A Well-founded Graph-based Summarization Framework for Description Logics

Cheikh-Brahim El Vaigh[0000−0002−9843−3001] and François Goasdoué[0000−0003−4532−7974]

Univ Rennes, Lannion, France
{cheikh-brahim.el-vaigh,fg}@irisa.fr

Abstract. The quotient operation from graph theory offers an elegant graph summarization framework that has been widely investigated in the literature, notably for the exploration and efficient management of large graphs; it consists in fusing equivalent vertices according to an equivalence relation.
In this paper, we study whether a similar operation may be used to summarize ABoxes. Towards this goal, we define and examine the quotient operation on an ABox: we establish that a quotient ABox is more specific than the ABox it summarizes, and characterize to which extent it is more specific. This preliminary investigation validates the interest of a quotient-based ABox summarization framework, and paves the way for further studies on it in the description logic setting, e.g., to devise equivalence relations suited to the optimization of typical data management and reasoning tasks on large ABoxes or to the visualization of large ABoxes, and on its utilization in related settings, e.g., Semantic Web.

Keywords: Data summaries · quotient graph · most specific concept

1 Introduction

Graph summarization has received considerable attention in the literature [14,5], in particular for the exploration and visualization of large graphs, as well as for the optimization of graph management systems (cardinality estimation, indexing, etc). A well-established tool used to summarize graphs into compact ones is the *quotient operation* from graph theory [12]; other kinds of summaries have also been considered like collections of graph statistics or of frequent graph patterns, or a mix of both as in [16].

The quotient operation offers an elegant graph summarization framework by decoupling the summarization method, which basically fuses equivalent vertices, from the high-level specification of equivalent vertices, defined by an equivalence relation, e.g., bisimilarity [1]. Applying the quotient operation to a graph results

in a (typically small) homomorphic approximation of it, called *quotient graph*; the equivalence relation to choose depends on the target summary usage.

In this paper, we study whether a similar operation may be used to summarize ABoxes. Towards this goal, we first *transfer the quotient operation from graphs to description logics (DLs)* [2] in order to summarize ABoxes as if they were merely graphs. Then, because ABoxes are not just graphs but are essentially first-order theories (or slight extensions of them), we examine the semantics of the quotient operation on an ABox. We establish that *a quotient ABox is more specific than the ABox it summarizes*, by showing that its structure entails the structure of the ABox, and we characterize *to which extent this quotient ABox is more specific than the ABox it summarizes*, by determining the precise nature of the approximation between the quotient ABox contents and the ABox contents. This latter characterization is the main result of this paper. It provides an in-depth understanding of the quotient operation itself as a tool for ABox summarization; the essence of quotient ABoxes is independent of the DL under consideration and of the chosen equivalence relation (that depends on the target summary usage).

The paper is organized as follows. We recall the basics of DLs in Sec. 2. Then, we transfer the quotient operation from graphs to ABoxes in Sec. 3 and we examine its semantics w.r.t. DLs in Sec. 4. Finally, we discuss related works in Sec. 5 and we conclude with salient perspectives in Sec. 6.

The proofs of our technical results can be found in the appendix of this paper.

2 Preliminaries

We recall the important aspects of DLs that are used in this paper [2], with a focus on the lightweight DL-lite$_\mathcal{R}$ DL [4] that we (just) use to illustrate our discussions throughout this paper; our results hold for (standard) DLs.

DL syntax and semantics Given a set N_C of *concept names* (unary predicates), a set N_R of *role names* (binary predicates), and a set N_I of *individuals* (constants), a DL allows expressing unary and binary formulae respectively called *concepts* and *roles*. They are used to build *TBoxes* (ontologies) made of *concept inclusions* of the form $\mathsf{C}_1 \sqsubseteq \mathsf{C}_2$ and *role inclusions* of the form $\mathsf{R}_1 \sqsubseteq \mathsf{R}_2$, as well as to build *ABoxes* (databases) storing *concept assertions* of the form $\mathsf{C}(a)$ and *role assertions* of the form $\mathsf{R}(a_1, a_2)$, where $a, a_1, a_2 \in N_I$. In DL-lite$_\mathcal{R}$, concepts and roles are built according to the following rules, where $A \in N_C$ and $R \in N_R$:

$$B := A \mid \exists Q,\ C := B \mid \neg B,\ Q := R \mid R^-,\ S := Q \mid \neg Q.$$

Also, TBox inclusions are of the forms $B \sqsubseteq C$ and $Q \sqsubseteq S$, while ABox assertions are of the forms $A(a)$ and $R(a_1, a_2)$.

First-order interpretations are used to define the semantics of DLs. They are of the form $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$, where $\Delta^\mathcal{I}$ is a *non-empty domain* and $\cdot^\mathcal{I}$ is a *function* that maps every $a \in N_I$ to $a^\mathcal{I} \in \Delta^\mathcal{I}$, every $A \in N_C$ to $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$, and every $R \in N_R$ to $R^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$. Further, the interpretation function $\cdot^\mathcal{I}$ is extended

to the concept and role formulae allowed in the DL under consideration. In DL-lite$_\mathcal{R}$, $\cdot^\mathcal{I}$ is extended as follows: $(R^-)^\mathcal{I} = \{(o_1, o_2) \mid (o_2, o_1) \in R^\mathcal{I}\}$, $(\exists Q)^\mathcal{I} = \{o_1 \mid \exists o_2\ (o_1, o_2) \in Q^\mathcal{I}\}$, $(\neg B)^\mathcal{I} = \Delta^\mathcal{I} \setminus B^\mathcal{I}$ and $(\neg Q)^\mathcal{I} = (\Delta^\mathcal{I} \times \Delta^\mathcal{I}) \setminus Q^\mathcal{I}$. Let $\mathcal{I}$ be an interpretation. $\mathcal{I}$ satisfies a DL concept C (resp. role R) iff it has a non-empty interpretation, i.e., $\mathsf{C}^\mathcal{I} \neq \emptyset$ (resp. $\mathsf{R}^\mathcal{I} \neq \emptyset$). $\mathcal{I}$ satisfies an inclusion $\mathsf{C}_1 \sqsubseteq \mathsf{C}_2$ (resp. $\mathsf{R}_1 \sqsubseteq \mathsf{R}_2$) iff $\mathsf{C}_1^\mathcal{I} \subseteq \mathsf{C}_2^\mathcal{I}$ (resp. $\mathsf{R}_1^\mathcal{I} \subseteq \mathsf{R}_2^\mathcal{I}$) holds, and it satisfies a TBox iff it satisfies *all* the TBox inclusions. $\mathcal{I}$ satisfies an assertion $\mathsf{C}(a)$ (resp. $\mathsf{R}(a_1, a_2)$) iff $a^\mathcal{I} \in \mathsf{C}^\mathcal{I}$ (resp. $(a_1^\mathcal{I}, a_2^\mathcal{I}) \in \mathsf{R}^\mathcal{I}$) holds, and it satisfies an ABox iff it satisfies *all* the ABox assertions. Finally, $\mathcal{I}$ is a *model* of a concept, role, inclusion, TBox, assertion or ABox iff $\mathcal{I}$ satisfies it.

