How to Select One Preferred Assertional-Based Repair from Inconsistent and Prioritized DL-Lite Knowledge Bases?

Salem Benferhat and Zied Bouraoui and Karim Tabia
Univ Lille Nord de France, F-59000 Lille, France
UArtois, CRIL - CNRS UMR 8188, F-62300 Lens, France
{benferhat,bouraoui,tabia}@cril.fr

Abstract
Managing inconsistency in DL-Lite knowledge bases where the assertional base is prioritized is a crucial problem in many applications. This is especially true when the assertions are provided by multiple sources having different reliability levels. This paper first reviews existing approaches for selecting preferred repairs. It then focuses on suitable strategies for handling inconsistency in DL-Lite knowledge bases. It proposes new approaches based on the selection of only one preferred repair. These strategies have as a starting point the so-called non-defeated repair and add one of the following principles: deductive closure, consistency, cardinality and priorities. Lastly, we provide a comparative analysis followed by an experimental evaluation of the studied approaches.

1 Introduction
Description Logics (DLs) are formal frameworks for representing and reasoning with ontologies. A DL knowledge base (KB) is built upon two distinct components: a terminological base (called TBox), representing generic knowledge, and an assertional base (called ABox), containing the facts that instantiate the TBox. DL-Lite [Calvanese et al., 2007] is a family of tractable DLs specifically tailored for applications that use huge volumes of data, in which query answering is the most important reasoning task. DL-Lite guarantees a low computational complexity. This fact makes DL-Lite especially well suited for Ontology-Based Data Access (OBDA).

A crucially important problem that arises in OBDA is how to manage inconsistency. In such setting, inconsistency is defined with respect to some assertions that contradict the terminology. Typically, a TBox is usually verified and validated while the assertions can be provided in large quantities by various and unreliable sources and may contradict the TBox. Moreover, it is often too expensive to manually check and validate all the assertions. This is why it is very important in OBDA to reason in the presence of inconsistency. Many works (e.g. [Lembo et al., 2010; Bienvenu and Rosati, 2013]), basically inspired by database approaches (e.g. [Bertossi, 2011]), tried to deal with inconsistency in DL-Lite by adapting several inconsistency-tolerant inference methods. These latter are based on the notion of assertional (or ABox) repair which is closely related to the notion of database repair. An ABox repair is simply a maximal assertional subbase which is consistent with respect to a given TBox [Lembo et al., 2010].

In many applications, assertions are often provided by several and potentially conflicting sources having different reliability levels. Moreover, a given source may provide different sets of uncertain assertions with different confidence levels. Gathering such sets of assertions gives a prioritized or a stratified assertional base. The role of priorities in handling inconsistency is very important and it is largely studied in the literature within propositional logic setting (e.g. [Baral et al., 1992; Benferhat et al., 1995]). Several works (e.g. [Martinez et al., 2008; Staworko et al., 2012; Du et al., 2013]) studied the notion of priority when querying inconsistent databases or DL KBs. Unfortunately, in the OBDA setting, there are only few works, such as the one given in [Bienvenu et al., 2014] for dealing with reasoning under prioritized DL-Lite ABox.

The main question addressed in this paper is how to select one preferred repair. Selecting only one preferred repair is important since it allows an efficient query answering once the repair is computed. In this paper, we first review main existing inconsistency-tolerant reasoning methods for prioritized KBs. It is important to note that some inference relations are specific to DL-Lite even if they are inspired by other formalisms. One of the main contributions of the paper consists in providing new strategies to define a single preferred repair based on the use of the so-called non-defeated assertional base, plus with one/several of the following four ingredients: priorities, deductive closure, cardinality and consistency. Interestingly enough, several of these strategies are suitable for the DL-Lite setting in the sense that they allow efficient handling of inconsistency, by producing a single preferred assertional repair. Our experimental results show the benefits of handling priorities when reasoning under inconsistency in DL-Lite.

2 DL-Lite Logic: A Brief Refresher
This section briefly recalls DL-Lite logics. For the sake of simplicity, we only consider DL-Lite\textsubscript{core} language [Calvanese et al., 2007] and we will simply use DL-Lite instead of DL-Lite\textsubscript{core}. However, results of this paper can be eas-
ily extended to any tractable DL-Lite where an ABox conflict involves at most two assertions, in particular DL-Lite$B$ and DL-Lite$F$. The DL-Lite language is defined as follows:

$$R \rightarrow P|P^- \quad B \rightarrow A|\exists R \quad C \rightarrow B|\neg B$$

where $A$ is an atomic concept, $P$ is an atomic role and $P^-$ is the inverse of $P$. $B$ (resp. $C$) is called basic (resp. complex) concept and role $R$ is called basic role. A DL-Lite knowledge base (KB) is a pair $K = \langle T, A \rangle$ where $T$ is the TBox and $A$ is the ABox. A TBox includes a finite set of inclusion axioms on concepts of the form: $B \sqsubseteq C$. The ABox contains a finite set of assertions on atomic concepts and roles respectively of the form $A(a)$ and $P(a, b)$ where $a$ and $b$ are two individuals.

