Understanding Inexpressibility of Model-Based ABox Evolution in DL-Lite

Evgeny Kharlamov, Dmitriy Zheleznyakov

Understanding Inexpressibility of Model-Based ABox Evolution in DL-Lite

Evgeny Kharlamov, and Dmitriy Zheleznyakov

KRDB Research Centre, Free University of Bozen-Bolzano, Italy
last_name@inf.unibz.it

Abstract. Evolution of Knowledge Bases (KBs) expressed in Description Logics (DLs) proved its importance. Recent studies of evolution in DLs mostly focussed on model-based approaches. They showed that evolution of KBs in tractable DLs, such as *DL-Lite*, suffers from inexpressibility, i.e., the result of evolution cannot be captured in *DL-Lite*. What is missing in these studies is *understanding*: in which *DL-Lite* fragments evolution can be captured, what causes the inexpressibility, which logics is sufficient to express evolution, and whether one can approximate it in *DL-Lite*. This paper provides some understanding of these issues for both update and revision. We found what *DL-Lite* formulas make evolution inexpressible and how to capture evolution in their absence. We introduce the notion of prototypes that gives an understanding of how to capture evolution for a rich *DL-Lite* fragment in $\mathsf{FO}[2]$. Decidability of $\mathsf{FO}[2]$ gives possibility for approximations.

1 Introduction

Description Logics (DLs) provide excellent mechanisms for representing structured knowledge by means of Knowledge Bases (KBs) $\mathcal{K}$ that are composed of two components: TBox (describes intensional or general knowledge about an application domain) and ABox (describes facts about individual objects). DLs constitute the foundations for various dialects of OWL, the Semantic Web ontology language.

Traditionally DLs have been used for modeling *static* and structural aspects of application domains [1]. Recently, however, the scope of KBs has broadened, and they are now used also for providing support in the maintenance and *evolution* phase of information systems. This makes it necessary to study *evolution of Knowledge Bases* [2], where the goal is to incorporate new knowledge $\mathcal{N}$ into an existing KB $\mathcal{K}$ so as to take into account changes that occur in the underlying application domain. In general, $\mathcal{N}$ is represented by a set of formulas denoting those properties that should be true after $\mathcal{K}$ has evolved, and the result of evolution, denoted $\mathcal{K} \diamond \mathcal{N}$, is also intended to be a set of formulas. In the case where $\mathcal{N}$ interacts with $\mathcal{K}$ in an undesirable way, e.g., by causing the KB or relevant parts of it to become unsatisfiable, $\mathcal{N}$ cannot simply be added to the KB. Instead, suitable changes need to be made in $\mathcal{K}$ so as to avoid this undesirable interaction. Different choices for changes are possible, corresponding to different approaches to semantics for KB evolution [3,4,5].

An important group of approaches to evolution semantics, that we focus in this paper, is called *model-based* (MBAs). Under MBAs the result of evolution $\mathcal{K} \diamond \mathcal{N}$ is a *set of* *models* of $\mathcal{N}$ that are minimally distanciated from models of $\mathcal{K}$. Depending on what the distance between models is and how to measure it, eight different MBAs were introduced (see Section 3 for details). Since $\mathcal{K} \diamond \mathcal{N}$ is a set of models, while $\mathcal{K}$ and $\mathcal{N}$ are logical theories, it is desirable to represent $\mathcal{K} \diamond \mathcal{N}$ as a logical theory using the same language as for $\mathcal{K}$ and $\mathcal{N}$. Thus, looking for representations of $\mathcal{K} \diamond \mathcal{N}$ is the main challenge in studies of evolution under MBAs. When $\mathcal{K}$ and $\mathcal{N}$ are propositional theories, representing $\mathcal{K} \diamond \mathcal{N}$ is well understood [5], while it becomes dramatically more complicated as soon as $\mathcal{K}$ and $\mathcal{N}$ are first-order, e.g., DL KBs [6].

Model based evolution of KBs where $\mathcal{K}$ and $\mathcal{N}$ are written in a language of the *DL-Lite* family [7] has been recently extensively studied [6,8,9]. The focus on *DL-Lite* is not surprising since *DL-Lite* is the basis of OWL 2 QL, a tractable OWL 2 profile. It has been shown that for every of the eight MBAs one can find *DL-Lite* $\mathcal{K}$ and $\mathcal{N}$ such that $\mathcal{K} \diamond \mathcal{N}$ cannot be expressed in *DL-Lite* [10,11], i.e., *DL-Lite* is *not closed* under MBA evolution. This phenomenon was also noted in [6,10] for some of the eight semantics. What is missing in all these studies of evolution for *DL-Lite* is *understanding* of:

(1) What fragments of *DL-Lite* are closed under model-based evolutions? What *DL-Lite* formulas are responsible for inexpressibility of model-based evolutions?
(2) What are sufficient extensions of *DL-Lite* to capture model-based evolutions of *DL-Lite* KBs? How to capture the evolutions of *DL-Lite* KBs in these extensions?
(3) For *DL-Lite* $\mathcal{K}$ and $\mathcal{N}$, is it possible and how to do "good" approximations of $\mathcal{K} \diamond \mathcal{N}$?

In this paper we study the problems *(1)-(3)* for so-called *ABox evolution*, i.e., $\mathcal{N}$ is a new ABox and the TBox of $\mathcal{K}$ should remain the same after the evolution. ABox evolution is important for areas, e.g., bioinformatics, where the structural knowledge TBox is well crafted and stable, while ABox facts about individuals may get changed. These ABox changes should be reflected in KBs in a way that the TBox is not affected. Our study covers both the case of ABox updates and ABox revision [4].

In Sections 2 and 3 we define *DL-Lite*$_\mathcal{R}$ and ABox evolution. In Section 4 we study relationships between different MBAs. In Section 5 we introduce *DL-Lite*$_\mathcal{R}^+$, a restriction on *DL-Lite*$_\mathcal{R}$ where disjointness of concepts with role projections is forbidden. We show that *DL-Lite*$_\mathcal{R}^+$ is closed under most of MBA evolutions and provide tractable algorithms computing $\mathcal{K} \diamond \mathcal{N}$. In Section 6 we study two important MBAs where distance between models is based on atoms and prove that *DL-Lite*$_\mathcal{R}^+$ is a sufficient borderline of expressibility for these semantics: the class of KBs that are in *DL-Lite*$_\mathcal{R}$ but not in *DL-Lite*$_\mathcal{R}^+$ is not closed under these semantics. We also introduce and study *DL-Lite*$_\mathcal{R}^{I}$ that restricts *DL-Lite*$_\mathcal{R}$, extends *DL-Lite*$_\mathcal{R}^+$ and not closed under MBAs. In order to capture evolution of *DL-Lite*$_\mathcal{R}^{I}$ KBs in FO, we introduce prototypes, based on which we show that the evolution can be expressed in $\mathsf{FO}[2]$ (restriction of first-order logics that only makes use of two variables). Finally, we argue that decidability of $\mathsf{FO}[2]$ gives possibilities for approximations of $\mathcal{K} \diamond \mathcal{N}$. Proofs can be found in [12].

2 DL-Lite$_\mathcal{R}$

We introduce some basic notions of DLs, (see [13] for more details). We consider a logic *DL-Lite*$_\mathcal{R}$ of *DL-Lite* family of DLs [7,14]. *DL-Lite*$_\mathcal{R}$ has the following constructs for (complex) *concepts* and *roles*: *(i)* $B ::= A \mid \exists R$, *(ii)* $C ::= B \mid \neg B$, *(iii)* $R ::= P \mid P^-$, where A and P stand for an *atomic concept* and *role*, respectively, which are just names. A *knowledge base* (KB) $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ is compound of two sets of *assertions*: *TBox* $\mathcal{T}$, and *ABox* $\mathcal{A}$. *DL-Lite*$_\mathcal{R}$ TBox assertions are *concept inclusion assertions* of the form $B \sqsubseteq C$ and *role inclusion assertions* $R_1 \sqsubseteq R_2$, while ABox assertions are *membership assertions* of the form $A(a)$, $\neg A(a)$, and $R(a, b)$. The *active domain* of $\mathcal{K}$, denoted *adom*$(\mathcal{K})$, is the set of all constants occurring in $\mathcal{K}$. The *DL-Lite* family has nice computational properties, for example, KB satisfiability has polynomial-time complexity in the size of the TBox and logarithmic-space in the size of the ABox [15,16].

