Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10)

A New Approach to Knowledge Base Revision in DL-Lite

Zhe Wang and **Kewen Wang** and **Rodney Topor**
School of ICT, Griffith University
Nathan, QLD 4111, Australia

Abstract

Revising knowledge bases (KBs) in description logics (DLs) in a syntax-independent manner is an important, nontrivial problem for the ontology management and DL communities. Several attempts have been made to adapt classical model-based belief revision and update techniques to DLs, but they are restricted in several ways. In particular, they do not provide operators or algorithms for general DL KB revision. The key difficulty is that, unlike propositional logic, a DL KB may have infinitely many models with complex (and possibly infinite) structures, making it difficult to define and compute revisions in terms of models. In this paper, we study general KBs in a specific DL in the DL-Lite family. We introduce the concept of features for such KBs, develop an alternative semantic characterization of KBs using features (instead of models), define two specific revision operators for KBs, and present the first algorithm for computing best approximations for syntax-independent revisions of KBs.

Introduction

Description logic (DL) has proved to be the most successful formalism for representing and reasoning about static knowledge in ontology applications. Such applications require a knowledge base (KB) consisting of a TBox (of terminological axioms) and an ABox (of data membership assertions). For example, the W3C ontology language framework OWL designed for Semantic Web applications is based on a particular set of description logics.

However, ontologies in Semantic Web applications are not static, but evolve over time. An important and nontrivial problem for such applications is thus how to effectively and efficiently revise/update KBs in a natural way. A typical scenario is the need for incremental ontology design to satisfy a changing environment.

Recently, there has been significant interest in revising/updating knowledge bases in description logics. In particular, several model-based revision/update approaches to DLs have been proposed (Liu et al. 2006; Giacomo et al. 2007; Qi et al. 2009). However, these revision/update operators for DLs are unable to deal with general KBs. Specifically, the first two approaches deal with only ABox update, while the third approach only considers revision of TBoxes.

In contrast to previous approaches, we focus on a specific DL, but address the problem of defining and computing revisions for general KBs, consisting of TBoxes *and* ABoxes. DL-Lite (Calvanese et al. 2007; Artale et al. 2007), which forms the basis of OWL 2 QL (one of the three profiles of OWL 2), is a family of lightweight DLs with efficient KB reasoning and query answering algorithms. We choose DL-Lite$^{\mathcal{N}}_{bool}$ (Artale et al. 2007), one of the most expressive members of the DL-Lite family, and define revision operators for DL-Lite$^{\mathcal{N}}_{bool}$ KBs in a way analogous to the model-based approaches in propositional logic. Although our approach is based on DL-Lite$^{\mathcal{N}}_{bool}$, we note that the definitions and algorithms can easily be adapted to other DL-Lite languages. We also note that update traditionally addresses changes of the actual state of the world (*e.g.*, that resulted from some action), whereas revision addresses the incorporation of new knowledge about the world. In this paper, we focus on revision.

The key issues in adapting classical model-based approaches to DLs are how to define the distance between models and how to construct the resulting KB (directly or indirectly) from selected models. However, such adaption is difficult for the following reasons. (1) DL interpretations have complex (possibly infinite) structures, which require a complex definition of the distance between two interpretations. (2) Unlike a propositional theory, a DL KB may have infinitely many models, making it impossible to compute the result effectively via models. (3) Given a collection $\mathbb{M}$ of interpretations, there may not exist a single KB $\mathcal{K}$ such that $\mathbb{M}$ is exactly the set of models for $\mathcal{K}$. These are also the reasons for the restrictions in previous approaches to DL revision.

In this paper, we first define *features* for DL-Lite$^{\mathcal{N}}_{bool}$, which precisely capture the most important semantic properties of DL-Lite$^{\mathcal{N}}_{bool}$ KBs, and (unlike models) are always finite. We adapt the techniques of model-based revision in propositional logic to the revision of DL-Lite$^{\mathcal{N}}_{bool}$ KBs, and define two specific revision operators based on two definitions of distance between features. We show that both revision operators possess desirable logical properties, and one of them preserves more knowledge from the original KB and thus yields a better result. As a set of features may not correspond exactly to any DL-Lite$^{\mathcal{N}}_{bool}$ KB, we also present syntactic algorithms for approximating the result of revision as a single DL-Lite$^{\mathcal{N}}_{bool}$ KB.

The DL-Lite Family

A *signature* is a finite set $\mathcal{S} = \mathcal{S}_C \cup \mathcal{S}_R \cup \mathcal{S}_I \cup \mathcal{S}_N$ where $\mathcal{S}_C$ is the set of atomic concepts, $\mathcal{S}_R$ is the set of atomic roles, $\mathcal{S}_I$ is the set of individual names and $\mathcal{S}_N$ is the set of natural numbers in $\mathcal{S}$. We assume 1 is always in $\mathcal{S}_N$. The special symbol $\top$ is neither an atomic concept nor an atomic role. Formally, given a signature $\mathcal{S}$, a DL-Lite$^{\mathcal{N}}_{bool}$ language has the following syntax:

$R \leftarrow P \mid P^-$

$B \leftarrow \top \mid A \mid \geqslant n\, R$

$C \leftarrow B \mid \neg C \mid C_1 \sqcap C_2$

where $n \in \mathcal{S}_N$, $A \in \mathcal{S}_C$ and $P \in \mathcal{S}_R$. B is called a *basic concept* and C is called a *general concept*. We write $\bot$ for $\neg\top$, $\exists R$ for $\geq 1\ R$, and $C_1 \sqcup C_2$ for $\neg(\neg C_1 \sqcap \neg C_2)$.

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

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

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

$(P^-)^{\mathcal{I}} = \{ (a^{\mathcal{I}}, b^{\mathcal{I}}) \mid (b^{\mathcal{I}}, a^{\mathcal{I}}) \in P^{\mathcal{I}} \}$,

$(\geqslant n\ R)^{\mathcal{I}} = \{ a^{\mathcal{I}} \mid |\{b^{\mathcal{I}} \mid (a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}\}| \geq n \}$,

$(\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$

$(C_1 \sqcap C_2)^{\mathcal{I}} = C_1^{\mathcal{I}} \cap C_2^{\mathcal{I}}$.

An interpretation $\mathcal{I}$ *satisfies* inclusion $C_1 \sqsubseteq C_2$ if $C_1^{\mathcal{I}} \subseteq C_2^{\mathcal{I}}$; $\mathcal{I}$ *satisfies* assertion $C(a)$ if $a^{\mathcal{I}} \in C^{\mathcal{I}}$; $\mathcal{I}$ *satisfies* assertion $R(a, b)$ if $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$. $\mathcal{I}$ *satisfies* TBox $\mathcal{T}$ (or ABox $\mathcal{A}$) if $\mathcal{I}$ satisfies each inclusion in $\mathcal{T}$ (resp., each assertion in $\mathcal{A}$). $\mathcal{I}$ is a *model* of a KB $\langle \mathcal{T}, \mathcal{A} \rangle$, if $\mathcal{I}$ satisfies both $\mathcal{T}$ and $\mathcal{A}$. We use $\mathsf{mod}(\mathcal{K})$ to denote the set of models of KB $\mathcal{K}$ and $\mathsf{sig}(\mathcal{K})$ to denote the signature of $\mathcal{K}$.