The semantics of a DL-Lite KB is given in term of interpretations. An interpretation $I = (\Delta_0, \ldots, \Delta_n)$ consists of a non-empty domain $\Delta_0$ and an interpretation function $\cdot|^I$ that maps each individual $a$ to $a^I \in \Delta_0$, each $A \subseteq \Delta_0$ and each role $P$ to $P^I \subseteq \Delta_0 \times \Delta_0$. Furthermore, the interpretation function $\cdot|^I$ is extended in a straightforward way for complex concepts and roles: $\neg B^I = \Delta_0 \setminus B^I$, $(P^-)^I = \{(y, x) \mid (x, y) \in P^I\}$ and $(\exists R)^I = \{x \mid \exists y. (x, y) \in R^I\}$. An interpretation $I$ is said to be a model of a concept inclusion axiom, denoted by $I \models B \sqsubseteq C$, iff $B^I \subseteq C^I$. Similarly, we say that $I$ satisfies a concept (resp. role) assertion, denoted by $I \models A(a)$ (resp. $I \models P(a, b)$), iff $a^I \in A^I$ (resp. $(a^I, b^I) \in P^I$). Note that we only consider DL-Lite with unique name assumption. A KB $K$ is said consistent if it admits at least one model, otherwise $K$ is said inconsistent. A TBox $T$ is said coherent if there exists at least a concept $C$ such that for each interpretation $I$ which is a model of $T$, we have $C^I \neq \emptyset$. In the rest of this paper, we denote by $q$ a query. The semantics of such query is given for instance in [Calvanese et al., 2007].

3 Existing Assertional-Based Preferred Repairs

This section reviews approaches dealing with inconsistent DL-Lite KB that either have been proposed in DLs setting or have been proposed in a propositional logic setting but need a slight adaptation to be suitable for DL-Lite.

A DL-Lite KB $K = \langle T, A \rangle$ with a prioritized assertional base is a DL-Lite KB where $A$ is partitioned into $n$ layers (or strata) of the form $A = S_1 \cup \ldots \cup S_n$ where each layer $S_i$ contains the set of assertions with the same level of priority $i$ and they are considered as more reliable than the ones present in a layer $S_{j}$ when $j > i$. Within the OBDA setting, we assume that $T$ is stable and hence its elements are not questionable in the presence of conflicts. Throughout this paper and when there is no ambiguity we simply use "prioritized DL-Lite KB $K = \langle T, A \rangle$" to refer to a DL-Lite KB with a prioritized assertional base of the form $A = S_1 \cup \ldots \cup S_n$.

Example 1. Let $K = \langle T, A \rangle$ such that $T = \{A \sqsubseteq \neg B\}$ and assume that assertional facts of $A$ come from three distinct sources $A = S_1 \cup S_2 \cup S_3$ such that: $S_1 = \{B(a), A(b)\}$, $S_2 = \{A(a)\}$ and $S_3 = \{B(c)\}$. $S_1$ contains the most reliable assertions. $S_3$ contains the least reliable assertions.

In Example 1, it is easy to check that the KB is inconsistent. Coping with inconsistency can be done by first computing the set of repairs, then using them to perform inference. In order to compute the repairs, we use the notion of conflict sets.

3.1 Conflict Sets

Within the OBDA setting, the inconsistency problem is always defined with respect to some ABox, since a TBox may be inconsistent but never inconsistent. We now introduce the notion of a conflict as a minimal inconsistent subset of assertions that contradict the TBox.

Definition 1. Let $K = \langle T, A \rangle$ be a DL-Lite KB. A subbase $C \subseteq A$ is said to be a conflict of $K$ iff $\langle T, C \rangle$ is inconsistent and $\forall f \in C, \langle T, C \setminus \{f\} \rangle$ is consistent.

From Definition 1, removing any fact $f$ from $C$ restores the consistency of $\langle T, C \rangle$. When the TBox is coherent, a conflict involves exactly two assertions. Roughly speaking, when priorities are available, restoring the consistency of a conflict comes down to ignoring the facts with the lowest level of priority.

3.2 Preferred Inclusion-Based Repair

In the flat case, one of the main strategies for handling inconsistency comes down to computing the ABox repair of an inconsistent DL-Lite KB. A repair is a maximal subbase of the ABox, denoted by MAR, that is consistent with the TBox.

