\mathcal{V}$, the intuition being that an axiom appearing in a core of large arity is more likely to be faulty. Both (Qi *et al.* 2008) and (Cóbe *et al.* 2013) implemented this strategy in the case the whole kernel is already known, by iteratively discarding axioms which appear in the largest number of elements of $\mathcal{J}$ not hit thus far, until consistency is reached. For instance, if $\mathcal{V} = \{\{\phi_1, \phi_2\}, \{\phi_1, \phi_3\}, \{\phi_4, \phi_5\}\}$, then ϕ_1 will be discarded first, because it appears in 2 elements of $\mathcal{V}$, against only 1 for each other axiom of $\bigcup \mathcal{V}$. Adopting this constraint, the set of minimal incisions in this example would be $\{\{\phi_1, \phi_4\}, \{\phi_1, \phi_5\}\}$. This approach has been criticized by (Friedrich and Shchekotykhin 2005), who showed that an incision D obtained this way is not guaranteed to be minimal wrt cardinality among all incisions, or in other words, that $K \setminus D \notin \mathcal{R}_{\leq}$. This should not be a surprise, as this strategy is actually the well-studied greedy algorithm to the (cardinality) minimal hitting set problem, and is known to be an (optimal) approximation. It should also be noted that $K \setminus D$ is not guaranteed to be minimal wrt set inclusion either, as illustrated by this counterexample : $\mathcal{J} = \{\{\phi_1, \phi_2, \phi_3\}, \{\phi_1, \phi_4, \phi_5\}, \{\phi_1, \phi_6, \phi_7\}, \{\phi_2, \phi_4\}, \{\phi_3, \phi_6\}, \{\phi_5, \phi_7\}\}$. No element of $\mathcal{J}$ is a subset of another, so this is a possible configuration. As ϕ_1 appears in 3 elements of $\mathcal{J}$, its removal will be prioritized, and $D = \{\phi_1, \phi_3, \phi_4, \phi_7\}$ is one of the resulting incisions. But $D \setminus \{\phi_1\}$ is itself an incision, so D is not a minimal incision, or in other words $D \notin \min_{\subseteq}(\mathrm{hs}(\mathcal{J}))$, and therefore $(K \setminus D) \notin \mathcal{R}_{\subseteq}$, i.e. $(K \setminus D) \cup \Theta$ could be extended with ϕ_1 without compromising consistency. So an additional (polynomial) verification is required, in order to identify the subset(s) of D which are indeed minimal incisions. But whether these minimal incisions are good candidates for removal (or equivalently, whether their respective complements in K should be among the selected remainders) is an open question.

(Qi *et al.* 2008) also investigated the usage of the kernel in order to perform prioritized base revision, which is the objective pursued in this article. Given a preference relation $\preceq_a$ over $\bigcup \mathcal{J}$, they propose to compute all minimal incisions over the elements of $\mathcal{J}$ reduced to their lower-ranked axioms wrt $\preceq_a$. In other words, if $X_1^{\preceq_a}$ is the equivalence class of lower ranked axioms of X wrt $\preceq_a$, and if $\mathcal{D} = \min_{\subseteq} \mathrm{hs}(\{J_1^{\preceq_a} \mid J \in \mathcal{J}\})$, then the procedure selects $\mathcal{R}_m \subseteq R$, defined by $\mathcal{R}_m = \{K \setminus D \,|D \in \mathcal{D}\}$. Each $R \in \mathcal{R}_m$ is indeed consistent, because all elements of $\mathcal{J}$ are hit, and within each J, the removal of axioms with lowest preference has also been prioritized. But once again, $\mathcal{R}_m \subseteq \mathcal{R}_{\subseteq}$ doesn't hold in general, as illustrated by the following example : $\mathcal{J} = \{J, J'\}$, with $J = \{\phi_1, \phi_2\}$, $J' = \{\phi_2, \phi_3\}$, and $\phi_1 \prec_a \phi_2 \prec_a \phi_3$. Then $J_1^{\preceq_a} = \{\phi_1\}$ and $J_1'^{\preceq_a} = \{\phi_2\}$, so the only possible incision is $D = \min_{\subseteq} \mathrm{hs}(\{J_1^{\preceq_a}, J_1'^{\preceq_a}\}) = \{\phi_1, \phi_2\}$, and therefore $\mathcal{R}_m = \{\{\phi_3\}\}$. But $\{\phi_1, \phi_3\} \cup \Theta \nvdash \bot$, so ϕ_1 has been unnecessarily discarded.