Generalization/specialization relations Concepts and roles are compared through *subsumption*, which may rely on the knowledge of a TBox $\mathcal{T}$. C_1 is subsumed by C_2 or C_2 subsumes C_1 w.r.t. $\mathcal{T}$, denoted $\mathsf{C}_1 \preceq_\mathcal{T} \mathsf{C}_2$, iff every model $\mathcal{I}$ of $\mathcal{T}$ is a model of the inclusion $\mathsf{C}_1 \sqsubseteq \mathsf{C}_2$; subsumption for roles is defined similarly.

ABoxes are compared through *entailment*, which may also rely on a TBox $\mathcal{T}$. $\mathcal{A}_1$ entails $\mathcal{A}_2$ or $\mathcal{A}_2$ is entailed by $\mathcal{A}_1$ w.r.t. $\mathcal{T}$, denoted $\mathcal{A}_1 \models_\mathcal{T} \mathcal{A}_2$, iff every model $\mathcal{I}$ of $\mathcal{A}_1$ *and* $\mathcal{T}$ is also a model of $\mathcal{A}_2$. As special case, an ABox $\mathcal{A}$ entails a concept assertion $\mathsf{C}(a)$ w.r.t. $\mathcal{T}$, denoted $\mathcal{A} \models_\mathcal{T} \mathsf{C}(a)$, if $\mathcal{A}$ entails w.r.t. $\mathcal{T}$ the ABox $\{\mathsf{C}(a)\}$; entailment of a role assertion is defined similarly.

Data management The main data management tasks are *consistency checking* and *query answering*. They involve an ABox $\mathcal{A}$ and may rely on the knowledge of a TBox $\mathcal{T}$. $\mathcal{A}$ is *consistent* w.r.t. $\mathcal{T}$ iff $\mathcal{A}$ has a model that is also a model of $\mathcal{T}$. To query ABoxes, we consider the first-order language of *unions of conjunctive queries* that is widely-considered for ontology-mediated query answering [3] and ontology-based data access [17]. A conjunctive query (CQ) is of the form $\exists \vec{Y} \phi$, where ϕ is a conjunction of the forms $A(t)$ or $R(t, t')$ where t, t' are variables or constants, and $\vec{Y}$ is a tuple of variable from ϕ. A CQ is Boolean if all its variables are existentially quantified. The answer to a Boolean CQ q is *true* if it is entailed by $\mathcal{A}$ *and* $\mathcal{T}$, denoted $\mathcal{A} \models_\mathcal{T} q$, i.e., q is true in *all* the models of $\mathcal{A}$ *and* $\mathcal{T}$; otherwise q is *false*. The answer to a non-Boolean CQ q with free variables $X_1, \ldots, X_n$, i.e., of arity n, is the set of all the tuples of constants of the form $\vec{a} = \langle a_1, \ldots, a_n \rangle$ such that $\mathcal{A} \models_\mathcal{T} q[\vec{a}]$, where $q[\vec{a}]$ is the Boolean CQ obtained by replacing each X_k by a_k for $1 \leq k \leq n$. We note $q^{\mathcal{A},\mathcal{T}}$ the answer set of q on $\mathcal{A}$ w.r.t. $\mathcal{T}$, with the convention that for Boolean queries $q^{\mathcal{A},\mathcal{T}} = \emptyset$ if q is $false$, and $q^{\mathcal{A},\mathcal{T}} = \{\langle\rangle\}$ if q is *true* where $\langle\rangle$ is the empty tuple. A union of CQs (UCQ) is a disjunction of CQs with same arity. The answer to a UCQ is the union of the answers to its individual CQs.

3 Quotient operation for ABoxes

An ABox $\mathcal{A}$ can be represented, and is frequently drawn, as a *multigraph with typed vertices and typed edges*: the vertices are the constants in $\mathcal{A}$, there is a vertex type C on the vertex a iff the concept assertion $\mathsf{C}(a)$ is in $\mathcal{A}$, and there

is an R-typed edge $a \xrightarrow{\mathsf{R}} a'$ iff the role assertion $\mathsf{R}(a, a')$ is in $\mathcal{A}$. Based on this analogy, we transfer the quotient operation from graphs to ABoxes to define *quotient ABoxes*.

Definition 1 (Quotient ABox). *Let $\mathcal{A}$ be an ABox, $\equiv$ be some equivalence relation between constants, and let $a^1_{\equiv}, \ldots, a^n_{\equiv}$ denote by a slight abuse of notation both the equivalence classes of the constants in $\mathcal{A}$ w.r.t. $\equiv$ and the names of these equivalence classes. The* quotient ABox of $\mathcal{A}$ w.r.t. $\equiv$ *is the ABox $\mathcal{A}_{\equiv}$ such that:*

- $\mathsf{C}(a^i_{\equiv}) \in \mathcal{A}_{\equiv}$ *iff there exists $a \in a^i_{\equiv}$ such that $\mathsf{C}(a) \in \mathcal{A}$, for $1 \leq i \leq n$,*
- $\mathsf{R}(a^i_{\equiv}, a^j_{\equiv}) \in \mathcal{A}_{\equiv}$ *iff there exist $a \in a^i_{\equiv}$ and $a' \in a^j_{\equiv}$ such that $\mathsf{R}(a, a') \in \mathcal{A}$, for $1 \leq i, j \leq n$.*

In this paper, we do not focus on particular equivalence relations between constants (i.e., binary relations that are reflexive, symmetric and transitive), though in the description logic setting we consider they should be defined w.r.t. an ABox and a TBox. Discussing DL-specific equivalence relations between constants is delegated to future work.

