Answering Conjunctive Queries with Inequalities in *DL-Lite*$_\mathcal{R}$

Gianluca Cima,[1] Maurizio Lenzerini,[1] Antonella Poggi[2,1]
[1]Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Sapienza Università di Roma
[2]Dipartimento di Lettere e Culture Moderne, Sapienza Università di Roma
{cima, lenzerini, poggi}@diag.uniroma1.it

Abstract

In the context of the Description Logic *DL-Lite*$_\mathcal{R}^{\neq}$, i.e., *DL-Lite*$_\mathcal{R}$ without UNA and with inequality axioms, we address the problem of adding to unions of conjunctive queries (UCQs) one of the simplest forms of negation, namely, inequality. It is well known that answering conjunctive queries with unrestricted inequalities over *DL-Lite*$_\mathcal{R}$ ontologies is in general undecidable. Therefore, we explore two strategies for recovering decidability, and, hopefully, tractability. Firstly, we weaken the ontology language, and consider the variant of *DL-Lite*$_\mathcal{R}^{\neq}$ corresponding to RDFS enriched with both inequality and disjointness axioms. Secondly, we weaken the query language, by preventing inequalities to be applied to existentially quantified variables, thus obtaining the class of queries named UCQ$^{\neq,b}$s. We prove that in the two cases, query answering is decidable, and we provide tight complexity bounds for the problem, both for data and combined complexity. Notably, the results show that answering UCQ$^{\neq,b}$s over *DL-Lite*$_\mathcal{R}^{\neq}$ ontologies is still in AC^0 in data complexity.

1 Introduction

Description Logics (DLs) (Baader et al. 2003; 2017) allow for defining ontologies in terms of two components, named TBox (general axioms on the concepts and relations in the domain of interest), and ABox (axioms about instances of concepts and relations). In this paper we consider *DL-Lite*$_\mathcal{R}$, which is the DL of the *DL-Lite* family (Calvanese et al. 2004; 2007) underpinning the `OWL 2` profile `OWL 2 QL` (Motik et al. 2012), and is arguably one of the most important formalisms in Ontology-Based Data Access (OBDA) (Poggi et al. 2008; Bienvenu 2016; Xiao et al. 2018; Ortiz 2018), where the aim is to access a typically huge amount of data represented as an ABox, either materialized or virtual. In particular, *DL-Lite*$_\mathcal{R}$ has been designed so that answering unions of conjunctive queries (UCQs) posed to an ontology expressed in this language can be reduced to evaluating first-order logic queries over the database corresponding to the ABox, and therefore the problem is in AC^0 in the size of the ABox, i.e., in the so-called *data complexity* (Vardi 1982).

Although UCQs constitute the most popular class of queries studied for both databases and ontologies, they have several limitations in expressive power. Notably, they do not allow any form of negation, not even the one expressed by the inequality (i.e., "not equal") predicate. For instance, the query computing all triangles in an undirected graph cannot be expressed as a conjunctive query (CQ), whereas it can be expressed as the following CQ with inequalities $\{(x, y, z) \mid \mathtt{edge}(x, y), \mathtt{edge}(x, z), \mathtt{edge}(y, z), x \neq y, x \neq z, y \neq z\}$, where the predicate `edge` represents the connections between nodes in the graph.

The above example shows that inequalities are indeed necessary for expressing even very simple properties, like triangle graphs. However, while answering UCQs in *DL-Lite*$_\mathcal{R}$ has been extensively studied in recent years (Xiao et al. 2018), the problem of answering CQs with inequalities (CQ$^{\neq}$s) and unions thereof (UCQ$^{\neq}$s) has been rarely investigated. To the best of our knowledge, the basic facts that are known about such problem can be summarized as follows.

- In stark contrast to the UCQ case, answering UCQ$^{\neq}$s is undecidable, even in the case of ontologies expressed in *DL-Lite*$_{core}$, which is the fragment of *DL-Lite*$_\mathcal{R}$ without role disjointness and role inclusion axioms (Gutiérrez-Basulto, Ibáñez-García, and Kontchakov 2012). For *DL-Lite*$_\mathcal{R}$ ontologies, undecidability holds already for CQ$^{\neq}$s (Gutiérrez-Basulto et al. 2015). Looking at these results, one can easily realize that the main of source of undecidability stems from both the ability of the ontology language to express incomplete information through existential quantifiers, and the possibility of imposing inequalities between existential variables in the query.
- In (Gutiérrez-Basulto et al. 2015) it is also proved that for the subclasses of CQ$^{\neq}$s and UCQ$^{\neq}$s named *local* CQ$^{\neq}$s and *local* UCQ$^{\neq}$s, respectively, query answering over *DL-Lite*$_\mathcal{R}$ ontologies is decidable, but with a coNExpTime upper bound in data complexity. Furthermore, it is provably intractable (in general coNP-hard in data complexity) already for *local* CQ$^{\neq}$s. Local (U)CQs are special (U)CQs with inequalities, designed in such a way that each inequality atom in the query that contributes to a certain answer with respect to a DL ontology has at least one of its terms bound by an individual in the ABox.

The goal of this paper is to investigate under which conditions, stronger than local UCQs, tractability of answering queries with inequalities is recovered, or at least the complexity is lowered with respect to the one of local UCQs. The basic idea to achieve this goal is to explore ontology languages and query languages ensuring the following property: each inequality atom $\alpha \neq \beta$ that contributes to the certain answer to a query with respect to a DL ontology, it does so with *both* terms α and β bounded by individuals in the ABox. In order to follow this path, we consider as basic language *DL-Lite*$_{\mathcal{R}}$ without the Unique Name Assumption (UNA)[1] and with inequality axioms, called *DL-Lite*$_{\mathcal{R}}^{\neq}$, and explore two alternative strategies.

- The first strategy is to weaken the ontology language, so as to eliminate all the constructs introducing incomplete information resulting from existentially quantified assertions. The outcome is a sublanguage of *DL-Lite*$_{\mathcal{R}}^{\neq}$, that we call *DL-Lite*$_{\text{RDFS}}^{\neg,\neq}$ because it extends *DL-Lite*$_{\text{RDFS}}$ (Cuenca Grau 2004; Rosati 2007; Cima, Lenzerini, and Poggi 2019) with both inequality axioms (in the ABox) and disjointness axioms (in the TBox).
- The second strategy is to keep *DL-Lite*$_{\mathcal{R}}^{\neq}$ as ontology language, but to weaken the query language by restricting the application of the inequality predicate to either individuals or distinguished variables (variables representing output values) only, as done in (Poggi 2016; Cima et al. 2017). The resulting query language is called "(U)CQs with bounded inequalities", and the corresponding class is denoted by (U)CQ$^{\neq,b}$. Observe that, although limited, the expressive power of (U)CQ$^{\neq,b}$ allow interesting queries to be expressed such as the one computing the triangles in a graph.

For the case of *DL-Lite*$_{\text{RDFS}}^{\neg,\neq}$, we show that answering UCQ$^{\neq}$s is decidable, and in particular coNP-complete in data complexity, and Π_2^p-complete in combined complexity (i.e., with respect to the size of the whole input, including the query). We also investigate if the number of inequalities in each disjunct plays a role in falling into intractability. We answer positively to this question, by showing that if the query has at most one inequality per disjunct, answering UCQ$^{\neq}$s is PTIME-complete in data complexity, and NP-complete in combined complexity (so, combined complexity class is the same as in the case without inequalities), while it is coNP-hard in data complexity if the query is conjunctive and has at most two inequalities. We also show that going from one to two inequalities causes the jump from NP-hardness to Π_2^p-hardness in combined complexity for (U)CQ$^{\neq}$s, and we conjecture that this holds already for CQ$^{\neq}$s.