The semantics of DL-Lite KBs is standard and based on first order interpretations, all over the same countable domain Δ. An *interpretation* $\mathcal{I}$ is a function $\cdot^\mathcal{I}$ that assigns to each C a subset $C^\mathcal{I}$ of Δ, and to R a binary relation $R^\mathcal{I}$ 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 $(P^-)^\mathcal{I} = \{(a_2, a_1) \mid (a_1, a_2) \in P^\mathcal{I}\}$. We assume that Δ contains the constants and that $c^\mathcal{I} = c$, i.e., we adopt *standard names*. Alternatively, we view an interpretation as a set of atoms and say that $A(a) \in \mathcal{I}$ iff $a \in A^\mathcal{I}$ and $P(a, b) \in \mathcal{I}$ iff $(a, b) \in P^\mathcal{I}$. An interpretation $\mathcal{I}$ is a *model* of a membership assertion $A(a)$ (resp., $\neg A(a)$) if $a \in A^\mathcal{I}$ (resp., $a \notin A^\mathcal{I}$), of $P(a, b)$ if $(a, b) \in P^\mathcal{I}$, and of an inclusion assertion $D_1 \sqsubseteq D_2$ if $D_1^\mathcal{I} \subseteq D_2^\mathcal{I}$.

As usual, we use $\mathcal{I} \models F$ to denote that $\mathcal{I}$ is a model of an assertion F, and $\mathcal{I} \models \mathcal{K}$ to denote that $\mathcal{I} \models F$ for each assertion F in $\mathcal{K}$. We use *Mod*$(\mathcal{K})$ to denote the set of all models of $\mathcal{K}$. A KB is *satisfiable* if it has at least one model. We use entailment on KBs $\mathcal{K} \models \mathcal{K}'$ in the standard sense. An ABox $\mathcal{A}$ $\mathcal{T}$*-entails* an ABox $\mathcal{A}'$, denoted $\mathcal{A} \models_\mathcal{T} \mathcal{A}'$, if $\mathcal{T} \cup \mathcal{A} \models \mathcal{A}'$, and $\mathcal{A}$ is $\mathcal{T}$*-equivalent* to $\mathcal{A}'$, denoted $\mathcal{A} \equiv_\mathcal{T} \mathcal{A}'$ if $\mathcal{A} \models_\mathcal{T} \mathcal{A}'$ and $\mathcal{A}' \models_\mathcal{T} \mathcal{A}$.

The deductive *closure of a TBox* $\mathcal{T}$ (of an ABox $\mathcal{A}$), denoted *cl*$(\mathcal{T})$ (resp., *cl*$_\mathcal{T}(\mathcal{A})$), is the set of all TBox (resp., positive ABox) assertions F such that $\mathcal{T} \models F$ (resp., $\mathcal{A} \models_\mathcal{T} F$). Clearly in *DL-Lite*$_\mathcal{R}$ *cl*$(\mathcal{T})$ (and *cl*$_\mathcal{T}(\mathcal{A})$) is computable in time quadratic in the number of assertions of $\mathcal{T}$, i.e., $|\mathcal{T}|$, (resp., $|\mathcal{T} \cup \mathcal{A}|$). In our work we assume that all TBoxes and ABoxes are closed, while results are extendable to arbitrarily KBs. For satisfiable KBs $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ a *full closure of* $\mathcal{A}$, denoted *fcl*$_\mathcal{T}(\mathcal{A})$, is the set of all membership assertions f (both positive and negative) over *adom*$(\mathcal{K})$ such that $\mathcal{A} \models_\mathcal{T} f$.

A *homomorphism* h from a model $\mathcal{I}$ to a model $\mathcal{J}$ is a mapping from Δ to Δ satisfying: *(i)* $h(a) = a$ for every constant a; *(ii)* if $\alpha \in A^\mathcal{I}$ (resp., $(\alpha, \beta) \in P^\mathcal{I}$), then $h(\alpha) \in A^\mathcal{J}$ (resp., $(h(\alpha), h(\beta)) \in P^\mathcal{J}$) for every A (resp., P). We write $\mathcal{I} \hookrightarrow \mathcal{J}$ if there is a homomorphism from $\mathcal{I}$ to $\mathcal{J}$. A canonical model $\mathcal{I}$ of $\mathcal{K}$, denoted as $\mathcal{I}_\mathcal{K}^{can}$ or just $\mathcal{I}^{can}$ when $\mathcal{K}$ is clear from the context, is a model of $\mathcal{K}$ which can be homomorphically embedded in every model of $\mathcal{K}$ [7].

3 Evolution of Knowledge Bases

This section is based on [10]. Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a *DL-Lite*$_\mathcal{R}$ KB and $\mathcal{N}$ a "new" ABox. We study how to incorporate $\mathcal{N}$'s assertions into $\mathcal{K}$, that is, how $\mathcal{K}$ *evolves* under $\mathcal{N}$ [2]. More practically, we study *evolution operators* that take $\mathcal{K}$ and $\mathcal{N}$ as input and return, possibly in *polynomial time*, a *DL-Lite*$_\mathcal{R}$ $\mathcal{K}' = (\mathcal{T}, \mathcal{A}')$ (with the same TBox as $\mathcal{K}$) that captures the evolution, and which we call *the (ABox) evolution of* $\mathcal{K}$ *under* $\mathcal{N}$. Based on the evolution principles of [10], we require $\mathcal{K}$ and $\mathcal{K}'$ to be satisfiable and coherent.

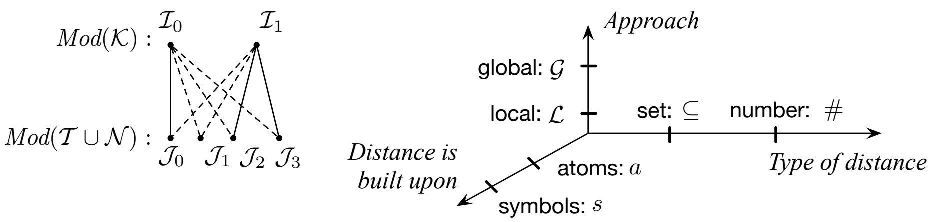

Fig. 1. Left: measuring distances between models and finding local minimums. Right: three-dimensional space of approaches to model-based evolution semantics.

Model-Based Semantics of Evolution. In model-based approaches (MBAs), the result of evolution of a KB $\mathcal{K}$ wrt new knowledge $\mathcal{N}$ is a set $\mathcal{K} \diamond \mathcal{N}$ of models. The idea of MBAs is to choose as $\mathcal{K} \diamond \mathcal{N}$ some models of $(\mathcal{T}, \mathcal{N})$ depending on their distance to $\mathcal{K}$'s models. Katsuno and Mendelzon [4] considered two ways, so called *local* and *global*, of choosing these models of $\mathcal{N}$, where the first choice corresponds to *knowledge update* and the second one to *knowledge revision*.

The idea of the local approaches is to consider all models $\mathcal{I}$ of $\mathcal{K}$ and for each $\mathcal{I}$ to take those models $\mathcal{J}$ of $(\mathcal{T}, \mathcal{N})$ that are minimally distant from $\mathcal{I}$. Formally,

$$\mathcal{K} \diamond \mathcal{N} = \bigcup_{\mathcal{I} \in \mathit{Mod}(\mathcal{K})} \mathcal{I} \diamond \mathcal{N}, \text{ where } \mathcal{I} \diamond \mathcal{N} = \underset{\mathcal{J} \in \mathit{Mod}(\mathcal{T} \cup \mathcal{N})}{\arg\min} \mathit{dist}(\mathcal{I}, \mathcal{J}).$$

where $\mathit{dist}(\cdot, \cdot)$ is a function whose range is a partially ordered domain and $\arg\min$ stands for the *argument of the minimum,* that is, in our case, the set of models $\mathcal{J}$ for which the value of $\mathit{dist}(\mathcal{I}, \mathcal{J})$ reaches its minimum value, given $\mathcal{I}$. The distance function *dist* varies from approach to approach and commonly takes as values either numbers or subsets of some fixed set. To get a better intuition of the local semantics, consider Figure 1, left, where we present two model $\mathcal{I}_0$ and $\mathcal{I}_1$ of a KB $\mathcal{K}$ and four models $\mathcal{J}_0, \ldots, \mathcal{J}_3$ of $(\mathcal{T}, \mathcal{N})$. We represent the distance between a model of $\mathcal{K}$ and a model of $\mathcal{T} \cup \mathcal{N}$ by the length of a line connecting them. Solid lines correspond to minimal distances, dashed lines to distances that are not minimal. In this figure $\{\mathcal{J}_0\} = \arg\min_{\mathcal{J} \in \{\mathcal{J}_0, \ldots, \mathcal{J}_3\}} \mathit{dist}(\mathcal{I}_0, \mathcal{J})$ and $\{\mathcal{J}_2, \mathcal{J}_3\} = \arg\min_{\mathcal{J} \in \{\mathcal{J}_0, \ldots, \mathcal{J}_3\}} \mathit{dist}(\mathcal{I}_1, \mathcal{J})$.