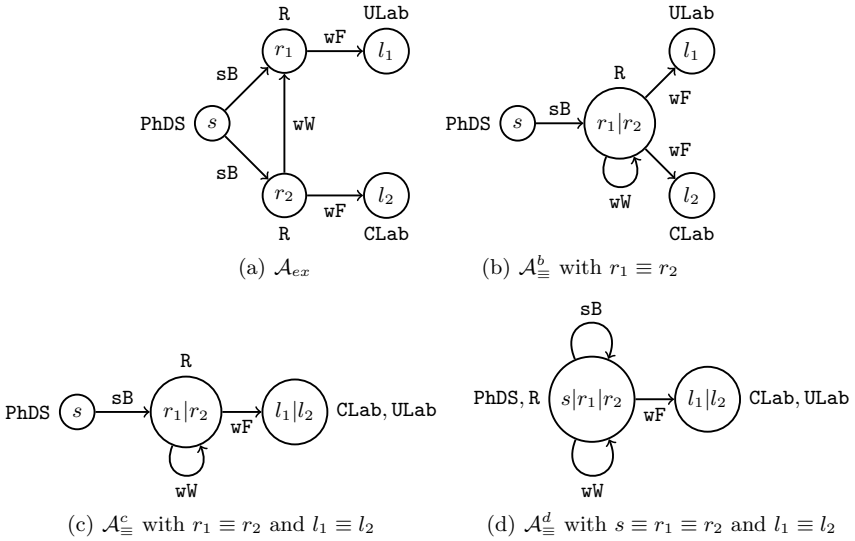
Example 1 (Running example). Let us consider the ABox
$\mathcal{A}_{ex} = \{\texttt{PhDS}(s), \texttt{sB}(s, r_1), \texttt{sB}(s, r_2), \texttt{R}(r_1), \texttt{R}(r_2), \texttt{wW}(r_2, r_1), \texttt{wF}(r_1, l_1), \texttt{wF}(r_2, l_2),$
$\texttt{ULab}(l_1), \texttt{CLab}(l_2)\}$
which states that s is a PhD student (PhDS), who is jointly supervized by (sB) the researchers (R) r_1, who works for (wF) the university lab (ULab) l_1, and r_2, who works for the company lab (CLab) l_2 and who also works with (wW) r_1.

The graph representations of $\mathcal{A}_{ex}$ and of three of quotient ABoxes obtained from it are depicted in Fig. 1. The equivalence relations used in this figure somehow reflect that equivalent constants are instances of the same concept for Fig. 1(b), or instances of non-disjoint concepts with a common generalizing concept if we assume that company labs and university labs are not necessarily disjoint kinds of labs for Fig. 1(c), i.e., $\texttt{CLab} \sqsubseteq \texttt{Lab} \in \mathcal{T}_{ex}$ and $\texttt{ULab} \sqsubseteq \texttt{Lab} \in \mathcal{T}_{ex}$, and if we further assume that PhD student and researcher are not necessarily disjoint kinds of employees for Fig. 1(d), i.e., $\texttt{PhDS} \sqsubseteq \texttt{Emp} \in \mathcal{T}_{ex}$ and $\texttt{R} \sqsubseteq \texttt{Emp} \in \mathcal{T}_{ex}$.

In the sequel, we term *summary* of an ABox $\mathcal{A}$ a quotient ABox of $\mathcal{A}$.

4 Characterization of ABox summaries

Although Definition 1 does enable the *quotient-based* summarization of an ABox, an ABox is not *just* a graph because it also has a first-order semantics. An important question that therefore arises is: *does there exist a semantic relationship between an ABox and a summary of it?* To answer this question we first point out that, in Definition 1, the *implicit function* that maps the constants in $\mathcal{A}$ to the constants in $\mathcal{A}_{\equiv}$ defines a *homomorphism* from $\mathcal{A}$ to $\mathcal{A}_{\equiv}$; from now, we name this function h and we say that an ABox constant a *is represented by* a summary



(a) $\mathcal{A}_{ex}$ (b) $\mathcal{A}^b_{\equiv}$ with $r_1 \equiv r_2$

(c) $\mathcal{A}^c_{\equiv}$ with $r_1 \equiv r_2$ and $l_1 \equiv l_2$ (d) $\mathcal{A}^d_{\equiv}$ with $s \equiv r_1 \equiv r_2$ and $l_1 \equiv l_2$

Fig. 1: The ABox $\mathcal{A}_{ex}$ and three quotient ABoxes of it: $\mathcal{A}^b_{\equiv}, \mathcal{A}^c_{\equiv}, \mathcal{A}^d_{\equiv}$.

constant $a_{\equiv}$ iff $h(a) = a_{\equiv}$ holds. Based on this ABox-to-summary homomorphism, we answer the above question by relating an ABox and a summary of it through *entailment between ABoxes w.r.t. a TBox*. The property below states that if we compare an ABox and a summary of it through entailment regardless of the constants they use (they are incomparable otherwise because they use different sets of constants), by just considering how the ABox (resp. summary) atoms *join according to these constants*, then the ABox is more general than its summary.

Property 1. Let $\mathcal{T}$ be a TBox, $a_1, \ldots, a_m$ be the constants in an ABox $\mathcal{A}$ and $a^1_{\equiv}, \ldots, a^n_{\equiv}$ be the constants in some summary $\mathcal{A}_{\equiv}$ of $\mathcal{A}$. If we consider these constants as existential variables, then $\exists a^1_{\equiv} \cdots \exists a^n_{\equiv}\ \mathcal{A}_{\equiv} \models_{\mathcal{T}} \exists a_1 \cdots \exists a_m\ \mathcal{A}$ holds.

Besides, it is also worth noting that, due to the existence of the ABox-to-summary homomorphism, an ABox and a summary of it are related through the essential data management tasks of *consistency checking* and *query answering*.

Property 2. Let $\mathcal{T}$ be a TBox, $\mathcal{A}$ be an ABox and $\mathcal{A}_{\equiv}$ be some summary of $\mathcal{A}$. If $\mathcal{A}_{\equiv}$ is consistent w.r.t. $\mathcal{T}$, then $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$.

Property 3. Let $\mathcal{T}$ be a TBox, $\mathcal{A}$ be an ABox, $\mathcal{A}_{\equiv}$ be some summary of $\mathcal{A}$, and let q be a UCQ. If $q^{\mathcal{A},\mathcal{T}} \neq \emptyset$ holds then $q_h^{\mathcal{A}_{\equiv},\mathcal{T}} \neq \emptyset$ holds, with q_h the query q in which every constant a is replaced by its image $h(a)$ through the $\mathcal{A}$-to-$\mathcal{A}_{\equiv}$ homomorphism h.

We remark that Property 2 and the contraposition of Property 3 (i.e., if $q_h^{\mathcal{A}_{\equiv},\mathcal{T}} = \emptyset$ holds then $q^{\mathcal{A},\mathcal{T}} = \emptyset$ holds) are of practical interest to check rapidly

if, *for sure*, an ABox is consistent and if a query has no answer, by using the typically (much) smaller ABox summary.

Example 2 (Cont.). From now, let us assume that $\mathcal{T}_{ex}$ just consists of the inclusions mentioned in Example 1.

To illustrate Property 1, it is easy to see on Fig. 1 that $\mathcal{A}^i_{\equiv} \models_\emptyset \mathcal{A}_{ex}$ holds for $i \in \{b, c, d\}$, hence holds with $\models_{\mathcal{T}_{ex}}$.

