Revisiting Controlled Query Evaluation in Description Logics

Domenico Lembo[1] , **Riccardo Rosati**[1] and **Domenico Fabio Savo**[2]
[1]Sapienza Università di Roma
[2]Università degli Studi di Bergamo
{lembo, rosati}@diag.uniroma1.it, domenicofabio.savo@unibg.it

Abstract

Controlled Query Evaluation (CQE) is a confidentiality-preserving framework in which private information is protected through a *policy*, and a (optimal) *censor* guarantees that answers to queries are maximized without violating the policy. CQE has been recently studied in the context of ontologies, where the focus has been mainly on the problem of the existence of an optimal censor. In this paper we instead consider *query answering* over all possible optimal censors. We study data complexity of this problem for ontologies specified in the Description Logics *DL-Lite*$_R$ and $\mathcal{EL}_\perp$ and for variants of the censor language, which is the language used by the censor to enforce the policy. In our investigation we also analyze the relationship between CQE and the problem of Consistent Query Answering (CQA). Some of the complexity results we provide are indeed obtained through mutual reduction between CQE and CQA.

1 Introduction

In Controlled Query Evaluation (CQE), a policy, i.e., a set of logical assertions, regulates the access to a database or knowledge base by specifying the information that must be kept secret, and a *censor* alters answers to queries so that confidential data cannot be inferred by the users on the basis of the queries they ask. The notion of censor traces back to [Sicherman *et al.*, 1983], and since then it has been investigated for propositional closed databases [Biskup and Bonatti, 2004a; Biskup and Bonatti, 2004b], incomplete databases [Biskup and Weibert, 2008], and, more recently, Description Logic (DL) ontologies [Bonatti and Sauro, 2013; Cuenca Grau *et al.*, 2013; Cuenca Grau *et al.*, 2015]. In this latter context, *optimal* censors are defined as those censors that modify query answers in a "minimal" way. Intuitively, such censors hide data to preserve confidentiality according to the policy, without restricting unnecessarily the ability of the system to return answers to users' queries.

In general, several optimal censors may exist for an instance of the CQE problem, since several incomparable ways of altering the answers may exist. For example, if the policy does not allow both facts hasName(01,John) and hasSalary(01,2000) to be divulged, an optimal censor discloses only the former, and another one only the latter (of course, censors hiding both facts are not optimal).

Previous work on CQE in DLs has mainly focused on the tasks of establishing the *existence* of an optimal censor and characterizing the complexity of computing it. In practice, however, considering only one such censor means making an arbitrary selection among several optimal censors. To avoid such a discretionary choice, in this paper we adopt a different approach and study *query answering over all optimal censors*. Intuitively, given a query q, the answers to q are the answers in the intersection of the answers computed by the optimal censors. A similar idea has been also previously discussed in [Cuenca Grau *et al.*, 2013].

Our approach has similarities with the work on Consistent Query Answering (CQA), a declarative framework for inconsistency management based on the notion of repair [Bertossi, 2011; Bienvenu and Bourgaux, 2016]. Roughly speaking, in DL, a repair of a possibly inconsistent ontology $\mathcal{O}$ is any ABox for $\mathcal{O}$ (i.e., the extensional component of the ontology) that is consistent with the TBox (i.e., the intensional component of $\mathcal{O}$), and that differs "minimally" from the original ABox. Then, computing query answers in CQA amounts to reasoning over all repairs and the TBox. The connection between CQE and CQA in DL is based on the intuition that the assertions in the policy in CQE seem to act as the class of assertions in $\mathcal{T}$ that may be violated by the data of the ABox.

Some connections between CQA and a declarative approach for privacy preservation had already been discussed in [Bertossi and Li, 2013]. The framework in that paper is similar to ours, with so-called secrecy views playing essentially the role of the policy. However, the setting considered there is relational and without intensional knowledge (TBox), and secrecy views are enforced through suitable virtual modifications of database values with SQL NULLs, so that this approach is incomparable with ours. Nonetheless, in this paper we elaborate on the intuition of [Bertossi and Li, 2013] and investigate in depth the relationship between our CQE framework and CQA in DLs. We provide some general conditions ensuring that the two problems are mutually reducible, and we show cases of practical interest for which such conditions are satisfied and cases for which they are not, which allows us to highlight similarities and differences between the two frameworks. We notice also that the connection between

CQA and CQE we explain in this paper was already discussed in a preliminary form in our extended abstract [Lembo *et al.*, 2018], in the context of a general formal framework aiming at capturing CQA, CQE, and update of DL ontologies.

