Revision of DL-Lite knowledge bases

Author
Wang, Zhe, Wang, Kewen, Topor, Rodney

Published
2009

Conference Title
Proceedings of the 22nd International Workshop on Description Logics (DL 2009)

Copyright Statement
Copyright remains with the authors 2009. The attached file is posted here with permission of the copyright owners for your personal use only. No further distribution permitted. For information about this conference please refer to the publisher's website or contact the authors.

Downloaded from
http://hdl.handle.net/10072/29805

Link to published version
http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-477/
Revision of DL-Lite Knowledge Bases

Zhe Wang, Kewen Wang, and Rodney Topor
Griffith University, Australia

Abstract. We address the revision problem for knowledge bases (KBs) in Description Logics (DLs). This problem has received much attention in the ontology management and DL communities, but the existing proposals are restricted in several ways. In this paper we develop a formal framework for revision of DL-Lite KBs, using techniques that are analogous to those for model-based revision in propositional logic. However, unlike propositional logic, a DL-Lite KB can have infinitely many models, which makes it hard to define and compute revision in terms of models. For this reason, we first develop an alternative semantic characterization for DL-Lite by introducing the concept of a feature and then define a specific revision operator for DL-Lite KBs based on features (instead of models). In contrast to previous approaches, we tackle the problem of revision between KBs, and the result of revision is always an unique DL-Lite KB. We also present an algorithm for computing KB revisions in DL-Lite.

1 Introduction

An ontology is a formal description of a common vocabulary for a domain of interest and the relationships between terms built from the vocabulary. Ontologies have been applied in a wide range of practical domains such as medical informatics, bio-informatics and, more recently, the Semantic Web. Description Logics (DLs) have been used as underlying formalisms of most ontology languages recently. A DL-based ontology is usually expressed as a DL-knowledge base (KB), which consists of a TBox (terminological box) and an ABox (set of assertions).

DL-Lite [1, 3] is specifically designed as a family of lightweight DLs with relatively restricted expressive power but efficient ontological reasoning algorithms. In particular, logics in the DL-Lite family have polynomial time computational complexity with respect to standard reasoning tasks, and LogSpace data complexity with respect to complex query answering. With reasonable expressive power and efficient reasoning support, especially on answering complex queries over large amounts of data, the DL-Lite family has proved to be an appealing formalism for ontology development.

A challenge in ontology maintenance is that ontologies are not static but may evolve over time. Because of limitations in the initial design, changes in users’ requirements, need to reflect changes in the real world, and so on, there is a need to revise ontologies in accordance with the new knowledge. Recently, intensive interest has been shown on revision/update in DLs. Earlier approaches concentrated on adapting revision postulates for propositional logic to DLs but no specific revision operators were provided, e.g. [4, 5]. Recently, there have been several attempts to define specific revision/update operators for DLs including [7, 9, 11–16]. In particular, an update operator for DL-Lite
ABoxes is defined in [7]. In fact, most specific revision/update operators for DLs can deal only with ABoxes, like [7, 9, 11], or only with TBoxes, like [13, 14]. As shown in [8, 14, 16], the kernel revision operator for propositional logic can be adapted to most DLs in a straightforward way, but the result of revision is not an unique KB in most cases and thus arbitrary choices by the user are required.

It is worth mentioning that update is traditionally distinguished from revision in that update addresses changes of the actual state of the world (e.g., that resulting from an action) whereas revision addresses the incorporation of new knowledge about the world. In this paper, we focus on the revision problem.

In propositional logic, the area of belief revision [6] is well developed, with numerous proposals and techniques for knowledge base revision. Among them, Satoh’s revision operator $\varphi *_s \mu$ [17] is well known. In his approach, the set of models of a revised knowledge base $K'$ are those models of $K$ that are ‘closest’ to those of the initial knowledge base $K$, where the distance between two models is based on their symmetric difference.

However, such model-based approaches cannot be applied directly to KB revision in DLs. The difficulties in defining and computing model-based revision of KBs in DLs are that: (1) unlike a propositional theory, a KB in a DL may have infinitely many models; (2) a model of a KB can be infinite; and (3) given a collection $M$ of models, there may not exist an unique KB $K$ such that $M$ is exactly the set of all models of $K$.

In this paper, we aim to define a rational revision operator for DL-Lite$_{bool}$ ontologies in a way analogous to the model-based approaches in propositional logic. In order to achieve this goal, we first define the concept of features for DL-Lite$_{bool}$, which precisely capture most important semantic properties of DL-Lite$_{bool}$ KBs, and are always finite. We adapt the techniques of model-based revision in propositional logic to the revision of DL-Lite$_{bool}$ KBs, and define a specific revision operator based on the distance between features. We also show that the revision operator possesses desirable properties, and can be computed through syntactic methods.

Although our approach is based on DL-Lite$_{bool}$ [1], we remark that the belief revision method we propose also applies to other logics in DL-Lite. In contrast to previous approaches, we address the problem of defining and computing revision of a DL-Lite KB by another KB. Also, the result of our revision is always an unique KB in DL-Lite.

Due to space limitation, proofs are omitted in this paper but a longer version of this paper with complete proofs can be found at http://www.cit.gu.edu.au/~kewen/revision_long.pdf.