In the global approach one choses models of $\mathcal{N}$ that are minimally distant from $\mathcal{K}$:

$$\mathcal{K} \diamond \mathcal{N} = \underset{\mathcal{J} \in \mathit{Mod}(\mathcal{N})}{\arg\min} \mathit{dist}(\mathit{Mod}(\mathcal{K}), \mathcal{J}), \tag{1}$$

where $\mathit{dist}(\mathit{Mod}(\mathcal{K}), \mathcal{J}) = \min_{\mathcal{I} \in \mathit{Mod}(\mathcal{K})} \mathit{dist}(\mathcal{I}, \mathcal{J})$. Consider again Figure 1, left, and assume that the distance between $\mathcal{I}_0$ and $\mathcal{J}_0$ is the global minimum, hence, $\{\mathcal{J}_0\} = \arg\min_{\mathcal{J} \in \{\mathcal{J}_0, \ldots, \mathcal{J}_3\}} \mathit{dist}(\{\mathcal{I}_0, \mathcal{I}_1\}, \mathcal{J})$.

Measuring Distance Between Interpretations. The classical MBAs were developed for propositional theories [5], where interpretations were sets of propositional atoms, two distance functions were introduced, respectively based on symmetric difference "$\ominus$" and on the cardinality of symmetric difference:

$$\mathit{dist}_{\subseteq}(\mathcal{I}, \mathcal{J}) = \mathcal{I} \ominus \mathcal{J} \qquad \text{and} \qquad \mathit{dist}_{\#}(\mathcal{I}, \mathcal{J}) = |\mathcal{I} \ominus \mathcal{J}|, \tag{2}$$

where $\mathcal{I} \ominus \mathcal{J} = (\mathcal{I} \setminus \mathcal{J}) \cup (\mathcal{J} \setminus \mathcal{I})$. Distances under $dist_{\subseteq}$ are sets and are compared by set inclusion, that is, $dist_{\subseteq}(\mathcal{I}_1, \mathcal{J}_1) \leq dist_{\subseteq}(\mathcal{I}_2, \mathcal{J}_2)$ iff $dist_{\subseteq}(\mathcal{I}_1, \mathcal{J}_1) \subseteq dist_{\subseteq}(\mathcal{I}_2, \mathcal{J}_2)$. Finite distances under $dist_{\#}$ are natural numbers and are compared in the standard way.

One can extend these distances to DL interpretations in two different ways. One way is to consider interpretations $\mathcal{I}$, $\mathcal{J}$ as sets of *atoms*. Then $\mathcal{I} \ominus \mathcal{J}$ is again a set of atoms and we can define distances as in Equation (2). We denote these distances as $dist^a_{\subseteq}(\mathcal{I}, \mathcal{J})$ and $dist^a_{\#}(\mathcal{I}, \mathcal{J})$. Another way is to define distances at the level of the concept and role *symbols* in the signature Σ underlying the interpretations:

$$dist^s_{\subseteq}(\mathcal{I}, \mathcal{J}) = \{S \in \Sigma \mid S^{\mathcal{I}} \neq S^{\mathcal{J}}\}, \quad \text{and} \quad dist^s_{\#}(\mathcal{I}, \mathcal{J}) = |\{S \in \Sigma \mid S^{\mathcal{I}} \neq S^{\mathcal{J}}\}|.$$

Summing up across the different possibilities, we have three dimensions, which give eight semantics of evolution according to MBAs by choosing: (1) the *local* or the *global* approach, (2) *atoms* or *symbols* for defining distances, and (3) *set inclusion* or *cardinality* to compare symmetric differences. In Figure 1, right, we depict these three dimensions. We denote each of these eight possibilities by a combination of three symbols, indicating the choice in each dimension, e.g. $\mathcal{L}^a_{\#}$ denotes the local semantics where the distances are expressed in terms of cardinality of sets of atoms.

Closure Under Evolution. Let $\mathcal{D}$ be a DL and M one of the eight MBAs introduced above. We say $\mathcal{D}$ is *closed under evolution wrt* M (or evolution wrt M is *expressible* in $\mathcal{D}$) if for any KBs $\mathcal{K}$ and $\mathcal{N}$ written in $\mathcal{D}$, there is a KB $\mathcal{K}'$ written in $\mathcal{D}$ such that $Mod(\mathcal{K}') = \mathcal{K} \diamond \mathcal{N}$, where $\mathcal{K} \diamond \mathcal{N}$ is the evolution result under semantics M.

We showed in [10,11] that *DL-Lite* is not closed under any of the eight model based semantics. The observation underlying these results is that on the one hand, the minimality of change principle intrinsically introduces implicit disjunction in the evolved KB. On the other hand, since *DL-Lite* is a slight extension of Horn logic [17], it does not allow one to express genuine disjunction (see Lemma 1 in [10] for details).

Let M be a set of models that resulted from the evolution of $(\mathcal{T}, \mathcal{A})$ with $\mathcal{N}$. A KB $(\mathcal{T}, \mathcal{A}')$ is a *sound approximation* of M if $M \subseteq Mod(\mathcal{T}, \mathcal{A}')$. A sound approximation $(\mathcal{T}, \mathcal{A}')$ is *minimal* if for every sound approximation $(\mathcal{T}, \mathcal{A}'')$ inequivalent to $(\mathcal{T}, \mathcal{A}')$, it holds $Mod(\mathcal{T}, \mathcal{A}'') \not\subset Mod(\mathcal{T}, \mathcal{A}')$, i.e. $(\mathcal{T}, \mathcal{A}')$ is minimal wrt "$\subseteq$".

4 Relationships Between Model-Based Semantics

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two evolution semantics and $\mathcal{D}$ a logic language. Then $\mathcal{S}_1$ *is subsumed by* $\mathcal{S}_2$ wrt $\mathcal{D}$, denoted $(\mathcal{S}_1 \preccurlyeq_{sem} \mathcal{S}_2)(\mathcal{D})$, or just $\mathcal{S}_1 \preccurlyeq_{sem} \mathcal{S}_2$ when $\mathcal{D}$ is clear from the context, if $\mathcal{K} \diamond_{\mathcal{S}_1} \mathcal{N} \subseteq \mathcal{K} \diamond_{\mathcal{S}_2} \mathcal{N}$ for all satisfiable KBs $\mathcal{K}$ and $\mathcal{N}$ written in $\mathcal{D}$, where $\mathcal{K} \diamond_{\mathcal{S}_i} \mathcal{N}$ denotes evolution under $\mathcal{S}_i$. Two semantics $\mathcal{S}_1$ and $\mathcal{S}_2$ are *equivalent* (wrt $\mathcal{D}$), denoted $(\mathcal{S}_1 \equiv_{sem} \mathcal{S}_2)(\mathcal{D})$, if $(\mathcal{S}_1 \preccurlyeq_{sem} \mathcal{S}_2)(\mathcal{D})$ and $(\mathcal{S}_2 \preccurlyeq_{sem} \mathcal{S}_1)(\mathcal{D})$. Further in this section we will consider $\mathcal{K}$ and $\mathcal{N}$ written in *DL-Lite*$_{\mathcal{R}}$. The following theorem shows the subsumption relation between different semantics, which we also depict with solid arrows in Figure 2. The figure is complete in the following sense: there is a solid path between any two semantics $\mathcal{S}_1$ and $\mathcal{S}_2$ iff there is a subsumption $\mathcal{S}_1 \preccurlyeq_{sem} \mathcal{S}_2$.

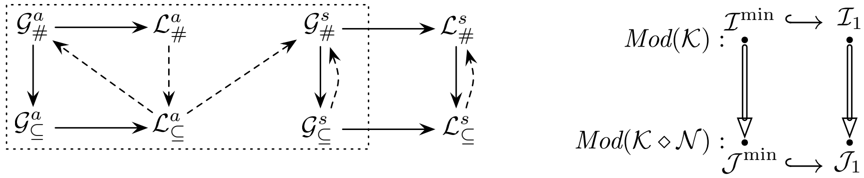

Fig. 2. Left: Subsumptions for evolution semantics.
" $\longrightarrow$ ": in *DL-Lite*$_\mathcal{R}$ (Theorem 1). " $\dashrightarrow$ ": in *DL-Lite*$^+_\mathcal{R}$ (Theorems 6, 7).
Dashed frame surrounds semantics under which *DL-Lite*$^+_\mathcal{R}$ is closed.
Right: commutation diagram for $\mathcal{L}^a_\subseteq$ semantics, "$\mathcal{I} \Rightarrow \mathcal{J}$" denotes that $\mathcal{J} \in \mathcal{I} \diamond \mathcal{N}$.

