Updating ABoxes in *DL-Lite*⋆

Diego Calvanese, Evgeny Kharlamov ⋆⋆, Werner Nutt, and Dmitriy Zheleznyakov

KRDB Research Centre, Free University of Bozen-Bolzano, Italy
calvanese,kharlamov,nutt,zheleznyakov@inf.unibz.it

Abstract. We study the problem of instance level (ABox) updates for Knowledge Bases (KBs) represented in Description Logics of the *DL-Lite* family. *DL-Lite* is at the basis of OWL 2 QL, one of the tractable fragments of OWL 2, the recently proposed revision of the Web Ontology Language. We examine known works on updates that follow the model-based approach and discuss their drawbacks. Specifically, the fact that model-based approaches intrinsically ignore the structural properties of KBs, leads to undesired properties of updates computed according to such semantics. Hence, we propose two novel formula-based approaches, and for each of them we develop a polynomial time algorithm to compute ABox updates for the Description Logic *DL-Lite*$_{\mathcal{FR}}$.

1 Introduction

Ontology languages, and hence Description Logics (DLs), provide excellent mechanisms for representing structured knowledge, and as such they have traditionally been used for modeling at the conceptual level the static and structural aspects of application domains [1]. In general, a DL ontology is structured in two parts, a *TBox* (where 'T' stands for *terminological*) containing general assertions about the domain (i.e., the intensional knowledge, or schema, or constraints), and an *ABox* (where 'A' stands for *assertional*) containing assertions about individual objects (i.e., the extensional knowledge, or data). A family of DLs that has received great attention recently, due to its tight connection with conceptual data models, such as the Entity-Relationship model and UML class diagrams, is the *DL-Lite* family [2]. This family of DLs exhibits nice computational properties, in particular when complexity is measured wrt the size of the data stored in the ABox of a DL ontology [2, 3]. It is also at the basis of the tractable profiles of OWL 2, the forthcoming edition of the W3C standard Web Ontology Language.

The reasoning services that have been investigated for the currently used DLs and implemented in state-of-the-art DL reasoners [4], traditionally focus on so-called standard reasoning, both at the TBox level (e.g., TBox satisfiability, concept satisfiability and subsumption wrt a TBox) and at the ABox level (e.g., knowledge base satisfiability, instance checking and retrieval, and more recently query answering) [5, 6]. Recently, however, the scope of ontologies has broadened, and they are now considered to be not only at the basis of the design and development of information systems, but also for

⋆ This research was partially funded by the European Research Council grant Webdam (under FP7), agreement n.226513.

⋆⋆ The author is co-affiliated with INRIA Saclay.

providing support in the maintenance and evolution phase of such systems. Moreover, ontologies are considered to be the premium mechanism through which services operating in a Web context can be accessed, both by human users and by other services. Supporting all these activities, makes it necessary to equip DL systems with additional kinds of inference tasks that go beyond the traditional ones, most notably that of *ontology update* operations [7, 8]. An update reflects the need of changing an ontology so as to take into account changes that occur in the domain of interest represented by the ontology. In general, such an update is represented by a set of formulas denoting those properties that should be true after the change. In the case where the update causes an undesirable interaction with the knowledge encoded in the ontology, e.g., by causing the ontology to become unsatisfiable, the update cannot simply be added to the ontology. Instead, suitable changes need to be made in the ontology so as to avoid the undesirable interaction, e.g., by deleting parts of the ontology that conflict with the update. Different choices are possible in general, corresponding to different update semantics, which in turn give rise to different update results [9]. Moreover, it is necessary to understand whether the desired update result can be represented at all as a KB in the DL at hand.

Previous work on updates in the context of DL ontologies has addressed ABox (or instance level) update [7, 8, 10], where the update consists of a set of ABox assertions. In [7] the problem is studied for DLs of the *DL-Lite* family, while [8] considers the case of expressive DLs. Both works show that it might be necessary to extend the ontology language with additional features/constructs in order to guarantee that the updated ontology can be represented.

In this paper we also study ABox updates, specifically, for the case of DLs of the *DL-Lite* family. There are two main families of approaches to updates: model and formula-based [11]. In [7], Winslett's semantics, which belongs to the former family is adopted. We reconsider this semantics and show that ABox updates of *DL-Lite* KBs cannot be expressed as a *DL-Lite* KB, in contrast to what claimed in [7]. In general, we argue that the fact that model-based approaches to update ignore structural properties of KBs may make them inappropriate for ABox updates. As a consequence, we explore formula-based approaches to update and propose two novel semantics for ABox updates, called Naive Semantics and Careful Semantics. Under both semantics, the properties of *DL-Lite* guarantee that there is a unique maximal set of assertions that are entailed by the original ABox (and TBox) and that do not conflict with the update. The Careful Semantics also allows us to capture models that are ruled out by the Naive Semantics but are natural. For both semantics, we develop polynomial time algorithms to compute updates for *DL-Lite*$_{\mathcal{FR}}$ KBs.

2 Preliminaries

Description Logics (DLs) [12] are knowledge representation formalisms, tailored for representing the domain of interest in terms of *concepts* and *roles*. In DLs, complex concept and role expressions (or simply, concepts and roles), denoting respectively unary and binary relation, are obtained starting from atomic concepts and roles by applying suitable constructs. Concepts and roles are then used in a DL knowledge base to model the domain of interest. Specifically, a DL *knowledge base* (KB) $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ is formed by two distinct parts, a *TBox* $\mathcal{T}$ representing the *intensional-level* of the KB and an *ABox* $\mathcal{A}$ providing information on the *instance-level* of the KB. In this paper we focus on a family of DLs called *DL-Lite* [2], that corresponds to the tractable OWL 2 QL profile of the Web Ontology Language OWL 2.

The logic of the *DL-Lite* family on which we focus here is *DL-Lite*$_{\mathcal{FR}}$, and when we write just *DL-Lite* we mean *any* language of this family. *DL-Lite*$_{\mathcal{FR}}$ has the following constructs: *(i)* $B ::= A \mid \exists R$, *(ii)* $C ::= B \mid \neg B$, *(iii)* $R ::= P \mid P^-$, where A denotes an *atomic concept*, B a *basic concept*, C a *general concept*, P an *atomic role*, and R a *basic role*. In the following, R^- denotes P^- when $R = P$, and P when $R = P^-$.

A *DL-Lite*$_{\mathcal{FR}}$ ABox $\mathcal{A}$ is a set of *membership assertions* of the form: $B(a)$ and $P(a, b)$, where a and b are constants. The *active domain* of $\mathcal{A}$, denoted *adom*$(\mathcal{A})$, is the set of all constants occurring in $\mathcal{A}$. A *DL-Lite*$_{\mathcal{FR}}$ TBox $\mathcal{T}$ may include *(i) concept inclusion assertions* $B \sqsubseteq C$, *(ii) role functionality assertions* $(\mathsf{funct}\ R)$, and *(iii) role inclusion assertions* $R_1 \sqsubseteq R_2$. The use of assertions *(ii)* and *(iii)* together leads in general to a high complexity of reasoning [13, 3]. It can be avoided by imposing the following restriction [2]: if $R_1 \sqsubseteq R_2$ appears in $\mathcal{T}$, then neither $(\mathsf{funct}\ R_2)$ nor $(\mathsf{funct}\ R_2^-)$ appears in $\mathcal{T}$. Hence, when talking about *DL-Lite*$_{\mathcal{FR}}$ KBs, we assume this syntactic restriction is satisfied. For *DL-Lite*$_{\mathcal{FR}}$, satisfiability of a KB can be checked in polynomial-time in (the size of the) TBox and in logarithmic-space in the size of the ABox [14, 3].

The semantics of a DL is given in terms of *FOL* interpretations. We consider interpretations over a fixed countably infinite set Δ. An *interpretation* $\mathcal{I}$ is a function $\cdot^{\mathcal{I}}$ assigning concepts C to subsets $C^{\mathcal{I}} \subseteq \Delta$, and roles R to binary relations $R^{\mathcal{I}}$ over Δ in such a way that $(\neg B)^{\mathcal{I}} = \Delta \setminus B^{\mathcal{I}}$, $(\exists R)^{\mathcal{I}} = \{a \mid \exists a'.(a, a') \in R^{\mathcal{I}}\}$, and $(P^-)^{\mathcal{I}} = \{(a_2, a_1) \mid (a_1, a_2) \in P^{\mathcal{I}}\}$. Moreover, $\mathcal{I}$ assigns to each constant a an elements of $a^{\mathcal{I}} \in \Delta$, and $a_1^{\mathcal{I}} \neq a_2^{\mathcal{I}}$ whenever $a_1 \neq a_2$, i.e., we adopt the *unique name assumption*. W.l.o.g., we consider each constant a to be interpreted as itself, i.e., $a^{\mathcal{I}} = a$, i.e., we adopt *standard names*. We also consider interpretations as sets of atoms, that is, $A(a) \in \mathcal{I}$ if and only if $a \in A^{\mathcal{I}}$ and $P(a, b) \in \mathcal{I}$ if and only if $(a, b) \in P^{\mathcal{I}}$.