Alternatively, assuming $\mathcal{J}$ is known still, as well as a ranking $\preceq_a$ over $\bigcup \mathcal{J}$ (or a confidence score for each $\phi \in \bigcup \mathcal{J}$), (Kalyanpur *et al.* 2006) proposed to compute all minimal incisions which maximize the sum of the rankings of axioms to be discarded (or which minimize the sum of their scores). Because the incisions are minimal (wrt set inclu-

sion), the output is this time an actual selection of remainders. But it is easy to show that this procedure does not comply in general to prioritized revision as defined in 3.2.

Another proposal is the so-called *lexicographic approach* (Benferhat *et al.* 1993), whose corresponding disjunctive KB can be computed with the *disjunctive maxi adjustment* procedure (Benferhat *et al.* 2004) (in the context of propositional logic). If $\preceq_a$ is a preference relation over $K \cup \Theta$ (with $\max_{\preceq_a}(K \cup \Theta) = (K \cup \Theta) \setminus \bigcup \mathcal{J}$), the procedure yields the set $\mathcal{R}_{\preceq_q} = \max_{\preceq_q} \mathcal{R}$ of candidate subbases,[7] with $\preceq_q$ a partial order over $\mathcal{R}$ defined by $R \prec_q R'$ iff there is an $i_{1 \leq i \leq n}$ such that :

$|R_i^{\preceq_a}| < |R_i'^{\preceq_a}|$ and $\forall j\ _{1 \leq j < i} : |R_j^{\preceq_a}| = |R_j'^{\preceq_a}|$.

In this case too, the output is (the disjunction of) a selection of remainders, or in other words :

Proposition 4.1. $\mathcal{R}_{\preceq_q} \subseteq \mathcal{R}_{\subseteq}$

But this option is not completely satisfying either, because it is based on cardinality just like $\mathcal{R}_{\leq}$ above (actually, $\mathcal{R}_{\leq}$ is a specific case of $\mathcal{R}_{\preceq_q}$, where $\preceq_a$ defines only two equivalence classes, namely $\bigcup \mathcal{J}$ and $(K \cup \Theta) \setminus \bigcup \mathcal{J}$), therefore yielding the same criticism.

5 Algorithm for prioritized base revision

Given a preference relation $\preceq_a$ over $\bigcup \mathcal{J}$, and assuming $\mathcal{J}$ is known, but not necessarily $\mathcal{R}_{\subseteq}$, this section provides an algorithm to perform prioritized revision as defined in section 3, i.e. it yields the set $\mathcal{R}_{\preceq_r} \subseteq \mathcal{R}_{\subseteq}$, without the need to compute the whole of $\mathcal{R}_{\subseteq}$ (unless $\mathcal{R}_{\preceq_r} = \mathcal{R}_{\subseteq}$). The procedure itself is just an improvement over the computation of $\mathcal{R}_{\subseteq}$ out of $\mathcal{J}$, taking advantage of $\preceq_a$. As explained in section 3.3, obtaining $\mathcal{R}_{\subseteq}$ out of $\mathcal{J}$ amounts to computing $\min_{\subseteq} \mathrm{hs}(\mathcal{J})$, i.e. all minimal incisions for $\mathcal{J}$, which is a well studied problem.[8] It is therefore assumed that some procedure to solve this problem is available (prototypically a search tree), which for any finite family of finite sets $\mathcal{X}$ returns $\min_{\subseteq} \mathrm{hs}(\mathcal{X})$.

Some definitions will be useful. Let $\mathrm{hit} : 2^{2^{\mathcal{L}}} \times \mathcal{L}^2 \mapsto 2^{2^{\mathcal{L}}}$ be defined by :

Definition 5.1. $\mathrm{hit}(\mathcal{X}, Y) = \{X \in \mathcal{X} \mid X \cap Y \neq \emptyset\}$

If $\mathcal{X} \subseteq 2^{\mathcal{L}}$, and $\preceq_x$ is a total preorder over $\bigcup \mathcal{X}$, let $(\bigcup \mathcal{X})^{\preceq_x} = ((\bigcup \mathcal{X})_1^{\preceq_x}, .., (\bigcup \mathcal{X})_n^{\preceq_x})$, and $1 \leq i \leq n$. Then the function $\mathrm{pHit} : 2^{2^{\mathcal{L}}} \times \mathcal{L}^2 \times n \mapsto 2^{2^{\mathcal{L}}}$ is defined by :

Definition 5.2. $\mathrm{pHit}(\mathcal{X}, \preceq_x, i) = \{X \in \mathcal{X} \mid X \in \mathrm{hit}(\mathcal{X}, (\bigcup \mathcal{X})_i^{\preceq_x})$ and $\forall 1 \leq j < i : X \notin \mathrm{hit}((\bigcup \mathcal{X})_j^{\preceq})\}$

[7] or a single base equivalent to the disjunction of these subbases in the case of *disjunctive maxi adjustment*.