2 DL-Lite Family

A signature is a finite set $S = S_C \cup S_R \cup S_O \cup S_N$ where $S_C$ is the set of atomic concepts, $S_R$ is the set of atomic roles, $S_O$ is the set of individual names (or, objects) and $S_N$ is the set of natural numbers in $S$. We assume $1$ is always in $S_N$, $\top$ and $\bot$ will not be considered as atomic concepts or atomic roles. Formally, given a signature $S$, a
DL-Lite$_{bool}$ language $\mathcal{L}$ has the following syntax:

\[
\begin{align*}
R & \leftarrow P \mid P^\shortmid \\
B & \leftarrow \top \mid \top \mid A \mid \geq n R \\
C & \leftarrow B \mid \lnot C \mid C_1 \cap C_2 \\
\end{align*}
\]

where $n \geq 1$, $A \in S_C$ and $P \in S_R$. $B$ is called a basic concept and $C$ is called a general concept. We write $\exists R$ as a shorthand for $\geq 1 R$, and let $R^+ = P$, where $P \in S_R$, whenever $R = P$ or $R = P^\shortmid$.

A literal is a basic concept or its negation. The set of all literals on $S$ is denoted by $\text{Lit}_S$, and the set of all basic concepts is denoted by $\text{Lit}^+_S$.

A DL-Lite$_{bool}$ TBox $T$ is a finite set of concept inclusions of the form $C_1 \sqsubseteq C_2$, where $C_1$ and $C_2$ are general concepts. A DL-Lite$_{bool}$ ABox $A$ is a finite set of membership assertions of the form $C(a)$, $R(a,b)$, where $a$, $b$ are individual names. We call $C(a)$ a concept assertion and $R(a,b)$ a role assertion. A DL-Lite$_{bool}$ knowledge base (KB) is a pair $\mathcal{K} = (T, A)$.

The semantics of a DL-Lite KB is given by interpretations. An interpretation $I$ is a pair $(\Delta^I, \mathcal{I})$, where $\Delta^I$ is a non-empty set called the domain and $\mathcal{I}$ is an interpretation function which associates each atomic concept $A$ with a subset $A^I$ of $\Delta^I$, each atomic role $R$ with a binary relation $P^I \subseteq \Delta^I \times \Delta^I$, and each individual name $a$ with an element $a^I$ of $\Delta^I$ such that $a^I \neq b^I$ for each pair of individual names $a$, $b$ (unique name assumption).

The interpretation function $\mathcal{I}$ can be extended to general concept descriptions:

\[
\begin{align*}
(P^\shortmid)^I & = \{ (a^I, b^I) \mid (b^I, a^I) \in P^I \} \\
(\geq n R)^I & = \{ a^I \mid | \{ b^I \mid (a^I, b^I) \in R^I \} | \geq n \} \\
(\lnot C)^I & = \Delta^I \setminus C^I \\
(C_1 \cap C_2)^I & = C_1^I \cap C_2^I \\
\end{align*}
\]

An interpretation $I$ is a model of inclusion $C_1 \sqsubseteq C_2$ if $C_1^I \subseteq C_2^I$; $I$ is a model of an assertion $C(a)$ if $a^I \in C^I$; $I$ is a model of an assertion $R(a,b)$ if $(a^I, b^I) \in R^I$. $I$ is a model of a TBox $T$ (or ABox $A$) if $I$ is a model of each inclusion in $T$ (resp., each assertion in $A$). Two inclusions (or resp., assertions, TBoxes, ABoxes) are said to be equivalent if they have exactly the same models. $I$ is a model of a KB $(T, A)$, if $I$ is a model of both $T$ and $A$. A KB $\mathcal{K}$ is consistent if it has at least one model. Two KBs $\mathcal{K}_1$, $\mathcal{K}_2$ that have the same models are said to be equivalent, denoted $\mathcal{K}_1 \equiv \mathcal{K}_2$. A KB $\mathcal{K}$ logically implies an inclusion or assertion $\alpha$, denoted $\mathcal{K} \models \alpha$, if all models of $\mathcal{K}$ are also models of $\alpha$. We will use $\text{Mod}(\mathcal{K})$ to denote the set of models of a KB $\mathcal{K}$. $\text{Sig}(\mathcal{K})$ is the signature of $\mathcal{K}$.

As in propositional logic, a DL-Lite$_{bool}$ concept $C$ can always be transformed into an equivalent disjunction of conjunctions of literal concepts (DNF).

## 3 Features in DL-Lite$_{bool}$

In this section, we introduce the concept of features in DL-Lite$_{bool}$, which provides an alternative semantic characterization for DL-Lite$_{bool}$. An advantage of semantic features over models is that the number of all features for a DL-Lite$_{bool}$ knowledge base
is finite and each feature is finite as we will see later. These finiteness properties make it possible to recast key approaches to revision for classical propositional logic into DL-Lite$_\text{bool}$.

Features for DL-Lite$_\text{bool}$ are based on the notion of types defined in [10].

In the following sections, we assume $S$ is a fixed signature. Recall that a $S$-literal (or, literal) is a basic concept on $S$. It is possible to recast key approaches to revision for classical propositional logic into $S$-literals. For example, let $S$ consists of only its basic concepts, following similar common practices for propositional models. For example, let $S_C = \{ A, B \}$, $S_R = \{ P \}$, and $S_N = \{ 1, 3 \}$. Then $\tau = \{ A, \exists P, \geq 3 P, \exists P^- \}$ is a type instead of $\tau = \{ T, A, \neg B, \exists P, \geq 3 P, \exists P^-; \neg(\geq 3 P^-) \}$.

A $S$-type is a subset $\tau$ of $\text{Lit}_S$ containing $\top$ and satisfying the following conditions:

- for every $S$-literal $L$, $L \in \tau$ iff $\neg L \not\in \tau$;
- for any $m, n \in S_N$ with $m < n$, $\geq n R \in \tau$ implies $\geq m R \in \tau$;
- for any $m, n \in S_N$ with $m < n$, $\neg(\geq m R) \in \tau$ implies $\neg(\geq n R) \in \tau$.