Theorem 1. *Let $\beta \in \{a, s\}$ and $\alpha \in \{\subseteq, \#\}$. Then*

$$\mathcal{G}^\beta_\alpha \preccurlyeq_{sem} \mathcal{L}^\beta_\alpha, \quad \mathcal{G}^\beta_\# \preccurlyeq_{sem} \mathcal{G}^\beta_\subseteq, \quad \textit{and} \quad \mathcal{L}^s_\# \preccurlyeq_{sem} \mathcal{L}^s_\subseteq.$$

Proof. Let *dist* be one of the four considered distances, $\mathcal{E}_\mathcal{G} = \mathcal{K} \diamond \mathcal{N}$ wrt $\mathcal{G}^\beta_\alpha$ and $\mathcal{E}_\mathcal{L} = \mathcal{K} \diamond \mathcal{N}$ wrt $\mathcal{L}^\beta_\alpha$ be corresponding global and local semantics based on *dist*. For given $\mathcal{K}$ and $\mathcal{N}$, let $\mathcal{J}' \in \mathcal{E}_\mathcal{G}$. Then, there is $\mathcal{I}' \models \mathcal{K}$ such that for every $\mathcal{I}'' \models \mathcal{K}$ and $\mathcal{J}'' \models \mathcal{T} \cup \mathcal{N}$ it *does not* hold $dist(\mathcal{I}'', \mathcal{J}'') \lneq dist(\mathcal{I}', \mathcal{J}')$. In particular, when $\mathcal{I}'' = \mathcal{I}'$, there is no $\mathcal{J}'' \models \mathcal{T} \cup \mathcal{N}$ such that $dist(\mathcal{I}', \mathcal{J}'') \lneq dist(\mathcal{I}', \mathcal{J}')$, which yields that $\mathcal{J}' \in \arg\min_{\mathcal{J} \in Mod(\mathcal{T} \cup \mathcal{N})} dist(\mathcal{I}', \mathcal{J})$, and $\mathcal{J}' \in \mathcal{E}_\mathcal{L}$. We conclude that: $\mathcal{G}^a_\# \preccurlyeq_{sem} \mathcal{L}^a_\#$, $\mathcal{G}^a_\subseteq \preccurlyeq_{sem} \mathcal{L}^a_\subseteq$, $\mathcal{G}^s_\# \preccurlyeq_{sem} \mathcal{L}^s_\#$, $\mathcal{G}^s_\subseteq \preccurlyeq_{sem} \mathcal{L}^s_\subseteq$.

Now consider $\mathcal{E}_\# = \mathcal{K} \diamond \mathcal{N}$ wrt $\mathcal{L}^\beta_\#$, which is based on $dist_\#$, and $\mathcal{E}_\subseteq = \mathcal{K} \diamond \mathcal{N}$ wrt $\mathcal{L}^\beta_\subseteq$, which is based on $dist_\subseteq$. Assume $\mathcal{J}' \in \mathcal{E}_\#$ and $\mathcal{J}' \notin \mathcal{E}_\subseteq$. Then, from the former assumption we conclude existence of $\mathcal{I}' \models \mathcal{K}$ such that $\mathcal{J}' \in \arg\min_{\mathcal{J} \in Mod(\mathcal{T} \cup \mathcal{N})} dist_\#(\mathcal{I}', \mathcal{J})$. From the latter assumption, $\mathcal{J}' \notin \mathcal{E}_\subseteq$, we conclude existence of a model $\mathcal{J}''$ such that $dist_\subseteq(\mathcal{I}', \mathcal{J}'') \subsetneq dist_\subseteq(\mathcal{I}', \mathcal{J}')$. This yields that $dist_\#(\mathcal{I}', \mathcal{J}'') \lneq dist_\#(\mathcal{I}', \mathcal{J}')$, which contradicts the fact that $\mathcal{J}' \in \mathcal{E}_\#$, assuming that $dist_\subseteq(\mathcal{I}', \mathcal{J}')$ is finite. Thus, $\mathcal{E}_\# \preccurlyeq_{sem} \mathcal{E}_\subseteq$ as soon as $dist_\subseteq(\mathcal{I}, \mathcal{J})$ is finite. This finiteness condition always holds for when $\beta = s$ since the signature of $\mathcal{K} \cup \mathcal{N}$ is finite. It is easy to check that $dist_\subseteq(\mathcal{I}, \mathcal{J})$ may not be finite when $\beta = a$, hence, $\mathcal{L}^a_\# \not\preccurlyeq_{sem} \mathcal{L}^a_\subseteq$.

Similarly one can show that $\mathcal{G}^s_\# \preccurlyeq_{sem} \mathcal{G}^s_\subseteq$. It also holds that $\mathcal{G}^a_\# \preccurlyeq_{sem} \mathcal{G}^a_\subseteq$ due to the finite model property in *DL-Lite*$_\mathcal{R}$. □

5 Evolution in *DL-Lite*$^+_\mathcal{R}$

Here we consider a restriction of *DL-Lite*$_\mathcal{R}$, which we call *DL-Lite*$^+_\mathcal{R}$. A KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ is in *DL-Lite*$^+_\mathcal{R}$ if it is in *DL-Lite*$_\mathcal{R}$ and $\mathcal{T} \not\models \exists R \sqsubseteq \neg B$ for any role R and any concept B. Intuitively, "$+$" emphasizes absence of disjointness for roles (only positive inclusions involving roles are permitted). *DL-Lite*$^+_\mathcal{R}$ is defined semantically, while one can syntactically check (in quadratic time), given a *DL-Lite*$_\mathcal{R}$ KB $\mathcal{K}$, whether $\mathcal{K}$ is in *DL-Lite*$^+_\mathcal{R}$: one computes a closure of $\mathcal{K}$, checks that no assertions of the form $\exists R \sqsubseteq \neg B$ are in the closure and if it is the case, then $\mathcal{K}$ is in *DL-Lite*$^+_\mathcal{R}$. *DL-Lite*$^+_\mathcal{R}$ is an extension of RDFS ontology language (of its FO fragment). *DL-Lite*$^+_\mathcal{R}$ adds to RDFS the ability of expressing disjointness of concepts ($A_1 \sqsubseteq \neg A_2$) and mandatory participation ($A \sqsubseteq \exists R$). Since *DL-Lite*$^+_\mathcal{R}$ restricts *DL-Lite*$_\mathcal{R}$, the semantics relations from Theorem 1) are also correct for *DL-Lite*$^+_\mathcal{R}$. We now study what further $\preccurlyeq_{sem}$-relations hold in *DL-Lite*$^+_\mathcal{R}$.

INPUT : consistent KBs $(\mathcal{T}, \mathcal{A})$ and $\mathcal{N}$
OUTPUT: a set $\mathcal{A}' \subseteq fcl_{\mathcal{T}}(\mathcal{A})$ of ABox assertions

$\mathcal{A}' := \emptyset$; $\mathcal{S} := fcl_{\mathcal{T}}(\mathcal{A})$;
repeat
  choose some $\phi \in \mathcal{S}$; $\mathcal{S} := \mathcal{S} \setminus \{\phi\}$;
  if $\{\phi\} \cup fcl_{\mathcal{T}}(\mathcal{N})$ *is consistent* **then** $\mathcal{A}' := \mathcal{A}' \cup \{\phi\}$
until $\mathcal{S} = \emptyset$;

Algorithm 1: Algorithm $AlignAlg((\mathcal{T}, \mathcal{A}), \mathcal{N})$ for $\mathcal{A}'$ deterministic computation

5.1 Capturing Atom-Based Evolution

We first study evolution under $\mathcal{L}^{a}_{\subseteq}$. Let $\mathcal{I}$ be an interpretation and $\mathcal{N}$ a knowledge base. An *alignment* of $\mathcal{I}$ with $\mathcal{N}$, denoted $Align(\mathcal{I}, \mathcal{N})$, is the interpretation $\mathcal{J} = \{f \mid f \in \mathcal{I}$ and f is satisfiable with $\mathcal{N}\}$. There is a tight connection (Lemma 2) between an alignment of a model and its evolution under $\mathcal{L}^{a}_{\subseteq}$ semantics. Moreover, alignment preserves homomorphic relationship on interpretations (Lemma 3). Let N be a set of membership assertions and $\mathcal{I}$ an interpretation. A union $\mathcal{I} \cup \mathcal{N}$ denotes the model $\mathcal{I} \cup \mathcal{J}_{\mathcal{N}}$, where $\mathcal{J}_{\mathcal{N}} = \{f \mid f \in \mathcal{N}$ and f is positive$\}$.