[8] To avoid confusions, the problem considered here, i.e. obtaining $\min_{\subseteq} \mathrm{hs}(\mathcal{J})$, differs from the canonical *minimal hitting set* problem (NP-complete in the size of $\mathcal{J}$), which consists in computing *one* hitting set minimal *wrt cardinality*, whereas $\min_{\subseteq} \mathrm{hs}(\mathcal{J})$ is the set of *all* hitting sets minimal *wrt inclusion*. In particular, the decision problem for $\min_{\subseteq} \mathrm{hs}(\mathcal{J})$ is not polynomial anymore, because $\min_{\subseteq} \mathrm{hs}(\mathcal{J})$ is exponential in $|\mathcal{J}|$ in the worst case.

In particular, if $(\bigcup \mathcal{J})^{\preceq_a} = ((\bigcup \mathcal{J})_1^{\preceq_a}, .., (\bigcup \mathcal{J})_n^{\preceq_a})$, then $\mathrm{pHit}(\mathcal{J}, \preceq_a, i)$ returns the elements of $\mathcal{J}$ which are hit by $(\bigcup \mathcal{J})_i^{\preceq_a}$, but not by any lower-ranked equivalence class of $(\bigcup \mathcal{J})^{\preceq_a}$.

The following observation gives the rationale behind algorithm 1 :

Proposition 5.1. $R \in \mathcal{R}_{\preceq_r}$ iff $R \subseteq K$ and $K \setminus R = D$, with $\forall i_{1 \leq i \leq n} : D_i^{\preceq_a} \in \min_{\subseteq} \mathrm{hs}(\mathrm{pHit}(\mathcal{J}, \preceq_a, i) \setminus (\mathrm{hit}(\mathcal{J}, \bigcup_{1 \leq j < i} D_j^{\preceq_a})))$

In other words, starting with the equivalence class $(\bigcup \mathcal{J})_n^{\preceq}$ of highest ranked axioms of $\bigcup \mathcal{J}$, the axioms discarded from each equivalence class $(\bigcup \mathcal{J})_i^{\preceq}$ must form a minimal hitting set for all elements of $\mathcal{J}$ not hit thus far, and which cannot be hit by any strictly lower ranked axiom.

Let $\mathrm{hitInt} : 2^{\mathcal{L}} \times 2^{2^{\mathcal{L}}} \mapsto 2^{2^{\mathcal{L}}}$ be the function which reduces the sets of an input family $\mathcal{X}$ to their respective intersections with an input set Δ, i.e. :

Definition 5.3. $\mathrm{hitInt}(\Delta, \mathcal{X}) = \{X \cap \Delta \mid X \in \mathcal{X}\}$

Then the following algorithm yields $\mathcal{R}_{\preceq_r}$:

Algorithm 1 Prioritized base revision with known kernel

```
1:  D ← {∅}
2:  for i ← n to 1 do
3:      D' ← {∅}
4:      PH ← pHit(J, ⪯a, i)
5:      for all D ∈ D do
6:          J' ← hitInt((⋃J)_i^⪯a, PH \ hit(J, D))
7:          for all H ∈ min⊆ hs(J') do
8:              D' ← D' ∪ {D ∪ H}
9:          end for
10:     end for
11:     D ← D'
12: end for
13: R_⪯r ← ∅
14: for all D ∈ D do
15:     R_⪯r ← R_⪯r ∪ {K \ D}
16: end for
```

The family $\mathcal{D}$ is the set of minimal incisions under construction, and the family $\mathcal{D}'$ is just a temporary variable to avoid concurrent modification of $\mathcal{D}$. The main loop (line 2) iterates over all equivalence classes defined by $\preceq_a$, starting with the best ranked equivalence class, i.e. $(\bigcup \mathcal{J})_n^{\preceq_a}$. Line 4, *PH* is the set of all elements of $\mathcal{J}$ hit by the current equivalence class $(\bigcup \mathcal{J})_i^{\preceq_a}$, and not by any lower ranked equivalence class. Line 5 starts an iteration over all incisions under construction. For each of these incisions D, line 6, $PH \setminus \mathrm{hit}(\mathcal{J}, D)$ is the set of elements of $\mathcal{J}$ hit by the current equivalence class, by no lower ranked equivalence class, and not hit by D yet. Then $\mathcal{J}'$ (temporarily) contains these, but reduced to their respective intersections with the current equivalence class. The minimal hitting set procedure $\min_{\subseteq} \mathrm{hs}$ is called line 7, and for each returned minimal hitting set H for $\mathcal{J}'$, $D \cup H$ is a possibly new minimal incision under construction. Finally, line 13 to 16, each $R \in \mathcal{R}_{\preceq_r}$ is

obtained as the complement in K of some minimal incision $D \in \mathcal{D}$.