An interpretation $\mathcal{I}$ is a *model* of an assertion $D_1 \sqsubseteq D_2$ if $D_1^{\mathcal{I}} \subseteq D_2^{\mathcal{I}}$, and of $(\mathsf{funct}\ R)$ if R is a function over Δ, i.e., for all $o, o_1, o_2 \in \Delta$ we have that $\{(o, o_1), (o, o_2)\} \subseteq R^{\mathcal{I}}$ implies $o_1 = o_2$. Also, $\mathcal{I}$ is a *model* of a membership assertion $B(a)$ if $a \in B^{\mathcal{I}}$, and of $P(a, b)$ if $(a, b) \in P^{\mathcal{I}}$. For an assertion F, the fact that $\mathcal{I}$ is a model of F is denoted by $\mathcal{I} \models F$. As usual, $\mathcal{I} \models \mathcal{K}$ if $\mathcal{I} \models F$ for every F of $\mathcal{K}$, and *Mod*$(\mathcal{T}, \mathcal{A})$ denotes the set of all such models. A KB is *satisfiable* if *Mod*$(\mathcal{T}, \mathcal{A}) \neq \emptyset$. We write $\mathcal{K} \models F$ if all models of $\mathcal{K}$ are also models of F. Similarly for a set $\mathcal{F}$ of assertions. We say that an ABox $\mathcal{A}$ $\mathcal{T}$*-entails* an ABox $\mathcal{A}'$, denoted $\mathcal{A} \models_{\mathcal{T}} \mathcal{A}'$ if $(\mathcal{T}, \mathcal{A}) \models \mathcal{A}'$.

The deductive *closure of a TBox* $\mathcal{T}$, denoted *cl*$(\mathcal{T})$, is the set of all inclusion and functionality assertions entailed by $\mathcal{T}$ (i.e., $F \in$ *cl*$(\mathcal{T})$ iff $\mathcal{T} \models F$). It is easy to see that in *DL-Lite* *cl*$(\mathcal{T})$ is of quadratic size and computable in quadratic time in the size of $\mathcal{T}$.

3 Problem Definition

Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a KB and $\mathcal{U}$ a set of (TBox and/or ABox) assertions, called an *update*. What we want to study is how to "incorporate" the assertions $\mathcal{U}$ into $\mathcal{K}$, that is, to perform an *update* of $\mathcal{K}$. In this paper we consider only updates at the ABox level (*ABox updates*), that is, when $\mathcal{U}$ consists of ABox assertions only.

Consider for example a registry office's KB, which contains information about marital status of people. Suppose that several married couples got divorced, and the update $\mathcal{U}$ is a list of these newly divorced people. The new data may conflict with some assertions in $\mathcal{A}$. For example, if $\mathcal{T}$ says that nobody can be married to a person who is single, $\mathcal{A}$ says that $John$ is married to $Mary$, and $Mary$ should become single due to $\mathcal{U}$, then $John$ can not be married to $Mary$ anymore. Therefore, in order to take $\mathcal{U}$ on board, one should resolve the conflicts between the old information in $\mathcal{K}$ and the new data in $\mathcal{U}$. This could be done in two ways: by modification of $\mathcal{T}$ or of $\mathcal{A}$. In our example, we can either relax constraints in $\mathcal{T}$, so that even singles keep their spouses (that is, we restructure the ontology, not to mention it is counterintuitive in our example), or delete from $\mathcal{A}$ the information about former spouses of people who are currently single (that is, we update the registry office's database). It seems unintuitive to restructure the ontology whenever new conflicting data arrives. Indeed, in most cases, to resolve conflicts that arise due to changes at the instance level, it is probably more appropriate to change the instance level rather than the intensional level. In other words, we assume that the updated KB $\mathcal{K}'$ should be of the form $\mathcal{K}' = (\mathcal{T}, \mathcal{A}')$ and computing an ABox update results in a new ABox $\mathcal{A}'$ that together with the original TBox $\mathcal{T}$ expresses the result of the update operation. This assumption is also made in [7].

When dealing with updates, both in the knowledge management and the AI communities, it is generally accepted that the updated KB $\mathcal{K}'$ should comply with the principle of *minimality of change* [9, 11], which states that the KB should change as little as possible if new information is incorporated. There are different approaches to updates, suitable for particular applications, and the current belief is there is no general notion of minimality that will "do the right thing" under all circumstances [11]. A number of candidate semantics for update operators have appeared in the literature [11, 15, **?**,17]; they can be classified into two groups: *model-based* and *formula-based*. In order to understand what semantics is more appropriate for ABox updates, we now review both model and formula-based approaches.

4 Model-Based Approach to Semantics

A number of model-based semantics for updates have been proposed in the literature [11]. Poggi et al. [7] proposed to use for ABox updates Winslett's semantics [17], and presented an algorithm, *ComputeUpdate*, to compute the updates for *DL-Lite*$_{\mathcal{F}}$. This work was extended to *DL-Lite*$_{\mathcal{FR}}$ [10]. We now recall the model-based semantics, discuss whether it is suitable for our needs, and reconsider the result produced by *ComputeUpdate*.

Under the model-based paradigm, the objects of change are individual models of $\mathcal{K}$. For a model $\mathcal{I}$ of $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, an update with $\mathcal{U}$ results in a set of models of $\mathcal{U}$ and $\mathcal{T}$. In order to update the KB $\mathcal{K}$ with $\mathcal{U}$, one has to *(i)* update every model $\mathcal{I}$ of $\mathcal{K}$ with $\mathcal{U}$ and then *(ii)* take the set of models that is the union of the sets of resulting models.

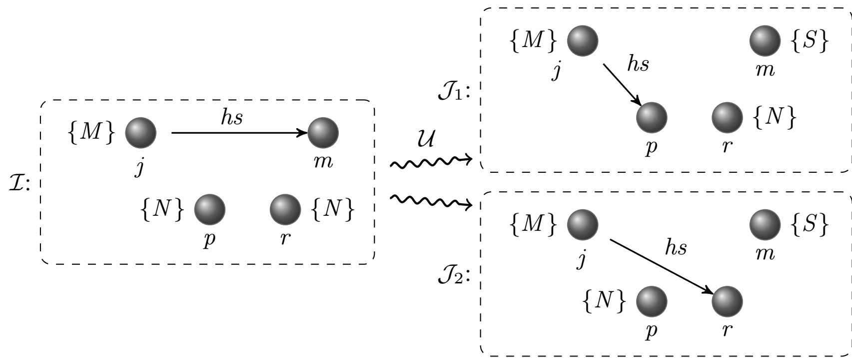

Fig. 1. Winslett's semantics is not expressible in DL-Lite

The *updated interpretation* $\mathcal{I}$ *with* $\mathcal{U}$ *wrt* $\mathcal{T}$, denoted $\textit{w-upd}_{\mathcal{T}}(\mathcal{I},\mathcal{U})$, where w indicates Winslett's semantics, is the set of interpretations defined as follows:

$$\{\mathcal{I}' \mid \mathcal{I}' \in \textit{Mod}(\mathcal{T},\mathcal{U}) \text{ and there is no } \mathcal{I}'' \in \textit{Mod}(\mathcal{T},\mathcal{U}) \text{ s.t. } \mathcal{I} \ominus \mathcal{I}'' \subsetneq \mathcal{I} \ominus \mathcal{I}'\},$$

where containment $\mathcal{I} \subseteq \mathcal{I}'$ and strict containment $\mathcal{I} \subsetneq \mathcal{I}'$ between interpretations are defined as usual (cf. [7]), and $\ominus$ denotes the symmetric difference between sets. Note that the non-existence of $\mathcal{I}''$ in the definition guarantees the minimality of change. Then, the *update of a KB* $(\mathcal{T}, \mathcal{A})$ *with* $\mathcal{U}$ is the set of interpretations

$$\textit{w-upd}(\mathcal{T}, \mathcal{A}, \mathcal{U}) \;=\; \bigcup_{\mathcal{I} \in \textit{Mod}(\mathcal{T},\mathcal{A})} \textit{w-upd}_{\mathcal{T}}(\mathcal{I},\mathcal{U}).$$

Returning to a user the result of an update as a (possibly infinite) set of models is in general not possible. What we want is to return a KB that describes exactly this set of models. We say that a KB $(\mathcal{T}, \mathcal{A}')$ *represents* the update $\textit{w-upd}(\mathcal{T}, \mathcal{A}, \mathcal{U})$ if $\textit{Mod}(\mathcal{T}, \mathcal{A}') = \textit{w-upd}(\mathcal{T}, \mathcal{A}, \mathcal{U})$.