A KB $\mathcal{K}$ is *consistent* if it has at least one model. A KB $\mathcal{K}$ is *coherent* if, for each concept name A in $\mathcal{K}$, there exists a model $\mathcal{I}$ of $\mathcal{K}$ such that $A^{\mathcal{I}} \neq \emptyset$. Two KBs $\mathcal{K}_1, \mathcal{K}_2$ are *equivalent*, written $\mathcal{K}_1 \equiv \mathcal{K}_2$, if they have the same models. A KB $\mathcal{K}$ *entails* an inclusion or assertion α, written $\mathcal{K} \models \alpha$, if all models of $\mathcal{K}$ satisfy α.

Given a set $\mathbb{M}$ of interpretations and a signature $\mathcal{S}$, in most cases there does not exist a KB $\mathcal{K}$ over $\mathcal{S}$ such that the set of models of $\mathcal{K}$ is exactly $\mathbb{M}$. To address this inexpressibility problem, a notion of best approximation is introduced in (Giacomo et al. 2007). A KB $\mathcal{K}$ is said to be a *maximal approximation* of $\mathbb{M}$ over $\mathcal{S}$ if (1) $\mathsf{sig}(\mathcal{K}) \subseteq \mathcal{S}$, (2) $\mathbb{M} \subseteq \mathsf{mod}(\mathcal{K})$, and (3) there exists no KB $\mathcal{K}'$ over $\mathcal{S}$ such that $\mathbb{M} \subseteq \mathsf{mod}(\mathcal{K}') \subset \mathsf{mod}(\mathcal{K})$.

A *disjunctive knowledge base* (DKB) (Meyer et al. 2005) is a set $\mathbb{K}$ of KBs, and $\mathsf{mod}(\mathbb{K}) = \bigcup_{\mathcal{K} \in \mathbb{K}} \mathsf{mod}(\mathcal{K})$.

Features in DL-Lite$^{\mathcal{N}}_{bool}$

In this section, we introduce the concept of a feature in DL-Lite$^{\mathcal{N}}_{bool}$, which provides an alternative semantic characterization for DL-Lite$^{\mathcal{N}}_{bool}$. An advantage of semantic features over models is that the number of all features for a DL-Lite$^{\mathcal{N}}_{bool}$ knowledge base is finite and each feature is finite. These finiteness properties provide a means to adapt previously used revision approaches for classical propositional logic to DL-Lite$^{\mathcal{N}}_{bool}$.

Features for DL-Lite$^{\mathcal{N}}_{bool}$ are based on the notion of *types* defined in (Kontchakov et al. 2008).

In the following sections, if not specified, we assume $\mathcal{S}$ is a (finite but large enough) fixed signature, *i.e.*, $\mathsf{sig}(E) \subseteq \mathcal{S}$ for any general concept, inclusion, assertion, or KB E used.

An $\mathcal{S}$*-type* τ is a set of basic concepts over $\mathcal{S}$, s.t. $\top \in \tau$, and for any $m, n \in \mathcal{S}_N$ with $m < n$, $R \in \mathcal{S}_R \cup \{ P^- \mid P \in \mathcal{S}_R \}$, $\geqslant n\ R \in \tau$ implies $\geqslant m\ R \in \tau$. When the signature $\mathcal{S}$ is clear from context, we will simply call an $\mathcal{S}$-type a *type*. As $\top \in \tau$ for any type τ, we omit it in examples.

For example, let $\mathcal{S}_C = \{A, B\}$, $\mathcal{S}_R = \{P\}$, and $\mathcal{S}_N = \{1, 3\}$. Then $\tau = \{ A, \exists P, \geqslant 3\ P, \exists P^- \}$ is a type.

Intuitively, if each concept C is viewed as a propositional atom, types correspond to propositional interpretations of C. But a type is different from a propositional interpretation in that an element in a type may be of complex form, *e.g.*, $\geqslant n\ R$ or $\exists P$. We say type τ *satisfies* basic concept B if $B \in \tau$, τ *satisfies* $\neg C$ if τ does not satisfy C, and τ *satisfies* $C_1 \sqcap C_2$ if τ satisfies both C_1 and C_2.

We also say type τ *satisfies* concept inclusion $C_1 \sqsubseteq C_2$ if τ satisfies concept $\neg C_1 \sqcup C_2$, and type τ *satisfies* TBox $\mathcal{T}$ if it satisfies every inclusion in $\mathcal{T}$.

Types are sufficient to capture the semantics of TBoxes, but as they do not refer to individuals, they are insufficient to capture the semantics of ABoxes. We need to extend the notion of types and thus define *Herbrand sets* for ABoxes.

Definition 1 *An $\mathcal{S}$-*Herbrand set *(or* Herbrand set *when $\mathcal{S}$ is clear from the context) $\mathcal{H}$ is a finite set of assertions of the form $B(a)$ or $P(a, b)$, where $a, b \in \mathcal{S}_I$, $P \in \mathcal{S}_R$ and B is a basic concept over $\mathcal{S}$, satisfying the following conditions*

1. *For each $a \in \mathcal{S}_I$, $\top(a) \in \mathcal{H}$, and $\geqslant n\ R(a) \in \mathcal{H}$ implies $\geqslant m\ R(a) \in \mathcal{H}$ for $m, n \in \mathcal{S}_N$ with $m < n$.*
2. *For each $P \in \mathcal{S}_R$, $P(a, b_i) \in \mathcal{H}$ ($i = 1, \ldots, n$) implies $\geqslant m\ P(a) \in \mathcal{H}$ for any $m \in \mathcal{S}_N$ such that $m \leq n$.*
3. *For each $P \in \mathcal{S}_R$, $P(b_i, a) \in \mathcal{H}$ ($i = 1, \ldots, n$) implies $\geqslant m\ P^-(a) \in \mathcal{H}$ for any $m \in \mathcal{S}_N$ such that $m \leq n$.*

By condition 1 in Definition 1, given all the concept assertions of a in $\mathcal{H}$, $B_1(a), \ldots, B_k(a)$ ($k \geq 1$), then $\tau = \{B_1, \ldots, B_k\}$ is a type. We call τ the *type of* a in $\mathcal{H}$. Conditions 2 and 3 preserve the consistency of a Herbrand set. Since $\top(a)$ is always in $\mathcal{H}$ for any $\mathcal{H}$ and $a \in \mathcal{S}_I$, for simplicity, we will omit it in examples.