A type $\tau$ can be equivalently represented by the conjunction of all literals in $\tau$, which we also call a type when the context is clear and no confusion is caused. An important observation is that every general concept $C$ over $S$ corresponds to a set $\text{Ts}(C)$ of $S$-types, in light of the fact that $C$ can be transformed into an equivalent disjunction of $S$-types. For example, concept $A \cap (B \cup \exists P)$ can be transformed into $(A \cap B \cap \exists P) \cup (A \cap \neg B \cap \exists P) \cup (A \cap B \cap \neg \exists P)$.

Informally speaking, types correspond to propositional models whereas $\text{Ts}(C)$ correspond to the set of models of formula $C$. For simplicity, we assume that a type $\tau$ consists of only its basic concepts, following similar common practices for propositional models. For example, let $S_C = \{ A, B \}$, $S_R = \{ P \}$, and $S_N = \{ 1, 3 \}$. Then $\tau = \{ A, \exists P, \geq 3 P, \exists P^- \}$ is a type instead of $\tau = \{ T, A, \neg B, \exists P, \geq 3 P, \exists P^-; \neg(\geq 3 P^-) \}$.

We say a type $\tau$ satisfies a concept $C$ if $\tau \in \text{Ts}(C)$, and $\tau$ satisfies concept inclusion $C_1 \sqsubseteq C_2$ if $\tau \in \text{Ts}(\neg C_1 \cup C_2)$. In this way, we define that $\tau$ satisfies a TBox $T$ if it satisfies each inclusion in $T$.

However, in order to capture the membership relations in a KB, we need to extend the notion of types and thus define Herbrand sets in DL-Lite.

**Definition 1.** A $S$-Herbrand set $\mathcal{H}$ is a finite set of assertions of the form $B(a)$ or $P(a, b)$, where $a, b \in S_O$, $B \in \text{Lit}_S^+$, $P \in S_R$, satisfying the following conditions

1. for each $a \in S_O$, if $B_1(a), \ldots, B_k(a)$ are all the concept assertions about $a$ in $\mathcal{H}$, then the set $\{ B_1, \ldots, B_k \}$ forms an $S$-type;
2. for each $P \in S_R$, if $P(a, b) (1 \leq i \leq n)$ are all the assertions in $\mathcal{H}$ of such form, then for any $m \in S_N$ with $m \leq n$, $(\geq m P)(a)$ is in $\mathcal{H}$;
3. for each $P \in S_R$, if $P(a, b) (1 \leq i \leq n)$ are all the assertions in $\mathcal{H}$ of such form, then for any $m \in S_N$ with $m \leq n$, $(\geq m P^-)(a)$ is in $\mathcal{H}$.

Informally, $B(a) \not\in \mathcal{H}$ with $B \in \text{Lit}_S^+$ means that $\neg B(a)$ holds in $\mathcal{H}$, and $P(a, b) \not\in \mathcal{H}$ with $P \in S_R$ means that $a, b$ are not related by $P$ in $\mathcal{H}$. Thus, Herbrand sets are different from ABoxes by adopting a close world assumption. It is not hard to see that, conditions 2 and 3 in Definition 1 preserve consistency of a Herbrand set. And condition 2 in Definition 1 requires that $(\geq m P) \in \mathcal{H}$ implies $(\geq m P) \in \mathcal{H}$ for all $m \in S_N$ with $m < n$. These conditions are introduced mainly for simplifying the definition of our revision operator in the following section.
For simplicity, we denote the group of all the concept assertions \( B_1(a), \ldots, B_k(a) \) about \( a \) in \( \mathcal{H} \) as \( \tau(a) \), where \( \tau = \{ B_1, \ldots, B_k \} \) is a type. We say \( \tau(a) \) is in \( \mathcal{H} \) if \( B_i(a) \) is in \( \mathcal{H} \) for all \( 1 \leq i \leq k \).

Herbrand sets can be used to capture membership relations in a KB. We say a Herbrand set \( \mathcal{H} \) satisfies a membership assertion \( C(a) \) if \( \tau(a) \) is in \( \mathcal{H} \) and \( \tau \in \text{Ts}(C) \); and \( \mathcal{H} \) satisfies assertion \( P(a,b) \) (resp., \( P^-(b,a) \)) if \( P(a,b) \) (resp., \( P^-(b,a) \)) is in \( \mathcal{H} \). \( \mathcal{H} \) satisfies an ABox \( \mathcal{A} \) if \( \mathcal{H} \) satisfies all the assertions in \( \mathcal{A} \).

To provide an alternative characterization for reasoning in a KB, we could use pairs \( \langle \tau, \mathcal{H} \rangle \), where \( \tau \) is a type and \( \mathcal{H} \) is a Herbrand set, to replace standard interpretations, such that \( \langle \tau, \mathcal{H} \rangle \) satisfies KB \( \langle T, \mathcal{A} \rangle \) if \( \tau \) satisfies \( T \) and \( \mathcal{H} \) satisfies \( \mathcal{A} \). To this end, we want it to guarantee that \( \mathcal{K} \) is consistent iff there exists some pair \( \langle \tau, \mathcal{H} \rangle \) satisfying \( \mathcal{K} \). However, the following example shows that this approach does not work.

**Example 1.** Let \( \mathcal{K} = \{ \exists P^- \subseteq \bot \} \) be a DL-Lite\(_{bool} \) KB. Obviously, \( \mathcal{K} \) is inconsistent and thus has no model. However, let \( S = \{ P, a, 1 \} \) and \( \tau = \{ \exists P \} \), then \( \langle \tau, \{ \tau(a) \} \rangle \) satisfies \( \mathcal{K} \).

The problem with using pairs \( \langle \tau, \mathcal{H} \rangle \) as alternative semantic characterization for DL-Lite is that a single type \( \tau \) is not sufficient to capture the semantic connection between the TBox and the ABox of a KB. Our investigation shows that it is plausible to use a pair of a set of types and a Herbrand set, as an alternative semantic characterization for a DL-Lite\(_{bool} \) KB. Using sets of types as semantic characterizations of DL-Lite TBoxes is also suggested in [10].