Unfortunately, updates according to Winslett's semantics are inexpressible in *DL-Lite*$_{\mathcal{FR}}$, as illustrated by the following example.

Example 1. Consider a KB that describes a registry office's ontology where a $\textit{married}$ person is a person who $\textit{has}$ a $\textit{spouse}$, and every person who is a spouse is neither a $\textit{single}$ person, nor a $\textit{nun}$. Moreover, $\textit{John}$ is $\textit{married}$ to $\textit{Mary}$, and $\textit{Patty}$ and $\textit{Rachel}$ are nuns. It can be expressed as a *DL-Lite*$_{\mathcal{FR}}$ KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$:

$\mathcal{T}$: $\quad M \sqsubseteq \exists hs, \quad \exists hs \sqsubseteq M, \quad \exists hs^- \sqsubseteq \neg S, \quad \exists hs^- \sqsubseteq \neg N, \quad M \sqsubseteq \neg S;$
$\mathcal{A}$: $\quad M(j), \quad hs(j, m), \quad N(p), \quad N(r);$

where the concept M stands for $\textit{married}$, S for $\textit{single}$, and N for $\textit{nun}$, the role hs for $\textit{hasspouse}$, the constant j for $\textit{John}$, m for $\textit{Mary}$, p for $\textit{Patty}$, and r for $\textit{Rachel}$. In the following we always use this convention for acronyms to save space.

Assume that $\textit{Mary}$ has decided to divorce her husband $\textit{John}$, so the update $\mathcal{U}$ is $\{S(m)\}$. We will show that for $\mathcal{K}$ and $\mathcal{U}$, the resulting update under Winslett's semantics (seen as a set of interpretations) satisfies the disjunction $N(p) \vee N(r)$, but not $N(p) \wedge N(r)$. Hence, each KB representing the update should entail the disjunction, but not any of the single atoms. But this is impossible for a *DL-Lite* KB, because each such KB that entails a disjunction of two atoms, also entails one of the disjuncts. This holds because *DL-Lite* KBs can be expressed by a slightly extended Horn logic (see [**?**]). Hence, this update cannot be represented by a *DL-Lite* ABox.

More precisely, each model $\mathcal{J}$ of the updated KB is of one of the three following kinds: *(i) John* is single, hence $j \notin M^{\mathcal{J}}$, and both *Patty* and *Rachel* are *nuns*, since they were nuns in every model of the original KB and, due to the principle of minimality of change, they have to remain nuns. In other words, $\mathcal{J} \models N(p) \land N(r)$. *(ii) John* is a married person, and his wife is not *Mary*, but another woman, say *Haley*. In this case, both *Patty* and *Rachel* are still *nuns*, as in *(i)*, that is, $\mathcal{J} \models N(p) \land N(r)$ again. *(iii) John* is married and his wife is either *Patty*, or *Rachel*. In this case, his new spouse cannot stay nun any longer, as it is in the models $\mathcal{J}_1$ and $\mathcal{J}_2$ in Figure 1, thus, $\mathcal{J} \models N(p) \lor N(r)$ and also $\mathcal{J}_1 \not\models N(p)$ and $\mathcal{J}_2 \not\models N(r)$. Note that it is not the case that *John* is married to both *Patty* and *Rachel*, since such a model $\mathcal{J}'$ is not minimally different from any model of the original KB. Consequently, *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U}) \models N(p) \lor N(r)$, and *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U}) \not\models N(p)$, and *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U}) \not\models N(r)$. □

From the example we conclude:

Theorem 1. *DL-Lite is not closed wrt ABox updates under Winslett's semantics.*

To discuss the consequences of Example 1 and Theorem 1, we need the following notions. Given a DL KB $\mathcal{K}_0$ and two sets $\mathcal{M}$, $\mathcal{M}^a$ of models of $\mathcal{K}_0$, we say that $\mathcal{M}^a$ is a *complete* (or *sound*) *approximation of* $\mathcal{M}$ if $\mathcal{M}^a \subseteq \mathcal{M}$ (resp., $\mathcal{M} \subseteq \mathcal{M}^a$). Intuitively, completeness (resp., soundness) means that for every FOL formula φ such that $\mathcal{M} \models \varphi$ (resp., $\mathcal{M}^a \models \varphi$), we also have $\mathcal{M}^a \models \varphi$ (resp., $\mathcal{M} \models \varphi$). According to this notion, we say that an algorithm that computes updates is complete (or sound), if *Mod*$(\mathcal{T}, \mathcal{A}')$ is a complete (resp., sound) approximation of *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$, where $\mathcal{A}'$ is the output of the algorithm and $\mathcal{A}' \models \mathcal{U}$. Note that both *Mod*$(\mathcal{T}, \mathcal{A}')$ and *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$ are sets of models of the KB $(\mathcal{T}, \mathcal{U})$.

A consequence of Theorem 1 is that the algorithm *ComputeUpdate*, which always outputs a *DL-Lite* ABox, cannot capture Winslett's semantics exactly. As shown below, *ComputeUpdate* is *neither* sound, *nor* complete. First, we briefly remind the reader of the algorithm.

The algorithm *ComputeUpdate* takes as input a TBox $\mathcal{T}$, an ABox $\mathcal{A}$, and an update $\mathcal{U}$, such that both $(\mathcal{T}, \mathcal{A})$ and $(\mathcal{T}, \mathcal{U})$ are satisfiable. Then, it computes first the set $\mathcal{N}$ of "contradictive" membership assertions F such that $\mathcal{U} \models_{\mathcal{T}} \neg F$. Using $\mathcal{N}$, it constructs a set $\mathcal{A}'$ as follows: *(i)* it initializes $\mathcal{A}'$ to $\mathcal{A} \cup \mathcal{U}$; *(ii)* it deletes from $\mathcal{A}'$ each assertion F in $\mathcal{N}$, and for each such F, it inserts into $\mathcal{A}'$ all assertions F' such that $F \models_{\mathcal{T}} F'$ and $F' \notin \mathcal{N}$; *(iii)* if $R(a, b)$ has been deleted, then $\exists R(a)$ and $\exists R^-(b)$ are deleted as well. The algorithm outputs $\mathcal{A}'$.

To show that *ComputeUpdate* is unsound, consider the KB and the update from Example 1, and compute the updated KB $\mathcal{K}'$ using the algorithm. It is easy to see that $\mathcal{K}'$ has the ABox consisting of $N(p)$, $N(r)$, $M(j)$, $S(m)$ and the old TBox $\mathcal{T}$, whereas, according to Winslett's semantics in every model of *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$ either $N(p)$ or $N(r)$ does not hold. Hence, *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U}) \not\subseteq$ *Mod*$(N(p), N(r))$, and we conclude that *ComputeUpdate* is unsound.

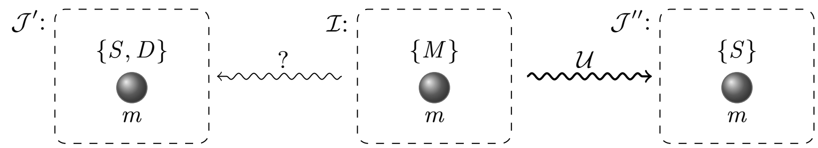

Fig. 2. Incompleteness of *ComputeUpdate* algorithm wrt Winslett's semantics

The following example shows the incompleteness.

Example 2. Consider the following *DL-Lite* KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ and update $\mathcal{U}$:

$$\mathcal{T}: \quad M \sqsubseteq \neg S, \quad D \sqsubseteq S; \qquad\qquad \mathcal{A}: \quad M(m); \qquad\qquad \mathcal{U}: \quad S(m),$$

where D stands for $delighted$ person. The call *ComputeUpdate*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$ returns the new ABox $\mathcal{A}' = \{S(m)\}$.