For the case of (U)CQ$^{\neq,b}$, we show that answering CQ$^{\neq,b}$s over *DL-Lite*$_{\mathcal{R}}^{\neq}$ ontologies has the same complexity of the UCQ case, i.e., it is in AC^0 in data complexity and NP-complete in combined complexity. However, perhaps surprisingly, answering UCQ$^{\neq,b}$s over *DL-Lite*$_{\mathcal{R}}^{\neq}$ ontologies is Π_2^p-complete in combined complexity. Therefore, unless NP $=$ coNP, the presence of union makes the problem of answering queries with inequalities over *DL-Lite*$_{\mathcal{R}}^{\neq}$ ontologies significantly different from the case of UCQs.

We argue that the above results considerably improve our understanding of the implication that the presence of inequalities in queries has in the context of lightweight ontologies. In particular, to the best of our knowledge, our investigation on *DL-Lite*$_{\text{RDFS}}^{\neg,\neq}$ provides the first results on reasoning with inequalities when querying *DL-Lite*$_{\text{RDFS}}$ ontologies, and they also contribute a new result on containment of UCQs with inequalities in databases (Kolaitis, Martin, and Thakur 1998; Koutris et al. 2017): the problem is in NP (and therefore NP-complete) in the case of at most one inequality for each disjunct, and Π_2^p-complete, in the case of at most two inequalities for each disjunct. On the other hand, our results on (U)CQ$^{\neq,b}$s posed to *DL-Lite*$_{\mathcal{R}}^{\neq}$ ontologies show that this class is currently the only class of queries with inequalities that can be answered with AC^0 data complexity. Indeed, the only previously result known for this class was the PTIME algorithm described in (Poggi 2016). We also observe that all the results on CQ$^{\neq,b}$s presented in this paper easily extend to UCQ$^{\neq}$s posed to OWL 2 QL ontologies interpreted under the *Direct Semantics Entailment Regime* (Glimm 2011), that is the regime usually adopted for SPARQL queries. So, we are improving on a result reported in (Cima et al. 2017), where it is shown that answering UCQ$^{\neq}$s over OWL 2 QL ontologies can be polynomially reduced to the evaluation of a Datalog program, and therefore is in PTIME in data complexity, and in EXPTIME in combined complexity.

The paper is organized as follows. In Section 2 we provide some details on the notions used in the paper. In Section 3 we illustrate the notion of chase that we use for *DL-Lite*$_{\mathcal{R}}^{\neq}$, which is the basis for some of the technical results presented in this paper. In Section 4 and Section 5 we present our results on *DL-Lite*$_{\text{RDFS}}^{\neg,\neq}$, and (U)CQ$^{\neq,b}$s, respectively. Finally, in Section 6 we conclude the paper with a discussion on future work.

2 Preliminaries

We define the syntax and the semantics of *DL-Lite*$_{\mathcal{R}}^{\neq}$, and present the query languages considered in the paper.

***DL-Lite*$_{\mathcal{R}}$ and its variants.** Essentially, *DL-Lite*$_{\mathcal{R}}^{\neq}$ generalizes *DL-Lite*$_{\mathcal{R}}$ by removing the UNA, and adding axioms asserting inequalities of individuals[2].

Formally, starting with an alphabet including symbols for *individuals*, *atomic concepts*, and *atomic roles*, and the binary relation symbol $\neq$, a *DL-Lite*$_{\mathcal{R}}^{\neq}$ *ontology*, or simply an ontology, is a pair $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, such that $\mathcal{T}$, called a *TBox*, and $\mathcal{A}$, called an *ABox*, are sets of axioms, that have, respec-

[1] Since the UNA can be enforced through suitable axioms in *DL-Lite*$_{\mathcal{R}}^{\neq}$, such a language is a generalization of the classical *DL-Lite*$_{\mathcal{R}}$, and it turns out to be much more interesting when dealing with queries with inequalities.

[2] In principle, the axiom $a \neq b$ can be expressed by the axioms $A_1(a), A_2(b), A_1 \sqsubseteq \neg A_2$, for two new atomic concepts A_1, A_2. However, we prefer to keep the inequality predicate, both for avoiding changing the TBox, and for being coherent with OWL 2 QL.

tively, the following forms:

$\mathcal{T}$:	$B_1 \sqsubseteq B_2$	$R_1 \sqsubseteq R_2$	(inclusion)
	$B_1 \sqsubseteq \neg B_2$	$R_1 \sqsubseteq \neg R_2$	(disjointness)
$\mathcal{A}$:	$A(a)$	$P(a,b)$	(membership)
		$a \neq b$	(inequality)

where a, b denote *individuals*, A and P denote an atomic concept and an atomic role, respectively, B_1, B_2 are *basic concepts*, i.e., expressions of the form A, $\exists P$, or $\exists P^-$, and R_1 and R_2 are *basic roles*, i.e., expressions of the form P, or P^-. In this paper we also consider other variants of *DL-Lite*$_\mathcal{R}$, obtained by progressively restricting *DL-Lite*$^{\neq}_{\mathcal{R}}$. The first one is obtained by disallowing existential quantifiers to occur in the right-hand side of inclusion axioms, and it is named *DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$, since it corresponds to *DL-Lite*$_{\mathrm{RDFS}}$ extended with both inequality axioms and disjointness axioms. Other variants are *DL-Lite*$^{\neg}_{\mathrm{RDFS}}$ and *DL-Lite*$^{\neq}_{\mathrm{RDFS}}$, which are obtained by further disallowing inequality and disjointness axioms, respectively.

As for the semantics of *DL-Lite*$^{\neq}_{\mathcal{R}}$, an *interpretation* for $\mathcal{O}$ is a pair $\mathcal{I} = \langle \Delta^\mathcal{I}, \cdot^\mathcal{I} \rangle$, where the *interpretation domain* $\Delta^\mathcal{I}$ is a non-empty set of objects, and the *interpretation function* $\cdot^\mathcal{I}$ assigns to each individual a an object $a^\mathcal{I} \in \Delta^\mathcal{I}$, to each atomic concept A a set of objects $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$, to each atomic role a set of pairs of objects $P^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$, and to the special predicate "$\neq$" the set of all pairs of distinct objects, i.e., $\neq^\mathcal{I} = \{(o_1, o_2) \mid o_1, o_2 \in \Delta^\mathcal{I} \wedge o_1 \neq o_2\}$ (so, we often write $(o_1, o_2) \in \neq^\mathcal{I}$ as $o_1^\mathcal{I} \neq^\mathcal{I} o_2^\mathcal{I}$, or even $o_1^\mathcal{I} \neq o_2^\mathcal{I}$). The interpretation function extends to the other basic concepts and the other other basic roles as follows: (*i*) $(\exists P)^\mathcal{I} = \{o \mid \exists o'.(o, o') \in P^\mathcal{I}\}$, (*ii*) $(\exists P^-)^\mathcal{I} = \{o \mid \exists o'.(o', o) \in P^\mathcal{I}\}$, and (*iii*) $(P^-)^\mathcal{I} = \{(o, o') \mid (o', o) \in P^\mathcal{I}\}$.

An interpretation $\mathcal{I}$ *satisfies* an axiom $\alpha \sqsubseteq \beta$ (resp., $\alpha \sqsubseteq \neg\beta$) if $\alpha^\mathcal{I} \subseteq \beta^\mathcal{I}$ (resp., $\alpha^\mathcal{I} \cap \beta^\mathcal{I} = \emptyset$), an axiom $A(a)$ (resp., $P(a, b)$) if $a^\mathcal{I} \in A^\mathcal{I}$ (resp., $(a^\mathcal{I}, b^\mathcal{I}) \in P^\mathcal{I}$), and an axiom $a \neq b$ if $a^\mathcal{I} \neq b^\mathcal{I}$). It satisfies a set γ of axioms if it satisfies all axioms in γ. Finally, $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ is *satisfiable* if there exists a *model* of $\mathcal{O}$, i.e., an interpretation for $\mathcal{O}$ that satisfies both the TBox $\mathcal{T}$ and the ABox $\mathcal{A}$.