Thus, we introduce the definition of features as follows.

**Definition 2.** Given a signature \( S \), a feature on \( S \), or simply \( S \)-feature, is defined as a pair \( \mathcal{F} = \langle \Xi, \mathcal{H} \rangle \), where \( \Xi \) is a non-empty set of \( S \)-types and \( \mathcal{H} \) a \( S \)-Herbrand set, such that the following conditions are satisfied:

1. for each \( P \in \mathcal{S}_\mathcal{O} \), \( \exists P \in \bigcup \Xi \) iff \( \exists P^- \in \bigcup \Xi \);
2. for each \( a \in \mathcal{S}_\mathcal{O} \) and \( \tau(a) \) in \( \mathcal{H} \), s. t. \( \tau \) is an \( S \)-type, \( \tau \in \Xi \).

Informally, the first condition in Definition 2 says that \( \exists P \) is a satisfiable concept in \( \mathcal{F} \) if and only if \( \exists P^- \) is also a satisfiable concept. And the second condition requires that the interpretation of the ABox must be in accordance to the TBox.

We use a feature \( \mathcal{F} = \langle \Xi, \mathcal{H} \rangle \) as an alternative for an interpretation in DL-Lite\(_{bool} \) as follows.

- \( \mathcal{F} \) satisfies an inclusion \( C_1 \subseteq C_2 \) over \( \mathcal{S} \), if \( \Xi \subseteq \text{Ts}(\neg C_1 \sqcup C_2) \).
- \( \mathcal{F} \) satisfies an assertion \( C(a) \) over \( \mathcal{S} \), if \( \tau(a) \) is in \( \mathcal{H} \) and \( \tau \in \text{Ts}(C) \).
- \( \mathcal{F} \) satisfies assertion \( P(a,b) \) (resp., \( P^-(b,a) \)) over \( \mathcal{S} \), if \( P(a,b) \) is in \( \mathcal{H} \).

We say \( \mathcal{F} \) is a model feature of DL-Lite\(_{bool} \) KB \( \mathcal{K} \) if \( \mathcal{F} \) satisfies each inclusion and each assertion in \( \mathcal{K} \). \( \Omega_\mathcal{K} \) denotes the set of all model features of \( \mathcal{K} \).

From the definition of features, a model feature is finite and the number of model features of a KB is also finite.

**Example 2.** Consider the KB \( \mathcal{K} = \langle T, \mathcal{A} \rangle \), where \( T = \{ A \subseteq \exists P, B \subseteq \exists P, \exists P^- \subseteq B, A \cap B \subseteq \bot, \geq 2 P^- \subseteq \bot \} \) and \( \mathcal{A} = \{ A(a), P(a,b) \} \). It is shown in [2] that \( \mathcal{K} \) is a KB having no finite model.
An infinite model \( I \) of \( K \) is defined as follows: \( \Delta^I = \{d_a, d_b, d_1, d_2, d_3 \ldots\} \), \( a^I = d_a \) and \( b^I = d_b \); the concept \( A \) is interpreted as a singleton \( \{d_a\} \) and \( B \) as \( \{d_b, d_1, d_2, d_3 \ldots\} \); and role \( P \) is interpreted as \( \{(d_a, d_b), (d_b, d_1), (d_1, d_2), \ldots, (d_i, d_{i+1}) \ldots\} \).

Take \( S = \text{Sig}(K) = \{A, B, P, 1, 2, a, b\} \). The (finite) model feature of \( K \) that corresponds to \( I \) is \( F = (\Xi, \mathcal{H}) \) where the Herbrand set \( \mathcal{H} = \{\tau_1(a), \tau_2(b), P(a, b)\} \) and \( \Xi = \{\tau_1, \tau_2\} \) with \( \tau_1 = \{A, \exists P\} \) and \( \tau_2 = \{B, \exists P, \exists P^-\} \).

Given an inclusion or assertion \( \alpha \) over \( S = \text{Sig}(K) \), we define \( K \models f \alpha \) if for each \( F \in \mathcal{O}_K \), \( F \) satisfies \( \alpha \). Given two KBs \( K_1 \) and \( K_2 \), let \( S = \text{Sig}(K_1 \cup K_2) \), define \( K_1 \models f K_2 \) if \( \mathcal{O}_{K_1} \subseteq \mathcal{O}_{K_2} \); and \( K_1 \equiv f K_2 \) if \( \mathcal{O}_{K_1} = \mathcal{O}_{K_2} \).

The following two results show that model features do capture the semantic properties of DL-Lite\textsubscript{bool} KBs.

**Proposition 1.** Let \( K \) be a DL-Lite\textsubscript{bool} KB and \( S = \text{Sig}(K) \). Then

1. \( K \) is consistent iff \( K \) has a model feature;
2. for any concept inclusion \( C_1 \subseteq C_2 \) over \( S \), \( K \models (C_1 \subseteq C_2) \) iff \( K \models f (C_1 \subseteq C_2) \);
3. for any membership assertion \( C(a) \) (resp., \( R(a, b) \)) over \( S \), \( K \models C(a) \) iff \( K \models f C(a) \) (resp., \( K \models R(a, b) \) iff \( K \models f R(a, b) \));

**Proposition 2.** Let \( K_1, K_2 \) be two DL-Lite\textsubscript{bool} KBs and \( S = \text{Sig}(K_1 \cup K_2) \). Then

1. \( K_1 \models f K_2 \) iff \( K_1 \models f K_2 \);
2. \( K_1 \equiv f K_2 \) iff \( K_1 \equiv f K_2 \).