We will show that there is a model $\mathcal{J}'$ of $(\mathcal{T}, \mathcal{A}')$ such that $\mathcal{J}' \notin$ *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$. We define $\mathcal{J}'$ as sketched in Figure 2, that is, $Mary$ is $single$ and $delighted$. Clearly, $\mathcal{J}'$ is a model of $\mathcal{T}$ and $\mathcal{A}'$.

Now, assume there is $\mathcal{I} \in$ *Mod*$(\mathcal{T}, \mathcal{A})$ such that $\mathcal{I} \ominus \mathcal{J}'$ is minimal, that is, there is no $\mathcal{J}'' \in$ *Mod*$(\mathcal{T}, \mathcal{U})$ such that $\mathcal{I} \ominus \mathcal{J}'' \subsetneq \mathcal{I} \ominus \mathcal{J}'$. We first conclude that $m \in M^{\mathcal{I}}$, since it is stated by the original ABox, so $m \notin S^{\mathcal{I}}$ and $m \notin D^{\mathcal{I}}$ because of $\mathcal{T}$. Additionally, we can conclude that $m \notin C^{\mathcal{I}}$ for every concept C, different from M, S, and D, otherwise m would be in $C^{\mathcal{J}'}$ according to the principle of minimality of change. Consider now the interpretation $\mathcal{J}'' := \mathcal{J}' \setminus \{D(m)\}$, which is obtained from $\mathcal{J}'$ by dropping m from D. Then we have that $\mathcal{I} \ominus \mathcal{J}'' = \{M(m), S(m)\} \subsetneq \{M(m), S(m), D(m)\} = \mathcal{I} \ominus \mathcal{J}'$. Notice that $\mathcal{J}'' \in$ *Mod*$(\mathcal{T}, \mathcal{U})$ as well. Hence, the symmetric difference $\mathcal{I} \ominus \mathcal{J}'$ is not minimal and $\mathcal{J}' \notin$ *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$. $\square$

Since *ComputeUpdate* is neither sound nor complete, we cannot use it for computing either sound or complete approximations of ABox updates for *DL-Lite* wrt Winslett's semantics.

To sum-up on Winsletts's semantics, we conclude that the semantics cannot be expressed by *DL-Lite* KBs and hence is inconvenient to adopt in the context of *DL-Lite*. In principle, since all model-based approaches focus on the models of a KB and not on the actual structure of the KB, they suffer from flexibility when dropping the assertions as far as some notion of minimality of change between models is kept (in Example 1, the assertion $N(p)$ was dropped from $\mathcal{J}_1$). These approaches are agnostic to the structural properties of KBs, which may lead to undesired results.

5 Formula-Based Approaches to Semantics

As a consequence of the observations made in the previous section, we consider now formula-based approaches for establishing the semantics of updates in *DL-Lite* KBs.

5.1 Naive Semantics for Updates

A straightforward way of proceeding is to try to keep as many assertions as possible that belong to the ABox or that are entailed by it via the TBox, without contradicting the update. We start with a crucial observation that is at the heart of this approach.

INPUT: TBox $\mathcal{T}$, and ABoxes $\mathcal{A}, \mathcal{N}$ each satisfiable with $\mathcal{T}$
OUTPUT: finite set of membership assertions $\mathcal{A}^w$
[1] $\mathcal{A}^w := \mathcal{A}$
[2] for each $B_1(c) \in \mathcal{N}$ **do**
[3] $\quad \mathcal{A}^w := \mathcal{A}^w \setminus \{B_1(c)\}$ **and**
[4] $\quad$ **for each** $B_2 \sqsubseteq B_1 \in \mathit{cl}(\mathcal{T})$ **do** $\mathcal{A}^w := \mathcal{A}^w \setminus \{B_2(c)\}$
[5] $\quad\quad$ **if** $B_2(c) = \exists R(c)$ **then**
[6] $\quad\quad\quad$ **for each** $R(c,d) \in \mathcal{A}^w$ **do** $\mathcal{N} := \mathcal{N} \cup \{R(c,d)\}$
[7] for each $R_1(a,b) \in \mathcal{N}$ **do**
[8] $\quad \mathcal{A}^w := \mathcal{A}^w \setminus \{R_1(a,b)\}$ **and**
[9] $\quad$ **for each** $R_2 \sqsubseteq R_1 \in \mathit{cl}(\mathcal{T})$ **do** $\mathcal{A}^w := \mathcal{A}^w \setminus \{R_2(a,b)\}$

Fig. 3. Algorithm *Weeding*$(\mathcal{T}, \mathcal{A}, \mathcal{N})$ for *DL-Lite*$_{\mathcal{FR}}$

Lemma 1. *Let $(\mathcal{T}, \mathcal{A})$ be a DL-Lite KB. If $(\mathcal{T}, \mathcal{A})$ is unsatisfiable, then there is a subset $\mathcal{A}_0 \subseteq \mathcal{A}$ with at most two elements, such that $(\mathcal{T}, \mathcal{A}_0)$ is unsatisfiable.*

In other words, in *DL-Lite*, unsatisfiability of an ABox wrt a TBox is caused either by a single ABox assertion, which will be a membership assertion for an unsatisfiable concept or role, or by a pair of assertions contradicting either a disjointness or a functionality assertion of the TBox.

Let $\mathcal{T}$ be a TBox, $\mathcal{A}$ an ABox, and $\mathcal{U}$ an update. We say that $\mathcal{A}$ *is $\mathcal{T}$-compatible with* $\mathcal{U}$ if $(\mathcal{T}, \mathcal{A} \cup \mathcal{U})$ is satisfiable. We also make use of the notion of *closure of an ABox $\mathcal{A}$ wrt a TBox $\mathcal{T}$*, denoted $\mathit{cl}_{\mathcal{T}}(\mathcal{A})$, which is the set of all membership assertions F over the constants in $\mathit{adom}(\mathcal{A})$ such that $(\mathcal{T}, \mathcal{A}) \models F$. Then, Lemma 2 implies the following result.

Theorem 2. *Let $\mathcal{T}$ be a DL-Lite TBox, $\mathcal{A}$ an ABox, and $\mathcal{U}$ an update, and suppose that both $(\mathcal{T}, \mathcal{A})$ and $(\mathcal{T}, \mathcal{U})$ are satisfiable. Then, the set*

$$\{\mathcal{A}_0 \subseteq \mathit{cl}_{\mathcal{T}}(\mathcal{A}) \mid \mathcal{A}_0 \textit{ is } \mathcal{T}\textit{-compatible with } \mathcal{U}\}$$

has a unique maximal element wrt set inclusion.

Thus, for every ABox $\mathcal{A}$ and every update $\mathcal{U}$, we can find a maximal set $\mathcal{A}^n_m$ of assertions (where m stands for maximal and n for naive) in $\mathit{cl}_{\mathcal{T}}(\mathcal{A})$, such that $\mathcal{A}^n_m$ is $\mathcal{T}$-compatible with $\mathcal{U}$. Based on this, we define the *naive update of* $(\mathcal{T}, \mathcal{A})$ *by* $\mathcal{U}$ as

$$\textit{n-upd}(\mathcal{T}, \mathcal{A}, \mathcal{U}) \;:=\; \mathcal{A}^n_m \cup \mathcal{U}.$$

We exhibit now the algorithm *NaiveUpdate*, which computes *n-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$. It exploits the algorithm *Weeding* (see Figure 3), which takes as inputs a TBox $\mathcal{T}$, an ABox $\mathcal{A}$, and a set $\mathcal{N}$ of membership assertions to be "deleted" from $\mathcal{A}$. It deletes from $\mathcal{A}$ (Lines [2]–[4]) all concept membership assertions $B_1(c) \in \mathcal{N}$ and also those membership assertions $B_2(c)$ that $\mathcal{T}$-entail $B_1(c)$. Then (Lines [5]–[6]), it detects role membership assertions that $\mathcal{T}$-entail $B_1(c)$ and adds them to $\mathcal{N}$. Finally (Lines [7]–[9]), it deletes from the remaining $\mathcal{A}$ all membership assertions $R_1(a,b) \in \mathcal{N}$ and also those