Queries over *DL-Lite*$^{\neq}_{\mathcal{R}}$. A *conjunctive query with inequalities (CQ$^{\neq}$)* over an ontology $\mathcal{O}$ is an expression of the form $q = \{\vec{x} \mid \phi(\vec{x}, \vec{y})\}$, where $\vec{x}$ and $\vec{y}$ are tuples of *variables*, called *distinguished* and *existential* variables of q, respectively, and $\phi(\vec{x}, \vec{y})$, called the *body* of q, is a finite conjunction of *DL-Lite*$^{\neq}_{\mathcal{R}}$ ABox assertions with variables that can appear in predicate arguments, i.e., atoms of the form $A(t_1)$, $P(t_1, t_2)$, or $t_1 \neq t_2$, where each t_j is either an individual of $\mathcal{O}$, or a variable in $\vec{x}$ or $\vec{y}$. We impose that every variable in $\vec{x}$ or $\vec{y}$ appears in some atom of $\phi(\vec{x}, \vec{y})$, as usual (Abiteboul, Hull, and Vianu 1995). If $\vec{x}$ is empty, then the query is called *boolean*. A CQ$^{\neq}$ q without atoms of the form $x_1 \neq x_2$ in its body is called a conjunctive query (CQ). An intermediate class of queries that lies between CQs and CQ$^{\neq}$s is the class of *conjunctive queries with bound inequalities (CQ$^{\neq,b}$)*. Specifically, a CQ$^{\neq,b}$ $q = \{\vec{x} \mid \phi(\vec{x}, \vec{y})\}$ is a CQ$^{\neq}$ whose inequalities involve only individuals or distinguished variables, i.e., for every atom $z_1 \neq z_2$ appearing in $\phi(\vec{x}, \vec{y})$, both z_1 and z_2 are not in $\vec{y}$. A UCQ (resp., UCQ$^{\neq,b}$, UCQ$^{\neq}$) is a union of a finite set of CQs (resp., CQ$^{\neq,b}$s, CQ$^{\neq}$s) with same arity.

The set $cert(q, \mathcal{O})$ of *certain answers* of a UCQ$^{\neq}$ q over $\mathcal{O}$ is the set of n-tuples $\vec{t} = \langle t_1, \ldots, t_n \rangle$ of individuals in $\mathcal{O}$ such that $\mathcal{O} \models q(\vec{t})$, i.e., $\langle t_1^\mathcal{I}, \ldots, t_n^\mathcal{I} \rangle \in q^\mathcal{I}$, also written $\mathcal{I} \models q(\langle t_1^\mathcal{I}, \ldots, t_n^\mathcal{I} \rangle)$, for every model $\mathcal{I}$ of $\mathcal{O}$, where $q^\mathcal{I}$ denotes the extension of q in $\mathcal{I}$. When q is a boolean query, we write $\mathcal{O} \models q$ if $q^\mathcal{I} = \{\langle\rangle\}$ (i.e., q is *true* in $\mathcal{I}$, also denoted by $\mathcal{I} \models q$) for every model $\mathcal{I}$ of $\mathcal{O}$. Observe that, when $\mathcal{I}$ is finite it can be seen as a relational database (Abiteboul, Hull, and Vianu 1995), and $q^\mathcal{I}$ simply denotes the evaluation of the UCQ q over $\mathcal{I}$.

When we talk about the problem of answering a query belonging to a class of queries $\mathcal{Q}$ over an $\mathcal{L}$-ontology, i.e., an ontology expressed in the DL $\mathcal{L}$, we implicitly refer to the following *decision problem*: Given a query $q \in \mathcal{Q}$, an $\mathcal{L}$-ontology $\mathcal{O}$, and an n-tuple $\vec{t}$ of individuals of $\mathcal{O}$, check whether $\vec{t} \in cert(q, \mathcal{O})$.

From results of (Calvanese et al. 2007), it is well known that checking whether a *DL-Lite*$_\mathcal{R}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ is satisfiable can be done by evaluating a suitable query over the ABox $\mathcal{A}$ seen as a relational database, in particular it can be done in AC^0 in data complexity and in PTIME in the size of the TBox $\mathcal{T}$. Furthermore, when a UCQ q is posed over a satisfiable *DL-Lite*$_\mathcal{R}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, it is possible to compute the set $cert(q, \mathcal{O})$ of certain answers by first reformulating q w.r.t. $\mathcal{T}$, and then by evaluating the reformulated query (which is again a UCQ) over the ABox $\mathcal{A}$ seen as a relational database. This yields the well-known result that answering UCQs over *DL-Lite*$_\mathcal{R}$ ontologies is in AC^0 in data complexity and NP-complete in combined complexity. Observe that, since *DL-Lite*$_\mathcal{R}$ is insensitive to the adoption of the UNA for UCQ answering (Artale et al. 2009), the same complexity results hold for the problem of answering UCQs over satisfiable *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontologies.

We end the section with the notion of homomorphism (Chandra and Merlin 1977), that will be used in the following. A *homomorphism* h from a CQ$^{\neq}$ q to a structure B is a function from variables and individuals of q to elements of B such that (*i*) $h(a) = a$ for each individual a occurring in q; (*ii*) for each atom of the form $A(t_1)$ (resp., $P(t_1, t_2)$), there is an atom $A(h(t_1))$ (resp., $P(h(t_1), h(t_2))$) occurring in B; and (*iii*) for each atom of the form $t_1 \neq t_2$, we have that $h(t_1) \neq h(t_2)$.

3 The chase

The conceptual tool that we use for addressing the problem of answering UCQ$^{\neq,b}$s over *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontologies is a modification of the chase used for *DL-Lite*$_\mathcal{R}$ (Calvanese et al. 2007). Specifically, given a *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, we build a (possibly infinite) structure, starting from $Ch^0(\mathcal{O}) = \mathcal{A}$, and repeatedly computing $Ch^{j+1}(\mathcal{O})$ from $Ch^j(\mathcal{O})$ by applying suitable rules, where each rule can be applied only if certain conditions hold. In doing so, we make use of a new infinite alphabet V of variables for introducing fresh unknown individuals, and we follow a deterministic strategy that is fair, i.e., it is such that if at some point a rule is applicable then it will be eventually applied. Finally, we set $Ch(\mathcal{O}) = \bigcup_{i\in\mathbb{N}} Ch^i(\mathcal{O})$. Note that we make use of the additional binary predicate symbol $ineq$, whose intended role is used to record all inequalities logically implied by $\mathcal{O}$.

The rules we use include all the ones illustrated in (Calvanese et al. 2007). For example, if $A_1 \sqsubseteq \exists P \in \mathcal{T}$, $A_1(e_1)$ is in $Ch^j(\mathcal{O})$, and no e_2 exists such that $P(e_1, e_2) \in Ch^j(\mathcal{O})$, then we set $Ch^{j+1}(\mathcal{O}) = Ch^j(\mathcal{O}) \cup \{P(e_1, s)\}$, where $s \in V$ does not appear in $Ch^j(\mathcal{O})$. There are, however, crucial additions related to the $ineq$ predicate. In what follows, when we say $R(e_1, e_2)$ holds in $Ch^j(\mathcal{O})$, where R is a basic role, we mean (i) $P(e_1, e_2) \in Ch^j(\mathcal{O})$, if $R = P$, or (ii) $P(e_2, e_1) \in Ch^j(\mathcal{O})$, if $R = P^-$. Also, when we say that $B(e_1)$ holds in $Ch^j(\mathcal{O})$, where B is a basic concept, we mean (i) $A(e_1) \in Ch^j(\mathcal{O})$ if $B = A$, and (ii) $R(e_1, e_2)$ holds in $Ch^j(\mathcal{O})$ for some e_2, if $B = \exists R$. The additional rules are as follows:

- If $e_1 \neq e_2$ is in $Ch^j(\mathcal{O})$, and $ineq(e_1, e_2)$ is not in $Ch^j(\mathcal{O})$, then $Ch^{j+1}(\mathcal{O}) = Ch^j(\mathcal{O}) \cup \{ineq(e_1, e_2)\}$;
- If $ineq(e_1, e_2)$ is in $Ch^j(\mathcal{O})$, and $ineq(e_2, e_1)$ is not in $Ch^j(\mathcal{O})$, then $Ch^{j+1}(\mathcal{O}) = Ch^j(\mathcal{O}) \cup \{ineq(e_2, e_1)\}$;
- if $B_1 \sqsubseteq \neg B_2 \in \mathcal{T}$, $B_1(e_1)$ and $B_2(e_2)$ hold in $Ch^j(\mathcal{O})$, and $ineq(e_1, e_2)$ is not in $Ch^j(\mathcal{O})$, then $Ch^{j+1}(\mathcal{O}) = Ch^j(\mathcal{O}) \cup \{ineq(e_1, e_2)\}$;
- if $R_1 \sqsubseteq \neg R_2 \in \mathcal{T}$, either $R_1(e_3, e_1)$ and $R_2(e_3, e_2)$, or $R_1(e_1, e_3)$ and $R_2(e_2, e_3)$ hold in $Ch^j(\mathcal{O})$, and $ineq(e_1, e_2)$ is not in $Ch^j(\mathcal{O})$, then $Ch^{j+1}(\mathcal{O}) = Ch^j(\mathcal{O}) \cup \{ineq(e_1, e_2)\}$.

From $Ch(\mathcal{O})$ it is immediate to define an interpretation $\mathcal{I}_{\mathcal{O}}$ for $\mathcal{O}$, extended in order to deal with predicate $ineq$:

- $\Delta^{\mathcal{I}_{\mathcal{O}}} = V_{\mathcal{O}} \cup V$, where $V_{\mathcal{O}}$ is the set of individuals occurring in $\mathcal{O}$;
- $e^{\mathcal{I}_{\mathcal{O}}} = e$ for every individual $e \in V_{\mathcal{O}}$;
- $A^{\mathcal{I}_{\mathcal{O}}} = \{e \mid A(e)$ occurs in $Ch(\mathcal{O})\}$ for every atomic concept A;
- $P^{\mathcal{I}_{\mathcal{O}}} = \{(e_1, e_2) \mid P(e_1, e_2)$ occurs in $Ch(\mathcal{O})\}$ for every atomic role P;
- $ineq^{\mathcal{I}_{\mathcal{O}}} = \{(e_1, e_2) \mid ineq(e_1, e_2)$ occurs in $Ch(\mathcal{O})\}$.

Note that, by definition, $\neq^{\mathcal{I}_{\mathcal{O}}}$ is the set of all pairs of distinct individuals in $V_{\mathcal{O}} \cup V$, i.e. $\neq^{\mathcal{I}_{\mathcal{O}}} = \{(e_1, e_2) \mid e_1, e_2 \in V_{\mathcal{O}} \cup V \wedge e_1 \neq e_2\}$.

Obviously, for a *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology, $Ch(\mathcal{O})$ can be infinite, due to the presence of existential quantifiers in the right-hand side of inclusion axioms, which, by introducing fresh unknown variables, can trigger an infinite number of rule applications. It is easy to see that, on the contrary, for a *DL-Lite*$^{\neg,\neq}_{\text{RDFS}}$ ontology $\mathcal{O}$, $Ch(\mathcal{O})$ is finite, and can be computed in polynomial time in the size of $\mathcal{O}$.

We next show that $\mathcal{I}_{\mathcal{O}}$ enjoys some crucial properties for *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontologies $\mathcal{O}$.

Proposition 1. *If $\mathcal{M} = \langle \Delta^{\mathcal{M}}, \cdot^{\mathcal{M}} \rangle$ is a model of a DL-Lite$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O}$, then there exists a function Ψ from $\Delta^{\mathcal{I}_{\mathcal{O}}}$ to $\Delta^{\mathcal{M}}$ such that:*

1. *for every $e \in \Delta^{\mathcal{I}_{\mathcal{O}}}$, if $e \in A^{\mathcal{I}_{\mathcal{O}}}$, then $\Psi(e) \in A^{\mathcal{M}}$;*
2. *for every pair $e_1, e_2 \in \Delta^{\mathcal{I}_{\mathcal{O}}}$, if $(e_1, e_2) \in P^{\mathcal{I}_{\mathcal{O}}}$, then $(\Psi(e_1), \Psi(e_2)) \in P^{\mathcal{M}}$;*
3. *for every pair $e_1, e_2 \in \Delta^{\mathcal{I}_{\mathcal{O}}}$, if $(e_1, e_2) \in ineq^{\mathcal{I}_{\mathcal{O}}}$, then $\Psi(e_1) \neq \Psi(e_2)$.*

The above proposition shows the importance of distinguishing between $\neq$ and $ineq$. Indeed, while by definition of $\mathcal{I}_{\mathcal{O}}$ two different individuals e_1, e_2 satisfy $e_1 \neq^{\mathcal{I}_{\mathcal{O}}} e_2$, it may happen that for some model $\mathcal{M}$ of $\mathcal{O}$, $e_1^{\mathcal{M}} = e_2^{\mathcal{M}}$, implying that no function Ψ exists from $\Delta^{\mathcal{I}_{\mathcal{O}}}$ to $\Delta^{\mathcal{M}}$ such that $\Psi(e_1) \neq \Psi(e_2)$. In other words, condition 3 in Proposition 1 does not hold if we replace $ineq$ with $\neq$.

Note that if $\mathcal{I}_{\mathcal{O}}$ satisfies all the axioms of $\mathcal{O}$, then it is a model of $\mathcal{O}$, and therefore $\mathcal{O}$ is satisfiable. Otherwise, it can be easily seen that $\mathcal{I}_{\mathcal{O}}$ violates at least one disjointness or one inequality axiom of $\mathcal{O}$. In particular, it can be proved that $\mathcal{I}_{\mathcal{O}}$ violates some inequality axiom if and only if $e \neq e$ occurs in $\mathcal{O}$ for some e in $V_{\mathcal{O}}$. As a result, we can devise a satisfiability checking algorithm for *DL-Lite*$^{\neq}_{\mathcal{R}}$ by slightly modifying the so-called violation query for *DL-Lite*$_{\mathcal{R}}$, and this shows that, similarly to the "canonical interpretation" of a *DL-Lite*$_{\mathcal{R}}$ ontology, $\mathcal{I}_{\mathcal{O}}$ is instrumental for checking the satisfiability of a *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O}$. In turn, this implies that checking the satisfiability of a *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ can be done in AC^0 in the size of $\mathcal{A}$ and in PTIME in the size of $\mathcal{T}$, exactly like in *DL-Lite*$_{\mathcal{R}}$.

A reasonable question to ask is whether $\mathcal{I}_{\mathcal{O}}$ is also the right tool for query answering. The next theorem provides a positive answer to this question for the class CQ$^{\neq,b}$. In what follows, $\delta(q)$ denotes the query obtained by replacing each inequality atom $t_1 \neq t_2$ in q with the atom $ineq(t_1, t_2)$.

Theorem 1. *Let $\vec{t}$ be a tuple of individuals of a satisfiable DL-Lite$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O}$, and let q be a CQ$^{\neq,b}$ over $\mathcal{O}$. We have that $\vec{t} \in cert(q, \mathcal{O})$ if and only if $\vec{t} \in \delta(q)^{\mathcal{I}_{\mathcal{O}}}$.*

The above theorem states that $\mathcal{I}_{\mathcal{O}}$ is instrumental also for answering CQ$^{\neq,b}$s over *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontologies. However, we will see in the next two sections that this theorem is no longer valid when we move from CQ$^{\neq,b}$s to either UCQ$^{\neq,b}$s, or CQ$^{\neq}$s.