In the limit case where all equivalence classes hit distinct elements of $\bigcup \mathcal{J}$, i.e. where all $\mathrm{hit}(\mathcal{J}, (\bigcup \mathcal{J})_i^{\preceq_a})$ are mutually disjoint for $1 \leq i \leq n$ (note that $n = 1$ is a particular subcase), computing $\mathcal{R}_{\preceq_r}$ amounts to computing $\mathcal{R}_{\subseteq}$. But in other cases, computing $\mathcal{R}_{\preceq_r}$ with algorithm 1 is strictly less expensive than computing $\mathcal{R}_{\subseteq}$. To see this, consider the worst possible scenario, where all elements of $\mathcal{J}$ are mutually disjoint, with $m = |\mathcal{J}|$, and for simplicity, let us assume that they all have the same cardinality k. Then $|\mathcal{R}_{\subseteq}| = k^m$. But if $f_i = \max_{J \in \mathrm{pHit}(\mathcal{J}, \preceq_a, i)} |J \cap (\bigcup \mathcal{J})_i^{\preceq_a}|$ and $q_i = |\mathrm{pHit}(\mathcal{J}, \preceq_a, i)|$, then inside each equivalence class $(\bigcup \mathcal{J})_i^{\preceq_a}$, the number of incisions to compute is at most $(f_i)^{q_i}$. Then by hypothesis, for some $1 \leq j < l \leq n$, $\mathrm{hit}(\mathcal{J}, (\bigcup \mathcal{J})_j^{\preceq_a})$ and $\mathrm{hit}(\mathcal{J}, (\bigcup \mathcal{J})_l^{\preceq_a})$ overlap, therefore not only $f_l < k$, but more importantly, $\sum_{i=1}^{n} q_i < m$. Therefore the total number of minimal incisions computed during the execution is :

$|\mathcal{D}| \leq \prod_{i=1}^{n} (f_i)^{q_i} < \prod_{i=1}^{n} (k)^{q_i} = (k)^{\sum_{i=1}^{n} q_i} < k^m$.

An additional (exponential) gain may come line 6 from the fact that reducing the elements of $\mathrm{pHit}(\mathcal{J}, \preceq_a, i) \setminus \mathrm{hit}(\mathcal{J}, D)$ to their respective intersections with $(\bigcup \mathcal{J})_i^{\preceq_a}$ can also reduce their overall number. Finally, for each i from n to 1, the hitting set tree procedure is called once per hitting set computed thus far, i.e. $|\mathcal{D}|$ times, which may turn out to be costly. Observing that the set $PH \setminus \mathrm{hit}(\mathcal{J}, D)$ may be identical for multiple $D \in \mathcal{D}$, a simple optimization consists in keeping track of the hitting sets computed for each D, which guarantees that the number of calls to the hitting set tree procedure line 7 is bounded by the smallest value between $|\mathcal{D}|$ and $2^{|\mathrm{hitInt}((\bigcup \mathcal{J})_i^{\preceq_a}, PH)|}$.

6 Obtaining a preference relation $\preceq_a$ over axioms

In the previous sections, it has been assumed that a preference relation $\preceq_a$ over the axioms of $\bigcup \mathcal{J}$ (i.e. a ranking of these axioms) was available, and the efforts have been centered on the computation of $R_{\preceq_r}$, guided by this relation. This assumption is often made in works dealing with prioritized base revision, where $\preceq_a$ is supposed to be obtained from external confidence scores for axioms, or from a manual review of these axioms. Syntactic criteria have also been proposed to define $\preceq_a$, for instance favoring TBox over ABox axioms (or the opposite), favoring axioms whose signatures contain elements with more syntactic occurrences within $K \cup \Theta$, or penalizing axioms based on some syntactic patterns (frequent modeling errors) ...For consistent but incoherent KBs, (Kalyanpur *et al.* 2006) also used as a ranking criterion the number of consequences of a given syntactic form which would be necessarily lost if an axiom ϕ was discarded. This section follows this last intuition, although it does not provide a specific $\preceq_a$ (possible concrete preference relations are discussed in section 7). Instead, it shows that if the kernel is known, for each uncertain axiom $\phi \in \bigcup \mathcal{J}$, two subbases of $K \cup \Theta$ can be computed in polynomial time, which respectively reflect what retaining or discarding ϕ necessarily implies.