INPUT: TBox $\mathcal{T}$, and ABoxes $\mathcal{A}, \mathcal{U}$ each satisfiable with $\mathcal{T}$
OUTPUT: finite set of membership assertions $\mathcal{A}^u$
[1] $\mathcal{A}^u := cl_{\mathcal{T}}(\mathcal{A} \cup \mathcal{U}), \mathcal{U} := cl^{\mathcal{T}}(\mathcal{U}),\ CA := \emptyset$
[2] **for each** $B \sqsubseteq \neg B' \in cl(\mathcal{T})$ **do**
[3]   **if** $\{B(c), B'(c)\} \subseteq \mathcal{A}^u$ **then**
[4]   **if** $B(c) \notin \mathcal{U}$ **then** $CA := CA \cup \{B(c)\}$
[5]   **otherwise** $CA := CA \cup \{B'(c)\}$
[6] **for each** $(\mathsf{funct}\ R) \in \mathcal{T}$ **do**
[7]   **if** $\{R(a,b), R(a,c)\} \subseteq \mathcal{A}^u$ **then**
[8]   **if** $R(a,b) \notin \mathcal{U}$ **then** $CA := CA \cup \{R(a,b)\}$
[9]   **otherwise** $CA := CA \cup \{R(a,c)\}$
[10] $\mathcal{A}^u :=$ *Weeding*$(\mathcal{T}, \mathcal{A}^u, CA)$

Fig. 4. Algorithm *NaiveUpdate*$(\mathcal{A},\mathcal{U},\mathcal{T})$ for *DL-Lite*$_{\mathcal{FR}}$

assertions $R_2(a,b)$ that $\mathcal{T}$-entail $R_1(a,b)$. As the result, *Weeding* returns the subset of $\mathcal{A}$ that does not contain assertions from $\mathcal{N}$ and assertions that $\mathcal{T}$-entail one of them.

The algorithm *NaiveUpdate* (see Figure 4) takes as inputs an ABox $\mathcal{A}$, an update $\mathcal{U}$, and a TBox $\mathcal{T}$. It detects (Lines [1]–[9]) conflicting assertions in the closure of $\mathcal{A} \cup \mathcal{U}$ and stores them in CA: it first detects conflicts of the form $B(c)$ and $\neg B(c)$ (Lines [2]–[5]), and then of the form $R(a,b)$ and $R(a,c)$ for a functional role R (Lines [7]–[9]). Finally, the algorithm resolves the detected conflicts by means of *Weeding* (Line [10]). We conclude with the correctness of the algorithm.

Theorem 3. *Let $(\mathcal{T},\mathcal{A})$ be a satisfiable DL-Lite$_{\mathcal{FR}}$ KB and $\mathcal{U}$ an update such that $(\mathcal{T},\mathcal{U})$ is satisfiable. Then NaiveUpdate$(\mathcal{T},\mathcal{A},\mathcal{U})$ runs in polynomial time in the size of $\mathcal{T} \cup \mathcal{A} \cup \mathcal{U}$, and NaiveUpdate$(\mathcal{T},\mathcal{A},\mathcal{U}) =$ n-upd$(\mathcal{T},\mathcal{A},\mathcal{U})$.*

5.2 Careful Semantics for Updates

We start with an example that illustrates some drawbacks of Naive Semantics.

Example 3. Continuing our example with $John$ and his spouse $Mary$, consider

$$\mathcal{T}: \quad \exists hs^- \sqsubseteq \neg S,\ \exists hs \sqsubseteq M,\ M \sqsubseteq \exists hs; \qquad \mathcal{A}: \quad M(j), hs(j,m).$$

Assume $Mary$ decided to divorce $John$, so the update $\mathcal{U}$ is $\{S(m)\}$. Consider the model $\mathcal{I}$ of $\mathcal{K}$ (see Figure 5) that precisely reflects what is in the ABox of the KB, and let us update $\mathcal{I}$. Since *Mary* is single now, the state of *John* is changed and he has not spouse *Mary* anymore. We do not have any explicit information about the situation with *John*, hence, we can make two assumptions. We can assume that *John* now is single too. We can also assume that *John* still $has\ a\ spouse$, some other woman, say *Haley* (denoted as h). The former situation is reflected by the model $\mathcal{J}_1$, while the latter one by the model $\mathcal{J}_2$ in Figure 5.

Note that the Naive Semantics (and, hence, the output of the algorithm *NaiveUpdate*) captures only the model $\mathcal{J}_2$, while we might be interested in a semantics that captures both possibilities. □

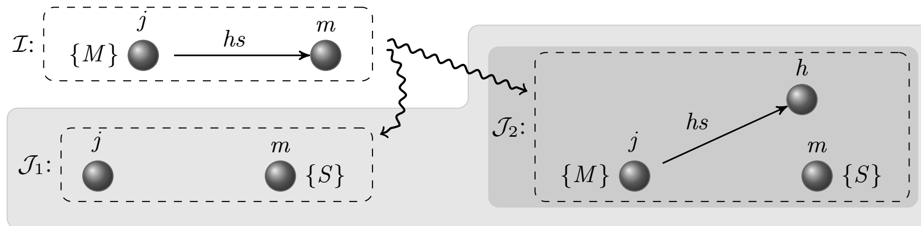

Fig. 5. Comparison of the *NaiveUpdate* and *CarefulUpdate* algorithms. Darker gray color: models captured by *NaiveUpdate*; lighter gray color: models captured by *CarefulUpdate*.

As we see from the example, the situation when the updated KB entails *unexpected information*, that is, information coming neither from the original KB, nor from the update, may be counterintuitive. In our example, the unexpected information is $\exists x(hs(j,x) \wedge (x \neq m))$, saying that $John$ has a spouse different from $Mary$. This information has a specific form: it restricts the possible values in the second component of the role hs. Our next semantics prohibits these role restrictions from being unexpectedly entailed from the updated KB.

We say that a formula is *role-constraining* if it is of the form $\exists x(R(a,x) \wedge (x \neq c_1) \wedge \cdots \wedge (x \neq c_n))$, where a and all c_is are constants. Let $\mathcal{T}$ be a TBox, $\mathcal{A}$ an ABox, and $\mathcal{U}$ an update. A subset $\mathcal{A}_1 \subseteq \mathcal{A}$ is *careful* if for every role-constraining formula φ, if $\mathcal{A}_1 \cup \mathcal{U} \models_{\mathcal{T}} \varphi$ holds, then either $\mathcal{A} \models_{\mathcal{T}} \varphi$ or $\mathcal{U} \models_{\mathcal{T}} \varphi$ holds.

Theorem 4. *Let $\mathcal{T}$ be a DL-Lite TBox, $\mathcal{A}$ an ABox, and $\mathcal{U}$ an update, and suppose that both $(\mathcal{T},\mathcal{A})$ and $(\mathcal{T},\mathcal{U})$ are satisfiable. Then, the set*

$$\{\mathcal{A}_0 \subseteq cl_{\mathcal{T}}(\mathcal{A}) \mid \mathcal{A}_0 \textit{ is careful and } \mathcal{A}_0 \cup \mathcal{U} \textit{ is satisfiable}\}$$

has a unique maximal element wrt set inclusion.

Similarly, as to what we did for the Naive Semantics, we can exploit the maximal set $\mathcal{A}^c_m$ of assertions (where c stands for careful), whose uniqueness is guaranteed by Theorem 4, to define the *careful update of* $(\mathcal{T},\mathcal{A})$ *by* $\mathcal{U}$ as

$$c\text{-}upd(\mathcal{T},\mathcal{A},\mathcal{U}) := \mathcal{A}^c_m \cup \mathcal{U}.$$

We exhibit now the algorithm *CarefulUpdate*, which computes $c\text{-}upd(\mathcal{T},\mathcal{A},\mathcal{U})$. The *prechase of an ABox $\mathcal{A}$ wrt a TBox $\mathcal{T}$*, denoted $Prechase_{\mathcal{T}}(\mathcal{A})$, is a subset of $cl_{\mathcal{T}}(\mathcal{A})$ obtained as follows: one removes from $cl_{\mathcal{T}}(\mathcal{A})$ all the assertions of the form $\exists R(a)$, whenever there is an assertion of the form $R(a,c)$ in $cl_{\mathcal{T}}(\mathcal{A})$, for some constant c. The prechase is needed to detect unexpected role-restricting formulas. The algorithm *CarefulUpdate* (see Figure 6) takes as input an ABox $\mathcal{A}$, an update $\mathcal{U}$, and a TBox $\mathcal{T}$. It first computes the update operation wrt Naive Semantics (Line [1]). Then, it computes the set *UF* of assertions, that cause unexpectedness in *NaiveUpdate*$(\mathcal{A},\mathcal{U},\mathcal{T})$, in two stages: first it adds to *UF* all $C(a)$ coming exclusively from $\mathcal{A}$ for which there is $\exists R(a)$ coming from $\mathcal{U}$, such that $C(a)$ and $\exists R(a)$ together yields unexpectedness (Lines [2]–[6]); second it adds to *UF* all $\exists R(a)$ coming exclusively from $\mathcal{A}$ for which there is