From now on, we implicitly assume to deal with satisfiable ontologies. Moreover, unless otherwise stated and without loss of generality, we consider only boolean UCQ$^{\neq}$s. Indeed, given an n-ary UCQ$^{\neq}$ q, a *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, and an n-tuple $\vec{t}$ of individuals of $\mathcal{O}$, checking whether $\vec{t} \in cert(q, \mathcal{O})$ is equivalent to checking whether $\mathcal{O} \models q(\vec{t})$, where $q(\vec{t})$ denotes the boolean UCQ$^{\neq}$ obtained by replacing appropriately the distinguished variables of each disjunct of q with the individuals of $\vec{t}$.

4 UCQ$^{\neq}$s over *DL-Lite*$^{\neg,\neq}_{\text{RDFS}}$ ontologies

We study the problem of answering UCQ$^{\neq}$s over satisfiable *DL-Lite*$^{\neg,\neq}_{\text{RDFS}}$ ontologies.

Theorem 1 tells us that the certain answers to a $\mathrm{CQ}^{\neq,b}$ q over a *DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$ ontology $\mathcal{O}$ coincide with $\delta(q)^{\mathcal{I}_{\mathcal{O}}}$. However, the following example shows that the problem drastically changes as soon as we consider general $\mathrm{CQ}^{\neq}$s.

Example 1. *Consider the DL-Lite*$^{\neg}_{\mathrm{RDFS}}$ *ontology* $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$*, where* $\mathcal{T} = \{A_1 \sqsubseteq \neg A_2\}$ *and* $\mathcal{A} = \{A_1(a_1), A_2(a_2), P(b, c_1), P(b, c_2), P(c_1, a_1), P(c_2, a_2)\}$. *For the boolean* $CQ^{\neq}$ $q = \{() \mid P(x, y_1) \wedge P(x, y_2) \wedge y_1 \neq y_2\}$*, we have that* $\delta(q)^{\mathcal{I}_{\mathcal{O}}}$ *is false because* $\mathit{ineq}(c_1, c_2)$ *is not in* $Ch(\mathcal{O})$*. However* $\mathcal{O} \models q$*, because in each model* $\mathcal{M}$ *where* $c_1^{\mathcal{M}} = c_2^{\mathcal{M}}$ *the query is true with the bindings* $x, y_1, y_2 \rightarrow c_1, a_1, a_2$*, whereas in each model* $\mathcal{M}$ *where* $c_1^{\mathcal{M}} \neq c_2^{\mathcal{M}}$*,* q *is true with the bindings* $x, y_1, y_2 \rightarrow b, c_1, c_2$. □

The above example provides a hint on how to design an algorithm for our problem. Intuitively, given a boolean $\mathrm{UCQ}^{\neq}$ q, and a *DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, we check whether $\mathcal{O} \not\models q$ by searching for a database that can be obtained from $Ch(\mathcal{O})$ by equating some of the individuals, and that falsifies q. We thus derive the upper bounds for the problem of answering $\mathrm{UCQ}^{\neq}$s in *DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$.

Theorem 2. *Answering* $UCQ^{\neq}$*s over DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$ *ontologies is in coNP in data complexity and in* Π_2^p *in combined complexity.*

We now provide matching lower bounds for both data and combined complexity, showing that they hold already for the case of $\mathrm{CQ}^{\neq}$s. We start with data complexity.

Theorem 3. *Answering* $CQ^{\neq}$*s over DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$ *ontologies is coNP-hard in data complexity.*

The proof of the above theorem has two interesting implications. (i) coNP-hardness in data complexity holds even for $\mathrm{CQ}^{2,\neq}$s over both *DL-Lite*$^{\neg}_{\mathrm{RDFS}}$ and *DL-Lite*$^{\neq}_{\mathrm{RDFS}}$, where $\mathrm{CQ}^{k,\neq}$ denotes the class of $\mathrm{CQ}^{\neq}$ including at most k inequalities (and $\mathrm{UCQ}^{k,\neq}$ the class of unions of finite sets of $\mathrm{CQ}^{k,\neq}$s with same arity). (ii) Answering $\mathrm{UCQ}^{2,\neq}$s over *DL-Lite*$_{\mathrm{RDFS}}$ ontologies is coNP-hard too, and this corrects an erroneous statement in (Rosati 2007, Theorem 11), where it is claimed that answering $\mathrm{UCQ}^{\neq}$s over *DL-Lite*$_{\mathrm{RDFS}}$ ontologies is in LOGSPACE in data complexity regardless of whether the UNA is adopted or not. It turns out that this latter statement is true only under the UNA.

The following theorem provides the matching lower bound for combined complexity.

Theorem 4. *Answering* $CQ^{\neq}$*s over DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$ *ontologies is* Π_2^p*-hard in combined complexity.*

By looking at the proof of the theorem, one can see that Π_2^p-hardness holds already in the case of both *DL-Lite*$^{\neg}_{\mathrm{RDFS}}$, and *DL-Lite*$^{\neq}_{\mathrm{RDFS}}$. However, the reduction builds a $\mathrm{CQ}^{\neq}$ whose number of inequalities depends on the input of the reduction, and therefore is not fixed a priori. It is thus natural to ask which is the minimum number of inequalities in $\mathrm{CQ}^{\neq}$s that makes the problem Π_2^p-hard in combined complexity. Similarly to the case of the coNP-hardness result in data complexity, we conjecture that such number is 2. Even though we have not been able to prove this conjecture, we show next that Π_2^p-hardness holds for $\mathrm{UCQ}^{2,\neq}$s.

Algorithm $\mathit{CheckGood}(\mathcal{O}, q, F)$
Input: *DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, $\mathrm{UCQ}^{1,\neq}$ q,
  sequence of functions $F = \{f_1, \ldots, f_m\}$
Output: *true* or *false*
begin
  Compute $B := Ch(\mathcal{O})$
  for each $i = 1$ to $m - 1$:
    if f_i is a homomorphism from a disjunct of q_1 to B
      let $t_1 \neq t_2$ be any inequality in any of such disjuncts
      if $\mathit{ineq}(f_i(t_1), f_i(t_2)) \in B$ **return** *true*
      else replace each occurrence of $f_i(t_1)$ appearing
        in B, in q, and in $\{f_{i+1}, \ldots, f_m\}$ with $f_i(t_2)$
    else return *false*
  return f_m is a homomorphism from a disjunct of q_2 to B
end

Figure 1: The algorithm $\mathit{CheckGood}(\mathcal{O}, q, F)$

Theorem 5. *Answering* $UCQ^{2,\neq}$*s over DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$ *ontologies is* Π_2^p*-hard in combined complexity.*

Interestingly, the proof of the previous theorem shows that Π_2^p-hardness holds even if the query is the union of a $\mathrm{CQ}^{2,\neq}$ and a CQ without inequalities, and the ontology is expressed in *DL-Lite*$_{\mathrm{RDFS}}$. Observe that, in this language, $\mathrm{CQ}^{\neq}$s containing even a single inequality have an empty set of certain answers. Thus, we are observing a surprising jump from constant time to Π_2^p-hardness if we add union to such $\mathrm{CQ}^{\neq}$s.

To complete the picture of answering $\mathrm{UCQ}^{\neq}$s in *DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$, it remains to study the case of $\mathrm{UCQ}^{1,\neq}$s. In what follows, without loss of generality, we assume that each $\mathrm{UCQ}^{1,\neq}$ is written as $q = q_1 \cup q_2$, where q_2 is a UCQ with no inequalities and q_1 is a $\mathrm{UCQ}^{1,\neq}$ having exactly one inequality per disjunct.