The intuition is simple, and can be summarized with only two questions. Let $\phi \in \bigcup \mathcal{J}$, i.e. ϕ is involved in the inconsistency of $K \cup \Theta$ but is not part of Θ, and is therefore a candidate for removal. The first question one may ask is, if ϕ was retained, which part of the initial base $K \cup \Theta$ would necessarily be retained with it. Let $\mathcal{S}_\phi$ be the set of all remainders which contain ϕ, i.e. :

Definition 6.1. $\mathcal{S}_\phi = \{R \in \mathcal{R}_{\subseteq} \mid \phi \in R\} \cup \Theta$

Then the knowledge necessarily retained together with ϕ can be conveniently represented by $\bigcap \mathcal{S}_\phi$. Whatever the selected remainders are, if they contain ϕ, then the output of the process is guaranteed to be at least as strong as $\bigcap \mathcal{S}_\phi$. Or in other words, if one adheres to the assumption made throughout this article that all selected subbases should be maximal wrt to set-inclusion, and if additionally one would like to retain ϕ, then $\bigcap \mathcal{S}_\phi$ must be retained as well. In particular, some properties of $\bigcap \mathcal{S}_\phi$ considered as a theory (i.e. $\mathrm{Cn}(\bigcap \mathcal{S}_\phi)$) may be exploited to obtain a score or ranking for ϕ, setting a basis for a semantically grounded computation of $\preceq_a$ (note that because all elements of $\mathcal{S}_\phi$ are remainders, they are also consistent, so $\bigcap \mathcal{S}_\phi$ is consistent as well).

A first legitimate objection to this proposition can be made, which is intuitively that $\bigcap \mathcal{S}_\phi$ as a theory does not fully reflect the impact of ϕ within $K \cup \Theta$. In particular, it misses consequences of some elements of $\mathcal{S}_\phi$. Arguably, if all elements of $\mathcal{S}_\phi$ could be studied individually, this would provide a much more accurate understanding of the impact of ϕ. But if all $\mathcal{S}_\phi$ were known for all ϕ, then the whole remainder set would be known as well, and it was assumed in section 3.3 that this is not necessarily the case. And even if the whole remainder was known, evaluating all remainders independently as theories can be simply prohibitive : if $n = |\bigcup \mathcal{J}|$, there are at most $\binom{n}{\lceil \frac{n}{2} \rceil}$ of them, whereas there is obviously just one $\bigcap \mathcal{S}_\phi$ for each ϕ, i.e. n of them. Actually, if the complete remainder set could be computed, and if its cardinality was reasonable as well, there would be reason to perform prioritized base revision in the first place. So $\bigcap \mathcal{S}_\phi$ should be viewed as a convenient computational compromise, altogether semantically motivated and easy to obtain if $\mathcal{J}$ is already known, as will be shown below.

A case could also be made for considering the disjunctive KB $\bigvee \mathcal{S}_\phi$ (defined in section 3.1) instead of $\bigcap \mathcal{S}_\phi$ for this purpose. Arguably, $\bigvee \mathcal{S}_\phi$ is a better representation of the knowledge being retained together with ϕ. But because disjunctive KBs cannot be natively represented in most DLs, $\bigvee \mathcal{S}_\phi$ must be manipulated as a family of KBs, namely $\mathcal{S}_\phi$ (and the corresponding theory is set to be $\bigcap_{\Delta \in \mathcal{S}_\phi} \mathrm{Cn}(\Delta)$), which leads back to the previous objection.

A second base can also be computed to answer a second dual question, which is what part of the input KB would necessarily remain if ϕ was discarded. This base is $\bigcap \mathcal{S}_{\setminus\phi}$, with $\mathcal{S}_{\setminus\phi}$ defined by :

Definition 6.2. $\mathcal{S}_{\setminus\phi} = \{R \in \mathcal{R}_{\subseteq} \mid \phi \notin R\} \cup \Theta$

This gives two KBs $\bigcap \mathcal{S}_\phi$ and $\bigcap \mathcal{S}_{\setminus\phi}$ for each uncertain axiom $\phi \in \bigcup \mathcal{J}$, i.e. $2*|\bigcup \mathcal{J}|$ KBs in total, which can serve as a basis to evaluate the axioms of $\bigcup \mathcal{J}$, and eventually compute the preference relation $\preceq_a$.