As for Property 2, it is also easy to see that $\mathcal{A}^i_{\equiv}$ is consistent w.r.t. $\mathcal{T}_{ex}$, for $i \in \{b, c, d\}$, as well as $\mathcal{A}_{ex}$. We remark that if PhD students and researchers would be disjoint, i.e., $\texttt{PhDS} \sqsubseteq \neg\texttt{R} \in \mathcal{T}_{ex}$, then $\mathcal{A}^d_{\equiv}$ would be inconsistent w.r.t. $\mathcal{T}_{ex}$ while $\mathcal{A}_{ex}$ would be consistent w.r.t. $\mathcal{T}_{ex}$. This observation indicates that DL-specific equivalence relations should avoid to the extent possible building inconsistent summaries out of consistent ABoxes.

Finally, let us illustrate Property 3 on Fig. 1. Consider the UCQ $q = \texttt{R}(X) \wedge \texttt{wF}(X, l_1) \cup \texttt{R}(X) \wedge \texttt{wF}(X, l_2)$ asking for the researchers who work for the lab l_1 or l_2. q has some answer on $\mathcal{A}_{ex}$ w.r.t. $\mathcal{T}_{ex}$ ($q^{\mathcal{A}_{ex}, \mathcal{T}_{ex}} = \{\langle r_1\rangle, \langle r_2\rangle\}$), thus its translation q_h has some answer w.r.t. $\mathcal{T}_{ex}$ on the summary $\mathcal{A}^i_{\equiv}$, $i \in \{b, c, d\}$. For the CQ $q = \texttt{PhDS}(X_1) \wedge \texttt{wW}(X_1, X_2)$ asking for the PhD students and people with whom they work, we note that because the q_h translation does not have answers on $\mathcal{A}^i_{\equiv}$ w.r.t. $\mathcal{T}_{ex}$, $i \in \{b, c\}$, i.e., $q_h^{\mathcal{A}^i_{\equiv}, \mathcal{T}} = \emptyset$, hence by contraposition of Property 3, q does not have answer on $\mathcal{A}_{ex}$ w.r.t. $\mathcal{T}_{ex}$ either, i.e., $q^{\mathcal{A}_{ex}, \mathcal{T}_{ex}} = \emptyset$. However, we observe that q_h has some answer on $\mathcal{A}^d_{\equiv}$ w.r.t. $\mathcal{T}_{ex}$, while q has no answer on $\mathcal{A}_{ex}$ w.r.t. $\mathcal{T}_{ex}$, i.e., this particular summary fails in detecting that q has no answer on $\mathcal{A}_{ex}$ w.r.t. $\mathcal{T}_{ex}$. This observation suggests that DL-specific equivalence relations should be carefully chosen depending on the target summary usage.

Resuming our initial discussion, though Property 1 states that a summary is *more specific than* the ABox it summarizes up to the set of constants they use, it immediately raises a second crucial question: *to which extent a summary is more specific than the original ABox it summarizes?* We answer this second question by characterizing to which extent the *knowledge* a summary has about a constant is more specific than the knowledge the ABox has about the equivalence class this summary constant represents, and similarly for a pair of summary constants and the pair of equivalence classes these summary constants represent in the ABox.

For an ABox $\mathcal{A}$ and a TBox $\mathcal{T}$, the most precise knowledge about a set of constants $a_1, \ldots, a_n$ corresponds to one or more concepts known as the *most specific concepts (MSCs)* $a_1, \ldots, a_n$ are all instances of [2]. A concept C is *a* MSC of $a_1, \ldots, a_n$ w.r.t. $\mathcal{A}$ and $\mathcal{T}$ iff (i) $\mathcal{A} \models_\mathcal{T} \mathsf{C}(a_i)$ holds for all $i \in [1, n]$, and (ii) for any concept D such that $\mathcal{A} \models_\mathcal{T} \mathsf{D}(a_i)$ holds for all $i \in [1, n]$, either $\mathsf{C} \preceq_\mathcal{T} \mathsf{D}$ holds or $\mathsf{C} \not\preceq_\mathcal{T} \mathsf{D}$ *and* $\mathsf{D} \not\preceq_\mathcal{T} \mathsf{C}$ hold. We remark that though a MSC may be of infinite size in some DLs, this is not an issue here as we just want a semantic characterization of it. In the sequel, we designate by $msc^{\mathcal{A}, \mathcal{T}}(a_1, \ldots, a_n)$ the *conjunction* ($\sqcap$) of the MSCs of $a_1, \ldots, a_n$ w.r.t. $\mathcal{A}$ and $\mathcal{T}$ [1] and we refer to it as *the* MSC of $a_1, \ldots, a_n$ w.r.t. $\mathcal{A}$ and $\mathcal{T}$, because it *always exists* (the

[1] $a_1, \ldots, a_n$ obviously have a *single* MSC in DLs equipped with conjunction ($\sqcap$).

empty conjunction is the universal concept $\top$) and it is obviously *unique* up to equivalence.

A first step towards answering our second question is Lemma 1, which provides an *upper approximation through subsumption* of the MSC of some summary constant w.r.t. the individual MSCs of the equivalent ABox constants it represents; Lemma 1 follows from the existence of the ABox-to-summary homomorphism.

Lemma 1. *Let $\mathcal{T}$ be a TBox, $\mathcal{A}$ be an ABox and $\mathcal{A}_{\equiv}$ be the summary of $\mathcal{A}$ w.r.t. the $\equiv$ equivalence relation. If $a_1, \ldots, a_n$ are all the constants in $\mathcal{A}$ that belong to the equivalence class $a_{\equiv}$ according to $\equiv$, then the following holds:*

$$msc^{\mathcal{A}_{\equiv},\mathcal{T}}(a_{\equiv}) \preceq_{\mathcal{T}} \prod_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i).$$

Example 3 (Cont.). Let us consider the summary $\mathcal{A}^d_{\equiv}$ of $\mathcal{A}_{ex}$ in Fig. 1 and illustrate Lemma 1 with the summary constant $l_1|l_2$ and the ABox constants l_1, l_2 it represents:
$msc^{\mathcal{A}^d_{\equiv},\mathcal{T}_{ex}}(l_1|l_2) = \texttt{Clab} \sqcap \texttt{ULab} \sqcap \exists \texttt{wF}^-$,
$msc^{\mathcal{A}_{ex},\mathcal{T}_{ex}}(l_1) = \texttt{ULab} \sqcap \exists \texttt{wF}^-$,
$msc^{\mathcal{A}_{ex},\mathcal{T}_{ex}}(l_2) = \texttt{CLab} \sqcap \exists \texttt{wF}^-$,
and $msc^{\mathcal{A}^d_{\equiv},\mathcal{T}_{ex}}(l_1|l_2) \preceq_{\mathcal{T}_{ex}} \sqcap_{x \in \{l_1,l_2\}} msc^{\mathcal{A}_{ex},\mathcal{T}_{ex}}(x)$, i.e., $\texttt{Clab} \sqcap \texttt{ULab} \sqcap \exists \texttt{wF}^- \preceq_{\mathcal{T}_{ex}} (\texttt{ULab} \sqcap \exists \texttt{wF}^-) \sqcap (\texttt{CLab} \sqcap \exists \texttt{wF}^-)$, clearly holds.