Definition 2. Let $K = \langle T, A \rangle$ be an flat DL-Lite KB. A subbase $\mathcal{C} \subseteq A$ is said to be a maximal assertional-based repair, denoted MAR, of $K$ if: i) $\langle T, \mathcal{C} \rangle$ is consistent, and ii) $\forall R' \subseteq \mathcal{C}, \langle T, R' \rangle$ is inconsistent.

Example 2. Consider $T = \{A \sqsubseteq -B\}$ and $A = \{A(a), B(a), A(b)\}$. We have $\mathcal{C}(A) = \{A(a), B(a)\}$. The set of MAR is: $\mathcal{R}_1 = \{A(a), A(b)\}$ and $\mathcal{R}_2 = \{B(a), A(b)\}$.

According to the definition of MAR, adding any assertion $f$ from $A \setminus R$ to $R$ entails the inconsistency of $\langle T, R \cup \{f\} \rangle$. Moreover, the maximality in MAR is used in the sense of set inclusion. We denote by MAR($A$) the set of MAR of $A$ with respect to $T$. The definition of MAR coincides with the definition of ABox repair proposed in [Lembo et al., 2010]. A query is said to be a universal consequence (or AR-consequence [Lembo et al., 2010]) iff it can be derived from every MAR. The following definition extends the definition of MAR when the DL-Lite ABox is prioritized.

Definition 3. Let $K = \langle T, A \rangle$ be a prioritized DL-Lite KB. A preferred inclusion-based repair (PAR) $P = P_1 \cup \ldots \cup P_n$ of $A$ is such that there is no a MAR $P' = P'_1 \cup \ldots \cup P'_n$ of $S_1 \cup \ldots \cup S_n$, and an integer $i$ where:

- $P_i$ is strictly included in $P'_i$, and
- $\forall j = 1..(i - 1), P_j$ is equal to $P'_j$.

A query $q$ is said to be a PAR-consequence of $K$, denoted $K \models_{PAR} q$, iff $\forall P \in PAR(A), \langle T, P \rangle = \models q$ where PAR($A$) denotes the set of PAR of $A$.

\footnote{By a flat knowledge base, we mean a base where all the assertions have the same priority.}
\footnote{\models denotes the standard entailment used from flat and consistent DL-Lite KB [Calvanese et al., 2007].}
This definition of PAR has been largely used in a propositional logic setting (e.g. [Brewka, 1989; Benferhat et al., 1998]) and has been recently used in a DL-Lite framework [Bienvenu et al., 2014]. Definition 3 states that a query $q$ is a universal consequence iff it can be deduced from every preferred inclusion-based repair. Note that the PAR-entailment extends the definition of AR-entailment proposed in [Lembo et al., 2010] when the ABox is prioritized. A PAR of $A$ is obtained by first computing the MAR of $S_1$, then enlarging this MAR as much as possible by assertions of $S_2$ while preserving consistency, and so on.

**Example 3.** Consider $T=\{A\sqcap\neg B\}$ and $A = S_1 \cup S_2$ where $S_1 = \{A(a)\}$ and $S_2 = \{B(a), A(b)\}$. There is exactly one PAR which is: $P_1 = \{A(a), A(b)\}$.

Priorities reduce the number of MAR as one can see in Example 3 in comparison with Example 2. Indeed, within a prioritized setting, the notion of PAR operates as a selection function among possible MAR. An important feature in restoring consistency in DL-Lite, when the ABox is layered, is that when there is no conflict in $A$ involving two assertions having the same priority level, then there exists only one PAR.

**Proposition 1.** Let $K=\langle T, A \rangle$ be a prioritized DL-Lite. Let $C(A)$ be the set of conflicts in $A$. Then if $\forall \subseteq (f, g) \in C(A)$ we have $f \in S_i, g \in S_j$ and $i \neq j$ then there exits exactly one PAR.

When a conflict involves two assertions having the same priority level, restoring consistency often leads to several PAR.

### 3.3 Lexicographic Preferred-Based Repair

This subsection rewrites the cardinality-based or lexicographic inference or prioritized removed set repair, defined in [Benferhat et al., 2014], to the context of inconsistency handling. The lexicographic inference has been widely used in the propositional setting (e.g. [Benferhat et al., 1998]). In fact, one of the major problems of PAR-entailment is the large number of PAR that can be computed from an inconsistent DL-Lite KB. In order to better choose a PAR, one can follow a lexicographic-based approach. We introduce a preferred lexicographic-based repair which is based on the cardinality criterion instead of the set inclusion criterion.

**Definition 4.** Let $PAR(A)$ be the set of PAR of $A$. Then $\mathcal{L} = L_1 \cup \ldots \cup L_n$ is said to be a lexicographical preferred-based repair, denoted by $PAR_{lex}$, iff:

i) $\forall \mathcal{P} = P_1 \cup \ldots \cup P_n \in PAR(A)$: there is no $i$ s.t $|P_i| > |L_i|$,  
ii) $\forall j < i, |P_j| = |L_j|$,  

where $|X|$ is the cardinality of the set $X$.