INPUT: TBox $\mathcal{T}$, and ABoxes $\mathcal{A}, \mathcal{U}$ each satisfiable with $\mathcal{T}$
OUTPUT: finite set of membership assertions $\mathcal{A}^u$
[1] $\mathcal{A}^u := \mathit{NaiveUpdate}(\mathcal{A}, \mathcal{U}, \mathcal{T})$, $\mathit{UF} := \emptyset$
[2] **for each** $\exists R(a) \in \mathit{Prechase}(\mathcal{U})$ **do**
[3]   **if** $R(a,b) \notin \mathcal{A}^u$ for any b **then**
[4]   **for each** $\exists R \sqsubseteq \neg C \in \mathit{cl}(T)$ **do**
[5]   **for each** $C(a) \in \mathit{cl}_{\mathcal{T}}(\mathcal{A}) \setminus \mathit{cl}_{\mathcal{T}}(\mathcal{U})$ **do**
[6] $\mathit{UF} := \mathit{UF} \cup \{C(a)\}$
[7] **for each** $\exists R(a) \in \mathcal{A}^u \setminus \mathit{Prechase}(\mathcal{U})$ **do**
[8]   **if** $R(a,b) \notin \mathcal{A}^u$ for every b **then**
[9]   **if** $(\exists R \sqsubseteq \neg C) \in \mathit{cl}(\mathcal{T})$ **and** $C(a) \in \mathit{cl}_{\mathcal{T}}(\mathcal{U}) \setminus \mathit{cl}_{\mathcal{T}}(\mathcal{A})$ for some a **then**
[10] $\mathit{UF} := \mathit{UF} \cup \{\exists R(a)\}$
[11] $\mathcal{A}^u := \mathit{Weeding}(\mathcal{T}, \mathcal{A}^u, \mathit{UF})$

Fig. 6. Algorithm *CarefulUpdate*$(\mathcal{A}, \mathcal{U}, \mathcal{T})$ for *DL-Lite*$_{\mathcal{FR}}$

$C(a)$ coming exclusively from $\mathcal{U}$, such that $\exists R(a)$ and $C(a)$ together yields unexpectedness (Lines [7]–[10]). Finally it removes *UF* from *NaiveUpdate*$(\mathcal{A}, \mathcal{U}, \mathcal{T})$ by means of *Weeding* (Line [11]). We conclude with the correctness of the algorithm.

Theorem 5. *Let $(\mathcal{T}, \mathcal{A})$ be a satisfiable DL-Lite$_{\mathcal{FR}}$ KB and $\mathcal{U}$ an update such that $(\mathcal{T}, \mathcal{U})$ is satisfiable. Then, CarefulUpdate$(\mathcal{T}, \mathcal{A}, \mathcal{U})$ runs in polynomial time in the size of $\mathcal{T} \cup \mathcal{A} \cup \mathcal{U}$, and CarefulUpdate$(\mathcal{T}, \mathcal{A}, \mathcal{U})$ = c-upd$(\mathcal{T}, \mathcal{A}, \mathcal{U})$.*

5.3 Comparison of Semantics

We now discuss the differences between the Naive and Careful Semantics. The crucial difference is in how the semantics "preserve" entailment of formulas. For disambiguation we denote an updated ABox $\mathcal{A}'$ computed under Naive Semantics as $\mathcal{A}'_n$ and under Careful Semantics as $\mathcal{A}'_c$.

Proposition 1. *Let $(\mathcal{T}, \mathcal{A})$ be a DL-Lite$_{\mathcal{FR}}$ KB and $\mathcal{U}$ an update. If $\mathcal{U} \models_{\mathcal{T}} \varphi$, then $\mathcal{A}'_n \models_{\mathcal{T}} \varphi$ and $\mathcal{A}'_c \models_{\mathcal{T}} \varphi$, for any FOL formula φ. If $\mathcal{U} \not\models_{\mathcal{T}} \varphi$ and $\mathcal{U} \not\models_{\mathcal{T}} \neg\varphi$, then*
(i) if $\mathcal{A} \models_{\mathcal{T}} \varphi$, then $\mathcal{A}'_n \models_{\mathcal{T}} \varphi$, when φ is an ABox assertion, and
(ii) if $\mathcal{A} \not\models_{\mathcal{T}} \varphi$, then $\mathcal{A}'_c \not\models_{\mathcal{T}} \varphi$, when φ is an ABox assertion or role-constraining.

Note that *Mod*$(\mathcal{T}, \mathcal{A}_n) \subseteq$ *Mod*$(\mathcal{T}, \mathcal{A}_c)$ and *Mod*$(\mathcal{T}, \mathcal{A}_n)$ intersects with *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$, but in general neither *Mod*$(\mathcal{T}, \mathcal{A}_c) \subseteq$ *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$, nor *w-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U}) \subseteq$ *Mod*$(\mathcal{T}, \mathcal{A}_c)$ holds.

Updating KBs Without Projections. We say that a *DL-Lite*$_{\mathcal{FR}}$ KB is *projection-free* if no concept inclusion and no concept membership assertion contains an occurrence of a role symbol, that is, domains and ranges of roles do not appear in the KB (though role inclusions may appear). The following theorem shows that for such KBs all semantics we have considered coincide and the *ComputeUpdate* algorithm of [7] is correct.

Theorem 6. *For projection-free DL-Lite$_{\mathcal{FR}}$ KBs, Winslett's, Naive, and Careful semantics of ABox updates coincide, and ComputeUpdate computes the updates correctly.*

6 Conclusion

We have studied ABox updates for Description Logics of the *DL-Lite* family. There are two main families of approaches to updates: model and formula-based. We examined the former family and concluded that these approaches are not fully appropriate for ABox updates, since the updates are in general not expressible in *DL-Lite*. Thus, we examined formula-based approaches and proposed two novel semantics for ABox updates, called Naive Semantics and Careful Semantics, which are closed for DL-Lite. For both, we developed polynomial time algorithms to compute updates in $\textit{DL-Lite}_{\mathcal{FR}}$.

References

1. Borgida, A., Brachman, R.J.: Conceptual modeling with description logics. [12] chapter 10 349–372
2. Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Rosati, R.: Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. J. of Automated Reasoning **39**(3) (2007) 385–429
3. Artale, A., Calvanese, D., Kontchakov, R., Zakharyaschev, M.: The *DL-Lite* family and relations. J. of Artificial Intelligence Research **36** (2009) 1–69
4. Möller, R., Haarslev, V.: Description logic systems. [12] chapter 8 282–305
5. Haarslev, V., Möller, R.: On the scalability of description logic instance retrieval. J. of Automated Reasoning **41**(2) (2008) 99–142
6. Calvanese, D., De Giacomo, G., Lenzerini, M.: Conjunctive query containment and answering under description logics constraints. ACM Trans. on Computational Logic **9**(3) (2008) 22.1–22.31
7. De Giacomo, G., Lenzerini, M., Poggi, A., Rosati, R.: On instance-level update and erasure in description logic ontologies. J. of Log. and Comp., Special Issue on Ontology Dynamics **19**(5) (2009) 745–770
8. Liu, H., Lutz, C., Milicic, M., Wolter, F.: Updating description logic ABoxes. In: Proc. of KR 2006. (2006) 46–56
9. Eiter, T., Gottlob, G.: On the complexity of propositional knowledge base revision, updates and counterfactuals. Artificial Intelligence **57** (1992) 227–270
10. Zheleznyakov, D.: Updating description logic knowledge bases. Master's thesis, Faculty of Computer Science, Free University of Bozen-Bolzano, Italy (2009)
11. Winslett, M.: Updating Logical Databases. Cambridge University Press (1990)
12. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P.F., eds.: The Description Logic Handbook: Theory, Implementation and Applications. Cambridge University Press (2003)
13. Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Rosati, R.: Data complexity of query answering in description logics. In: Proc. of KR 2006. (2006) 260–270
14. Poggi, A., Lembo, D., Calvanese, D., De Giacomo, G., Lenzerini, M., Rosati, R.: Linking data to ontologies. J. on Data Semantics **X** (2008) 133–173
15. Abiteboul, S., Grahne, G.: Update semantics for incomplete databases. In: Proc. of VLDB'85. (1985)
16. Ginsberg, M.L., Smith, D.E.: Reasoning about action I: A possible worlds approach. Technical Report KSL-86-65, Knowledge Systems, AI Laboratory (1987)
17. Winslett, M.: A model-based approach to updating databases with incomplete information. ACM Trans. on Database Systems **13**(2) (1988) 167–196
18. Calvanese, D., Kharlamov, E., Nutt, W.: A Proof Theory for DL-Lite. In: Proc. DL Workshop, Brixen, Italy (2007) Conference without formal proceedings.