At the same time, in the above Lemma 1, to the set of (equivalent) constants $a_1, \ldots, a_n$ corresponds its most specific concept $msc^{\mathcal{A},\mathcal{T}}(a_1, \ldots, a_n)$, as recalled above. As the next step towards answering our second question, Lemma 2 provides a *lower approximation through subsumption* of the MSC of $a_1, \ldots, a_n$ w.r.t. the individual MSCs of these constants; it follows from the definition of MSC.

Lemma 2. *Let $\mathcal{T}$ be a TBox and $\mathcal{A}$ be an ABox. If $a_1, \ldots, a_n$ are constants in $\mathcal{A}$, then the following holds:*

$$\bigsqcup_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i) \preceq_{\mathcal{T}} msc^{\mathcal{A},\mathcal{T}}(a_1, \ldots, a_n).$$

Example 4 (Cont.). Let us consider again the summary $\mathcal{A}^d_{\equiv}$ of $\mathcal{A}_{ex}$ in Fig. 1 and illustrate Lemma 2 with the summary constant $l_1|l_2$ and the ABox constants l_1, l_2 it represents like in the preceding example:
$msc^{\mathcal{A}_{ex},\mathcal{T}_{ex}}(l_1, l_2) = \texttt{Lab} \sqcap \exists \texttt{wF}^-$,
and $\sqcup_{x \in \{l_1,l_2\}} msc^{\mathcal{A}_{ex},\mathcal{T}_{ex}}(x) \preceq_{\mathcal{T}} msc^{\mathcal{A}_{ex},\mathcal{T}_{ex}}(l_1, l_2)$, i.e., $(\texttt{ULab} \sqcap \exists \texttt{wF}^-) \sqcup (\texttt{CLab} \sqcap \exists \texttt{wF}^-) \preceq_{\mathcal{T}_{ex}} \texttt{Lab} \sqcap \exists \texttt{wF}^-$, clearly holds.

We are now ready to answer our second question with Theorem 1, which is a direct corollary of Lemma 1 and Lemma 2. It relates $msc^{\mathcal{A},\mathcal{T}}(a_{\equiv})$ and $msc^{\mathcal{A},\mathcal{T}}(a_1, \ldots, a_n)$ through their respective upper and lower approximations.

Theorem 1. *Let $\mathcal{T}$ be a TBox, $\mathcal{A}$ be an ABox and $\mathcal{A}_{\equiv}$ be the summary of $\mathcal{A}$ w.r.t. the $\equiv$ equivalence relation. If $a_1, \ldots, a_n$ are all the constants in $\mathcal{A}$ that belong to the equivalence class $a_{\equiv}$ according to $\equiv$, then the following holds:*

$$msc^{\mathcal{A}_{\equiv},\mathcal{T}}(a_{\equiv}) \preceq_{\mathcal{T}} \prod_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i) \preceq_{\mathcal{T}} \bigsqcup_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i) \preceq_{\mathcal{T}} msc^{\mathcal{A},\mathcal{T}}(a_1, \ldots, a_n).$$

From Theorem 1, it is now clear that the *semantic distance through subsumption* between the knowledge about each constant $a_{\equiv}$ in the summary, i.e., $msc(a_{\equiv})$, and the knowledge about the set of equivalent constants that $a_{\equiv}$ represents in the ABox, i.e., $msc^{\mathcal{A},\mathcal{T}}(a_1, \ldots, a_n)$, corresponds *at least* to approximate the *union* of individual knowledge about $a_1, \ldots, a_n$ in the ABox by a *conjunction*, i.e., approximating $\bigsqcup_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i)$ by $\prod_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i)$.

We establish similar results (omitted due to space limitations) by adapting the notion of MSC to *roles*.

Finally, we stress that all the above results just follow from the definition of quotient ABox (Definition 1). In particular, they are *independent* of the equivalence relation ($\equiv$) and of the DL under consideration.

5 Related work

We present below the works from the DL literature that are closely related to ABox summarization.

An *ABox summarization* approach for the $\mathcal{SHIN}$ DL has been studied to optimize *consistency checking* [8] and *instance retrieval* in the subsequent work [6], i.e., finding the instances of a given $\mathcal{SHIN}$ concept; these two summary-based optimizations are implemented in the SHER system [7]. This approach can be seen as a particular instantiation for the $\mathcal{SHIN}$ DL of our general ABox summarization framework, in which equivalent constants are non-distinct constants[2] stored in the ABox as instances of the same $\mathcal{SHIN}$ concepts; our framework allows using this definition of equivalent constants for the equivalence relation $\equiv$ or any other relevant to the target summary usage. We therefore provide a better understanding of the $\mathcal{SHIN}$ summaries of [8,6,7] by establishing that they are more specific than the $\mathcal{SHIN}$ ABoxes they summarize, and notably to which extent they are more specific. Further, we remark that our ABox summarization approach, because it is based on the quotient operation, is more *declarative* than that of [8,6,7], by elegantly decoupling the summarization means (i.e., fusing equivalent constants) from the high-level specification of equivalent constants, i.e., the equivalence relation $\equiv$ to use. By contrast, $\mathcal{SHIN}$ summaries are defined by an *explicit canonical function* that maps ABox constants to summary one; our function h has similar goal but is *implicit* because derived from the ABox and equivalence relation at hand.

Another form of ABox summarization is *ABox abstraction*, which has been studied for *materialization* in the Horn-$\mathcal{ALCHOI}$ and Horn-$\mathcal{SHOIF}$ DLs [9,10],

[2] In $\mathcal{SHIN}$, constants are stated distinct by using assertions of the built-in role $\neq$.

i.e., for pre-computing and storing entailed assertions. The idea is to represent several assertions of an ABox by a single one in the ABox abstraction. To this aim, an ABox $\mathcal{B}$ is defined as an abstraction of an ABox $\mathcal{A}$ if there exists a *mapping* from the constants in $\mathcal{B}$ to the constants in $\mathcal{A}$ that defines a homomorphism from $\mathcal{B}$ to $\mathcal{A}$. Hence, by contrast with our summaries that are more specific than original ABoxes, abstractions are *more general* than original ABoxes, thus capture a subset of their contents.