If $\mathcal{R}_\subseteq$ is known, obtaining $\bigcap \mathcal{S}_\phi$ and $\bigcap \mathcal{S}_{\setminus\phi}$ is trivial from their definitions. But a more interesting observation is that even if only $\mathcal{J}$ is known, and not $\mathcal{R}_\subseteq$, then the intersections $\bigcap \mathcal{S}_\phi$ and $\bigcap \mathcal{S}_{\setminus\phi}$ can still be obtained in time polynomial in the size of $\mathcal{J}$, without the need to compute $\mathcal{S}_\phi$ and $\mathcal{S}_{\setminus\phi}$.

A few additional definitions are useful. K_s will designate the "safe" part of K, i.e. the axioms of K which are not involved in the inconsistency of $K \cup \Theta$, or equivalently, which do not appear in any element of $\mathcal{J}$.

Definition 6.3. $K_s = K \setminus \bigcup \mathcal{J}$

Finally, given a family of sets $\mathcal{X}$, and two elements x_1 and x_2, the function $\mathrm{hitDiff} : 2^{2^L} \times L \times L \mapsto 2^{2^L}$ returns the elements of $\mathcal{X}$ to which x_1 belongs, but not x_2, i.e. :

Definition 6.4. $\mathrm{hitDiff}(\mathcal{X}, x_1, x_2) = \mathrm{hit}(\mathcal{X}, \{x_1\}) \setminus \mathrm{hit}(\mathcal{X}, \{x_2\})$

Then the two following equalities give two straightforward procedures to compute $\bigcap \mathcal{S}_\phi$ and $\bigcap \mathcal{S}_{\setminus\phi}$ for each $\phi \in \bigcup \mathcal{J}$, provided $\mathcal{J}$ is known :[9]

Proposition 6.1. Let $\phi \in \bigcup \mathcal{J}$.
If $\{\phi\} \cup \Theta \vdash \bot$, then $\bigcap \mathcal{S}_\phi = \emptyset$.
Otherwise, $\bigcap \mathcal{S}_\phi = K_s \cup \{\phi\} \cup \Theta \cup$
$\{\psi \in \bigcup \mathcal{J} \setminus \bigcup \mathrm{hit}(\mathcal{J}, \{\phi\}) \mid$
$\forall J_1 \in \mathrm{hit}(\mathcal{J}, \{\psi\}), \exists J_2 \in \mathrm{hit}(\mathcal{J}, \{\phi\}) : J_2 \setminus \{\phi\} \subseteq J_1\}$

The first precaution is just a limit case, where no remainder contains ϕ. Otherwise, trivially, $K_s \cup \{\phi\} \cup \Theta \subseteq \bigcap \mathcal{S}_\phi$. Then given any other $\psi \in \bigcup \mathcal{J}$ which does not appear in an element of $\mathcal{J}$ hit by $\{\phi\}$, in order to decide whether $\psi \in \bigcap \mathcal{S}_\phi$, it is sufficient to check for each element J_1 of $\mathcal{J}$ which contains ψ if there is another element J_2 of $\mathcal{J}$ which contains ϕ, and such that $J_2 \setminus \{\phi\} \subseteq J_1$. Such a verification remains polynomial in $|\mathcal{J}|$.

Proposition 6.2. Let $\phi \in \bigcup \mathcal{J}$.
Then $\bigcap \mathcal{S}_{\setminus\phi} = K_s \cup \Theta \cup \{\psi \in ((\bigcup \mathcal{J}) \setminus \{\phi\}) \mid$
if $\mathrm{hitDiff}(\mathcal{J}, \psi, \phi) \neq \emptyset$ and $\mathrm{hitDiff}(\mathcal{J}, \phi, \psi) \neq \emptyset$,
then $\forall J_1 \in \mathrm{hitDiff}(\mathcal{J}, \psi, \phi), \forall J_2 \in \mathrm{hitDiff}(\mathcal{J}, \phi, \psi) :$
$((J_1 \cup J_2) \setminus \{\phi, \psi\}) \cup \Theta \vdash \bot \}$

Again, trivially, $K_s \cup \Theta \subseteq \bigcap \mathcal{S}_{\setminus\phi}$. Then any $\psi \neq \phi$ such that $\psi \in \bigcup \mathcal{J}$ and either $\mathrm{hit}(\mathcal{J}, \{\phi\}) \subseteq \mathrm{hit}(\mathcal{J}, \{\psi\})$ or $\mathrm{hit}(\mathcal{J}, \{\psi\}) \subseteq \mathrm{hit}(\mathcal{J}, \{\phi\})$ will be retained as well. Otherwise, for ψ to be retained, for each element J_1 only hit by $\{\psi\}$, for each element J_2 only hit by $\{\phi\}$, $((J_1 \cup J_2) \setminus \{\phi, \psi\}) \cup \Theta \vdash \bot$ must hold. This very last condition may suggest a consistency check, but it is actually not required. Because $\mathcal{J}$ is known, it is sufficient instead to check whether there is a $J_3 \in \mathcal{J}$ such that $J_3 \subseteq ((J_1 \cup J_2) \setminus \{\phi, \psi\}))$.