A query $q$ is said to be $Lex$-consequence of $K$, denoted by $K \models_{lex} q$, iff $\forall \subseteq \in PAR_{lex}(A) \langle T, L \rangle \models q$ where $PAR_{lex}(A)$ is the set of $PAR_{lex}$ of $A$.

Clearly, using a lexicographic-based approach comes down to select among the set of repairs in $PAR(A)$ the ones having the maximal number of elements. We propose in the two next subsections inconsistency-tolerant inferences based only on selecting one preferred repair.

### 3.4 Possibilistic-Based Repair

One of the interesting aspects of possibilistic KBs, and more generally weighted KBs, is the ability of reasoning with partially inconsistent knowledge [Dubois and Prade, 1991]. As shown in [Benferhat and Bouraoui, 2013], the entailment in possibilistic DL-Lite, an adaptation of DL-Lite entailment within a possibility theory setting, is based on the selection of one consistent, but not necessarily maximal, subbase of $K$. This subbase is induced by a level of priority called the inconsistency degree. The following definition reformulates the definition of inconsistency degree to fit the case where $A$ is prioritized.

**Definition 5.** Let $K=\langle T, A \rangle$ be an inconsistent prioritized DL-Lite KB.

- The inconsistency degree of $K$, denoted $Inc(K)$, is defined as follows: $Inc(K)=i+1$ iff $\langle T, S_1 \cup \ldots \cup S_i \rangle$ is consistent and $\langle T, S_1 \cup \ldots \cup S_{i+1} \rangle$ is inconsistent.
- A query $q$ is said to be a $\pi$-consequence of $K$, denoted $K \models_{\pi} q$, iff $\langle T, \pi(A) \rangle \models q$ where $\pi(A)$ is the repair of $A$ defined by $\pi(A)=S_1 \cup \ldots \cup S_{Inc(K)-1}$.

The subbase $\pi(A)$ is made of the assertions having priority levels that are strictly less than $Inc(K)$. If $K$ is consistent then we simply let $\pi(A)=A$. The $\pi$-entailment is cautious in the sense that assertions from $A\setminus \pi(A)$ that are not involved in any conflict are inhibited because of their low priority levels.

### 3.5 Linear-Based Repair

One way to recover the inhibited assertions by the possibilistic entailment is to define the linear-based repair from $A$. The following definition introduces the notion of linear subset. Linear entailment has been used in a propositional logic setting in [Nebel, 1994] and has been applied for a DL setting (e.g. [Qi et al., 2011]).

**Definition 6.** Let $K=\langle T, A \rangle$ be a prioritized DL-Lite KB. The linear-based repair of $A$, denoted $\ell(A)$, is defined as follows:

i) For $i=1$: $\ell(S_1)=S_1$ if $\langle T, S_1 \rangle$ is consistent. Otherwise $\ell(S_1)=\emptyset$.  
ii) For $i>1$: $\ell(S_1 \cup \ldots \cup S_i)=\ell(S_1 \cup \ldots \cup S_{i-1}) \cup S_i$ if $\langle T, \ell(S_1 \cup \ldots \cup S_{i-1}) \cup S_i \rangle$ is consistent. Otherwise $\ell(S_1 \cup \ldots \cup S_i)=\ell(S_1 \cup \ldots \cup S_{i-1})$.

A query $q$ is a linear consequence ($\ell$-consequence) from $K$, denoted $K \models_{\ell} q$, iff $\langle T, \ell(A) \rangle \models q$.

Clearly, $\ell(A)$ is obtained by discarding a layer $S_i$ when its facts conflict with the ones involved in the previous layer. The subbase $\ell(A)$ is unique and consistent with $T$. The following proposition gives the complexity of $\pi$-entailment and $\ell$-entailment which are in $P$.

**Proposition 2.** The computational complexity of $\pi$-entailment is in $O(cons)$ where cons is the complexity of consistency checking of standard DL-Lite. The complexity of $\ell$-entailment is in $O(n \times cons)$ where $n$ is the number of strata in the KB.

The $\ell$-entailment is more productive than $\pi$-entailment, but incomparable with $PAR$-entailment and $Lex$-entailment. However from Definitions 5 and 6, both $\pi(A)$ and $\ell(A)$ are not guaranteed to be maximal.
4 Sensitivity to the Prioritized Closure

Before presenting new strategies that only select one preferred repair, we briefly introduce the concept of a prioritized closure and check which among existing approaches is sensitive to the use of the deductive closure. In fact, the inference relations given in the previous section can be either defined on \( \langle T, A \rangle \) or on \( \langle T, cl(A) \rangle \) where \( cl \) denotes the deductive closure of a set of assertions. Let us first define the notion of a deductive closure in DL-Lite.