Also related to ABox summarization is *ABox modularization*, which has been studied for *instance retrieval* in the $\mathcal{SHIF}$ [13] and $\mathcal{SHI}$ [15] DLs. It consists in extracting a subset of an ABox that entails all the assertions in which a given constant is involved.

Finally, less related works on ABox summarization *do not* summarize an ABox by another, e.g., [16] which computes statistics about the concept and roles instances, and *Abstract Knowledge Patterns* of the form $(\mathsf{C}_1, \mathsf{r}, \mathsf{C}_2)$ indicating that the role r relates instances of the concept C_1 to instances of the concept C_2.

6 Conclusion and perspectives

We devised a well-founded ABox summarization framework for DLs, by applying the quotient operation from graph theory to ABoxes and examining the quotient ABoxes, i.e., summaries, it produces.

With this now well-understood ABox summarization framework in place, the next necessary step is to study *DL-specific equivalence relations* between constants, to explore, visualize or optimize reasoning on and the management of ABoxes. These equivalence relations might be sought either from scratch or derived from the ones used for graph summarization if meaningful w.r.t. the DL semantics, e.g., the well-established and widely adopted ones based on bisimilation [14] or the recent ones based on the cooccurrence of types and edges [11]. Also, ideally, such relations should be devised to avoid to the extent possible building inconsistent summaries out of consistent ABoxes (recall Example 2).

Acknowledgments

This work was partially supported by the ANR project CQFD (ANR-18-CE23-0003).

Appendix

Proof (Proof of Property 1). In case $\exists a^1_{\equiv}, \ldots, a^n_{\equiv}\, \mathcal{A}_{\equiv}$ is inconsistent w.r.t. $\mathcal{T}$, then Property 1 trivially holds. Otherwise, let $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ be a model of $\mathcal{T}$ and $\exists a^1_{\equiv}, \ldots, a^n_{\equiv}\, \mathcal{A}_{\equiv}$ and let us show that $\mathcal{I}$ is also a model of $\exists a_1, \ldots, a_m\, \mathcal{A}$.

Because $\mathcal{I}$ is a model of $\exists a^1_{\equiv}, \ldots, a^n_{\equiv}\, \mathcal{A}_{\equiv}$, there exists a function ψ from $a^1_{\equiv}, \ldots, a^n_{\equiv}$ to the interpretation domain $\Delta^{\mathcal{I}}$ of $\mathcal{I}$ such that $\psi(a^i_{\equiv}) \in \mathsf{C}^{\mathcal{I}}$ for every

$\mathsf{C}(a^i_{\equiv}) \in \mathcal{A}_{\equiv}$ and $1 \leq i \leq n$, and such that $(\psi(a^i_{\equiv}), \psi(a^j_{\equiv})) \in \mathsf{R}^{\mathcal{I}}$ for every $\mathsf{R}(a^i_{\equiv}, a^j_{\equiv}) \in \mathcal{A}_{\equiv}$ and $1 \leq i, j \leq n$.

Let us consider the composite function $\psi \circ h$, where ψ is the above function and h is the implicit function that defines our ABox-to-summary homomorphism. This function thus maps $a_1, \ldots, a_m$ to the interpretation domain of $\mathcal{I}$ and clearly $\psi(h(a_i)) \in \mathsf{C}^{\mathcal{I}}$ for every $\mathsf{C}(a_i) \in \mathcal{A}$ and $1 \leq i \leq m$, because by construction of $\mathcal{A}_{\equiv}$ $\mathsf{C}(h(a_i)) \in \mathcal{A}_{\equiv}$, and such that $(\psi(h(a_i)), \psi(h(a_j))) \in \mathsf{R}^{\mathcal{I}}$ for every $\mathsf{R}(a_i, a_j) \in \mathcal{A}$ and $1 \leq i, j \leq m$, because by construction of $\mathcal{A}_{\equiv}$ $\mathsf{R}(h(a_i), h(a_j)) \in \mathcal{A}_{\equiv}$. Therefore, $\mathcal{I}$ is also a model of $\exists a_1, \ldots, a_m\, \mathcal{A}$. □

Proof (Proof of Property 2). If $\mathcal{A}_{\equiv}$ is consistent w.r.t. $\mathcal{T}$, there must exist some model $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ of $\mathcal{A}_{\equiv}$. Let us show that we can build a model of $\mathcal{A}$ out of $\mathcal{I}$, hence that $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$.

Consider the model $\mathcal{J} = (\Delta^{\mathcal{J}}, \cdot^{\mathcal{J}})$ defined as follows: $\Delta^{\mathcal{J}} = \Delta^{\mathcal{I}}$, $A^{\mathcal{J}} = A^{\mathcal{I}}$ for every concept name $A \in N_C$, $R^{\mathcal{J}} = R^{\mathcal{I}}$ for every role name $R \in N_R$, and for every constant a in $\mathcal{A}$, $a^{\mathcal{J}} = h(a)^{\mathcal{I}}$.

For every concept assertion $\mathsf{C}(a) \in \mathcal{A}$, $\mathsf{C}(h(a)) \in \mathcal{A}_{\equiv}$ by definition of a quotient ABox, where h is the implicit function defining the ABox-to-summary homomorphism. Since $\mathcal{I}$ is a model of $\mathcal{A}_{\equiv}$, $\mathcal{I}$ is a model of $\mathsf{C}(h(a))$, i.e., $h(a)^{\mathcal{I}} \in \mathsf{C}^{\mathcal{I}}$. By definition of $\mathcal{J}$, $\mathsf{C}^{\mathcal{J}} = \mathsf{C}^{\mathcal{I}}$ and $a^{\mathcal{J}} = h(a)^{\mathcal{I}}$, thus $a^{\mathcal{J}} \in \mathsf{C}^{\mathcal{J}}$, i.e., $\mathcal{J}$ is a model of $\mathsf{C}(a)$. $\mathcal{J}$ is therefore a model of all the concept assertions in $\mathcal{A}$.

Similarly, for every role assertion $\mathsf{R}(a_k, a_l) \in \mathcal{A}$, $\mathsf{R}(h(a_k), h(a_l)) \in \mathcal{A}_{\equiv}$ by definition of a quotient ABox, where h is the implicit function defining the ABox-to-summary homomorphism. Since $\mathcal{I}$ is a model of $\mathcal{A}_{\equiv}$, $\mathcal{I}$ is a model of $\mathsf{R}(h(a_k), h(a_l))$, i.e., $(h(a_k)^{\mathcal{I}}, h(a_l)^{\mathcal{I}}) \in \mathsf{R}^{\mathcal{I}}$. By definition of $\mathcal{J}$, $\mathsf{R}^{\mathcal{J}} = \mathsf{R}^{\mathcal{I}}$, $a_k^{\mathcal{J}} = h(a_k)^{\mathcal{I}}$ and $a_l^{\mathcal{J}} = h(a_l)^{\mathcal{I}}$, thus $(a_k^{\mathcal{J}}, a_l^{\mathcal{J}}) \in \mathsf{R}^{\mathcal{J}}$, i.e., $\mathcal{J}$ is a model of $\mathsf{R}(a_k, a_l)$. $\mathcal{J}$ is therefore a model of all the role assertions in $\mathcal{A}$.