In summary, the advantages of model features over standard models can be described as follows. First, each model feature of a KB is finite in structure. Second, there are a finite number of model features for each KB. Third, each non-empty set of model features determines a unique KB up to KB equivalence. In particular, given a non-empty set \( \mathcal{O} \) of \( S \)-features, we can always construct an unique DL-Lite\textsubscript{bool} KB \( \overline{K} \) over \( S \) such that \( \mathcal{O} \subseteq \mathcal{O}_K \) and \( \mathcal{O}_K \) is minimal among all supersets \( \mathcal{O}_{\overline{K}} \) of \( \mathcal{O} \). In Section 5, we will introduce a method for constructing such a DL-Lite\textsubscript{bool} KB from a set of features.

## 4 Feature-based Revision

In this section, we introduce a specific revision operator for DL-Lite KBs based on model features rather than models.

Before introducing the definition of our feature-based revision operator, we first extend the definition of symmetric difference \( \triangle \) so that it can be used for model features.

Recall that \( S_1 \triangle S_2 = (S_1 - S_2) \cup (S_2 - S_1) \) for any two sets \( S_1 \) and \( S_2 \). Given two \( S \)-features \( F_1 = (\Xi_1, \mathcal{H}_1) \) and \( F_2 = (\Xi_2, \mathcal{H}_2) \), the distance between \( F_1 \) and \( F_2 \), denoted \( F_1 \triangle F_2 \), is a pair \( (\Xi_1 \triangle \Xi_2, \mathcal{H}_1 \triangle \mathcal{H}_2) \). Note that \( \mathcal{H}_1 \triangle \mathcal{H}_2 \) may not be a Herbrand set anymore. For example, let \( S = \{P, a, b_1, b_2, b_3, 1, 2, 3\} \), \( \mathcal{H}_1 = \{\exists P(a), P(a, b_1)\} \) and \( \mathcal{H}_2 = \{\exists P(a), (\geq 2 P)(a), P(a, b_2), P(a, b_3)\} \). Then we get \( \mathcal{H}_1 \triangle \mathcal{H}_2 = \{\geq 2 P(a), P(a, b_1), P(a, b_2), P(a, b_3)\} \), which is not a Herbrand set because it does not contain \( \exists P(a) \) and \( (\geq 3 P)(a) \).
To compare two distances, we could define \( F_1 \triangle F_2 \subseteq_d F_3 \triangle F_4 \) if \( \Xi_1 \triangle \Xi_2 \subseteq_d \Xi_3 \triangle \Xi_4 \) and \( H_1 \triangle H_2 \subseteq_d H_3 \triangle H_4 \), where \( F_i = \langle \Xi_i, H_i \rangle \) for \( i = 1, 2, 3, 4 \); and \( F_1 \triangle F_2 \subseteq_d F_3 \triangle F_4 \) if \( F_1 \triangle F_2 \subseteq_d F_3 \triangle F_4 \) and \( F_3 \triangle F_4 \not\subseteq_d F_1 \triangle F_2 \). However, we will show in an example later that such a measure is too weak for KB revision and cannot preserve enough information. Instead, we set a preference on Herbrand sets over type sets: \( F_1 \triangle F_2 \subseteq_d F_3 \triangle F_4 \) if

1. \( H_1 \triangle H_2 \subset H_3 \triangle H_4 \), or
2. \( H_1 \triangle H_2 = H_3 \triangle H_4 \) and \( \Xi_1 \triangle \Xi_2 \subset \Xi_3 \triangle \Xi_4 \).

Note that here we set a preference on Herbrand set over type set, such that assertions represent more specific information than inclusions and thus from the view point of knowledge representation, should have higher priority.

According to our new semantic characterization for DL-Litebool KBs, the semantics of each KB \( \mathcal{K} \) is determined by the set \( \Omega_\mathcal{K} \) of all model features of \( \mathcal{K} \).

Given two KBs \( \mathcal{K} \) and \( \mathcal{K}' \), to define our revision operator \( \mathcal{K} \circ \mathcal{K}' \), we need to specify the subset of \( \Omega_\mathcal{K} \) that is closest to \( \Omega_\mathcal{K} \).

Let \( \Omega_1 \) and \( \Omega_2 \) be two sets of features. Then we define

\[
\sigma(\Omega_1, \Omega_2) = \{ F_1 \in \Omega_1 \mid \exists F_2 \in \Omega_2 \text{ s.t. } (F_1 \triangle F_2) \subseteq_d (F_1 \triangle F_2) \text{ for all } F_1 \in \Omega_1 \text{ and } F_2 \in \Omega_2 \}
\]

Now we are ready to define our revision operator in terms of model features.

**Definition 3.** Let \( \mathcal{K}, \mathcal{K}' \) be two DL-Litebool KBs and \( \mathcal{S} = \text{Sig}(\mathcal{K} \cup \mathcal{K}') \). The (feature-based) revision of \( \mathcal{K} \) by \( \mathcal{K}' \) is defined as the DL-Litebool KB \( \mathcal{K} \circ \mathcal{K}' \) such that if \( \Omega_\mathcal{K} = \emptyset \), then \( \Omega_{\mathcal{K} \circ \mathcal{K}'} = \Omega_{\mathcal{K}'} \); otherwise

1. \( \sigma(\Omega_{\mathcal{K}'}, \Omega_\mathcal{K}) \subseteq \Omega_{\mathcal{K} \circ \mathcal{K}'} \), and
2. \( \Omega_{\mathcal{K} \circ \mathcal{K}'} \) is minimal in the sense that, for any DL-Litebool KB \( \mathcal{K}'' \) satisfying the above condition 1, \( \Omega_{\mathcal{K} \circ \mathcal{K}'} \subseteq \Omega_{\mathcal{K}''} \).