We say Herbrand set $\mathcal{H}$ *satisfies* concept assertion $C(a)$ if the type of a in $\mathcal{H}$ satisfies concept C. $\mathcal{H}$ *satisfies* role assertions $P(a, b)$ and $P^-(b, a)$ if $P(a, b)$ is in $\mathcal{H}$. $\mathcal{H}$ *satisfies* ABox $\mathcal{A}$ if $\mathcal{H}$ satisfies every assertion in $\mathcal{A}$.

To provide an alternative characterization for reasoning in a KB, we could use pairs $\langle \tau, \mathcal{H} \rangle$, where τ is a type and $\mathcal{H}$ is a Herbrand set, to replace standard interpretations, such that $\langle \tau, \mathcal{H} \rangle$ satisfies KB $\langle \mathcal{T}, \mathcal{A} \rangle$ if τ satisfies $\mathcal{T}$ and $\mathcal{H}$ satisfies $\mathcal{A}$. The resulting satisfaction relation should guarantee that $\mathcal{K}$ is consistent iff there is a pair $\langle \tau, \mathcal{H} \rangle$ satisfying $\mathcal{K}$. However, the following example shows that this is not the case.

Example 1 *Let $\mathcal{K} = \langle \{ \exists P^- \sqsubseteq \bot \}, \{ \exists P(a) \} \rangle$. Obviously, $\mathcal{K}$ is inconsistent and thus has no model. However, if $\mathcal{S} = \{P, a, 1\}$, then $\langle \{\exists P\}, \{\exists\, P(a)\} \rangle$ satisfies $\mathcal{K}$.*

The problem with using pairs $\langle \tau, \mathcal{H} \rangle$ is that they are not sufficient to capture the semantic connection between the TBox and the ABox of a KB. Our investigation shows that it is necessary to use a set of types (instead of a single type) in such a pair, as an alternative semantic characterization for a DL-Lite$^{\mathcal{N}}_{bool}$ KB. Using sets of types as semantic characterizations of DL-Lite TBoxes is also suggested in (Kontchakov et al. 2008).

Thus, we introduce the definition of a *feature* as follows.

Definition 2 (Features) *Given a signature $\mathcal{S}$, an $\mathcal{S}$-*feature *(or simply* feature *when $\mathcal{S}$ is clear) is a pair $\mathcal{F} = \langle \Xi, \mathcal{H} \rangle$, where Ξ is a non-empty set of $\mathcal{S}$-types and $\mathcal{H}$ an $\mathcal{S}$-Herbrand set, satisfying the following conditions:*

1. $\exists P \in \bigcup \Xi$ iff $\exists P^- \in \bigcup \Xi$, for each $P \in \mathcal{S}_R$.

2. $\tau \in \Xi$, for each $a \in \mathcal{S}_I$ and τ the type of a in $\mathcal{H}$.

The intuition behind the two conditions in Definition 2 can be easily seen after we explain how features can be used as an alternative for DL-Lite interpretations later.

We define the satisfaction relation of an inclusion or assertion *w.r.t.* a feature $\mathcal{F} = \langle \Xi, \mathcal{H} \rangle$ as follows:

- $\mathcal{F}$ satisfies $C_1 \sqsubseteq C_2$ if τ satisfies $\neg C_1 \sqcup C_2$ for all $\tau \in \Xi$.
- $\mathcal{F}$ satisfies assertion $C(a)$ or $R(a, b)$ if $\mathcal{H}$ satisfies it.

We can see that the first condition in Definition 2 guarantees that $\exists P$ is unsatisfiable (*i.e.*, $\exists P \sqsubseteq \bot$ is satisfied) *w.r.t.* the TBox if and only if $\exists P^-$ is also unsatisfiable *w.r.t.* the TBox. The second condition in Definition 2 requires that if $\mathcal{F}$ satisfies assertion $C(a)$, then C must be satisfiable *w.r.t.* the TBox.

We call $\mathcal{F}$ a *model feature* of KB $\mathcal{K}$ if $\mathcal{F}$ satisfies every inclusion and assertion in $\mathcal{K}$. We use $\mathsf{M}_{\mathcal{F}}(\mathcal{K})$ to denote the set of all model features of $\mathcal{K}$.

From the definition of features, as we only consider finite signatures, a model feature is always finite in structure, and the number of model features of a KB is also finite.

Example 2 *Consider the KB $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{T} = \{ A \sqsubseteq \exists P,\ B \sqsubseteq \exists P,\ \exists P^- \sqsubseteq B,\ A \sqcap B \sqsubseteq \bot,\ \geq 2\ P^- \sqsubseteq \bot \}$ and $\mathcal{A} = \{ A(a),\ P(a, b) \}$. It is shown in (Calvanese et al. 2006) that $\mathcal{K}$ is a KB having no finite model. In fact, an infinite model $\mathcal{I}$ of $\mathcal{K}$ can be defined as follows:*

$\Delta^{\mathcal{I}} = \{d_a, d_b, d_1, d_2, d_3 \ldots\}$, $a^{\mathcal{I}} = d_a$ and $b^{\mathcal{I}} = d_b$; 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 $\mathcal{S} = \mathsf{sig}(\mathcal{K}) = \{A, B, P, 1, 2, a, b\}$. The (finite) model feature of $\mathcal{K}$ that corresponds to $\mathcal{I}$ is $\mathcal{F} = \langle \Xi, \mathcal{H} \rangle$, where $\Xi = \{\tau_1, \tau_2\}$ with $\tau_1 = \{A, \exists P\}$ and $\tau_2 = \{B, \exists P, \exists P^-\}$, and $\mathcal{H} = \{ A(a),\ \exists P(a),\ B(b),\ \exists P(b),\ \exists P^-(b),\ P(a, b) \}$.

Given an inclusion or assertion α, define $\mathcal{K} \models_f \alpha$ if all features in $\mathsf{M}_{\mathcal{F}}(\mathcal{K})$ satisfy α. Given two KBs $\mathcal{K}_1, \mathcal{K}_2$ and $\mathcal{S} = \mathsf{sig}(\mathcal{K}_1 \cup \mathcal{K}_2)$, define $\mathcal{K}_1 \models_f \mathcal{K}_2$ if $\mathsf{M}_{\mathcal{F}}(\mathcal{K}_1) \subseteq \mathsf{M}_{\mathcal{F}}(\mathcal{K}_2)$, and $\mathcal{K}_1 \equiv_f \mathcal{K}_2$ if $\mathsf{M}_{\mathcal{F}}(\mathcal{K}_1) = \mathsf{M}_{\mathcal{F}}(\mathcal{K}_2)$.

The following two results show that model features do capture the semantic properties of DL-Lite KBs.

Proposition 1 *Let $\mathcal{K}$ be a DL-Lite$^{\mathcal{N}}_{bool}$ KB and $\mathcal{S} = \mathsf{sig}(\mathcal{K})$. Then we have*

- $\mathcal{K}$ is consistent iff $\mathcal{K}$ has a model feature.