It therefore follows that $\mathcal{J}$ is a model of all the assertions in $\mathcal{A}$, hence a model of $\mathcal{A}$. □

Proof (Proof of Property 3). We remark that if a query is asked on an ABox that is inconsistent w.r.t. its associated TBox, then both *true* and $false$ if the query is Boolean, and all the tuples of the arity of the query otherwise, are answers.

If $\mathcal{A}$ is not consistent w.r.t. $\mathcal{T}$, then $q^{\mathcal{A},\mathcal{T}} \neq \emptyset$, and by the contraposition of Property 2, $\mathcal{A}_{\equiv}$ is not consistent w.r.t. $\mathcal{T}$ too, therefore $q_h^{\mathcal{A}_{\equiv},\mathcal{T}} \neq \emptyset$ and Property 3 holds in this first case.

If $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$, while $\mathcal{A}_{\equiv}$ is not consistent w.r.t. $\mathcal{T}$ (see Example 2), then Property 3 holds in this second case: $q_h^{\mathcal{A}_{\equiv},\mathcal{T}} \neq \emptyset$ holds whenever $q^{\mathcal{A},\mathcal{T}} \neq \emptyset$ or $q^{\mathcal{A},\mathcal{T}} = \emptyset$ holds.

Otherwise, $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$ *and* $\mathcal{A}_{\equiv}$ is consistent w.r.t. $\mathcal{T}$. Let us show that Property 3 holds in this third and last case. If the UCQ q has a non-empty answer set, then at least one of its CQs has a non-empty answer set. Let q' be such a CQ in q. Let us assume, without loss of generality, that q' is Boolean as the definition of non-Boolean query answers reduces to the Boolean case. Let us consider some model $\mathcal{J}$ of $\mathcal{A}_{\equiv}$ and $\mathcal{T}$. Let $\mathcal{I}$ be the interpretation such that $\Delta^{\mathcal{I}} = \Delta^{\mathcal{J}}$, $A^{\mathcal{I}} = A^{\mathcal{J}}$ for every concept name $A \in N_C$, $R^{\mathcal{I}} = R^{\mathcal{J}}$ for

every role name $R \in N_R$, and $a^{\mathcal{I}} = h(a)^{\mathcal{J}}$ for every constant a in $\mathcal{A}$. Clearly, by construction, $\mathcal{I}$ is homomorphic to $\mathcal{J}$ and is a model of $\mathcal{A}$ and $\mathcal{T}$. Thus, $\mathcal{I}$ is also a model of q'. Further, $\mathcal{J}$ is obviously a model of q' just in case q' does not contain contants, because the constants in $\mathcal{A}$ do not have interpretations in $\mathcal{J}$: $\Delta^{\mathcal{I}} = \Delta^{\mathcal{J}}$, $A^{\mathcal{I}} = A^{\mathcal{J}}$ for every concept name $A \in N_C$, $R^{\mathcal{I}} = R^{\mathcal{J}}$ for every role name $R \in N_R$. Otherwise, $\mathcal{J}$ is a model of q'_h, in which constants from $\mathcal{A}$ are replaced by the constants in $\mathcal{A}_{\equiv}$ they are represented by: $a^{\mathcal{I}} = h(a)^{\mathcal{J}}$ for every constant a in $\mathcal{A}$. Therefore, every model of $\mathcal{A}_{\equiv}$ and $\mathcal{T}$ is a model of q'_h, i.e., q'_h is entailed by $\mathcal{A}_{\equiv}$ and $\mathcal{T}$, thus has a non-empty answer set, hence q_h has a non-empty answer set too. □

Proof (Proof of Lemma 1). If $\mathcal{A}$ is inconsistent w.r.t. $\mathcal{T}$, then $\mathcal{A}_{\equiv}$ is also inconsistent w.r.t. $\mathcal{T}$ (by contraposition of Property 2), and the Lemma clearly holds.

If $\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$ while $\mathcal{A}_{\equiv}$ is inconsistent w.r.t. $\mathcal{T}$, then the Lemma also clearly holds.

Otherwise, both $\mathcal{A}$ and $\mathcal{A}_{\equiv}$ are consistent w.r.t. $\mathcal{T}$, and let us consider some model $\mathcal{J}$ of $\mathcal{A}_{\equiv}$ and $\mathcal{T}$. Let $\mathcal{I}$ be the interpretation such that $\Delta^{\mathcal{I}} = \Delta^{\mathcal{J}}$, $A^{\mathcal{I}} = A^{\mathcal{J}}$ for every concept name $A \in N_C$, $R^{\mathcal{I}} = R^{\mathcal{J}}$ for every role name $R \in N_R$, and $a^{\mathcal{I}} = h(a)^{\mathcal{J}}$ for every constant a in $\mathcal{A}$. Clearly, by construction, $\mathcal{I}$ is homomorphic to $\mathcal{J}$ and is a model of $\mathcal{A}$ and $\mathcal{T}$. Because $\mathcal{I}$ is a model of $\mathcal{A}$ and $\mathcal{T}$, $\mathcal{I}$ is a model of the concept assertion $(msc^{\mathcal{A},\mathcal{T}}(a_i))(a_i)$ for $1 \leq i \leq n$. By construction of $\mathcal{I}$, $\mathcal{J}$ is a model of $(msc^{\mathcal{A},\mathcal{T}}(a_i))(h(a_i))$ for $1 \leq i \leq n$, where h is the ABox-to-summary homomorphism, i.e., $\mathcal{J}$ is a model of $(msc^{\mathcal{A},\mathcal{T}}(a_i))(a_{\equiv})$ for $1 \leq i \leq n$.