Note that the definition is slightly different from Satoh’s as \( \sigma(\Omega_{\mathcal{K}'}, \Omega_\mathcal{K}) \) may not correspond exactly to a DL-Litebool KB. By requiring \( \Omega_{\mathcal{K} \circ \mathcal{K}'} \) to be the minimal superset of \( \sigma(\Omega_{\mathcal{K}'}, \Omega_\mathcal{K}) \), we define \( \mathcal{K} \circ \mathcal{K}' \) in a conservative manner. Another option is to define \( \Omega_{\mathcal{K} \circ \mathcal{K}'} \) to be a maximal subset of \( \sigma(\Omega_{\mathcal{K}'}, \Omega_\mathcal{K}) \). However, as \( \Omega_{\mathcal{K} \circ \mathcal{K}'} \) can be very small if defined this way, there is a risk that unexpected logical consequences are introduced. Thus, it is more plausible to use supersets instead of subsets.

For any two DL-Litebool KBs \( \mathcal{K}_1, \mathcal{K}_2 \) that are both revisions of \( \mathcal{K} \) by \( \mathcal{K}' \), we have \( \Omega_{\mathcal{K}_1} = \Omega_{\mathcal{K}_2} \). By Proposition 2, this definition uniquely defines the result of \( \mathcal{K} \circ \mathcal{K}' \) up to equivalence. We will show in Section 5 that \( \mathcal{K} \circ \mathcal{K}' \) always exists.

**Example 3.** Consider a DL-Litebool KB

\[
\mathcal{K} = \langle \{ \text{PhD} \sqsubseteq \text{Student}, \text{Student} \sqsubseteq \neg \exists \text{teaches} \}, \{ \text{PhD}(\text{Tom}) \} \rangle.
\]
The TBox of $K$ specifies that PhD students are students, and students are not allowed to teach any courses, while the ABox states that Tom is a PhD student. Suppose PhD students are actually allowed to teach, and we want to revise $K$ with $K' = \{ \{ \text{PhD} \subseteq \exists \text{teaches} \}, \emptyset \}$.

Then $F' = \{ \{ \tau_1 \}, \{ \tau_1(\text{Tom}) \} \}$, where $\tau_1 = \{ \text{Student} \}$ is a model feature of $K'$. Take $F = \{ \{ \tau_1, \tau_2 \}, \{ \tau_2(\text{Tom}) \} \}$ where $\tau_2 = \{ \text{PhD}, \text{Student} \}$. It can be verified that $F$ is a model feature of $K$ and it is closest to $F'$.

The distance between $F$ and $F'$ is a minimal one among those between model features of $K$ and $K'$. Thus, $F'$ is a model feature of $K \circ K'$.

In fact, we can show that $K \circ K'$ is $\{ \{ \text{PhD} \subseteq \text{Student}, \text{PhD} \subseteq \exists \text{teaches}, \text{Student} \land \exists \text{teaches} \subseteq \text{PhD} \}, \{ \text{Student}(\text{Tom}) \} \}$.

In Example 3, the new inclusion added causes concept PhD to be unsatisfiable, i.e., PhD must be interpreted as an empty set, and thus contradicts $\text{PhD}(\text{Tom})$. The contradiction is resolved through revision by replacing $\text{Student} \subseteq \neg \exists \text{teaches}$ with $\text{Student} \cap \exists \text{teaches} \subseteq \text{PhD}$. Also, $\text{PhD}(\text{Tom})$ is weaken to be $\text{Student}(\text{Tom})$, because $\neg \exists \text{teaches}(\text{Tom})$ is also (implicitly) expressed in $K$. However, if we treat type sets and Herbrand sets equally in measuring distances, as we can show, $\text{Student}(\text{Tom})$ will be lost from the result of revision.

The standard AGM postulates (R1) – (R6) for propositional belief revision have been adapted to DLs in [15]. However, the authors present the postulates in terms of models of KBs. In the following, we reformulate them using KB combinations and entailments, in a manner analogous to classical AGM postulates.

(R1) $K \circ K' \models K'$;
(R2) if $K \cup K'$ is consistent, then $K \circ K' \equiv K \cup K'$;
(R3) if $K'$ is consistent, then $K \circ K'$ is consistent;
(R4) if $K_1 \equiv K_2$ and $K_1' \equiv K_2'$, then $K_1 \circ K_1' \equiv K_2 \circ K_2'$;
(R5) if $\Omega_{K \circ K'} = \sigma(\Omega_K, \Omega_{K'})$, then $(K \circ K') \cup \Omega'' \models K \circ (K' \cup \Omega'')$;
(R6) if $(K \circ K') \cup \Omega''$ is consistent, then $K \circ (K' \cup \Omega'') \equiv (K \circ K') \cup \Omega''$.

The following theorem states the first five postulates are satisfied by our revision operator.

**Theorem 1.** The revision operator defined in Definition 3 satisfies postulates (R1) – (R5).

Note that (R5) is conditional, and in this sense varies from its corresponding classical AGM postulate. We remark that the condition in (R5) cannot be removed. That is, $(K \circ K') \cup \Omega'' \models K \circ (K' \cup \Omega'')$ is not always satisfied. Also, (R6) is not always satisfied by our revision operator, as with Satoh’s revision operator in propositional logic. This can be seen from the following example.

**Example 4.** Let $K = \{ \{ A \subseteq B, B \subseteq C \}, \{ A(a) \} \}$, $K' = \{ \{ B \subseteq A, A \subseteq \neg C \}, \{ (A \cup C)(a) \} \}$, and $\Omega'' = \{ \emptyset, \{ \neg B(a) \} \}$. Then $K \circ K' = \{ \{ A \subseteq B, B \subseteq A, A \subseteq \neg C \}, \{ (A \cup C)(a) \} \}$, and after combined with $K''$, $(A \cup C)(a)$ in the ABox is replaced with $C(a)$. However, $K \circ (K' \cup K'') = \{ \{ B \subseteq C, B \subseteq A, A \subseteq \neg C \}, \{ (A \cup C)(a) \} \}$. That is, $(K \circ K') \cup \Omega'' \not\models K \circ (K' \cup \Omega'')$ and $K \circ (K' \cup \Omega'') \not\models (K \circ K') \cup \Omega''$. 