Lemma 2. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ and $(\mathcal{T}, \mathcal{N})$ be satisfiable DL-Lite$^{+}_{\mathcal{R}}$ KBs, and $\mathcal{I} \models \mathcal{K}$. Then $\mathcal{I} \diamond \mathcal{N} = \{Align(\mathcal{I}, \mathcal{N}) \cup \mathcal{N}\}$, under $\mathcal{L}^{a}_{\subseteq}$ semantics.*

Lemma 3. *Let $\mathcal{K}$ and $\mathcal{N}$ be satisfiable DL-Lite$^{+}_{\mathcal{R}}$ KBs and $\mathcal{I}$ be a model of $\mathcal{K}$. Then $Align(\mathcal{I}^{can}_{\mathcal{K}}, \mathcal{N}) \hookrightarrow Align(\mathcal{I}, \mathcal{N})$.*

Lemmas 3 and 2 imply that $(\mathcal{I}^{can} \diamond \mathcal{N}) \hookrightarrow (\mathcal{I} \diamond \mathcal{N})$ for every $\mathcal{I} \in Mod(\mathcal{K})$ under $\mathcal{L}^{a}_{\subseteq}$, where we apply $\hookrightarrow$ to the sets $(\mathcal{I}^{can} \diamond \mathcal{N})$ and $(\mathcal{I} \diamond \mathcal{N})$, instead of their single models. We depict this phenomenon in Figure 2, right: evolution preserves homomorphic embeddability of $\mathcal{I}^{can}$ into $\mathcal{I}$. This fundamental property implies that $\mathcal{I}^{can} \diamond \mathcal{N}$ consists of the universal model of $\mathcal{K} \diamond \mathcal{N}$ wrt $\mathcal{L}^{a}_{\subseteq}$.

Consider an algorithm *AlignAlg* (see Algorithm 1) that inputs $\mathcal{K}$, $\mathcal{N}$, and returns the alignment $Align(\mathcal{I}^{can}, \mathcal{N})$: it drops all the assertions of $fcl_{\mathcal{T}}(\mathcal{A})$ contradicting $\mathcal{N}$ and keeps the rest. Using *AlignAlg* we can compute $\mathcal{K} \diamond \mathcal{N}$:

Theorem 4. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ and $(\mathcal{T}, \mathcal{N})$ be satisfiable DL-Lite$^{+}_{\mathcal{R}}$ KBs. Then:*

$$\mathcal{K} \diamond \mathcal{N} = Mod(\mathcal{T}, AlignAlg(\mathcal{K}, \mathcal{N}) \cup \mathcal{N}) \text{ (wrt } \mathcal{L}^{a}_{\subseteq}\text{)}.$$

Note that since computation of $AlignAlg(\mathcal{T}, \mathcal{A}, \mathcal{N})$ is polynomial in $\mathcal{A}$ and $\mathcal{N}$, computation of $\mathcal{K} \diamond \mathcal{N}$ is also polynomial.

Example 5. Consider $\mathcal{T} = \{B_0 \sqsubseteq B,\ B \sqsubseteq \neg C\}$, $\mathcal{A} = \{C(a)\}$, and $\mathcal{N} = B(a)$. Then, $fcl_{\mathcal{T}}(\mathcal{A}) = \{C(a), \neg B_0(a), \neg B(a)\}$. The alignment of $(\mathcal{T}, \mathcal{A})$ with $\mathcal{N}$ is the set $AlignAlg((\mathcal{T}, \mathcal{A}), \mathcal{N}) = \{B(a), \neg B(a)\}$. Hence, the result of evolution under every the model-based semantics with atom-based distance is $(\mathcal{T}, \{B(a), \neg B_0(a)\})$. ■

Relationships Between Atom-Based Semantics. Next theorem shows that evolutions wrt all four model-based semantics on atoms coincide wrt *DL-Lite*$^{+}_{\mathcal{R}}$. We depict these relations between semantics in Figure 2 using a dashed arrow, e.g., between $\mathcal{L}^{a}_{\subseteq} \preccurlyeq_{sem} \mathcal{G}^{a}_{\#}$. Note that there is a path with solid or dashed lines between any two semantics if and only if there is a subsumption wrt *DL-Lite*$^{+}_{\mathcal{R}}$.

Theorem 6. $\mathcal{L}^{a}_{\#} \equiv_{sem} \mathcal{L}^{a}_{\subseteq} \equiv_{sem} \mathcal{G}^{a}_{\#} \equiv_{sem} \mathcal{G}^{a}_{\subseteq}$.

Theorems 6 and 4 imply that in *DL-Lite*$^{+}_{\mathcal{R}}$ one can use *AlignAlg* to compute evolution under all MBAs on atoms.

5.2 Capturing Symbol-Based Evolution

Observe that symbol-based semantics behave differently from atom-based ones: two local semantics (on set inclusion and cardinality) coincide, as well as two global ones, while there is no subsumption between local and global ones, as depicted in Figure 2.

Theorem 7. *The following relations hold:*
(i) $\mathcal{L}^{a}_{\subseteq} \preccurlyeq_{sem} \mathcal{G}^{s}_{\#}$, *while* $\mathcal{G}^{s}_{\#} \not\preccurlyeq_{sem} \mathcal{L}^{a}_{\subseteq}$;
(ii) $\mathcal{L}^{s}_{\subseteq} \equiv_{sem} \mathcal{L}^{s}_{\#}$, *and* $\mathcal{G}^{s}_{\subseteq} \equiv_{sem} \mathcal{G}^{s}_{\#}$, *while* $\mathcal{L}^{s}_{\subseteq} \not\preccurlyeq_{sem} \mathcal{G}^{s}_{\#}$.

As a corollary of Theorem 7, the algorithm of computing $\mathcal{K} \diamond \mathcal{N}$ of Theorem 4 in general does not work for computing evolution under any of the symbol-based semantics. At the same time what it outputs is a complete approximation of all symbol-based semantics, while it approximates global semantics better than the local ones.

Consider the algorithm *SymAlg* in Figure 2 that will be used for evolutions on symbols. It works as follows: for every atom ϕ in $\mathcal{N}$ it checks whether ϕ satisfies a condition Π (Line 4). If it does, the algorithm deletes all those literals that share their concept name with ϕ. Both local and global semantics have their own Π: $\Pi_{\mathcal{G}}$ and $\Pi_{\mathcal{L}}$.

Capturing Global Semantics. $\Pi_{\mathcal{G}}(\phi)$ checks whether ϕ of $\mathcal{N}$ $\mathcal{T}$-contradicts $\mathcal{A}$: $\Pi_{\mathcal{G}}(\phi)$ is true iff $\neg\phi \in \mathit{fcl}_{\mathcal{T}}(\mathcal{A}) \setminus \mathit{AlignAlg}((\mathcal{T},\mathcal{A}),\mathcal{N})$. Intuitively, *SymAlg* for global semantics works as follows: having contradiction between $\mathcal{N}$ and $\mathcal{A}$ on $\phi = B(c)$, the change of B's interpretation is inevitable. Since the semantics traces changes on symbols only, and B is already changed, one can drop from $\mathcal{A}$ all the assertions of the form $B(d)$. Clearly, $\mathit{SymAlg}(\mathcal{K},\mathcal{N},\Pi_{\mathcal{G}})$ can be computed in time polynomial in $|\mathcal{A} \cup \mathcal{K}|$. The following theorem shows correctness of this algorithm.

Theorem 8. *Let* $\mathcal{K} = (\mathcal{T},\mathcal{A})$ *and* $(\mathcal{T},\mathcal{N})$ *be satisfiable DL-Lite*$^{+}_{\mathcal{R}}$ *KBs. Then DL-Lite*$^{+}_{\mathcal{R}}$ *is closed under* $\mathcal{G}^{s}_{\subseteq}$ *and* $\mathcal{G}^{s}_{\#}$*, and for both* $\mathcal{K} \diamond \mathcal{N} = \mathit{Mod}(\mathcal{T}, \mathit{SymAlg}(\mathcal{K},\mathcal{N},\Pi_{\mathcal{G}}))$,

Capturing Local Semantics. Observe that $\mathcal{L}^{s}_{\subseteq}$ and $\mathcal{L}^{s}_{\#}$ are not expressible in *DL-Lite*$^{+}_{\mathcal{R}}$.

Theorem 9. *DL-Lite*$^{+}_{\mathcal{R}}$ *is not closed under* $\mathcal{L}^{s}_{\subseteq}$ *and* $\mathcal{L}^{s}_{\#}$ *semantics.*

To compute the minimal sound approximations under local semantics on symbols, we use algorithm *SymAlg* in Figure 2 with the following $\Pi_{\mathcal{L}}$: $\Pi_{\mathcal{L}}(\phi)$ is true iff $\phi \notin \mathcal{S}_1$. That is, $\Pi_{\mathcal{L}}$ checks whether the ABox $\mathcal{T}$-entails $A(c) \in \mathit{fcl}_{\mathcal{T}}(\mathcal{N})$, and if it does, the

INPUT : consistent *DL-Lite*$^{+}_{\mathcal{R}}$ KB $(\mathcal{T}, \mathcal{A})$ and ABox $\mathcal{N}$, a formula property Π
OUTPUT: a set $\mathcal{A}' \subseteq fcl_{\mathcal{T}}(\mathcal{A}) \cup fcl_{\mathcal{T}}(\mathcal{N})$ of ABox assertions