- $\mathcal{K} \models (C_1 \sqsubseteq C_2)$ iff $\mathcal{K} \models_f (C_1 \sqsubseteq C_2)$ for any $C_1 \sqsubseteq C_2$ over $\mathcal{S}$.

- $\mathcal{K} \models C(a)$ iff $\mathcal{K} \models_f C(a)$ for any $C(a)$ over $\mathcal{S}$.

- $\mathcal{K} \models R(a, b)$ iff $\mathcal{K} \models_f R(a, b)$ for any $R(a, b)$ over $\mathcal{S}$.

Theorem 1 *Let $\mathcal{K}_1, \mathcal{K}_2$ be two DL-Lite$^{\mathcal{N}}_{bool}$ KBs and $\mathcal{S} = \mathsf{sig}(\mathcal{K}_1 \cup \mathcal{K}_2)$. Then $\mathcal{K}_1 \models \mathcal{K}_2$ iff $\mathcal{K}_1 \models_f \mathcal{K}_2$, and $\mathcal{K}_1 \equiv \mathcal{K}_2$ iff $\mathcal{K}_1 \equiv_f \mathcal{K}_2$.*

As with interpretations, given a set $\mathbb{F}$ of $\mathcal{S}$-features, there may be no KB $\mathcal{K}$ such that $\mathbb{F} = \mathsf{M}_{\mathcal{F}}(\mathcal{K})$. Hence, we define the *maximal approximation* of a set $\mathbb{F}$ of $\mathcal{S}$-features to be the KB $\mathcal{K}$ such that: (1) $\mathsf{sig}(\mathcal{K}) \subseteq \mathcal{S}$, (2) $\mathbb{F} \subseteq \mathsf{M}_{\mathcal{F}}(\mathcal{K})$, and (3) there exists no KB $\mathcal{K}'$ over $\mathcal{S}$ such that $\mathbb{F} \subseteq \mathsf{M}_{\mathcal{F}}(\mathcal{K}') \subset \mathsf{M}_{\mathcal{F}}(\mathcal{K})$. The maximal approximation of any given set $\mathbb{F}$ always exists in DL-Lite$^{\mathcal{N}}_{bool}$, and is unique up to KB equivalence.

Feature Distance and Revision

In the following two sections, we define two notions of distance between features, in the spirit of Hamming distance for propositional models. The first distance is defined as the set of concept and role names interpreted differently in the two features. The second distance is based on a generalized notion of symmetric difference. Based on these two distances, we define two specific revision operators for DL-Lite KBs in an analogous way to Satoh's (Satoh 1988), and show that they have desirable properties.

Given a set Σ of concept and role names and $\mathcal{S}$-types τ_1, τ_2, denote $\tau_1 \sim_\Sigma \tau_2$ if for all basic concepts B over $\mathcal{S} - \Sigma$, $B \in \tau_1$ iff $B \in \tau_2$.

Let $\mathcal{F}_1 = \langle \Xi_1, \mathcal{H}_1 \rangle$ and $\mathcal{F}_2 = \langle \Xi_2, \mathcal{H}_2 \rangle$ be two $\mathcal{S}$-features, and $\Sigma \subseteq \mathcal{S}_C \cup \mathcal{S}_R$. Define $\mathcal{F}_1 \leftrightarrow_\Sigma \mathcal{F}_2$ if the following conditions are satisfied:

1. For each $\tau_1 \in \Xi_1$, there exists $\tau_2 \in \Xi_2$ s.t. $\tau_1 \sim_\Sigma \tau_2$; and for each $\tau_2 \in \Xi_2$, there exists $\tau_1 \in \Xi_1$ s.t. $\tau_1 \sim_\Sigma \tau_2$.

2. For each $a \in \mathcal{S}_I$, $\tau_1 \sim_\Sigma \tau_2$, where τ_i ($i = 1, 2$) is the type of a in $\mathcal{H}_i$; and $P(a, b) \in \mathcal{H}_1$ iff $P(a, b) \in \mathcal{H}_2$ for each $P \in \mathcal{S}_R - \Sigma$ and $a, b \in \mathcal{S}_I$.

Intuitively, the minimal sets Σ such that $\mathcal{F}_1 \leftrightarrow_\Sigma \mathcal{F}_2$ are the sets of concept and role names on whose interpretations $\mathcal{F}_1, \mathcal{F}_2$ disagree.

Given two KBs $\mathcal{K}_1, \mathcal{K}_2$ and $\mathcal{S} = \mathsf{sig}(\mathcal{K}_1 \cup \mathcal{K}_2)$, define the distance between $\mathcal{K}_1$ and $\mathcal{K}_2$ as the set of all minimal distances between model features of $\mathcal{K}_1$ and $\mathcal{K}_2$:

$d_f(\mathcal{K}_1, \mathcal{K}_2) = \min_{\subseteq}(\{ \Sigma \subseteq \mathcal{S}_C \cup \mathcal{S}_R \mid \exists \mathcal{F}_1 \in \mathsf{M}_{\mathcal{F}}(\mathcal{K}_1), \exists \mathcal{F}_2 \in \mathsf{M}_{\mathcal{F}}(\mathcal{K}_2)$ *s.t.* $\mathcal{F}_1 \leftrightarrow_\Sigma \mathcal{F}_2 \})$.

To define a revision operator $\mathcal{K} \circ \mathcal{K}'$ in analogy to classical model-based revision, we need to specify the subset of $\mathsf{M}_{\mathcal{F}}(\mathcal{K}')$ that is *closest* to $\mathsf{M}_{\mathcal{F}}(\mathcal{K})$ (*w.r.t.* feature distance).

Definition 3 (S-Revision) *Let $\mathcal{K}, \mathcal{K}'$ be two DL-Lite$^{\mathcal{N}}_{bool}$ KBs and $\mathcal{S} = \mathsf{sig}(\mathcal{K} \cup \mathcal{K}')$. Define the* s-revision *of $\mathcal{K}$ by $\mathcal{K}'$, denoted $\mathcal{K} \circ_s \mathcal{K}'$, such that $\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ_s \mathcal{K}') = \mathsf{M}_{\mathcal{F}}(\mathcal{K}')$ if $\mathsf{M}_{\mathcal{F}}(\mathcal{K}) = \emptyset$, and otherwise,*

$$\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ_s \mathcal{K}') = \{\, \mathcal{F}' \in \mathsf{M}_{\mathcal{F}}(\mathcal{K}') \mid \exists \mathcal{F} \in \mathsf{M}_{\mathcal{F}}(\mathcal{K}) \text{ s.t. } \mathcal{F} \leftrightarrow_{\Sigma} \mathcal{F}' \text{ and } \Sigma \in d_f(\mathcal{K}, \mathcal{K}') \,\}.$$

Example 3 *Consider the following KB,*

$$\begin{aligned}\mathcal{K} = \langle\ \{\ & PhD \sqsubseteq Student \sqcap Postgrad,\\ & Student \sqsubseteq \neg\exists teaches,\ \exists teaches^- \sqsubseteq Course,\\ & Student \sqcap Course \sqsubseteq \bot\ \},\ \{\ PhD(Tom)\ \}\ \rangle.\end{aligned}$$

The TBox of $\mathcal{K}$ specifies that PhD students are postgraduate 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 $\mathcal{K}$ with $\mathcal{K}' = \langle\, \{\, PhD \sqsubseteq \exists teaches \,\},\, \emptyset \,\rangle$.