7 Proofs for Section 5.1

Lemma 2. *Let $(\mathcal{T}, \mathcal{A})$ be a DL-Lite KB. If $(\mathcal{T}, \mathcal{A})$ is unsatisfiable, then there is a subset $\mathcal{A}_0 \subseteq \mathcal{A}$ with at most two elements, such that $(\mathcal{T}, \mathcal{A}_0)$ is unsatisfiable.*

Proof (Sketch). To see this, we consider first TBoxes without functionality assertions. Then $\mathcal{K}$ is satisfiable if and only if the Skolemized version of $\mathcal{K}$, say $\mathcal{K}'$, is satisfiable. The set $\mathcal{K}'$ is a set of Horn clauses. It is known that, if such a set is unsatisfiable, there is a subset with exactly one negative clause that is unsatisfiable. The only negative clauses in $\mathcal{K}'$ are denials that stem from disjointness axioms of the form $A \sqsubseteq \neg B$. Suppose that $\mathcal{K}'$ is unsatisfiable and that $\mathcal{K}'_0$ is an unsatisfiable subset with at most one negative clause stemming from $B \sqsubseteq \neg C$. Then, because of the form of the assertions in $\mathcal{T}$, there are membership assertions $B_0(a), C_0(a)$ in $\mathcal{K}'_0$, for unary concepts B_0 and C_0 such that $\mathcal{T} \models B0 \sqsubseteq B$ and $\mathcal{T} \models C_0 \sqsubseteq C$. (Actually, instead of $B_0(a)$ and $C_0(a)$, we can also have assertions of the form $P_1(a, x)$, $P_2(a, y)$.)

If we consider also functionality assertions, the claim still holds. In this case, the only relevant negative clauses are the ones of the form $a \neq b$, for distinct constants a and b, which are given implicitly, due to the unique name assumption. They can lead to a contradiction by the interplay with functionality and role membership assertions. Again, no more than two role membership assertions are needed to give rise to a contradiction. □

Theorem 7. *Let $\mathcal{T}$ be a DL-Lite TBox, $\mathcal{A}$ an ABox, and $\mathcal{U}$ an update, and both $(\mathcal{T}, \mathcal{A})$ and $(\mathcal{T},\mathcal{U})$ are satisfiable. Then, the set*

$$\{\mathcal{A}_0 \subseteq cl_{\mathcal{T}}(\mathcal{A}) \mid \mathcal{A}_0 \textit{ is } \mathcal{T}\textit{-compatible with } \mathcal{U}\}$$

has a unique maximal element wrt set inclusion.

Proof. We can construct the unique maximal element $\mathcal{A}_m$ as

$$\mathcal{A}_m = \{F \in cl_{\mathcal{T}}(\mathcal{A}) \mid \{F\} \text{ is } \mathcal{T}\text{-compatible with } \mathcal{U}\}.$$

To prove that $\mathcal{A}_m$ is $\mathcal{T}$-compatible with $\mathcal{U}$, by Lemma 2, we only have to show that for each $\mathcal{F} \subseteq \mathcal{A}_m \cup \mathcal{U}$ of at most two elements, $(\mathcal{T}, \mathcal{F})$ is satisfiable. If $\mathcal{F} = \{F\}$, then either $F \in cl_{\mathcal{T}}(\mathcal{A})$ or $F \in \mathcal{U}$, and since both $(\mathcal{T}, \mathcal{A})$ and $(\mathcal{T},\mathcal{U})$ are satisfiable, also $(\mathcal{T}, \{F\})$ is satisfiable. Let's consider the case where $\mathcal{F} = \{F_1, F_2\}$. If $\mathcal{F} \subseteq cl_{\mathcal{T}}(\mathcal{A})$ or $\mathcal{F} \subseteq \mathcal{U}$, we can argue as before. Instead, if $F_1 \in cl_{\mathcal{T}}(\mathcal{A})$ and $F_2 \in \mathcal{U}$, then $(\mathcal{T}, \mathcal{F})$ is satisfiable because by definition of $\mathcal{A}_m$ we have that $\{F_1\}$ is $\mathcal{T}$-compatible with $\mathcal{U}$.

To see that $\mathcal{A}_m$ is also maximal, assume that $A_m \subseteq \mathcal{A}' \subseteq \mathcal{A}$ and that also $(\mathcal{T}, \mathcal{A}')$ is satisfiable. Then, for every $F \in \mathcal{A}'$, we have that $\{F\}$ is $\mathcal{T}$-compatible with $\mathcal{U}$, hence $F \in \mathcal{A}_m$. Thus $\mathcal{A}' \subseteq \mathcal{A}_m$. □

Theorem 8. *Let $(\mathcal{T}, \mathcal{A})$ be a satisfiable DL-Lite$_{\mathcal{FR}}$ KB and $\mathcal{U}$ an update such that $(\mathcal{T},\mathcal{U})$ is satisfiable. Then NaiveUpdate$(\mathcal{T}, \mathcal{A},\mathcal{U})$ runs in polynomial time in the size of $\mathcal{T} \cup \mathcal{A} \cup \mathcal{U}$, and NaiveUpdate$(\mathcal{T}, \mathcal{A},\mathcal{U}) \equiv_{\mathcal{T}}$ n-upd$(\mathcal{T}, \mathcal{A},\mathcal{U})$.*

Proof (Sketch). The fact that *NaiveUpdate* returns a maximal non-contradicting set of assertions follows from the construction of the algorithm. In Lines [1]–[9] of *NaiveUpdate*, all atoms that are responsible for unsatisfiability of $(\mathcal{T}, \mathcal{A} \cup \mathcal{U})$ are detected and deleted from $\mathcal{A}$ by means of *Weeding*. The resulting set, say $\mathcal{S}$, is non-contradicting by construction. The set is also maximal, because if one adds to $\mathcal{S}$ any of the assertions of $cl_\mathcal{T}(\mathcal{A})$ deleted by *Weeding*, then this assertion, say F, will $\mathcal{T}$-entail either an assertion $B(c)$ such that $\neg B(c)$ is $\mathcal{T}$-entailed by $\mathcal{U}$, or $R(a,b)$ such that $R(a,c)$ is $\mathcal{T}$-entailed by $\mathcal{U}$ and R is functional in $\mathcal{T}$. Hence, we would get a set $\mathcal{S} \cup \{F\}$ such that $(\mathcal{T}, \mathcal{S} \cup \{F\})$ is unsatisfiable. Polynomiality of *NaiveUpdate* follows from polynomiality of computing $cl(\mathcal{T})$ and $cl_\mathcal{T}(\mathcal{A})$ and of *Weeding*. □

8 Proofs for Section 5.2

Definition 1. *A formula φ is* unexpected *for $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{T}$ if $\mathcal{A}_1 \cup \mathcal{A}_2 \models_\mathcal{T} \varphi$ and $\mathcal{A}_1 \not\models_\mathcal{T} \varphi$ and $\mathcal{A}_2 \not\models_\mathcal{T} \varphi$.*

Theorem 9. *Let $(\mathcal{T}, \mathcal{A})$ be a satisfiable DL-Lite KB and $\mathcal{U}$ be an update. Then the set*

$$\{\mathcal{A}_0 \subseteq cl_\mathcal{T}(\mathcal{A}) \mid \mathcal{A}_0 \text{ is careful and } \mathcal{A}_0 \cup \mathcal{U} \text{ is satisfiable}\}$$

has a unique maximal element wrt set inclusion.