$\mathcal{A}' := \emptyset$; $\mathcal{S}_1 := AlignAlg((\mathcal{T}, \mathcal{A}), \mathcal{N})$; $\mathcal{S}_2 := fcl_{\mathcal{T}}(\mathcal{N})$;
repeat
  choose some $\phi \in \mathcal{S}_2$; $\mathcal{S}_2 := \mathcal{S}_2 \setminus \{\phi\}$;
  if $\Pi(\phi) = TRUE$ **then** $\mathcal{S}_1 := \mathcal{S}_1 \setminus \{\phi' \mid \phi$ and ϕ' have the same concept name$\}$
until $\mathcal{S}_2 = \emptyset$;
$\mathcal{A}' := \mathcal{S}_1 \cup fcl_{\mathcal{T}}(\mathcal{N})$

Algorithm 2: Algorithm $SymAlg((\mathcal{T}, \mathcal{A}), \mathcal{N}, \Pi)$ for deterministic computation of $\mathcal{K} \diamond \mathcal{N}$ under $\mathcal{G}^s_{\subseteq}$ and $\mathcal{G}^s_{\#}$ semantics and minimal sound approximation under $\mathcal{L}^s_{\subseteq}$ and $\mathcal{L}^s_{\#}$ semantics

algorithm deletes all the assertions from $fcl_{\mathcal{T}}(\mathcal{A})$ that share the concept name with $A(c)$. This property is weaker than one for global semantics, since it is easier to get changes in interpretation of A by choosing a model of $\mathcal{K}$ which does not include $A(c)$. Clearly, $SymAlg(\mathcal{K}, \mathcal{N}, \Pi_{\mathcal{L}})$ can be computed in time polynomial in $|\mathcal{A} \cup \mathcal{K}|$. The following theorem shows correctness of the algorithm.

Theorem 10. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ and $(\mathcal{T}, \mathcal{N})$ be satisfiable DL-Lite$^{+}_{\mathcal{R}}$ KBs. Then the KB $(\mathcal{T}, SymAlg(\mathcal{K}, \mathcal{N}, \Pi_{\mathcal{L}}))$ is a minimal sound approximation of $\mathcal{K} \diamond \mathcal{N}$ under $\mathcal{L}^s_{\subseteq}$ and $\mathcal{L}^s_{\#}$.*

6 Evolution in *DL-Lite*$^{I}_{\mathcal{R}}$

In the previous section we considered evolution in *DL-Lite*$^{+}_{\mathcal{R}}$. Here we show that *DL-Lite*$^{+}_{\mathcal{R}}$ is essentially a maximal fragment of *DL-Lite*$_{\mathcal{R}}$ closed under atom-based evolution, and present a wide fragment of *DL-Lite*$_{\mathcal{R}}$, namely *DL-Lite*$^{I}_{\mathcal{R}}$, that is not closed under atom-based evolution while the evolution can be captured in $\mathsf{FO}[2]$.

The following theorem shows that the restriction not to have roles involved into negation relation is essential, and even a minimal violation leads to inexpressibility of evolution. The minimal way to violate conditions of *DL-Lite*$^{+}_{\mathcal{R}}$ is to have two assertions: $A \sqsubseteq \exists R$ and $\exists R^- \sqsubseteq \neg C$ entailed from a TBox.

Theorem 11. *Let assertions $\{A \sqsubseteq \exists R,\ \exists R^- \sqsubseteq \neg C\}$ be entailed from a DL-Lite$_{\mathcal{R}}$ TBox $\mathcal{T}$, for some R and C, Then there exist ABoxes $\mathcal{A}$ and $\mathcal{N}$ such that $(\mathcal{T}, \mathcal{A}) \diamond \mathcal{N}$ is inexpressible in DL-Lite$_{\mathcal{R}}$ under $\mathcal{L}^a_{\subseteq}$ and under $\mathcal{L}^a_{\#}$.*

The following example provides reasons why *DL-Lite*$^{+}_{\mathcal{R}}$ restrictions are important for expressibility of atom-based MBAs. We will further use this example to show how to capture $\mathcal{L}^a_{\subseteq}$ evolution in FO for KBs violating *DL-Lite*$^{+}_{\mathcal{R}}$ restrictions.

Example 12. Consider the following *DL-Lite* KBs $\mathcal{K}_1 = (\mathcal{T}_1, \mathcal{A}_1)$ and $\mathcal{N}_1 = \{C(b)\}$:

$$\mathcal{T}_1 = \{A \sqsubseteq \exists R, \exists R^- \sqsubseteq \neg C\}; \quad \mathcal{A}_1 = \{A(a), C(e), C(d), R(a, b)\}.$$

Consider the following model $\mathcal{I}$ of $\mathcal{K}_1$:

$$\mathcal{I}: \quad A^{\mathcal{I}} = \{a, x\}, \quad C^{\mathcal{I}} = \{d, e\}, \quad R^{\mathcal{I}} = \{(a, b), (x, b)\},$$

where $x \in \Delta \setminus adom$. We now show that the following models belong to $\mathcal{I} \diamond \mathcal{N}_1$, this will be used further to capture $\mathcal{K}_1 \diamond \mathcal{N}_1$.

$$\begin{array}{llll}
\mathcal{J}_0: & A^{\mathcal{I}} = \emptyset, & C^{\mathcal{I}} = \{d, e, b\}, & R^{\mathcal{I}} = \emptyset, \\
\mathcal{J}_1: & A^{\mathcal{I}} = \{x\}, & C^{\mathcal{I}} = \{e, b\}, & R^{\mathcal{I}} = \{(x, d)\}, \\
\mathcal{J}_2: & A^{\mathcal{I}} = \{x\}, & C^{\mathcal{I}} = \{d, b\}, & R^{\mathcal{I}} = \{(x, e)\}, \\
\mathcal{J}_3: & A^{\mathcal{I}} = \{a, x\}, & C^{\mathcal{I}} = \{b\}, & R^{\mathcal{I}} = \{(a, d), (x, e)\}, \\
\mathcal{J}_4: & A^{\mathcal{I}} = \{x\}, & C^{\mathcal{I}} = \{d, e, b\}, & R^{\mathcal{I}} = \{(x, \beta_1)\},
\end{array}$$

where $\beta_1 \in \Delta \setminus adom(\mathcal{K}_1)$. Clearly all five models satisfy $\mathcal{N}_1$ and $\mathcal{T}_1$. To see that they are in $\mathcal{I} \diamond \mathcal{N}_1$ observe that every model $\mathcal{J}(I) \in (\mathcal{I} \diamond \mathcal{N}_1)$ can be obtained from $\mathcal{I}$ by making modifications that guarantee that $\mathcal{J}(I) \models (\mathcal{N}_1 \cup \mathcal{T}_1)$ and that the distance between $\mathcal{I}$ and $\mathcal{J}(I)$ is minimal. What are these modifications? Since in every such $\mathcal{J}(I)$ the new knowledge $C(b)$ holds and $(C \sqsubseteq \neg\exists R^-) \in \mathcal{T}_1$, there should be no R-arc incoming into b in $\mathcal{J}(\mathcal{I})$, hence the necessary modification of $\mathcal{I}$ is to forbid the R-arcs going from a and from x to point into b. Where in $\mathcal{J}(I)$ should this R-arcs point in? There are only three mutually exclusive cases for each of a and x: *(i)* there is no R-arc in $\mathcal{J}(I)$ that points from a (resp., x), we simply drop it from $\mathcal{I}$, *(ii)* it points to some element $\beta \in \Delta \setminus \{d, e, b\}$, that is, $\mathcal{J}(I) \models R(a, \beta)$. *(iii)* it points to an element γ of $\{d, e\}$, that is, $\mathcal{J}(I) \models R(a, \gamma)$. Case (i) for both a and x corresponds to $\mathcal{J}_0$, hence, $\mathcal{J}_0 \in \mathcal{I} \diamond \mathcal{N}_1$. Case (ii) for both a and x corresponds to $\mathcal{J}_4$, hence, $\mathcal{J}_4 \in \mathcal{I} \diamond \mathcal{N}_1$. Case (i) and Case (iii) for a corresponds to $\mathcal{J}_1$ and $\mathcal{J}_2$, hence, both $\mathcal{J}_1$ and $\mathcal{J}_2$ are in $\mathcal{I} \diamond \mathcal{N}_1$. Case (iii) for both a and x corresponds to $\mathcal{J}_3$, hence, $\mathcal{J}_3 \in \mathcal{I} \diamond \mathcal{N}_1$. There are clearly more models $\mathcal{I} \diamond \mathcal{N}_1$ than these five.