7 Applications

As explained in the previous section, this article does not define a unique preference relation $\preceq_a$ over the set $\bigcup \mathcal{J}$ of potentially undesired axioms, but instead provides a way to compute two bases $\bigcap \mathcal{S}_\phi$ and $\bigcap \mathcal{S}_{\setminus\phi}$ for each ϕ, which respectively reflect the impact of retaining or discarding ϕ. A wide array of existing strategies to evaluate the quality of a KB on a semantic basis can then be used to evaluate $\bigcap \mathcal{S}_\phi$ and $\bigcap \mathcal{S}_{\setminus\phi}$, and rank all $\phi \in \bigcup \mathcal{J}$ accordingly. This section only lists a few of them.

From an engineering perspective, a traditional approach to evaluate a knowledge base is through requirements or so-called *competency questions* (Pinto and Martins 2004), which can generally be viewed as a set Ψ of statements the KB should entail (or sometimes not entail) for a given application. Let us assume that a base debugging algorithm, whichever it is (for instance the one proposed in section 5), produces as an output some selection of remainders (or their intersection, or their disjunction). And let $\Psi^+_\phi = \Psi \cap \mathrm{Cn}(\bigcap \mathcal{S}_\phi)$. Then $\psi \in \Psi^+_\phi$ iff ψ is guaranteed to be a consequence of any output KB containing ϕ. In other words, if ϕ is retained, then the requirements in Ψ^+_ϕ will all necessarily be met, and there is no other requirement in Ψ for which this is guaranteed. Similarly, if $\Psi^-_\phi = \Psi \cap \mathrm{Cn}(\bigcap \mathcal{S}_{\setminus\phi})$, then $\psi \in \Psi^-_\phi$ iff ψ is guaranteed to be entailed by any output KB which does not contain ϕ. So $\preceq_a$ may be computed on this simple basis, depending on the importance given to the respective requirements in Ψ. The main limitation here is that it cannot be guaranteed that some requirement ψ will *not* be met if ϕ is retained (resp. discarded).

As an alternative, the evaluation of $\bigcap \mathcal{S}_\phi$ (resp. $\bigcap \mathcal{S}_{\setminus\phi}$) may be based on the complete derivation of some syntactically defined subset of $\mathrm{Cn}(\bigcap \mathcal{S}_\phi)$ (resp. $\mathrm{Cn}(\bigcap \mathcal{S}_{\setminus\phi})$). This has been experimented by (Kalyanpur *et al.* 2006; Cóbe *et al.* 2013) or (Corman *et al.* 2015a).

But more theoretically motivated approaches are conceivable as well. In particular, model-theoretic distances used for belief set revision (Qi *et al.* 2006; Qi and Du 2009) or update (Liu *et al.* 2006) may be used here to obtain $\preceq_a$, considering for each $\phi \in \bigcup \mathcal{J}$ the distance between the set of models of K and the set of models of $\bigcap \mathcal{S}_\phi$ (resp. $\bigcap \mathcal{S}_{\setminus\phi}$).

8 Conclusion

This article investigates the applicability of prioritized base revision/debugging in Description Logics, requiring that the output be a selection of base remainders, and that this selection be independent from the cardinality of these remainders. It is shown that propositions made in the literature do not satisfy both these requirements, and an effective algorithm for this task is provided, assuming the kernel is known.

The second main contribution concerns the prior computation of a preference relation $\preceq_a$ over the axioms of the input base, which is necessary for prioritized base revision/debugging. Assuming once again that the kernel is known, it is shown that for each candidate axiom ϕ for removal, two bases $\bigcap \mathcal{S}_\phi$ and $\bigcap \mathcal{S}_{\setminus\phi}$ can immediately be obtained, which respectively represent what part of the initial knowledge will necessarily be preserved if ϕ is retained or discarded, allowing the computation of $\preceq_a$ on a semantic basis, thus offering a good compromise between computational and more theoretical requirements.

[9] Again, all proofs are available at the abovementioned url.

An interesting continuation of this work would be to determine whether and/or to what extent the computation of the whole kernel beforehand could be avoided, for both tasks, i.e. obtaining $\bigcap \mathcal{S}_\phi$ and $\bigcap \mathcal{S}_{\setminus\phi}$ on the one hand, and debugging on the other hand.