$\mathcal{F}' = \langle\, \{\tau_1\}, \{Student(Tom), Postgrad(Tom)\} \,\rangle$, where $\tau_1 = \{Student, Postgrad\}$ is a model feature of $\mathcal{K}'$. From the model features of $\mathcal{K}$, take $\mathcal{F} = \langle\, \{\tau_1, \tau_2\}, \{PhD(Tom), Student(Tom), Postgrad(Tom)\} \,\rangle$ where $\tau_2 = \{PhD, Student, Postgrad\}$. Then $\mathcal{F} \leftrightarrow_{\{PhD\}} \mathcal{F}'$ and $\{PhD\} \in d_f(\mathcal{K}, \mathcal{K}')$. Thus, $\mathcal{F}'$ is a model feature of $\mathcal{K} \circ_s \mathcal{K}'$.

Finally, $\mathcal{K} \circ_s \mathcal{K}'$ is the DKB $\{\mathcal{K}_1, \mathcal{K}_2\}$, where

$$\begin{aligned}\mathcal{K}_1 = \langle\ \{\ & PhD \sqsubseteq \exists teaches,\ Student \sqsubseteq \neg\exists teaches,\\ & \exists teaches^- \sqsubseteq Course,\ Student \sqcap Course \sqsubseteq \bot\ \},\\ & \{\ Student(Tom),\ Postgrad(Tom)\ \}\ \rangle, \text{ and}\\ \mathcal{K}_2 = \langle\ \{\ & PhD \sqsubseteq \exists teaches,\ PhD \sqsubseteq Student \sqcap Postgrad,\\ & Student \sqcap Course \sqsubseteq \bot\ \},\ \{\ PhD(Tom)\ \}\ \rangle.\end{aligned}$$

This definition improves on the more coarse-grained, concept-based definition of the distance between models in (Qi et al. 2009), which cannot reflect the difference between the two models on their interpretations on roles.

We note that corresponding feature-based revision operators can be defined for all three revision operators in (Qi et al. 2009), and can be used to resolve incoherence, as they are there, but, as resolving incoherence is not the focus of this paper, we omit the details here.

An important observation is that $\mathcal{K} \circ_s \mathcal{K}'$ can be computed by query-based forgetting. In particular, if $\mathsf{forget}(\mathcal{K}, \Sigma)$ denotes a result of $\mathcal{Q}^u_{\mathcal{L}}$-forgetting about Σ in $\mathcal{K}$ (Wang et al. 2010), then we have the following connection between revision and forgetting.

Theorem 2 *Let $\mathcal{K}, \mathcal{K}'$ be two consistent DL-Lite$^{\mathcal{N}}_{bool}$ KBs and $\mathcal{S} = \mathsf{sig}(\mathcal{K} \cup \mathcal{K}')$. Then*

$$\mathcal{K} \circ_s \mathcal{K}' = \{\, \mathsf{forget}(\mathcal{K}, \Sigma) \cup \mathcal{K}' \mid \Sigma \in d_f(\mathcal{K}, \mathcal{K}') \,\},$$

where $\mathsf{forget}(\mathcal{K}, \Sigma)$ is a result of $\mathcal{Q}^u_{\mathcal{L}}$-forgetting about Σ in $\mathcal{K}$.

As shown in (Wang et al. 2010), the result of $\mathcal{Q}^u_{\mathcal{L}}$-forgetting is always expressible in DL-Lite$^{u}_{bool}$, an extension of DL-Lite$^{\mathcal{N}}_{bool}$ (Kontchakov et al. 2008). Thus we have shown that $\mathcal{K} \circ_s \mathcal{K}'$ is always expressible as a DKB in DL-Lite$^{u}_{bool}$. However, $\mathcal{K} \circ_s \mathcal{K}'$ may not be expressible as a single KB in DL-Lite$^{u}_{bool}$, *e.g.*, the revision defined in Example 3 can not be so expressed.

Revision under Approximation

For many applications, it is desirable to have the revision as a single DL-Lite$^{\mathcal{N}}_{bool}$ KB rather than a DKB, *e.g.*, in the case of iterative revision. That is, the maximal approximation of the revision is desired. However, in most cases, $\mathcal{K} \circ_s \mathcal{K}'$ does not preserve enough knowledge of the original KB $\mathcal{K}$.

Example 4 *In Example 3, $\mathcal{K} \circ_s \mathcal{K}'$ is a DKB, whose maximal approximation is the following KB,*

$$\begin{aligned}\langle\ \{\ & PhD \sqsubseteq \exists teaches,\ Student \sqcap Course \sqsubseteq \bot\ \},\\ \{\ & (Student(Tom),\ Postgrad(Tom),\\ & \big(PhD \sqcup (\neg\exists teaches \sqcap \neg\exists teaches^-)\big)(Tom)\ \}\ \rangle.\end{aligned}$$

Note that in the above example, knowledge in $\mathcal{K}$ about concept PhD and about role $teaches$ are totally lost after revision and approximation. In particular, $PhD \sqsubseteq Postgrad$ and $\exists teaches^- \sqsubseteq Course$ are eliminated, though they have nothing to do with the inconsistency.

We argue that the reason why the revision operator $\circ_s$ performs poorly under approximation is that the distance defined on concept and role names is too simple to reflect differences between model features, as can be seen from the following example. Let $\mathcal{F} = \langle\, \{\tau_1, \tau_3\}, \{A(a), B(a)\} \,\rangle$, $\mathcal{F}' = \langle\, \{\tau_1, \tau_2, \tau_3\}, \{A(a), A(b)\} \,\rangle$, and $\mathcal{F}'' = \langle\, \{\tau_2\}, \{A(a), A(b), A(c), A(d)\} \,\rangle$, where $\tau_1 = \{A, B\}$, $\tau_2 = \{A\}$, and $\tau_3 = \emptyset$. Obviously, $\mathcal{F}$ is closer to $\mathcal{F}'$ than to $\mathcal{F}''$. However, such difference cannot be measured using only concept names, as $\mathcal{F} \leftrightarrow_{\{A,B\}} \mathcal{F}'$ and $\mathcal{F} \leftrightarrow_{\{A,B\}} \mathcal{F}''$.

From the above discussion, we can see that it is insufficient to measure the distance between two features or models using only a set of concepts (and roles). To obtain a better definition of KB revision, we need to introduce a more complex notion of feature distance, by extending the definition of symmetric difference $\triangle$.