To show that $\mathcal{K}_1 \diamond \mathcal{N}_1$ is inexpressible in *DL-Lite*$_{\mathcal{R}}$ consider a model $\mathcal{J}_5 \in$ *Mod*$(\mathcal{T}_1, \mathcal{N}_1)$ which is not in $\mathcal{K}_1 \diamond \mathcal{N}_1$, while it gives a property of models from $\mathcal{K}_1 \diamond \mathcal{N}_1$ that is responsible for inexpressibility of $\mathcal{K}_1 \diamond \mathcal{N}_1$.

$$\mathcal{J}_5: \quad A^{\mathcal{I}} = \{x\}, \quad C^{\mathcal{I}} = \{e, b\}, \quad R^{\mathcal{I}} = \{(x, d), (x, \beta_1)\}.$$

Observe that $\mathcal{J}_4$ is closer to $\mathcal{I}$ that $\mathcal{J}_5$, hence $\mathcal{J}_5 \notin (\mathcal{I} \diamond \mathcal{N}_1)$. Indeed, since for every model $\mathcal{I}' \in$ *Mod*$(\mathcal{K}_1)$ it holds that $\mathcal{I}' \models C(d)$ and $\mathcal{I}' \not\models R(x, d)$, we have that $\{C(d), R(x, d)\} \subseteq \mathcal{I}' \ominus \mathcal{J}_5$. At the same time $\mathcal{J}_4 \models C(d)$ and $\mathcal{J}_4 \not\models R(x, d)$, while $\mathcal{J}_4$ and $\mathcal{J}_5$ agree on all the atoms but $C(d)$ and $R(x, d)$. Hence, $(\mathcal{I}' \ominus \mathcal{J}_4) \subsetneq (\mathcal{I}' \ominus \mathcal{J}_5)$ and $\mathcal{J}_5 \notin (\mathcal{I} \diamond \mathcal{N}_1)$. Moreover, since $\mathcal{I}'$ is an arbitrary model of $\mathcal{K}_1$, $\mathcal{J}_5 \notin (\mathcal{K}_1 \diamond \mathcal{N}_1)$.

Let us make a closer look at $\mathcal{J}_5$: it extends $\mathcal{J}_1$ with the atom $R(x, \beta_1)$, and this is the reason why it is not in the result of evolution $\mathcal{K}_1 \diamond \mathcal{N}_1$. This observation gives a restriction on all the models $\mathcal{J}$ on $\mathcal{K}_1 \diamond \mathcal{N}_1$: if in $\mathcal{J}$ there is one R-arc from some x into d, then $\mathcal{J}$ has no other arcs from x. In a way, we have a functionality restriction on the role R, when it connects two specific elements of the domain: x and d. The same kind of functionality holds for x and e, and both of them are not expressible in *DL-Lite*. This functionality for x and d, and for x and e can be formally written as the following ϕ and ψ, respectively:

$$\phi = \forall x \forall y. \left[R(x, d) \wedge R(x, y) \rightarrow y = d \right], \quad \psi = \forall x \forall y. \left[R(x, e) \wedge R(x, y) \rightarrow y = e \right].$$

Since every model of $\mathcal{K}_1 \diamond \mathcal{N}_1$ should satisfy ϕ and ψ, and *DL-Lite* is a slight extension of Horn logics [17], $\mathcal{K}_1 \diamond \mathcal{N}_1$ cannot be captured in *DL-Lite*. ■

This example shows that *DL-Lite*$_{\mathcal{R}}$ is not closed under model-based evolution. Besides the fact that both ϕ and ψ are inexpressible in *DL-Lite*$_{\mathcal{R}}$, while should be entailed by the result of evolution, there is another observation also responsible for inexpressibility of $\mathcal{K}_1 \diamond \mathcal{N}_1$: it has no canonical model. At the same time, one can see that the set $\mathcal{K}_1 \diamond \mathcal{N}_1$ can be divided in four subsets, where each has a canonical model. We now show that these four canonical models, which we call *prototypes* for $\mathcal{K}_1 \diamond \mathcal{N}_1$, can be used to capture $\mathcal{K}_1 \diamond \mathcal{N}_1$ in $\mathsf{FO}[2]$, in a similar way as we used the unique canonical model of the evolution result in *DL-Lite*$^+_{\mathcal{R}}$ to capture the evolution in *DL-Lite*$^+_{\mathcal{R}}$.

Definition 13. *Let $\mathcal{K}$ be a DL-Lite$_{\mathcal{R}}$ KB and $\mathcal{N}$ be an ABox. A* prototypal set for $\mathcal{K} \diamond \mathcal{N}$ *is a minimal subset $\daleth = \{\mathcal{J}_1, \dots, \mathcal{J}_n\}$ of $\mathcal{K} \diamond \mathcal{N}$ satisfying the property: for every $\mathcal{J} \in \mathcal{K} \diamond \mathcal{N}$ there is $\mathcal{J}_i \in \daleth$ such that $\mathcal{J}_i \hookrightarrow \mathcal{J}$. Every $\mathcal{J}_i \in \daleth$ is a* prototype for $\mathcal{K} \diamond \mathcal{N}$.

Continuing with Example 12, one can check that the prototypal set $\daleth$ for $\mathcal{K}_1 \diamond \mathcal{N}_1$ consists of four models: $\{\mathcal{J}_0, \mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_6\}$, where

$\mathcal{J}_6$: $\quad A^{\mathcal{I}} = \{x_1, x_2\}, \quad C^{\mathcal{I}} = \{b\}, \quad R^{\mathcal{I}} = \{(x_1, d), (x_2, e)\}.$

Note that $\mathcal{J}_6 \hookrightarrow \mathcal{J}_3$ and $\mathcal{J}_0 \hookrightarrow \mathcal{J}_4$.

We now introduce a restriction of *DL-Lite*$_{\mathcal{R}}$ for which the result of evolution under $\mathcal{L}^a_{\subseteq}$ can be captured in FO using prototypes. *DL-Lite*$^I_{\mathcal{R}}$ (where *I* stands for (mutual) *independence* of roles) is a restriction of *DL-Lite*$_{\mathcal{R}}$ in which TBoxes $\mathcal{T}$ satisfy: for any two roles R and R', $\mathcal{T} \not\models \exists R \sqsubseteq \exists R'$ and $\mathcal{T} \not\models \exists R \sqsubseteq \neg\exists R'$. That is, we forbid direct role interaction (subsumption and disjointness) between role projections. The interaction is still possible but in a simple manner only, e.g., projections may contain the same concept. This restriction allows us to analyze evolution that affects roles independently for every role. For the ease of exhibition of the following procedure constructing $\daleth$, we further assume that in *DL-Lite*$^I_{\mathcal{R}}$ for each role R there is exactly one concept A_R such that $\mathcal{K} \models A_R \sqsubseteq \exists R$. As for *DL-Lite*$^+_{\mathcal{R}}$, one can syntactically check (in quadratic time), given a *DL-Lite*$_{\mathcal{R}}$ KB $\mathcal{K}$, whether $\mathcal{K}$ is in *DL-Lite*$^I_{\mathcal{R}}$ using the closure of $\mathcal{K}$.

If f is an ABox assertion, then $root_{\mathcal{T}}(f)$ is a set of all the ABox assertions, that $\mathcal{T}$-entail f. The procedure $BP(\mathcal{K}, \mathcal{N})$ (where BP stands for build prototypes) of constructing $\daleth$ for the case of *DL-Lite*$^I_{\mathcal{R}}$, generalizes what was done in Example 12 in the following way: In Items 1 and 2 it takes into account all the roles and concepts that may be interpreted differently in different resulted models (R, A and C in our example) and all the constants that could not be present at the second coordinate of those roles in the models of the original KB (e and d in our example). In Items 3, 4, and 5 it builds the prototype $\mathcal{J}_0$. Finally in Item 6 it builds the rest of prototypes basing on $\mathcal{J}_0$. A pseudocode of the procedure $BP(\mathcal{K}, \mathcal{N})$ is the following:

1. $\mathcal{S}_R = \{R_1, \dots, R_n\}$ is a set of all roles such that
   (i) $\mathcal{N} \models_{\mathcal{T}} \neg R_i^-(b_i)$ for some b_i and
   (ii) for every $R_i(a_i, b_i)$ in $fcl_{\mathcal{T}}(\mathcal{A})$, an atom $A_R(a_i) \in fcl_{\mathcal{T}}(\mathcal{A})$ and $R(a_i, b_i') \notin Align(fcl_{\mathcal{T}}(\mathcal{A}), \mathcal{N})$, where $b_i' \notin \{b_j\}_{j=1}^n$.
   $\mathcal{S}_{at} = \bigcup_{j=1}^n \{R_j(a, b_j) \mid R_j(a, b_j) \in fcl_{\mathcal{T}}(\mathcal{A})\}$.
2. $\mathtt{FA}(R_i)$ is equal to the following set
   $\{D(c) \in fcl_{\mathcal{T}}(\mathcal{A}) \mid \exists R_i^-(c) \wedge D(c) \models_{\mathcal{T}} \bot$ and $\mathcal{N} \not\models_{\mathcal{T}} D(c),$ and $\mathcal{N} \not\models_{\mathcal{T}} \neg D(c)\}$.
   Then $\quad \mathtt{FA} = \bigcup_{R_i \in \mathcal{S}_R} \mathtt{FA}(R_i), \quad \mathtt{FC} = \{c \mid D(c) \in \mathtt{FA}\}$,
   where the acronyms $\mathtt{FA}$ and $\mathtt{FC}$ stand for forbidden atoms and constants, respectively.