Definition 7. Let \( K = \langle T, A \rangle \) be a flat DL-Lite KB. Let \( T_p \) be the set of all positive inclusion axioms of \( T \). We define the deductive closure of \( A \) with respect to \( T \) as follows: \( cl(A) = \{ B(a) : (T_p, A) = B(a) \} \) where, \( B \) is a concept of \( T \) and \( a \) is an individual of \( A \) if \( \{ R(a, b) : (T_p, A) = R(a, b) \} \) is a role of \( T \) and \( a, b \) are individuals of \( A \).

The use of a deductive closure of an ABox fully makes sense in DL languages, while for instance in propositional logic the closure of an inconsistent KB trivially leads to produce the whole language. The following definition extends Definition 7 to the prioritized case.

Definition 8. Let \( K = \langle T, A \rangle \) be a prioritized DL-Lite KB. We define the prioritized closure of \( A \) with respect to \( T \), simply denoted \( cl(A) \), as follows: \( cl(A) = S_1^1 \cup \ldots \cup S_n^\ell \) where:

\[
S_1^1 = cl(S_1),
\]

\[
\forall i = 2, \ldots, n : S_i^\ell = cl(S_1 \cup \ldots \cup S_i \setminus (S_1 \cup \ldots \cup S_{i-1})).
\]

Example 4. Consider \( T = \{ A \sqsubseteq B, B \sqsubseteq C, C \sqsubseteq \neg D \} \) and \( A = S_1 \cup S_2 \) where \( S_1 = \{ A(a), D(a) \} \) and \( S_2 = \{ B(b) \} \). Using Definition 8, we have \( cl(A) = S_1^1 \cup S_2^1 \) where \( S_1^1 = \{ A(a), B(a), C(a), D(a) \} \) and \( S_2^1 = \{ B(b), C(b) \} \).

An important feature of \( \pi \)-inference and \( \ell \)-inference is that they are insensitive to the deductive closure. This is not the case with \( PAR \)-entailment or \( Lex \)-entailment, more precisely:

Proposition 3. Let \( K = \langle T, A \rangle \) be a prioritized DL-Lite KB. Then \( \forall q \vdash \langle T, cl(A) \rangle \models q \iff \langle T, A \rangle \models q \) and \( \forall q \vdash \langle T, cl(A) \rangle \models q \iff \langle T, A \rangle \models q \).

Proposition 4 (Counterexample for \( PAR \)-entailment). Let \( T = \{ E \} \) and \( A = S_1 \cup S_2 \) where \( S_1 = \{ A(a), B(a) \} \) and \( S_2 = \{ E(a) \} \). We have \( P_1 = \{ A(a) \} \) and \( P_2 = \{ B(a), E(a) \} \). Consider now the deductive closure: we have \( cl(S_1) = \{ A(a), B(a), D(a) \} \) and \( S_1 \cup S_2 = \{ E(a) \} \). We also have: \( P_1 = \{ A(a), D(a) \} \) and \( P_2 = \{ B(a), D(a) \} \). One can check that \( D(a) \) is a \( PAR \)-entailment of \( \langle T, cl(A) \rangle \) while it does not follow from \( \langle T, A \rangle \) and \( E(a) \vee A(a) \) is a \( PAR \)-entailment of \( \langle T, A \rangle \) while it does not follow from \( \langle T, cl(A) \rangle \).

Example 6 (Counterexample \( Lex \)-entailment). Let us consider the following cases:

1. \( T = \{ A \sqsubseteq B, A \sqsubseteq C \} \) and \( A = S_1 = \{ A(a), B(a) \} \). We have \( \langle T, A \rangle \not\models \forall x \exists z C(x) \) and \( \langle T, cl(A) \rangle \models \forall x \exists z C(x) \).

2. \( T = \{ A \sqsubseteq B, B \sqsubseteq F, F \sqsubseteq A, C \sqsubseteq \neg B \} \) and \( S_1 = \{ A(a), B(a) \} \) and \( S_2 = \{ C(a) \} \). We only have a lexicographic subbase of \( \langle T, cl(A) \rangle \) which is \( cl(S_1) = \{ A(a), B(a), F(a) \} \) and \( cl(S_2) = \{ C(a) \} \). We also have one lexicographic subbase of \( \langle T, cl(A) \rangle \) which is \( cl(S_1) = \{ B(a), F(a) \} \).

5 New Strategies for Selecting One Repair

This section presents new strategies that only select one preferred repair. Selecting only one repair is important since it allows efficient query answering once the preferred repair is computed. These strategies are based on the so-called non-defeated entailment, described in the next section, by adding different criteria: deductive closure, cardinality, consistency and priorities.