References

[Baader *et al.* 2003] Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter Patel-Schneider. *The description logic handbook: theory, implementation and applications*. Cambridge university press, 2003.

[Benferhat *et al.* 1993] Salem Benferhat, Claudette Cayrol, Didier Dubois, Jerome Lang, and Henri Prade. Inconsistency management and prioritized syntax-based entailment. In *IJCAI proceedings*, 1993.

[Benferhat *et al.* 2004] Salem Benferhat, Souhila Kaci, Daniel Le Berre, and Mary-Anne Williams. Weakening conflicting information for iterated revision and knowledge integration. *Artificial Intelligence*, 153(1), 2004.

[Corman *et al.* 2015a] Julien Corman, Nathalie Aussenac-Gilles, and Laure Vieu. Distributional semantics for ontology verification. In *StarSem proceedings*, 2015.

[Corman *et al.* 2015b] Julien Corman, Laure Vieu, and Nathalie Aussenac-Gilles. Ontological analysis for practical knowledge base debugging in Description Logics. In *CommonSense proceedings*, 2015.

[Cóbe *et al.* 2013] Raphael Cóbe, Fillipe Resina, and Renata Wassermann. Merging Ontologies via Kernel Contraction. *ONTOBRAS proceedings*, pages 94–105, 2013.

[Flouris 2006] Giorgos Flouris. *On Belief Change and Ontology Evolution*. PhD thesis, University of Crete, 2006.

[Friedrich and Shchekotykhin 2005] Gerhard Friedrich and Kostyantyn Shchekotykhin. A general diagnosis method for ontologies. In *ISWC proceedings*, 2005.

[Horrocks *et al.* 2006] Ian Horrocks, Oliver Kutz, and Ulrike Sattler. The Even More Irresistible SROIQ. *KR proceedings*, 2006.

[Kalyanpur *et al.* 2006] Aditya Kalyanpur, Bijan Parsia, Evren Sirin, and Bernardo Cuenca-Grau. Repairing unsatisfiable concepts in OWL ontologies. In *ESWC proceedings*, 2006.

[Liu *et al.* 2006] Hongkai Liu, Carsten Lutz, Maja Milicic, and Frank Wolter. Updating Description Logic ABoxes. In *KR proceedings*, 2006.

[Mendes *et al.* 2012] Pablo N Mendes, Max Jakob, and Christian Bizer. DBpedia: A Multilingual Cross-domain Knowledge Base. In *LREC proceedings*, 2012.

[Meyer *et al.* 2005] Thomas Meyer, Kevin Lee, and Richard Booth. Knowledge integration for description logics. In *AAAI proceedings*, 2005.

[Nebel 1992] Bernhard Nebel. Syntax-based approaches to belief revision. *Belief revision*, 29:52–88, 1992.

[Pinto and Martins 2004] Helena Sofia Pinto and João P Martins. Ontologies: How can they be built? *Knowledge and Information Systems*, 6(4):441–464, 2004.

[Qi and Du 2009] Guilin Qi and Jianfeng Du. Model-based Revision Operators for Terminologies in Description Logics. In *IJCAI proceedings*, 2009.

[Qi *et al.* 2006] Guilin Qi, Weiru Liu, and David Bell. A revision-based approach to handling inconsistency in description logics. *Artificial Intelligence Review*, 26(1-2), 2006.

[Qi *et al.* 2008] Guilin Qi, Peter Haase, Zhisheng Huang, Qiu Ji, Jeff Z Pan, and Johanna Völker. A kernel revision operator for terminologies - algorithms and evaluation. In *ISWC proceedings*, 2008.

[Ribeiro and Wassermann 2008] Márcio Moretto Ribeiro and Renata Wassermann. The Ontology Reviser Plug-In for Protégé. In *WONTO proceedings*, 2008.

[Ribeiro and Wassermann 2009] Márcio M Ribeiro and Renata Wassermann. Base revision for ontology debugging. *Journal of Logic and Computation*, 19(5), 2009.

[Ribeiro 2013] Márcio Moretto Ribeiro. *Belief revision in non-classical logics*. Springer, 2013.

[Schlobach and Cornet 2003] Stefan Schlobach and Ronald Cornet. Non-standard reasoning services for the debugging of description logic terminologies. In *IJCAI proceedings*, 2003.

[Schlobach 2005] Stefan Schlobach. Diagnosing terminologies. In *AAAI proceedings*, 2005.

[Wang *et al.* 2010] Zhe Wang, Kewen Wang, and Rodney W Topor. A New Approach to Knowledge Base Revision in DL-Lite. In *AAAI proceedings*, 2010.