3. $\mathcal{I}' := Align(\mathcal{I}^{can}, \mathcal{N}) \cup \mathcal{N}$, where $\mathcal{I}^{can}$ is the canonical model of $\mathcal{K}$.
4. For each $R(a,b) \in \mathcal{S}_{at}$, do $\mathcal{I}' := \mathcal{I}' \setminus \{A_R(a)\}$.
5. $\mathcal{J}_0 := \mathcal{I}'$, $\mathfrak{J} := \{\mathcal{J}_0\}$.
6. For each subset $\mathcal{D} = \{D_1(c_1), \dots, D_k(c_k)\} \subseteq \texttt{FA}$ do
   for each $\mathcal{R} = (R_{i_1}, \dots, R_{i_k})$ such that $D_j(c_j) \in \texttt{FA}(R_{i_j})$ for $j = 1, \dots, k$ do
   $\mathcal{J}[\mathcal{D}, \mathcal{R}] := \left[\mathcal{J}_0 \setminus \bigcup_{i=1}^{k} \mathit{root}_{\mathcal{T}}(D_i(c_i))\right] \cup \bigcup_{i=1}^{k} \left[cl_{\mathcal{T}}(R_i'(x_i, c_i)) \cup \{A_{R_i'}(x_i)\}\right]$,
   where all x_i's are different constants from $\Delta \setminus \mathit{adom}(\mathcal{K})$, fresh for $\mathcal{I}^{can}$.
   $\mathfrak{J} := \mathfrak{J} \cup \{\mathcal{J}[\mathcal{D}, \mathcal{R}]\}$.

Theorem 14. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a DL-Lite$^{I}_{\mathcal{R}}$ KB, and $\mathcal{N}$ a DL-Lite$_{\mathcal{R}}$ ABox consistent with $\mathcal{T}$. Then the set $BP(\mathcal{K}, \mathcal{N})$ is a prototypal set for $\mathcal{K} \diamond \mathcal{N}$.*

We proceed to correctness of BP in capturing evolution in *DL-Lite*$^{I}_{\mathcal{R}}$.

Theorem 15. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be a DL-Lite$^{I}_{\mathcal{R}}$ KB, $\mathcal{N}$ a DL-Lite$_{\mathcal{R}}$ ABox consistent with $\mathcal{T}$, and $BP(\mathcal{K}, \mathcal{N}) = \{\mathcal{J}_0, \dots, \mathcal{J}_n\}$ is a prototypal set for $\mathcal{K} \diamond \mathcal{N}$. Then*

$$\mathcal{K} \diamond \mathcal{N} = \mathit{Mod}(\mathcal{T}) \cap \mathit{Mod}(\mathcal{A}_1 \vee \ldots \vee \mathcal{A}_n) \cap \mathit{Mod}(\Phi \wedge \Psi),$$

where A_i is a DL-Lite$_{\mathcal{R}}$ ABox such that $\mathcal{J}_i$ is a canonical model for $(\mathcal{T}, \mathcal{A}_i)$, and

$$\Phi = \bigwedge_{c_j \in \texttt{FC}} \forall x. \left[(R(x, c_j) \to A_R(x)) \wedge \forall y. (R(x, c_j) \wedge R(x, y) \to y = c_j)\right],$$
$$\Psi = \bigwedge_{R(a,b) \in \mathcal{S}_{at}} \exists R(a) \to A_R(a).$$

Note that Theorem 4 about computation of $\mathcal{K} \diamond \mathcal{N}$ for *DL-Lite*$^{+}_{\mathcal{R}}$ is a particular case of Theorem 15 when $\texttt{FC} = \emptyset$ and there is just one prototype. Next theorem allows us to approximate result of evolution by a *DL-Lite*$_{\mathcal{R}}$ KB, since $\mathsf{FO}[2]$ is decidable.

Theorem 16. *$\mathcal{K} \diamond \mathcal{N}$ under $\mathcal{L}^{a}_{\subseteq}$ for KBs in DL-Lite$^{I}_{\mathcal{R}}$ is in* $\mathsf{FO}[2]$.

7 Conclusion

We studied expressibility of ABox evolution (for both update and revision) over two subfamilies of *DL-Lite*$_{\mathcal{R}}$: *DL-Lite*$^{+}_{\mathcal{R}}$ and *DL-Lite*$^{I}_{\mathcal{R}}$, that both extend RDFS. The first one is closed under most of the evolution semantics, while the second one is not closed even under MBAs were distance is based on atoms. We isolated conditions on TBox assertions that lead to inexpressibility: pairs of assertions of the form $\mathcal{A} \sqsubseteq \exists R$ and $\exists R^- \sqsubseteq \neg C$ bring inexpressibility. Note that this condition is similar to the notion of unexpected facts introduces in [10], for formula-based semantics of evolutions. For *DL-Lite*$^{+}_{\mathcal{R}}$ we provided algorithms how to compute semantics that are expressible, and how to approximate those that are not. For *DL-Lite*$^{I}_{\mathcal{R}}$ we captured local model-based semantics, where the distance between models defined on atoms, in $\mathsf{FO}[2]$. For this purpose we introduced prototypes and showed how they can be used to capture results of evolution.

It is the first attempt to provide an understanding of inexpressibility of MBAs for *DL-Lite* evolution. Without this understanding it is unclear how to proceed with the study of evolution in more expressive DLs and what to expect from MBAs in such logics. Moreover, isolating the smallest fragment of FO that is sufficient to captured *DL-Lite* evolution is important in understanding whether it is possible and how to do reasonable approximations of evolution results. We also believe that our techniques of capturing semantics based on prototypes give a better understanding of how MBAs behave on FO theories. We are currently working on extending the results to capture evolution for general $DL\text{-}Lite_{\mathcal{R}}$ KBs.

Acknowledgements We are thankful to Diego Calvanese, Balder ten Cate, and Werner Nutt for discussions. The authors are supported by EU projects ACSI (FP7-ICT-257593), Ontorule (FP7-ICT-231875), and ERC FP7 grant Webdam (agreement n. 226513).

References

1. Borgida, A., Brachman, R.J.: Conceptual modeling with description logics. [13] chapter 10
2. Flouris, G., Manakanatas, D., Kondylakis, H., Plexousakis, D., Antoniou, G.: Ontology change: Classification and survey. Knowledge Engineering Review **23**(2) (2008) 117–152
3. Abiteboul, S., Grahne, G.: Update semantics for incomplete databases. (In: Proc. of VLDB'85)
4. Katsuno, H., Mendelzon, A.: On the difference between updating a knowledge base and revising it. In: Proc. of KR'91. (1991) 387–394
5. Eiter, T., Gottlob, G.: On the complexity of propositional knowledge base revision, updates and counterfactuals. Artificial Intelligence **57** (1992) 227–270
6. Liu, H., Lutz, C., Milicic, M., Wolter, F.: Updating description logic ABoxes. In: Proc. of KR 2006. (2006) 46–56
7. Calvanese, D., Giacomo, G.D., Lembo, D., Lenzerini, M., Rosati, R.: Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. JAR **39**(3) (2007)
8. Giacomo, G.D., Lenzerini, M., Poggi, A., Rosati, R.: On the update of description logic ontologies at the instance level. In: Proc. of AAAI 2006. (2006) 1271–1276
9. Wang, Z., Wang, K., Topor, R.W.: A new approach to knowledge base revision in DL-Lite. In: AAAI. (2010)
10. Calvanese, D., Kharlamov, E., Nutt, W., Zheleznyakov, D.: Evolution of *DL-Lite* knowledge bases. In: International Semantic Web Conference (1). (2010) 112–128
11. Calvanese, D., Kharlamov, E., Nutt, W., Zheleznyakov, D.: Evolution of DL-Lite knowledge bases (extended version). Technical Report KRDB11-3, (KRDB Research Centre)
12. Kharlamov, E., Zheleznyakov, D.: Understanding inexpressibility of model-based ABox evolution in DL-Lite (extended version). (Technical Report KRDB11-4)
13. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P.F., eds.: (The Description Logic Handbook)
14. Artale, A., Calvanese, D., Kontchakov, R., Zakharyaschev, M.: The *DL-Lite* family and relations. J. Artif. Intell. Res. (JAIR) **36** (2009) 1–69
15. Artale, A., Calvanese, D., Kontchakov, R., Zakharyaschev, M.: The *DL-Lite* family and relations. J. of Artificial Intelligence Research **36** (2009) 1–69
16. Poggi, A., Lembo, D., C., D., Giacomo, G.D., Lenzerini, M., Rosati, R.: Linking data to ontologies. J. on Data Semantics **X** (2008) 133–173
17. Calvanese, D., Kharlamov, E., Nutt, W.: A proof theory for *DL-Lite*. In: Proc. of DL 2007. (Volume 250.) 235–242