5.1 Non-Defeated Repair

One way to get one preferred repair is to iteratively apply, layer by layer, the intersection of maximally assertional-based repairs (i.e. \( MAR \)).

Definition 9. Let \( K = \langle T, A \rangle \) be a prioritized DL-Lite KB. We define the non-defeated repair, denoted by \( ndl(A) = S'_1 \cup \ldots \cup S'_n \) as follows:

\[
\forall i = 1, \ldots, n : S'_i = \bigcap_{R_i \in \text{MAR}(S'_1 \cup \ldots \cup S'_n)} R_i
\]

A query \( q \) is a non-defeated consequence (\( nd \)-consequence) of \( K \), denoted \( K \models nd q \iff \langle T, ndl(A) \rangle \models q \).

As will be shown below, the non-defeated entailment corresponds to the definition of non-defeated subbase proposed in [Benferhat et al., 1998] within a propositional logic setting. However, contrarily to the propositional setting i) the non-defeated repair can be applied on \( A \) or its deductive closure \( cl(A) \) which leads to two different inference relations, ii) the non-defeated repair is computed in polynomial time in a DL-Lite setting while its computation is hard in a propositional logic setting. Let us now rephrase non-defeated repair (Equation 1) using the concept of free inference. First, we recall the notion of non-conflicting or free elements.

Definition 10. Let \( K = \langle T, A \rangle \) be DL-Lite KB. An assertion \( f \in A \) is said to be free iff \( \forall C \in C(A) : f \notin C \).

Intuitively, \emph{free} assertions are those assertions that are not involved in any conflict. Let \( S \in A \) be a set of assertions, we denote by \( free(S) \) the set of \emph{free} assertions in \( S \). The notions of \emph{free} elements were originally proposed in [Benferhat et al., 1992] in a propositional logic setting. The definition of \emph{free}-entailment is also equivalent to the \( MAR \)-entailment given in [Lembo et al., 2010] for flat DL-Lite KBs. The following proposition shows that the notion of \( free(A) \), extended to the prioritized case, leads to a non-defeated repair.

Proposition 4. The non-defeated repair, given in Definition 9, is equivalent to:

\[
ndl(A) = \text{free}(S_1) \cup \text{free}(S_1 \cup S_2) \cup \ldots \cup \text{free}(S_1 \cup \ldots \cup S_n)
\]

where \( \forall i : \text{free}(S_1 \cup \ldots \cup S_i) \) denotes the set of \emph{free} assertions in \( (S_1 \cup \ldots \cup S_i) \).

The following proposition shows that the non-defeated repair is consistent and its computation is in \( P \).
5.3 Combining Linear Entailment and Non-Defeated entailment: Adding consistency

We present a new way to select a single preferred assertion-based repair. It consists in slightly improving both linear entailment and non-defeated entailment, where rather to ignore a flat stratum, in case of inconsistency, one can only ignore conflicting elements. The linear-based non-defeated repair, denoted by $\ell\text{nd}(A)$, is given by Algorithm 1.

Example 8. Let $T = \{A \models B, \neg B \models C\}$ and $A = S_1 \cup S_2$ where $S_1 = \{A(a)\}$ and $S_2 = \{B(a)\}$. We have $\ell\text{nd}(A) = \{A(a), C(b), \{A(a), C(b)\}\}$. Hence $C(a)$ is an $\ell$-consequence of $\langle T, A \rangle$ but it is not an $\ell\text{nd}$-consequence of $\langle T, A \rangle$.

5.4 Introducing Consistency and Cardinality Criterion

A natural question is whether one can introduce a cardinality criterion, instead of set inclusion criterion, in the definition of non-defeated repair given by Equation 1. Namely, we define the cardinality-based non-defeated repair as follows:

**Definition 11.** Let $K = \langle T, A \rangle$ be a prioritized DL-Lite KB. The cardinality-based non-defeated repair, denoted by $nd(A)$, is defined as follows:

$$\forall i = 1, \ldots, n : S'_i = \bigcup_{R \in \text{MAR}_{\text{card}}(S_1 \cup \ldots \cup S_i)} R \quad (3)$$

where $\text{MAR}_{\text{card}}(S) = \{R : R \in \text{MAR}(S) \text{ and } |R' > |R|\}$.

One main advantage of this approach is that it produces more conclusions that the standard non-defeated inference relation. Namely, $nd(A)$ and $nd(A)$ are respectively given by Equations 1 and 3. The converse is false:

Example 9. Consider $T = \{A \models B, \neg B \models C\}$ and $A = S_1 \cup S_2$ where $S_1 = \{A(a), B(a)\}$ and $S_2 = \{C(a)\}$. We have $nd(A) = \emptyset$ while $nd(A)$ is not an $\ell\text{nd}$-consequence of $\langle T, A \rangle$.