Lemma 3. *Let $(\mathcal{T}, \mathcal{A})$ be a KB and $\exists x (R(a,x) \wedge \bigwedge_i^n x \neq c_i)$ be a role-constraining formula. Then,*

$$(\mathcal{T}, \mathcal{A}) \models \exists x (R(a,x) \wedge \bigwedge_i^n x \neq c_i)$$

if an only if at least one of the following holds

(i) $R(a,b) \models_\mathcal{T} \exists x (R(a,x) \wedge \bigwedge_i^n x \neq c_i)$, where $\mathcal{A} \models_\mathcal{T} R(a,b)$ and all c_is are different from b, or

(ii) $\exists R(a), \neg \exists R^-(c_1), \ldots, \neg \exists R^-(c_n) \models \exists x (R(a,x) \wedge \bigwedge_i^n x \neq c_i)$, where $\mathcal{A} \models_\mathcal{T} \exists R(a)$ and $\mathcal{A} \models_\mathcal{T} \neg \exists R^-(c_i)$ for $i \in \{1, \ldots, n\}$.

Proof. The second case means that the KB forbids certain constants to be in the range of R. The first case is a simple consequence of the unique name assumption and it means that the KB allows the constant b to be in the range of R. ■

Lemma 4. *Let $\mathcal{A}_1$, $\mathcal{A}_2$ be two TBoxes, $\mathcal{T}$ be a TBox and $R(a,b)$ be an ABox assertion. Then, $\mathcal{A}_1 \cup \mathcal{A}_2 \models_\mathcal{T} R(a,b)$ if and only if $\mathcal{A}_1 \models_\mathcal{T} R(a,b)$ or $\mathcal{A}_2 \models_\mathcal{T} R(a,b)$.*

Proof. Follows from the definition of *DL-Lite*. ■

Lemma 5. *If for some of $i = \{1, \ldots, n\}$ it holds $\mathcal{A} \not\models_\mathcal{T} \exists x (R(a,x) \wedge x \neq c_i)$, then*

$$\mathcal{A} \not\models_\mathcal{T} \exists x (R(a,x) \wedge \bigwedge_i^n x \neq c_i).$$

Proof. . . . ∎

Corollary 1 (Two Witnesses Property). *Let $(\mathcal{T}, \mathcal{A})$ be a DL-Lite KB and $\mathcal{U}$ an update. If there is an unexpected role-constraining formula, then there are two assertions $C_1(a)$ and $C_2(b)$ such that $\mathcal{A} \models_{\mathcal{T}} C_1(a)$ and $\mathcal{U} \models_{\mathcal{T}} C_2(b)$, and $\{C_1(a), C_2(b)\}$ entails an unexpected role-constraining formula of the form $\exists x(R(a,x) \wedge (x \neq b))$ or $\exists x(R(b,x) \wedge (x \neq a))$ for some role name R occurring in $\mathcal{A}$ or $\mathcal{U}$.*

Lemma 6. *For a given KB and update, there is only finitely many unexpected role-constraining formulas.*

Proof. We need it to show existence of max-careful subset.

Since only the formulas of type *(ii)* are unexpected, one can forbid constants to appear on roles' components by ABox assertions only, and every ABox is finite. ∎

Proof (of Theorem 4, Existence of Maximal A_0).

Wlg we assume that $\mathcal{A} = cl_{\mathcal{T}}(\mathcal{A})$ and $\mathcal{U} = cl_{\mathcal{T}}(\mathcal{U})$.

Take a maximal satisfiable subset $\mathcal{A}_m$ of $\mathcal{A}$. We now construct $\mathcal{A}_0$. Let $\mathcal{A}_0 := \mathcal{A}_m$. Assume there is a role-constraining formula $\varphi = \exists x(R(a,x) \wedge \bigwedge_i^n (x \neq c_i))$.

Then, due to Lemma 3, φ is entailed from some $R(a,b)$, and/or from $S = \{\exists R(a), \neg\exists R^-(c_1), \ldots, \neg\exists R^-(c_n)\}$.

In the former case, by Lemma 4, φ is also entailed either from $\mathcal{A}_0$, or from $\mathcal{U}$. Thus, φ is not unexpected.

If the latter but not the former case holds, then: *(i)* $S \subseteq \mathcal{A}_0$ or *(ii)* $S \subseteq U$ or *(iii)* $S \not\subseteq \mathcal{A}_0$ and $S \not\subseteq U$. In (i) and (ii), φ is not unexpected. In (iii) it is unexpected and we modify $\mathcal{A}_0$ in order to prohibit the entailment of φ. Let $\varphi_i = \{\exists R(a), \neg\exists R^-(c_i)\}$.

Consider two cases how to modify $\mathcal{A}_0$:

1. Let $\exists R(a)$ is in $\mathcal{U}$. Let also $\neg\exists R(c_{i_j})$ for $j = \{1, \ldots, k\}$ be all elements of S in $\mathcal{A}_0 \setminus \mathcal{U}$. Then, set $\mathcal{A}_0 := \mathcal{A}_0 \setminus \bigcup_{j=1}^k \{\neg\exists R(c_{i_j})\}$ and $\mathcal{A}_0 \cup \mathcal{U} \not\models \varphi$, due to Lemma 5.
2. Let $\exists R(a)$ is in $\mathcal{A}_0 \setminus \mathcal{U}$. Then, set $\mathcal{A}_0 := \mathcal{A}_0 \setminus \{\exists R(a)\}$ and $\mathcal{A}_0 \cup \mathcal{U} \not\models \varphi$, since it does not entail $\exists R(a)$.

Afterwords, delete from $\mathcal{A}_0$ all the assertions $\mathcal{T}$-entail the deleted ones.

By iterating the procedure described above for every unexpected role-constraining (due to Lemma 6 there are finitely many of them), we obtain a careful $\mathcal{A}_0$.

We now show that the constructed $\mathcal{A}_0$ is maximal. Assume it is not. Therefore, there is an assertion $C(d)$ in $\mathcal{A} \setminus \mathcal{A}_0$ such that $\mathcal{A}_0 \cup \{C(d)\}$ is careful and $\mathcal{U} \cup \mathcal{A}_0 \cup \{C(d)\}$ is satisfiable.

If $C(d) \in \mathcal{A} \setminus \mathcal{A}_m$, then due to maximality of $\mathcal{A}_m$ and Lemma 2 $\mathcal{U} \cup \{C(d)\}$ is unsatisfiable. We got a contradiction.

If $C(d) \in \mathcal{A}_m \setminus \mathcal{A}_0$, then $C(d)$ was dropped from A_m during its modification in one of two steps. If it was dropped in Step 1, then $C(d) = \neg\exists R^-(d)$ and there is $\exists R(c)$ in $\mathcal{U} \setminus \mathcal{A}_m$. Then, $\exists x(R(c,x) \wedge (x \neq d))$ is unexpected. We got a contradiction. If it was dropped in Step 2, then $C(d) = \exists R(d)$ and there is $\neg\exists R^-(c)$ in $\mathcal{U} \setminus \mathcal{A}_m$. Then, $\exists x(R(d,x) \wedge (x \neq c))$ is unexpected. We got a contradiction.

We conclude that $\mathcal{A}_0$ is maximal.

To show uniqueness of $\mathcal{A}_0$, assume there is another maximal careful satisfiable subset $\mathcal{A}_0'$ of $\mathcal{A}$. Then, consider $C(a)$ that is in $\mathcal{A}_0' \setminus \mathcal{A}_0$. Since both $\mathcal{A}_0'$ and $\mathcal{A}_0$ are satisfiable, they are subsets of $\mathcal{A}_m$. Hence, $C(a) \in \mathcal{A}_m \setminus \mathcal{A}_0$ and it was dropped from $\mathcal{A}_m$ while constructing $\mathcal{A}_0$. Due to Corollary 1, there is an assertion $C'(b)$ in $\mathcal{U}$, such that $\{C(a), C'(b)\}$ $\mathcal{T}$-entails an unexpected role-constraining formula. Hence, $\mathcal{A}_0' \setminus \mathcal{A}_0$ is empty. Therefore, $\mathcal{A}_0' \subseteq \mathcal{A}_0$, which contradicts maximality of $\mathcal{A}_0'$. ∎

Theorem 10. *Let* $(\mathcal{T}, \mathcal{A})$ *be a satisfiable DL-Lite*$_{\mathcal{FR}}$ *KB and* $\mathcal{U}$ *an update such that* $(\mathcal{T}, \mathcal{U})$ *is satisfiable. Then, CarefulUpdate*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$ *runs in polynomial time in the size of* $\mathcal{T} \cup \mathcal{A} \cup \mathcal{U}$*, and CarefulUpdate*$(\mathcal{T}, \mathcal{A}, \mathcal{U}) \equiv_{\mathcal{T}}$ *c-upd*$(\mathcal{T}, \mathcal{A}, \mathcal{U})$.

Proof. Follows from the the proof of Theorem 5, since it exactly mimics the construction of the proof.