5 Algorithm for Computing Revision

In this section, we investigate the problem of computing revision and as a result, present an algorithm that can compute the result of revision syntactically.

The revision operator is based on model features of KBs. In what follows, we introduce a method for computing all the model features of a DL-Lite \( \text{bool} \) KB \( K \), by first computing the assembling of all the model features, which can be obtained directly from \( K \) through syntactic transformations.

Given two features \( F_1 = (\Xi_1, H_1) \) and \( F_2 = (\Xi_2, H_2) \), we define the assembling of \( F_1 \) and \( F_2 \) as \( F_1 \oplus F_2 = (\Xi_1 \cup \Xi_2, H_1 \oplus H_2) \), where \( H_1 \oplus H_2 \) is called a multiple Herbrand set, consisting of:

- \( \{\tau_1, \tau_2\}(a) \), for each \( a \in S_O \), \( \tau_1(a) \in H_1 \) and \( \tau_2(a) \in H_2 \); and
- \( P(a, b) \), for \( a, b \in S_O \), \( P(a, b) \in H_1 \) and \( P(a, b) \in H_2 \).

In a multiple Herbrand set, \( \{\tau_1, \tau_2\} \) can be viewed as a disjunction of types \( \tau_1 \) and \( \tau_2 \). Since role disjunction is not allowed in DL-Lite \( \text{bool} \), we only need to consider those role assertions appearing in both \( H_1 \) and \( H_2 \).

Given a DL-Lite \( \text{bool} \) KB \( K \), we use \( \bigoplus \Omega_K = (\Xi_K, M_K) \) to denote the assembling of all the model features of \( K \), where \( \Xi_K \) is a type set and \( M_K \) a multiple Herbrand set. For simplicity, given \( \{\tau_1, \ldots, \tau_n\} \) in a multiple Herbrand set \( M \), we also represent it as \( \Xi(a) \) with \( \Xi = \{\tau_1, \ldots, \tau_n\} \).

**Example 5.** Recall the KB \( K \) in Example 2. \( \Xi_K \) consists of nine types: \( \{\emptyset, \tau_1, \tau_2, \{A, \exists P, \geq 2 P\}, \{B, \exists P\}, \{B, \exists P, \geq 2 P, (\exists P^-)\}, \{\exists P, (\geq 2 P)\}\} \), where a type containing a basic concept in brackets, e.g., \( (\exists P^-) \), represents two types, one containing \( \exists P^- \) and the other not. The multiple Herbrand set \( M_K \) consists of \( \{\tau_1, \{A, \exists P, \geq 2 P\}\} \)(a), \( \{\tau_2, \{B, \exists P, \geq 2 P, \exists P^-\}\}\)(b), and \( P(a, b) \).

To compute the model features of a KB \( K \), a naive method is to check for each \( S \)-feature \( F \) whether \( F \) satisfies \( K \). However, we can show that \( \bigoplus \Omega_K \) can be computed directly from \( K \) through syntactic transformations, and \( \Omega_K \) can be easily obtained from \( \bigoplus \Omega_K \).

Now we show how to compute \( \bigoplus \Omega_K \) for a DL-Lite \( \text{bool} \) KB \( K \) through syntactic method (ref. Figure 1). Note that 2 and 3 of Step 3 ensure that \( M \) is a multiple Herbrand set, and thus, together with Step 2, ensure \( (\Xi, M) \) is an assembling of features.

**Theorem 2.** Given a consistent DL-Lite \( \text{bool} \) KB \( K \), Algorithm 1 always returns \( \bigoplus \Omega_K \).

The following result shows that all the model features of \( K \) can be easily obtained from \( \bigoplus \Omega_K \).

**Proposition 3.** Given a DL-Lite \( \text{bool} \) KB \( K \), if \( \bigoplus \Omega_K = (\Xi_K, M_K) \), then for any feature \( F = (\Xi, H) \), \( F \) is a model feature of \( K \) iff the following conditions are satisfied:

1. \( \Xi \subseteq \Xi_K \);
2. \( \tau_a \in \Xi_a \), for each \( a \in S_O \), \( \tau_a(a) \in H \) and \( \Xi_a(a) \in M_K \); and
Algorithm 1 (Compute the assembling of all the model features of a DL-Lite\_bool KB)

**Input:** A DL-Lite\_bool KB $K = (T, A)$ and a signature $S$.

**Output:** $\bigoplus \Omega_K$.

**Method:**

**Step 1.** Compute $\mathcal{S}$-type set $\mathcal{E} = \bigcap \{ C_1 \sqsubseteq C_2 | \quad T \mathcal{S}(\neg C_1 \sqcup C_2) \}.$

**Step 2.** Eliminate from $\mathcal{E}$ all the types $\tau$ such that $\exists R \in \tau$ for some $R$ but $\exists R^{-} \not\in \bigcup \mathcal{E}$.