In principle, for answering $\mathrm{UCQ}^{1,\neq}$s over *DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$ ontologies it is possible to use the algorithm provided in (Fagin et al. 2005, Theorem 5.12) in the context of data exchange. However, this would result in an exponential time algorithm with respect to the size of the query. On the contrary, by elaborating on the idea of (Fagin et al. 2005, Theorem 5.12), we have devised an algorithm that runs in PTIME in data complexity and in NP in combined complexity. We start with the following definition.

Definition 1. *Let* $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ *be a DL-Lite*$^{\neg,\neq}_{\mathrm{RDFS}}$ *ontology, and let* q *be a boolean* $UCQ^{1,\neq}$ *over* $\mathcal{O}$*. A sequence* $F = \{f_1, \ldots, f_m\}$ *of functions from variables and individuals of* q *to individuals of* $\mathcal{A}$ *is* good *w.r.t.* $\mathcal{O}$ *and* q *if the algorithm* $\mathit{CheckGood}(\mathcal{O}, q, F)$ *provided in Figure 1 returns* true.

Roughly speaking, starting from $B := Ch(\mathcal{O})$, in each step i from 1 to $m-1$ such that f_i is a homomorphism from a disjunct of q_1 to B, the algorithm $\mathit{CheckGood}(\mathcal{O}, q, F)$ replaces everywhere the individual $f_i(t_1)$ with the individual $f_i(t_2)$, to consider the models in which $f_i(t_1) = f_i(t_2)$ (since there is a homomorphism, q is true in the models where $f_i(t_1) \neq f_i(t_2)$). Afterwards, the algorithm sanctions that F is a good sequence if and only if either it is not possible to equate two individuals without contradicting an ineq

atom of B, or the resulting B and q_2 are such that $B \models q_2$. Using the above notion of good sequence, it is possible to derive the following characterization ($n_{\mathcal{A}}$ denotes the number of individuals occurring in the ABox $\mathcal{A}$).

Proposition 2. *Let $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a DL-Lite$^{\neg,\neq}_{\text{RDFS}}$ ontology, and let q be a boolean UCQ$^{1,\neq}$ over $\mathcal{O}$. We have that $\mathcal{O} \models q$ if and only if there exists a sequence $F = \{f_1, \ldots, f_m\}$ of $m \leq n_{\mathcal{A}}$ functions that is good w.r.t. $\mathcal{O}$ and q.*

We are now ready to establish our result on answering UCQ$^{1,\neq}$s in *DL-Lite*$^{\neg,\neq}_{\text{RDFS}}$.

Theorem 6. *Answering UCQ$^{1,\neq}$s over DL-Lite$^{\neg,\neq}_{\text{RDFS}}$ ontologies is* PTIME*-complete in data complexity and NP-complete in combined complexity.*

Proof. (Sketch) NP-hardness in combined complexity follows from NP-hardness of CQ evaluation over relational databases (Chandra and Merlin 1977). In the rest of this proof sketch, we discuss only the upper bounds. By Proposition 2 it is possible to decide whether $\mathcal{O} \models q$ as follows. We guess a sequence $F = \{f_1, \ldots, f_m\}$ (with $m \leq n_{\mathcal{A}}$) of functions from disjuncts of q_1 to $Ch(\mathcal{O})$ (note that this can be done in PTIME in the size of $\mathcal{A}$). Then, by exploiting the algorithm $CheckGood(\mathcal{O}, q, F)$, we check whether F is a good sequence w.r.t. $\mathcal{O}$ and q using: (*i*) a PTIME step in the size of $\mathcal{O}$ for computing $B = Ch(\mathcal{O})$; (*ii*) for each $i \in [1, m-1]$, a PTIME step for checking whether f_i is a homomorphism from a disjunct of q_1 to B, $ineq(f_i(t_1), f_i(t_2)) \in B$, and for replacing each occurrence of $f_i(t_1)$ with $f_i(t_2)$; finally, (*iii*) a PTIME step for checking whether f_m is a homomorphism from some disjunct of q_2 to B. $\square$

Again, from the proof of the theorem we can derive interesting observations. (i) PTIME-hardness in data complexity holds even for the problem of answering CQ$^{1,\neq}$s over both *DL-Lite*$^{\neq}_{\text{RDFS}}$ and *DL-Lite*$^{\neg}_{\text{RDFS}}$ ontologies. (ii) The result holds even for ontologies with disjointness on concepts only, and therefore it strengthens the PTIME-hardness result of (Gutiérrez-Basulto et al. 2015, Theorem 15) for the case of *DL-Lite*$_{core}$ ontologies. (iii) Answering UCQ$^{1,\neq}$s over *DL-Lite*$_{\text{RDFS}}$ ontologies is PTIME-hard in data complexity, too.

5 The case of CQ$^{\neq,b}$s and UCQ$^{\neq,b}$s

In this section, we consider answering queries with bounded inequalities over satisfiable *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontologies. In particular, we first deal with the case of CQ$^{\neq,b}$s, and then we address the case of UCQ$^{\neq,b}$s.

As for CQ$^{\neq,b}$s, we start by introducing some notations. Given an inequality atom $x_1 \neq x_2$ and a disjointness axiom γ, $\rho(x_1 \neq x_2, \gamma)$ denotes the following formula:

- $\rho(x_1 \neq x_2, A_1 \sqsubseteq \neg A_2) = A_1(x_1) \wedge A_2(x_2)$,
- $\rho(x_1 \neq x_2, A \sqsubseteq \neg \exists R) = \rho(x_1 \neq x_2, \exists R \sqsubseteq \neg A) = A(x_1) \wedge R(x_2, z)$, where z is a fresh variable,
- $\rho(x_1 \neq x_2, \exists R_1 \sqsubseteq \neg \exists R_2) = R_1(x_1, z) \wedge R_2(x_2, w)$, where z and w are fresh variables, and
- $\rho(x_1 \neq x_2, R_1 \sqsubseteq \neg R_2) = R_1(x_1, z) \wedge R_2(x_2, z) \vee R_1(z, x_1) \wedge R_2(z, x_2)$, where z is a fresh variable,

where an atom of the form $R(x, y)$ stands for either $P(x, y)$ if R denotes an atomic role P, or $P(y, x)$ if R denotes the inverse of an atomic role, i.e., $R = P^-$.

Given an inequality atom $x_1 \neq x_2$ and a *DL-Lite*$^{\neq}_{\mathcal{R}}$ TBox $\mathcal{T}$, we denote by $\sigma(x_1 \neq x_2, \mathcal{T})$ the disjunction

$$ineq(x_1, x_2) \vee ineq(x_2, x_1) \vee \bigvee_{i=1}^{m} (\rho(x_1 \neq x_2, \gamma_i) \vee \rho(x_2 \neq x_1, \gamma_i)),$$

where $\gamma_1, \ldots, \gamma_m$ are all the disjointness axioms of $\mathcal{T}$.

Finally, we denote by $\tau(q, \mathcal{T})$ the query obtained from q by substituting every inequality $x_1 \neq x_2$ by $\sigma(x_1 \neq x_2, \mathcal{T})$, and then turning the resulting query into an equivalent UCQ. We next illustrate the query $\tau(q, \mathcal{T})$ by an example.