One way to overcome such limitation is to only select $\text{MAR}_{\text{card}}(S_1 \cup \ldots \cup S_i)$ that are consistent with $S'_i \cup \ldots \cup S'_i$. Namely: $S'_i = \bigcap_{R \in \text{MAR}_{\text{card}}(S_1 \cup \ldots \cup S_i)} R \quad (4)$

Contrary to $nd(A)$, $\text{MAR}_{\text{card}}(A)$ is always consistent with $T$. Clearly, $\text{MAR}_{\text{card}}(A)$ is consistent with $T$.

5.5 Adding Priorities

In the definition of $\ell\text{nd}$-inference, given by Equation 1, a flat notion of $\text{MAR}$ (maximally inclusion-based repair) has been used. A natural way to extend the $\ell\text{nd}$-inference is to use a prioritized version of $\text{MAR}$, namely:

**Definition 13.** Let $K = \langle T, A \rangle$ be a prioritized DL-Lite KB. We define the prioritized inclusion-based non-defeated repair, denoted by $\text{pind}(A)$, as follows:

$$\forall i = 1, \ldots, n : S'_i = \bigcup_{P \in \text{PAR}(S_1 \cup \ldots \cup S_i)} P \quad (4)$$

where $\text{PAR}(S) = \{P : P \in \text{PAR}(S) \text{ and } |P' > |P|\}$.
The following proposition shows that there is no need to consider all \( S_i \) for \( i < n \) when computing \( pind(A) \), namely:

**Proposition 7.** Let \( K=\langle T, A \rangle \) be a prioritized DL-Lite KB. Then \( pind(A) = \bigcap\{P \in PAR(S_1 \cup \ldots \cup S_n) : P \} \).

Besides, a cardinality-based version of Equation 4, denoted by \( pind(A)_{\leq} = S_1 \cup \ldots \cup S_n \), can be defined as follows:

\[
\forall i = 1, \ldots, n : S_i = \bigcap_{L \in PAR_{\leq}(S_1 \cup \ldots \cup S_n)} \mathcal{L}
\]

Lastly, both \( pind(A) \) and \( pind(A)_{\leq} \) can be defined on \( cl(A) \) instead of \( A \) or be defined on closed repairs instead of repairs themselves. This leads to new inferences strategies that only select one preferred subbase.

6 Comparative Analysis and Experimental Evaluation

From a computational complexity point of view, \( \pi \)-entailment, \( \ell \)-entailment, \( nd \)-entailment and \( \ell nd \)-entailment and the entailments based on their closures, are the most promising ones since both computing the repair and query answering are tractable. For other strategies based on the \( nd \)-inference, computing the repairs is a hard task, but it is done \( ONCE \). Answering queries, when the single repair is computed, is efficiently computed since it has the same complexity as in standard DL-Lite.

![Figure 1: Relationships between inferences where \( n1 \rightarrow n2 \) means that each conclusion that can be universally derived from repairs in \( n1 \) is also a conclusion using repairs in \( n2 \).](image)

From productivity, Figure 1 summarizes the relationships between main entailments considered in the paper when the ABox is prioritized. Note that for the sake of simplicity, we do not make reference to inferences defined on \( cl(A) \). From Figure 1, \( \pi \)-entailment is the most cautious relation. Adding priorities, cardinality and consistency to the definition of \( nd \)-entailment allow to provide more productive inference relations. However \( \ell \)-entailment remains incomparable with the \( nd \)-entailment, since layers including non free assertions can be present in \( \ell(A) \). Moreover, \( \ell nd(A) \) is incomparable with other approaches. Within the prioritized setting, \( nd(A) \) plays the same role with respect to \( PAR \) as \( free(A) \) for \( MAR \) in the flat case. As a consequence, each \( nd \)-consequence of \( A \) is also a \( PAR \)-consequence of \( A \). The converse is false. Moreover, it is well-known that each \( PAR \)-entailment is also a \( Lex \)-entailment and the converse is false, since the \( Lex \)-entailment only uses subsets of prioritized repair (PAR).

We now provide an experimental evaluation where we considered a TBox containing 100 negative inclusion axioms with a proportion of conflicts at least equal to 1/5 per assertion. This TBox is adapted from the DL-Lite\( R \) university benchmark proposed in [Lutz et al., 2013]. We use the Extended University Data Generator (EUDG) \(^4\) to generate the ABox assertions. Once the ABox is produced, we fit it to our setting using 4 strata until 7 strata. Moreover the computation of conflicts is performed layer per layer. Note that the time used for computing the conflicts is not included in the time used for computing the repairs, since this is done in a polynomial time. Said differently, computing conflicts is negligible with respect to computing repairs.