**Step 3.** Initially, take $M = \{ P(a, b) | P(a, b) or P^{-}(b, a) \in A \}$. Then, for each $a \in S_O$, perform the following procedure:
1. add $\top(a)$ into $A$, and compute $\mathcal{E}_a = \bigcap \{ C(a) \in A \quad T \mathcal{S}(C) \} \cap \mathcal{E}$;
2. for each $P \in S_R$, if $n$ is the number of $P$-successors of $a$ in $M$ and $m$ the largest number in $S_N$ with $m \leq n$, then eliminate from $\mathcal{E}_a$ all the types not containing $\geq m P$;
3. for each $P \in S_R$, if $n$ is the number of $P$-predecessors of $a$ in $M$ and $m$ the largest number in $S_N$ with $m \leq n$, then eliminate from $\mathcal{E}_a$ all the types not containing $\geq m P^{-}$;
4. add $\mathcal{E}_a(a)$ into $M$.

**Step 4.** Return $\langle \mathcal{E}, M \rangle$ as $\bigoplus \Omega_K$.

**Fig. 1.** Compute assembling of model features.

3. $P(a, b) \in H$, for each $P(a, b) \in M_K$.

From Proposition 3, $\Omega_K$ is the set consisting of all the features satisfying 1 – 3 of Proposition 3.

Given $\Omega_K$ and $\Omega_{K'}$, we can select those features in $\Omega_{K'}$ that are closest to $\Omega_K$, i.e., $\sigma(\Omega_{K'}, \Omega_K)$, through the definition. To provide a complete algorithm for our revision operator, we only need a method to construct $K \circ K'$ from $\sigma(\Omega_{K'}, \Omega_K)$.

Each KB uniquely determines an assembling of features, that is, the assembling of all its model features. The converse is also true, as shown in the following proposition.

**Proposition 4.** Given a set $\Omega$ of features, there always exists a DL-Lite\_bool KB $K$ such that

1. $\bigoplus \Omega_K = \bigoplus \Omega$;
2. $\Omega \subseteq \Omega_K$; and for any DL-Lite\_bool KB $K'$ with $\Omega \subseteq \Omega_{K'}$, $\Omega_K \subseteq \Omega_{K'}$.

The above proposition suggests that, $K \circ K'$ can be constructed from $\bigoplus \Omega_K \otimes \Omega_{K'}$. Also, to compute $\bigoplus \Omega_K \otimes \Omega_{K'}$, we can avoid computing all the features in $\Omega_K \otimes \Omega_{K'}$, but obtain it directly from $\sigma(\Omega_{K'}, \Omega_K)$. Indeed, we have $\bigoplus \Omega_K \otimes \Omega_{K'} = \bigoplus \sigma(\Omega_{K'}, \Omega_K)$.

Finally, we present an algorithm for DL-Lite\_bool KB revision in Figure 2, which follows from the definition.

**Theorem 3.** Given two consistent DL-Lite\_bool KBs $K$ and $K'$, Algorithm 2 always returns $K \circ K'$.

The problem of computing the revision is decidable although it may be exponential in the worst case.
Algorithm 2 (Compute the result of revising a DL-Lite_true KB)

Input: Two DL-Lite_true KBs \( K \) and \( K' \), \( S = \text{Sig}(K \cup K') \).

Output: \( K \circ K' \).

Method:

Step 1. Use Algorithm 1 to compute \( \bigoplus \Omega_K \) and \( \bigoplus \Omega_{K'} \).

Step 2. Compute \( \Omega_K \) and \( \Omega_{K'} \) through Proposition 3.

Step 3. Obtain \( \sigma(\Omega_{K'}, \Omega_K) \) and compute \( \bigoplus \Omega_K \otimes \bigoplus \Omega_{K'} = \bigoplus \sigma(\Omega_K \otimes \Omega_{K'}) \). Suppose \( \bigoplus \Omega_K \otimes \bigoplus \Omega_{K'} = \langle \Xi, M \rangle \).

Step 4. Construct from \( \langle \Xi, M \rangle \) the DL-Lite_true KB \( \langle T, A \rangle \) with

\[
T = \{ \bigcap_{B \in \tau} B \cap \bigcap_{B \notin \tau} \neg B \sqsubseteq \perp \mid B \in Lit^+_S; \tau \text{ is an } S\text{-type s.t. } \tau \notin \Xi \},
\]

and

\[
A = \{ ( \bigcap_{\tau \in \Xi}(a) \bigcap_{B \in \tau} B \cap \bigcap_{B \notin \tau} \neg B))(a) \mid a \in S_A, \Xi_a(a) \in M \} \cup \{ P(a, b) \in M \}.
\]

Step 5. Return \( \langle T, A \rangle \) as \( K \circ K' \).

Fig. 2. Compute revision.

6 Conclusion

Approaches on adapting classical model-based revision to DLs have proposed in [13, 15]. In [15], the author discuss a special case in KB revision, that is, when inconsistency is caused by ABox assertions. The distances between two models are defined through the number of individuals whose membership differ in the two models. A TBox revision operator that can handle incoherence is introduced in [13], in which the distance between two models are defined through the number of atomic concepts they disagree. In both of these two approaches, the results of revision are not single KBs, but disjunctive KBs. Also, when applying to Example 3 (with only the TBox revised in the second approach), neither of these two approaches has Student(Tom) as a logical consequence of the revision. Same problem also occurs in those kernel revision operators with disjunctive KBs as the results of revision.

We have developed a formal framework for revising general KBs (with no specific restriction) in DL-Lite_true, based on the notion of features. We have defined a specific revision operator, which is a natural adaptation of revision approaches from propositional logic. We have also developed algorithms for computing KB revisions in DL-Lite_true, and the result of our revision is always an unique DL-Lite_true KB. We note that other propositional revision operators, e.g., Dalal’s revision, belief contraction and update can also be easily adopted in our framework.

There are still several problems remaining for future work. One is to extend our approach to KB revision in other DLs. The major difficulty here is that a general concept in most expressive DLs cannot be transformed into an equivalent disjunction of conjunctions of literal concepts. Another problem is to look at applications of our revision operator in nonmonotonic reasoning problems in DLs.
References