Example 2. *Consider the DL-Lite$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ with $\mathcal{T} = \{P_1 \sqsubseteq P_2, A_1 \sqsubseteq \neg A_2\}$, and the CQ$^{\neq,b}$*

$$q = \{(x_1, x_2) \mid P_2(x_1, x_2) \wedge x_1 \neq c\}$$

over $\mathcal{O}$. It is easy to see that $\sigma(x_1 \neq c, \mathcal{T})$ is the formula $ineq(x_1, c) \vee ineq(c, x_1) \vee A_1(x_1) \wedge A_2(c) \vee A_2(x_1) \wedge A_1(c)$. Then, $\tau(q, \mathcal{T})$ is the UCQ$^{\neq,b}$ whose disjuncts are the following: $\{(x_1, x_2) \mid P_2(x_1, x_2) \wedge ineq(x, c)\}$, $\{(x_1, x_2) \mid P_2(x_1, x_2) \wedge ineq(c, x_1)\}$, $\{(x_1, x_2) \mid P_2(x_1, x_2) \wedge A_1(x_1) \wedge A_2(c)\}$, and $\{(x_1, x_2) \mid P_2(x_1, x_2) \wedge A_1(c) \wedge A_2(x)\}$. $\square$

For a *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, we denote by $\mathcal{O}^{ineq} = \langle \mathcal{T}, \mathcal{A}^{ineq} \rangle$ the *DL-Lite*$_{\mathcal{R}}$ ontology where $ineq$ is a new atomic role, and $\mathcal{A}^{ineq}$ is the *DL-Lite*$_{\mathcal{R}}$ ABox obtained from $\mathcal{A}$ by replacing each assertion $c_1 \neq c_2$ appearing in $\mathcal{A}$ with the assertion $ineq(c_1, c_2)$.

The next proposition, whose proof relies on an extension of (Calvanese et al. 2007, Lemma 39) and on Theorem 1, states that computing $cert(q, \mathcal{O})$ for a given *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, and a CQ$^{\neq,b}$ q over $\mathcal{O}$, can be reduced to computing the certain answers of the UCQ $\tau(q, \mathcal{T})$ over the *DL-Lite*$_{\mathcal{R}}$ ontology $\mathcal{O}^{ineq}$.

Proposition 3. *Let $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a DL-Lite$^{\neq}_{\mathcal{R}}$ ontology, and let q be a CQ$^{\neq,b}$ over $\mathcal{O}$. Then, we have that $cert(q, \mathcal{O}) = cert(\tau(q, \mathcal{T}), \mathcal{O}^{ineq})$.*

From the above proposition, we immediately derive that answering CQ$^{\neq,b}$s over *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontologies has the same data and combined complexity as answering UCQs over *DL-Lite*$_{\mathcal{R}}$ ontologies.

Theorem 7. *Answering CQ$^{\neq,b}$s over DL-Lite$^{\neq}_{\mathcal{R}}$ ontologies is in AC^0 in data complexity, and NP-complete in combined complexity.*

Looking at the proof of the two above statements, one realizes the importance of Theorem 1, stating that, similarly to *DL-Lite*$_{\mathcal{R}}$, *DL-Lite*$^{\neq}_{\mathcal{R}}$ admits a model $\mathcal{I}_{\mathcal{O}}$ that is representative of all the models of $\mathcal{O}$ w.r.t. answering CQ$^{\neq,b}$s. One might therefore think that, analogously to *DL-Lite*$_{\mathcal{R}}$, this property extend to UCQ$^{\neq,b}$s. The following example shows that, surprisingly, this is not the case.

Example 3. *Consider the DL-Lite$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{T} = \emptyset$ and $\mathcal{A} = \{P(a,b)\}$. For the UCQ$^{\neq,b}$ $Q = q_1 \cup q_2$, where $q_1 = \{() \mid P(a,a)\}$ and $q_2 = \{() \mid a \neq b\}$, it is easy to see that $\delta(Q)^{\mathcal{I}_{\mathcal{O}}}$ is false. However, one can verify that $\mathcal{O} \models Q$. Indeed, for any model $\mathcal{M}$ of $\mathcal{O}$, either $a^{\mathcal{M}} = b^{\mathcal{M}}$ and $\mathcal{M} \models q_1$, or $a^{\mathcal{M}} \neq b^{\mathcal{M}}$ and $\mathcal{M} \models q_2$.* □

What the example tells us is that answering UCQ$^{\neq,b}$s over a *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O}$ cannot be done simply using the interpretion $\mathcal{I}_{\mathcal{O}}$. Nevertheless, we are able to prove that the problem of answering UCQ$^{\neq,b}$s over *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontologies is still in AC^0 in data complexity, although, unless the polynomial hierarchy collapses to the first level, it does not have the same combined complexity as the UCQ and the CQ$^{\neq,b}$ cases.

We start by introducing the notions of e-satisfiability and e-entailment for an equivalence relation e[3]. In what follows, we write $c_1 \sim_e c_2$ to denote $(c_1, c_2) \in e$.

Definition 2. *Let $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a DL-Lite$^{\neq}_{\mathcal{R}}$ ontology, e be an equivalence relation on a set $\mathcal{C}$ of individuals of $\mathcal{O}$, and $\mathcal{I}$ be a model of $\mathcal{O}$. Then, we say that $\mathcal{I}$ is an* e-model *of $\mathcal{O}$ if, for any pair of individuals c_1, c_2 of $\mathcal{O}$, we have that $c_1^{\mathcal{I}} = c_2^{\mathcal{I}}$ if and only if $c_1 \sim_e c_2$. Furthermore, we say that $\mathcal{O}$ is* e-satisfiable *if it has an e-model.*

Definition 3. *Let $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a DL-Lite$^{\neq}_{\mathcal{R}}$ ontology, e be an equivalence relation on a set $\mathcal{C}$ of individuals of $\mathcal{O}$, and q be a boolean UCQ$^{\neq}$ over $\mathcal{O}$. Then, we say that $\mathcal{O}$ e-entails q, denoted by $\mathcal{O} \models_e q$, if $\mathcal{I} \models q$ for each e-model $\mathcal{I}$ of $\mathcal{O}$.*

Note that for a *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, if e is the equivalence relation on the set of all individuals appearing in $\mathcal{A}$ such that $e = \{(c,c) \mid c \text{ occurs in } \mathcal{A}\}$, then the notions of e-satisfiability and e-entailment coincide with the usual notion of satisfiability and entailment, respectively, when the UNA is enforced. In such cases: (*i*) the e-satisfiability can be checked in AC^0 in the size of $\mathcal{A}$ and in PTIME in the size of $\mathcal{T}$ (Calvanese et al. 2007), and (*ii*) it can be readily seen that checking whether $\mathcal{O} \models_e q$ for a UCQ$^{\neq,b}$ q over $\mathcal{O}$ can be done in AC^0 in the size of $\mathcal{A}$ and it is NP-complete in combined complexity. The next proposition shows that the complexity of both the above computational problems remains the same even when e is any arbitrary equivalence relation on a set $\mathcal{C}$ of individuals of $\mathcal{O}$.

Proposition 4. *Let $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a DL-Lite$^{\neq}_{\mathcal{R}}$ ontology, and let e be an equivalence relation on a set of individuals $\mathcal{C}$ of $\mathcal{O}$. We have that:*

- *checking whether $\mathcal{O}$ is e-satisfiable can be done in AC^0 in the size of $\mathcal{A}$ and in* PTIME *in the size of $\mathcal{T}$;*
- *if q is a boolean UCQ$^{\neq,b}$ over $\mathcal{O}$, then checking whether $\mathcal{O} \models_e q$ can be done in AC^0 in the size of $\mathcal{A}$ and it is NP-complete in combined complexity.*

Based on this result, we now characterize when $\mathcal{O} \not\models q$ for a boolean UCQ$^{\neq,b}$ and a *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology $\mathcal{O}$.

[3]An equivalence relation e on a set of individuals $\mathcal{C}$ is a binary relation over $\mathcal{C}$ that is reflexive, symmetric, and transitive.