Table 1a gives the experimental results of the computation of \( MAR \) and \( MAR_{\text{card}} \). One can see that using the cardinality criterion instead of the set inclusion one refines the result and improves the computation time of the repairs. Moreover, an important influential parameter when computing the repairs is the number of occurrences of an assertion in conflicts. Namely, the more an assertion is recurring in conflicts the more the conflict resolution has better chances to be achieved. For instance, in Table 1a considering the case of 37 conflicts, by increasing the percentage of occurrences of some assertions in conflicts, we obtain 23082 \( MAR \) in 136ms instead of 16815986 in 206089ms. In such case, the number of \( Lex \) decreases also where we compute only 24 \#\( MAR_{\text{card}} \) having cardinality equal to 14 assertions. Similar results on the effect of the number of occurrences of assertions in conflicts are provided [Pivert and Prade, 2010; D.Deagustini et al., 2014].

Now, concerning \( PAR_{\leq} \), we also use the notion of minimal inconsistent subsets where the minimality refers to a lexicographic ordering. Table 1b gives the results on the computation of \( PAR_{\leq} \) and the main repairs given in this paper. One can first observe that given an ABox \( A \) whatever is its size, computing \( \pi \) or \( \ell \) does not need long computation time as needed by inconsistency checking. Regarding now the computation of the non-defeated repair, it depends on the number of conflicts in the ABox. Another parameter that also influences the results is the number of layers. This can be clearly seen when computing \#\( PAR_{\leq} \). Indeed, the number of \( PAR_{\leq} \) decreases as the number of layers increases. Clearly, more the stratification of the ABox is important more the conflicts resolution has better chances to be achieved.

7 Conclusion

This paper focuses on how to produce a single preferred repair from a prioritized inconsistent DL-Lite KB based on the notion of the non-defeated inference relation. We first reviewed some well-known approaches that select one repair (such as possibilistic repair or linear-based repair) or several repairs (such as preferred inclusion-based repairs or lexicographic-based repairs). Then, we presented different

\(^4\) available at https://code.google.com/p/combo-obda/
strategies for selecting one preferred repair. These strategies have as starting point the non-defeated repair and mainly add one/several of the four main criteria: priorities, deductive closure, cardinality and consistency.

### Acknowledgment
This work has been supported by the French National Research Agency. ASPIQ project ANR-12-BS02-0003

### References


### Table 1: Experimental evaluation of main inferences proposed in this paper.

<table>
<thead>
<tr>
<th># Conflicts</th>
<th># MAR</th>
<th>time # MAR</th>
<th># MAR&lt;sub&gt;card&lt;/sub&gt;</th>
<th>time # MAR&lt;sub&gt;card&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>28080</td>
<td>105ms</td>
<td>192</td>
<td>65ms</td>
</tr>
<tr>
<td>25</td>
<td>688128</td>
<td>2268ms</td>
<td>256</td>
<td>789ms</td>
</tr>
<tr>
<td>37</td>
<td>1681598</td>
<td>206089ms</td>
<td>56</td>
<td>5422ms</td>
</tr>
<tr>
<td>75</td>
<td>20160000</td>
<td>272830ms</td>
<td>96</td>
<td>216236ms</td>
</tr>
<tr>
<td>105</td>
<td>-</td>
<td>Time-out</td>
<td>2034</td>
<td>8259s</td>
</tr>
</tbody>
</table>

(a) Number of conflicts, number of MAR, time taken to compute MAR in ms (milliseconds) or s (seconds), number of #MAR<sub>card</sub>, time taken to compute #MAR<sub>card</sub>.

<table>
<thead>
<tr>
<th># Conflicts</th>
<th># Strata</th>
<th>time π</th>
<th>time ℓ</th>
<th>time nd</th>
<th># PAR&lt;sub&gt;lex&lt;/sub&gt;</th>
<th>time PAR&lt;sub&gt;lex&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
<td>4</td>
<td>4ms</td>
<td>7ms</td>
<td>16</td>
<td>17ms</td>
<td>17ms</td>
</tr>
<tr>
<td>123</td>
<td>4</td>
<td>5ms</td>
<td>8ms</td>
<td>6ms</td>
<td>2</td>
<td>11ms</td>
</tr>
<tr>
<td>502</td>
<td>4</td>
<td>5ms</td>
<td>9ms</td>
<td>24ms</td>
<td>2024</td>
<td>1072ms</td>
</tr>
<tr>
<td>1562</td>
<td>4</td>
<td>4ms</td>
<td>8ms</td>
<td>129ms</td>
<td>1392</td>
<td>128/475</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5ms</td>
<td>8ms</td>
<td>64ms</td>
<td>232</td>
<td>34/52s</td>
</tr>
</tbody>
</table>

(b) Number of conflicts, number of strata, time taken to compute π, ℓ, nd and #PAR<sub>lex</sub> and number of computed PAR<sub>lex</sub>.

---