It follows that all the models of $\mathcal{A}_{\equiv}$ and $\mathcal{T}$ are models of $(msc^{\mathcal{A},\mathcal{T}}(a_i))(a_{\equiv})$ for $1 \leq i \leq n$, and of $(msc^{\mathcal{A}_{\equiv},\mathcal{T}}(a_{\equiv}))(a_{\equiv})$ $(*)$. By definition of the MSC of $a_{\equiv}$, either (i) $msc^{\mathcal{A}_{\equiv},\mathcal{T}}(a_{\equiv}) \preceq_{\mathcal{T}} \bigsqcap_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i)$ holds or (ii) $msc^{\mathcal{A}_{\equiv},\mathcal{T}}(a_{\equiv}) \not\preceq_{\mathcal{T}} \bigsqcap_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i)$ and $\bigsqcap_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i) \not\preceq_{\mathcal{T}} msc^{\mathcal{A}_{\equiv},\mathcal{T}}(a_{\equiv})$ hold. In the latter case (ii), $msc^{\mathcal{A}_{\equiv},\mathcal{T}}(a_{\equiv}) \sqcap \bigsqcap_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i)$ would be strictly subsumed by $msc^{\mathcal{A}_{\equiv},\mathcal{T}}(a_{\equiv})$ w.r.t. $\mathcal{T}$, and because of $(*)$ above, $msc^{\mathcal{A}_{\equiv},\mathcal{T}}(a_{\equiv}) \sqcap \bigsqcap_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i)$ would be the MSC of $a_{\equiv}$, a contradiction. Therefore, we must be in case (i) and $msc^{\mathcal{A}_{\equiv},\mathcal{T}}(a_{\equiv}) \preceq_{\mathcal{T}} \bigsqcap_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i)$ holds. □

Proof (Proof of Lemma 2). Lemma 2 follows the well-known relationship between the MSC of constants and the notion of *least common subsumer (LCS) of concepts* [2], i.e., a most specific concept generalizing all of them. A concept C is a LCS of the concepts $\mathsf{C}_1, \ldots, \mathsf{C}_n$ w.r.t. a TBox $\mathcal{T}$ iff (i) $\mathsf{C}_i \preceq_{\mathcal{T}} \mathsf{C}$ holds for all $i \in [1, n]$, and (ii) for any concept D such that $\mathsf{C}_i \preceq_{\mathcal{T}} \mathsf{D}$ holds for all $i \in [1, n]$, either $\mathsf{C} \preceq_{\mathcal{T}} \mathsf{D}$ holds or $\mathsf{C} \not\preceq_{\mathcal{T}} \mathsf{D}$ and $\mathsf{D} \not\preceq_{\mathcal{T}} \mathsf{C}$ hold. We denote by $lcs^{\mathcal{T}}(\mathsf{C}_1, \ldots, \mathsf{C}_n)$ the *conjunction of the LCSs* of $\mathsf{C}_1, \ldots, \mathsf{C}_n$ w.r.t. $\mathcal{T}$ and we refer to it as *the* LCS of $\mathsf{C}_1, \ldots, \mathsf{C}_n$ w.r.t. $\mathcal{T}$. In particular, it is well-known that $msc^{\mathcal{A},\mathcal{T}}(a_1, \ldots, a_n)$ and $lcs^{\mathcal{T}}(msc^{\mathcal{A},\mathcal{T}}(a_1), \ldots, msc^{\mathcal{A},\mathcal{T}}(a_n))$ are *equivalent* [2]. Further, it is also known that $lcs^{\mathcal{T}}(msc^{\mathcal{A},\mathcal{T}}(a_1), \ldots, msc^{\mathcal{A},\mathcal{T}}(a_n))$ *cannot* be more specific than $\bigsqcup_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i)$, notably they are equivalent in DLs allowing disjunction ($\sqcup$). Thus, $\bigsqcup_{i=1}^{n} msc^{\mathcal{A},\mathcal{T}}(a_i) \preceq_{\mathcal{T}} msc^{\mathcal{A},\mathcal{T}}(a_1, \ldots, a_n)$ holds. □

Proof (Proof of Theorem 1). The proof trivially follows from Lemma 1, Lemma 2, and the standard first-order semantics of disjunction $\sqcup$ and conjunction $\sqcap$ in DLs. □

References

1. Abriola, S., Barceló, P., Figueira, D., Figueira, S.: Bisimulations on data graphs. In: KR (2016)
2. Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F. (eds.): The Description Logic Handbook. Cambridge University Press (2003)
3. Bienvenu, M.: Ontology-mediated query answering: Harnessing knowledge to get more from data. In: IJCAI (2016)
4. Calvanese, D., Giacomo, G.D., Lembo, D., Lenzerini, M., Rosati, R.: Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. J. Autom. Reason. **39**(3) (2007)
5. Cebiric, S., Goasdoué, F., Kondylakis, H., Kotzinos, D., Manolescu, I., Troullinou, G., Zneika, M.: Summarizing semantic graphs: a survey. VLDB J. **28**(3) (2019)
6. Dolby, J., Fokoue, A., Kalyanpur, A., Kershenbaum, A., Schonberg, E., Srinivas, K., Ma, L.: Scalable semantic retrieval through summarization and refinement. In: AAAI (2007)
7. Dolby, J., Fokoue, A., Kalyanpur, A., Schonberg, E., Srinivas, K.: Scalable highly expressive reasoner (SHER). J. Web Semant. **7**(4) (2009)
8. Fokoue, A., Kershenbaum, A., Ma, L., Schonberg, E., Srinivas, K.: The summary abox: Cutting ontologies down to size. In: ISWC (2006)
9. Glimm, B., Kazakov, Y., Liebig, T., Tran, T., Vialard, V.: Abstraction refinement for ontology materialization. In: ISWC (2014)
10. Glimm, B., Kazakov, Y., Tran, T.: Ontology materialization by abstraction refinement in horn SHOIF. In: Singh, S.P., Markovitch, S. (eds.) AAAI (2017)
11. Goasdoué, F., Guzewicz, P., Manolescu, I.: RDF graph summarization for first-sight structure discovery. VLDB J. **29**(5) (2020)
12. Gross, J.L., Yellen, J., Zhang, P. (eds.): Handbook of Graph Theory. Discrete Mathematics and Its Applications, Chapman & Hall / CRC Press, Taylor & Francis (2013)
13. Guo, Y., Heflin, J.: A scalable approach for partitioning owl knowledge bases. In: International Workshop on Scalable Semantic Web Knowledge Base Systems (2006)
14. Liu, Y., Safavi, T., Dighe, A., Koutra, D.: Graph summarization methods and applications: A survey. ACM Comput. Surv. **51**(3) (2018)
15. Möller, R., Neuenstadt, C., Özçep, Ö.L., Wandelt, S.: Advances in accessing big data with expressive ontologies. In: International Workshop on Description Logics (2013)
16. Palmonari, M., Rula, A., Porrini, R., Maurino, A., Spahiu, B., Ferme, V.: ABSTAT: linked data summaries with abstraction and statistics. In: ESWC (2015)
17. Xiao, G., Calvanese, D., Kontchakov, R., Lembo, D., Poggi, A., Rosati, R., Zakharyaschev, M.: Ontology-based data access: A survey. In: IJCAI (2018)