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 $\mathcal{S}$-features $\mathcal{F}_1 = \langle \Xi_1, \mathcal{H}_1 \rangle$ and $\mathcal{F}_2 = \langle \Xi_2, \mathcal{H}_2 \rangle$, we define the *distance* between $\mathcal{F}_1$ and $\mathcal{F}_2$, denoted $\mathcal{F}_1 \triangle \mathcal{F}_2$, as a pair $\langle\, \Xi_1 \triangle \Xi_2,\, \mathcal{H}_1 \triangle \mathcal{H}_2 \,\rangle$. Note that we do not require $\mathcal{H}_1 \triangle \mathcal{H}_2$ to be a Herbrand set.

To compare two distances, given $\mathcal{F}_i = \langle \Xi_i, \mathcal{H}_i \rangle$ for $i = 1, 2, 3, 4$, we could define $\mathcal{F}_1 \triangle \mathcal{F}_2 \subseteq_f \mathcal{F}_3 \triangle \mathcal{F}_4$ if $\Xi_1 \triangle \Xi_2 \subseteq \Xi_3 \triangle \Xi_4$ and $\mathcal{H}_1 \triangle \mathcal{H}_2 \subseteq \mathcal{H}_3 \triangle \mathcal{H}_4$; and $\mathcal{F}_1 \triangle \mathcal{F}_2 \subset_f \mathcal{F}_3 \triangle \mathcal{F}_4$ if $\mathcal{F}_1 \triangle \mathcal{F}_2 \subseteq \mathcal{F}_3 \triangle \mathcal{F}_4$ and $\mathcal{F}_3 \triangle \mathcal{F}_4 \not\subseteq \mathcal{F}_1 \triangle \mathcal{F}_2$. However, our research shows that such a measure is still too weak to preserve enough knowledge of the original KB, as many features are still incomparable under such measure. Instead, we define a preference for Herbrand sets over type sets: $\mathcal{F}_1 \triangle \mathcal{F}_2 \subset_f \mathcal{F}_3 \triangle \mathcal{F}_4$ iff

- $\mathcal{H}_1 \triangle \mathcal{H}_2 \subset \mathcal{H}_3 \triangle \mathcal{H}_4$, or
- $\mathcal{H}_1 \triangle \mathcal{H}_2 = \mathcal{H}_3 \triangle \mathcal{H}_4$ and $\Xi_1 \triangle \Xi_2 \subset \Xi_3 \triangle \Xi_4$.

The intuition behind such a preference for Herbrand sets over type sets is that when both TBox inclusions and ABox assertions contribute to the inconsistency, assertions in the original ABox have priority to be preserved whereas TBox inclusions are candidates for revision. The reason for this is two-fold: First, due to the nature of revision and update, revision is more suitable for DL TBox change whereas update is more suitable for ABox change. This claim is justified in the literature (Liu et al. 2006; Giacomo et al. 2007;

Qi et al. 2009). Second, if inconsistency is caused by incoherence of the combination of the TBoxes (as in Example 3), modifying only ABox assertions can help restore consistency but will leave the resulting TBox incoherent. Giving TBox inclusions priority for change has the merit that inconsistency and incoherence can be resolved at the same time, so that no extra mechanism is needed to restore coherence. We will demonstrate this point with an example. Before that, we first introduce our revision operator in terms of the new distance measure.

Definition 4 (F-Revision) *Let $\mathcal{K}, \mathcal{K}'$ be two DL-Lite$^{\mathcal{N}}_{bool}$ KBs and $\mathcal{S} = \mathsf{sig}(\mathcal{K} \cup \mathcal{K}')$. Define the* f-revision *of $\mathcal{K}$ by $\mathcal{K}'$, denoted $\mathcal{K} \circ_f \mathcal{K}'$, such that $\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ_f \mathcal{K}') = \mathsf{M}_{\mathcal{F}}(\mathcal{K}')$ if $\mathsf{M}_{\mathcal{F}}(\mathcal{K}) = \emptyset$, and otherwise*

$$\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ_f \mathcal{K}') = \{ \mathcal{F}' \in \mathsf{M}_{\mathcal{F}}(\mathcal{K}') \mid \exists \mathcal{F} \in \mathsf{M}_{\mathcal{F}}(\mathcal{K}) \text{ s.t. } \forall \mathcal{F}_i \in \mathsf{M}_{\mathcal{F}}(\mathcal{K}), \forall \mathcal{F}'_j \in \mathsf{M}_{\mathcal{F}}(\mathcal{K}'),\ (\mathcal{F}_i \triangle \mathcal{F}'_j) \not\subset_f (\mathcal{F} \triangle \mathcal{F}') \}.$$

The next example shows that $\circ_f$ performs better $\circ_s$ under maximal approximation.

Example 5 *Consider the KBs $\mathcal{K}, \mathcal{K}'$ in Example 3. The maximal approximation of $\mathcal{K} \circ_f \mathcal{K}'$ is*

$$\langle\ \{ \mathit{PhD} \sqsubseteq \mathit{Student} \sqcap \mathit{Postgrad},\ \mathit{PhD} \sqsubseteq \exists \mathit{teaches},\ \mathit{Student} \sqcap \exists \mathit{teaches} \sqsubseteq \mathit{PhD},\ \exists \mathit{teaches}^- \sqsubseteq \mathit{Course},\ \mathit{Student} \sqcap \mathit{Course} \sqsubseteq \bot \},\ \{ \mathit{Student}(\mathit{Tom}),\ \mathit{Postgrad}(\mathit{Tom}) \} \rangle.$$

Note that $\mathit{Student} \sqsubseteq \neg \exists \mathit{teaches}$ is revised (and weakened) to $\mathit{Student} \sqcap \exists \mathit{teaches} \sqsubseteq \mathit{PhD}$. In this way, consistency and coherence are both restored. Also, more of the knowledge in $\mathcal{K}$ is preserved than in Example 4.

The AGM postulates are widely used to value the rationality of belief revision operators, which we adapt to DLs as follows.

(R1) $\mathcal{K} \circ \mathcal{K}' \models_f \mathcal{K}'$;

(R2) if $\mathcal{K} \cup \mathcal{K}'$ is consistent, then $\mathcal{K} \circ \mathcal{K}' = \mathcal{K} \cup \mathcal{K}'$;

(R3) if $\mathcal{K}'$ is consistent, then $\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ \mathcal{K}') \neq \emptyset$;

(R4) if $\mathcal{K}_1 \equiv \mathcal{K}_2$ and $\mathcal{K}'_1 \equiv \mathcal{K}'_2$, then $\mathcal{K}_1 \circ \mathcal{K}'_1 \equiv_f \mathcal{K}_2 \circ \mathcal{K}'_2$;

(R5) $(\mathcal{K} \circ \mathcal{K}') \cup \mathcal{K}'' \models_f \mathcal{K} \circ (\mathcal{K}' \cup \mathcal{K}'')$;

(R6) if $(\mathcal{K} \circ \mathcal{K}') \cup \mathcal{K}''$ is consistent, then $\mathcal{K} \circ (\mathcal{K}' \cup \mathcal{K}'') \models_f (\mathcal{K} \circ \mathcal{K}') \cup \mathcal{K}''$.