Proposition 5. *Let $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a DL-Lite$^{\neq}_{\mathcal{R}}$ ontology, and let q be a boolean UCQ$^{\neq,b}$ over $\mathcal{O}$. We have that $\mathcal{O} \not\models q$ if and only if there exists an equivalence relation e on the set $\mathcal{C}_q$ of all individuals appearing in q such that $\mathcal{O} \not\models_e q$.*

Intuitively, to decide $\mathcal{O} \not\models q$, it is sufficient to guess an equivalence relation e between the individuals of q for which there exists an e-model $\mathcal{I}$ of $\mathcal{O}$ such that $\mathcal{I} \not\models q$. Observe that, by definition, such model exists if and only if $\mathcal{O} \not\models_e q$.

The following theorem characterizes the complexity of answering UCQ$^{\neq,b}$s over *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontologies.

Theorem 8. *Answering UCQ$^{\neq,b}$s over DL-Lite$^{\neq}_{\mathcal{R}}$ ontologies is in AC^0 in data complexity and Π_2^p-complete in combined complexity.*

Proof. (Sketch) As for the upper bounds, we now show how to decide whether $\mathcal{O} \not\models q$ in AC^0 in data complexity and in Σ_2^p in combined complexity. In particular, observe that by Proposition 5 it is sufficient to: (*i*) guess an equivalence relation e; (*ii*) and check whether $\mathcal{O} \not\models_e q$, where this last step, due to Proposition 4, can be done in AC^0 in the size of $\mathcal{A}$, and with an NP-oracle in the size of the input. □

The proof of the above theorem allows us to conclude that the same complexity results hold even for the problem of answering UCQ$^{\neq,b}$s over *DL-Lite*$_{\text{RDFS}}$ ontologies.

6 Conclusion

We have carried a thorough analysis of the problem of answering UCQs with inequalities posed to a *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontology. The results presented in this paper greatly contribute to clarify how inequalities impact on the problem of answering queries over *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontologies. In particular, we have presented the first results on dealing with inequalities in queries posed to *DL-Lite*$^{\neg,\neq}_{\text{RDFS}}$ ontologies, and we have deeply investigated a specific class of queries, namely UCQ$^{\neq,b}$s, for which query answering over *DL-Lite*$^{\neq}_{\mathcal{R}}$ ontologies is still in AC^0 in data complexity. We have also mentioned the connection between the problems studied here and two other problems, namely containment of conjunctive queries with inequalities in databases, and answering UCQ$^{\neq}$s over OWL 2 QL ontologies under the direct semantics, although we could not elaborate on these aspects for the lack of space.

There are several issues to consider for continuing the work presented in this paper, the most obvious being trying to decide which is the minimum number of inequalities that makes query answering over *DL-Lite*$^{\neg,\neq}_{\text{RDFS}}$ Π_2^p-hard in combined complexity. Another interesting future work is to look for extensions of both *DL-Lite*$^{\neq}_{\mathcal{R}}$, and UCQ$^{\neq,b}$s for which query answering is still decidable/tractable. Finally, we observe that it is still open whether answering CQ$^{\neq}$s over *DL-Lite*$_{core}$ ontologies is decidable.

Acknowledgments

This work has been supported by Sapienza under the project "PRE-O-PRE" and by MIUR, under the SIR project "MODEUS" - grant n. RBSI14TQHQ, and under the PRIN 2017 project "HOPE".

References

Abiteboul, S.; Hull, R.; and Vianu, V. 1995. *Foundations of Databases*. Addison Wesley Publ. Co.

Artale, A.; Calvanese, D.; Kontchakov, R.; and Zakharyaschev, M. 2009. The *DL-Lite* family and relations. *Journal of Artificial Intelligence Research* 36:1–69.

Baader, F.; Calvanese, D.; McGuinness, D.; Nardi, D.; and Patel-Schneider, P. F., eds. 2003. *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press.

Baader, F.; Horrocks, I.; Lutz, C.; and Sattler, U. 2017. *An Introduction to Description Logic*. Cambridge University Press.

Bienvenu, M. 2016. Ontology-mediated query answering: Harnessing knowledge to get more from data. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI 2016)*, 4058–4061.

Calvanese, D.; De Giacomo, G.; Lenzerini, M.; Rosati, R.; and Vetere, G. 2004. *DL-Lite*: Practical reasoning for rich DLs. In *Proceedings of the Seventeenth International Workshop on Description Logic (DL 2004)*, volume 104 of *CEUR Electronic Workshop Proceedings,* http://ceur-ws.org/.

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

Chandra, A. K., and Merlin, P. M. 1977. Optimal implementation of conjunctive queries in relational data bases. In *Proceedings of the Ninth ACM Symposium on Theory of Computing (STOC 1977)*, 77–90.

Cima, G.; De Giacomo, G.; Lenzerini, M.; and Poggi, A. 2017. On the SPARQL metamodeling semantics entailment regime for OWL 2 QL ontologies. In *Proceedings of the Seventh International Conference on Web Intelligence, Mining and Semantics (WIMS 2017)*, 10:1–10:6.

Cima, G.; Lenzerini, M.; and Poggi, A. 2019. Semantic characterization of data services through ontologies. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI 2019)*, 1647–1653.

Cuenca Grau, B. 2004. A possible simplification of the semantic web architecture. In *Proceedings of the Thirteenth International World Wide Web Conference (WWW 2004)*, 704–713.

Fagin, R.; Kolaitis, P. G.; Miller, R. J.; and Popa, L. 2005. Data exchange: Semantics and query answering. *Theoretical Computer Science* 336(1):89–124.

Glimm, B. 2011. Using SPARQL with RDFS and OWL entailment. In *Reasoning Web. Semantic Technologies for the Web of Data – Seventh International Summer School Tutorial Lectures (RW 2011)*. 137–201.

Gutiérrez-Basulto, V.; Ibáñez-García, Y. A.; Kontchakov, R.; and Kostylev, E. V. 2015. Queries with negation and inequalities over lightweight ontologies. *Journal of Web Semantics* 35:184–202.

Gutiérrez-Basulto, V.; Ibáñez-García, Y. A.; and Kontchakov, R. 2012. An update on query answering with restricted forms of negation. In *Web Reasoning and Rule Systems - Sixth International Conference (RR 2012)*, 75–89.

Kolaitis, P. G.; Martin, D. L.; and Thakur, M. N. 1998. On the complexity of the containment problem for conjunctive queries with built-in predicates. In *Proceedings of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS 1998)*, 197–204.

Koutris, P.; Milo, T.; Roy, S.; and Suciu, D. 2017. Answering conjunctive queries with inequalities. *Theory of Computing Systems* 61(1):2–30.

Motik, B.; Cuenca Grau, B.; Horrocks, I.; Wu, Z.; Fokoue, A.; and Lutz, C. 2012. OWL 2 Web Ontology Language profiles (second edition). W3C Recommendation, World Wide Web Consortium. Available at http://www.w3.org/TR/owl2-profiles/.

Ortiz, M. 2018. Improving data management using domain knowledge. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI 2018)*, 5709–5713.

Poggi, A.; Lembo, D.; Calvanese, D.; De Giacomo, G.; Lenzerini, M.; and Rosati, R. 2008. Linking data to ontologies. *Journal on Data Semantics* X:133–173.

Poggi, A. 2016. On the SPARQL direct semantics entailment regime for OWL 2 QL. In *Proceedings of the Twenty-Ninth International Workshop on Description Logics (DL 2014)*, CEUR Electronic Workshop Proceedings, http://ceur-ws.org/.

Rosati, R. 2007. The limits of querying ontologies. In *Proceedings of the Eleventh International Conference on Database Theory (ICDT 2007)*, volume 4353 of *Lecture Notes in Computer Science*, 164–178.

Vardi, M. Y. 1982. The complexity of relational query languages. In *Proceedings of the Fourteenth ACM SIGACT Symposium on Theory of Computing (STOC 1982)*, 137–146.

Xiao, G.; Calvanese, D.; an Domenico Lembo, R. K.; Poggi, A.; Rosati, R.; and Zakharyaschev, M. 2018. Ontology-based data access: A survey. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI 2018)*, 5511–5519.