As with Satoh's (Satoh 1988) revision operator, both $\circ_s$ and $\circ_f$ satisfy the first five postulates.

Proposition 2 *The revision operators defined in Definitions 3 and 4 both satisfy postulates $(R1)$ – $(R5)$.*

If using its maximal approximation to replace the revision, postulates $(R1)$ – $(R4)$ are always satisfied.

However, $(R5)$ and $(R6)$ together require a total order over feature distances, which we argue is too strong for DL revision. For this reason, like Satoh, we do not require our revision operators to satisfy $(R6)$.

Algorithm for Computing Revision

In this section, we introduce an algorithm for computing the maximal approximation of revision syntactically. As the two revision operators are based on model features of KBs, we first introduce a method for computing all the model features of a KB, and then show how the maximal approximation of revision can be constructed via model features.

We first introduce a method that computes $\mathsf{M}_{\mathcal{F}}(\mathcal{K})$ from $\mathcal{K}$ through syntactic transformations (ref. Figure 1).

Algorithm 1
Input: A DL-Lite$^{\mathcal{N}}_{bool}$ KB $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ and a signature $\mathcal{S}$.
Output: $\mathsf{M}_{\mathcal{F}}(\mathcal{K})$.
Method: Initially, let $\mathbb{F} = \emptyset$.
Step 1. Compute the set $\Xi_{\mathcal{T}}$ of all $\mathcal{S}$-types satisfying $\mathcal{T}$.
Let $\mathcal{P} = \{ P(a,b) \mid P(a,b) \text{ or } P^-(b,a) \in \mathcal{A} \}$.
Step 2. Add into $\mathbb{F}$ all pairs $\langle \Xi, \mathcal{H} \rangle$ such that:
1. $\Xi \subseteq \Xi_{\mathcal{T}}$, and $\exists P \in \bigcup \Xi$ iff $\exists P^- \in \bigcup \Xi$ for all $P \in \mathcal{S}_R$.
2. $\mathcal{P} \subseteq \mathcal{H}$, and for each $a \in \mathcal{S}_I$, the type τ of a in $\mathcal{H}$ satisfies the following conditions:
(1) $\tau \in \Xi$, and τ satisfies C for every $C(a) \in \mathcal{A}$.
(2) $\geqslant m\ P \in \tau$, for each $P \in \mathcal{S}_R$ *s.t.* $P(a, b_i) \in \mathcal{H}$ with $i = 1, \ldots, n$, and $m \in \mathcal{S}_N$ with $m \leq n$.
(3) $\geqslant m\ P^- \in \tau$, for each $P \in \mathcal{S}_R$ *s.t.* $P(b_i, a) \in \mathcal{H}$ with $i = 1, \ldots, n$, and $m \in \mathcal{S}_N$ with $m \leq n$.
Step 3. Return $\mathbb{F}$ as $\mathsf{M}_{\mathcal{F}}(\mathcal{K})$.

Figure 1: Compute model features.

Note that each pair $\langle \Xi, \mathcal{H} \rangle$ added to $\mathbb{F}$ is a feature, and satisfies both $\mathcal{T}$ and $\mathcal{A}$. Algorithm 1 returns $\emptyset$ if and only if $\mathcal{K}$ is inconsistent.

Proposition 3 *Given a DL-Lite$^{\mathcal{N}}_{bool}$ KB $\mathcal{K}$ and a signature $\mathcal{S}$, Algorithm 1 always returns $\mathsf{M}_{\mathcal{F}}(\mathcal{K})$.*

Given that $\mathsf{M}_{\mathcal{F}}(\mathcal{K})$ and $\mathsf{M}_{\mathcal{F}}(\mathcal{K}')$ can be computed by Algorithm 1, and they are finite, it is straightforward to obtain $\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ_s \mathcal{K}')$ and $\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ_f \mathcal{K}')$ through Definitions 3 and 4. Now we show that the maximal approximation of $\mathcal{K} \circ \mathcal{K}'$ can be constructed from $\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ \mathcal{K}')$. Indeed, the maximal approximation of $\mathbb{F}$ can be constructed for any set $\mathbb{F}$ of model features in the same way. Given a $\mathcal{S}$-type τ, we denote the concept $C_\tau = \bigsqcap_{B \in \tau} B \sqcap \bigsqcap_{B \notin \tau} \neg B$, where B is a basic concept over $\mathcal{S}$. In what follows, we present an algorithm for DL-Lite$^{\mathcal{N}}_{bool}$ KB revision (ref. Figure 2).

Algorithm 2
Input: Two DL-Lite$^{\mathcal{N}}_{bool}$ KBs $\mathcal{K}$ and $\mathcal{K}'$, $\mathcal{S} = \mathsf{sig}(\mathcal{K} \cup \mathcal{K}')$.
Output: $\mathcal{K} \circ_f \mathcal{K}'$.
Method: Initially, let $\mathcal{T} = \emptyset$ and $\mathcal{A} = \emptyset$.
Step 1. Use Algorithm 1 to compute $\mathsf{M}_{\mathcal{F}}(\mathcal{K})$ and $\mathsf{M}_{\mathcal{F}}(\mathcal{K}')$.
Step 2. Obtain $\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ_f \mathcal{K}')$ from $\mathsf{M}_{\mathcal{F}}(\mathcal{K})$ and $\mathsf{M}_{\mathcal{F}}(\mathcal{K}')$ by Definition 4.
Step 3. For each $\mathcal{S}$-type τ not occurring in any type set in $\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ_f \mathcal{K}')$, add inclusion $C_\tau \sqsubseteq \bot$ into $\mathcal{T}$.
Step 4. For each individual $a \in \mathcal{S}_I$, add concept assertion $(\bigsqcup_{\tau \in \Xi_a} C_\tau)(a)$ into $\mathcal{A}$, where $\Xi_a = \{ \tau \mid \exists \langle \Xi, \mathcal{H} \rangle \in \mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ_f \mathcal{K}')$ *s.t.* τ is the type of a in $\mathcal{H} \}$.
Step 5. For each role assertion $P(a,b)$ occurring in every Herbrand set in $\mathsf{M}_{\mathcal{F}}(\mathcal{K} \circ_f \mathcal{K}')$, add $P(a,b)$ into $\mathcal{A}$.
Step 6. Return $\langle \mathcal{T}, \mathcal{A} \rangle$ as $\mathcal{K} \circ_f \mathcal{K}'$.

Figure 2: Compute f-revision.

Note that Algorithm 2 does not rely on the definition of a specific revision operator, as long as it is defined in terms of feature distance. Hence, Algorithm 2 can be used to compute $\mathcal{K} \circ_s \mathcal{K}'$ if Definition 3 is adopted in Step 2. Such an algorithm based on features gives us the flexibility to compute other possible revisions proposed according different application needs.

Theorem 3 *Given two consistent DL-Lite$^{\mathcal{N}}_{bool}$ KBs $\mathcal{K}$ and $\mathcal{K}'$, Algorithm 2 always returns the maximal approximation of $\mathcal{K} \circ_f \mathcal{K}'$.*

In the worst case, the maximal approximation of $\mathcal{K} \circ_f \mathcal{K}'$ is exponential in size *w.r.t.* $\mathcal{K}$ and $\mathcal{K}'$. We argue that there exists no tractable algorithm for DL revision, as propositional logic revision has been shown to have high complexity, and DL revision is even more complex.

However, in most applications, only a small part of the whole knowledge base needs to be revised. As a result, the algorithms can be optimized. In particular, only the subset of $\mathcal{K}$ that is *relevant* to $\mathcal{K}'$ needs to be revised, whereas the other irrelevant inclusions and assertions in $\mathcal{K}$ can be added directly into the result of revision.

A notion of (signature) relevance is formally defined in propositional logic via *language splitting* (Parikh 1996), and is adapted to DLs for decomposing TBoxes (Konev et al. 2010). In what follows, we generalize this notion and show its application in KB revision. Given a KB $\mathcal{K}$, we say that a set of signatures $\mathbb{S} = \{\mathcal{S}_1 \dots, \mathcal{S}_n\}$ $(n \geq 1)$ is a *splitting* of $\mathcal{K}$ if $\mathcal{S}_i \cap \mathcal{S}_j \subseteq \mathcal{S}_N$ for $1 \leq i < j \leq n$, where $\mathcal{S}_N$ consists of all the numbers in $\mathbb{S}$, and there exist KBs $\mathcal{K}_1 \dots, \mathcal{K}_n$ such that $\mathsf{sig}(\mathcal{K}_i) \subseteq \mathcal{S}_i$ for $1 \leq i \leq n$ and $\mathcal{K} \equiv \bigcup_{1 \leq i \leq n} \mathcal{K}_i$.

The following result states that our revision operators enjoy a decomposition property regarding signature relevance. We use $\circ$ to denote either one of the revision operators defined in Definitions 3 and 4.

Proposition 4 *Let $\{\mathcal{S}_r, \mathcal{S}_{ir}\}$ be a splitting of $\mathcal{K}$, with $\mathcal{K} \equiv \mathcal{K}_r \cup \mathcal{K}_{ir}$, $\mathsf{sig}(\mathcal{K}_r) \subseteq \mathcal{S}_r$ and $\mathsf{sig}(\mathcal{K}_{ir}) \subseteq \mathcal{S}_{ir}$. Suppose $\mathsf{sig}(\mathcal{K}') \cap \mathcal{S}_{ir} \subseteq \mathcal{S}_N$ with $\mathcal{S}_N$ consisting of all the numbers in $\{\mathcal{S}_r, \mathcal{S}_{ir}\}$, then $\mathcal{K} \circ \mathcal{K}' = (\mathcal{K}_r \circ \mathcal{K}') \cup \mathcal{K}_{ir}$.*

In ontology revision applications, the new knowledge $\mathcal{K}'$, which represents changes to the knowledge of the world, is often small both in size and in signature, compared to the large existing KB $\mathcal{K}$. The subset $\mathcal{K}_r$ of $\mathcal{K}$, that is relevant to $\mathcal{K}'$, is usually also small. Thus, it is reasonable to expect efficient revision computation in realistic applications.

Conclusion

We have developed a formal framework for revising general KBs (consisting of both TBoxes and ABoxes) in DL-Lite$^{\mathcal{N}}_{bool}$, based on the notion of features. We have defined two specific revision operators, which are natural adaption of model-based revision approaches from propositional logic. We have shown that the second operator performs better *w.r.t.* maximal approximation. We have also developed algorithms for computing maximal approximations of KB revisions in DL-Lite$^{\mathcal{N}}_{bool}$. We note that other propositional revision operators, e. g., Dalal's revision, belief contraction and update can also be easily defined in our framework.

In contrast to previous revision/update approaches in DLs, we consider revisions of general KBs, and a KB can be revised by another KB with both (nonempty) ABox and (nonempty) TBox. Therefore, the situation in our case is much more complex than previous works. As a result, some tasks such as developing an algorithm for revision become more difficult.

There are several interesting issues remaining for future work. One is to extend our approach to KB revision in other DLs. However, for more expressive DLs, the structure of features will need to be more complex. It would be useful to develop more efficient algorithms for each specific revision operator. A final ongoing problem is to consider applications of our revision operator for ontology evolution in the Semantic Web.

Acknowledgements: This work was partially supported by the Australia Research Council (ARC) under DP0666107 and DP1093652.

References

Artale, A.; Calvanese, D.; Kontchakov, R.; and Zakharyaschev, M. 2007. DL-Lite in the light of first-order logic. In *Proc. AAAI*, 361–366.

Calvanese, D.; Giacomo, G. D.; Lembo, D.; Lenzerini, M.; and Rosati, R. 2006. Data complexity of query answering in description logics. In *Proc. KR*, 260–270.

Calvanese, D.; Giacomo, G. D.; Lembo, D.; Lenzerini, M.; and Rosati, R. 2007. Tractable reasoning and efficient query answering in description logics: The DL-Lite family. *J. Autom. Reasoning* 39(3):385–429.

Giacomo, G. D.; Lenzerini, M.; Poggi, A.; and Rosati, R. 2007. On the approximation of instance level update and erasure in description logics. In *Proc. AAAI*, 403–408.

Konev, B.; Lutz, C.; Ponomaryov, D.; and Wolter, F. 2010. Decomposing description logic ontologies. In *Proc. KR*. To appear.

Kontchakov, R.; Wolter, F.; and Zakharyaschev, M. 2008. Can you tell the difference between DL-Lite ontologies? In *Proc. KR*, 285–295.

Liu, H.; Lutz, C.; Milicic, M.; and Wolter, F. 2006. Updating description logic ABoxes. In *Proc. KR*, 46–56.

Meyer, T.; Lee, K.; and Booth, R. 2005. Knowledge integration for description logics. In *Proc. AAAI*, 645–650.

Parikh, R. 1996. Beliefs, belief revision, and splitting languages. In *Logic, Language and Computation*, 266–278.

Qi, G., and Du, J. 2009. Model-based revision operators for terminologies in description logics. In *Proc. IJCAI*, 891–897.

Satoh, K. 1988. Nonmonotonic reasoning by minimal belief revision. In *Proc. FGCS*, 455–462.

Wang, Z., and Wang, K., and Topor, R., and Pan, J. Z. 2010. Forgetting for knowledge bases in DL-Lite. *Ann. Math. Artif. Intell.* To appear.