The ultimate goal of this paper is to investigate data complexity of answering conjunctive queries (CQs) in CQE. In our analysis we consider ontologies specified in *DL-Lite*$_R$ [Calvanese *et al.*, 2007] and $\mathcal{EL}_\bot$ [Baader *et al.*, 2005], two popular lightweight ontology languages, which are at the basis of two OWL tractable profiles[1]. We also consider some variants of the *censor language* $\mathcal{L}_\mathcal{C}$, which is the language used by the censor to enforce the policy. Roughly speaking, $\mathcal{L}_\mathcal{C}$ is the language in which the censor expresses the sentences implied by the ontology that can be disclosed to the users without violating the policy. We provide data complexity results for the cases when: (i) $\mathcal{L}_\mathcal{C}$ is the ABox of the ontology (i.e., the censor can enforce the policy only by selecting facts in the ABox); (ii) $\mathcal{L}_\mathcal{C}$ coincides with the set of facts expressed over the signature of the ontology; and (iii) $\mathcal{L}_\mathcal{C}$ is the language of CQs expressed over the signature of the ontology (for $\mathcal{EL}_\bot$ we in fact limit to the language of CQs whose maximum length is k). Some of the complexity results follow from the correspondence between CQA and CQE; we devise novel techniques for the cases in which the CQE problem does not have a CQA counterpart.

Confidentiality issues in DLs have been previously studied in [Calvanese *et al.*, 2012], under authorization views, an approach to some extent complementary to ours. Provable data privacy on views has been considered in [Stouppa and Studer, 2009], for concept retrieval and subsumption queries over $\mathcal{ALC}$ ontologies. Properties of censors for Boolean $\mathcal{ALC}$ ontologies have been investigated in [Studer and Werner, 2014], for concept subsumption only. Secrecy preserving reasoning in the presence of several agents has been instead studied in [Tao *et al.*, 2014], for propositional Horn logics and the DL $\mathcal{AL}$. Privacy-preserving query answering as a reasoning problem has been considered in [Cuenca Grau and Horrocks, 2008], whereas instance checking for $\mathcal{EL}$ has been studied in [Tao *et al.*, 2010], in both cases under frameworks different from CQE. Then, the problem of establishing whether an ontology-based data integration system discloses a source query has been recently studied in [Benedikt *et al.*, 2018].

In the rest of the paper, after some preliminaries (Sec. 2), we introduce our CQE framework (Sec. 3), and study the relationship between CQE and CQA (Sec. 4). We then establish complexity of query answering (both instance checking and entailment of CQs) for restricted censor languages (Sec. 5), and for the full censor language considered in this paper, namely CQs (Sec. 6). We conclude the paper in Sec. 7.

2 Preliminaries

We consider a signature Σ of predicates and constants, and a countably infinite alphabet of variables $\mathcal{V}$. To simplify the presentation, we consider only languages containing FO sentences, i.e., formulas without free variables (our results apply to open formulas as well, modulo standard encoding of open formulas into closed ones). We use $\mathbf{FO}$ to indicate the language of all function-free FO sentences over Σ and $\mathcal{V}$. Every language considered in this paper is a subset of $\mathbf{FO}$.

Given a set $\mathcal{K} \subseteq \mathbf{FO}$, $Mod(\mathcal{K})$ indicates the set of models of $\mathcal{K}$, i.e., the FO interpretations $\mathcal{I}$ such that $\phi^\mathcal{I}$ (i.e., the interpretation of ϕ in $\mathcal{I}$) evaluates to true, for each $\phi \in \mathcal{K}$. A set $\mathcal{K}$ is consistent if it has at least one model, i.e., if $Mod(\mathcal{K}) \neq \emptyset$, inconsistent otherwise, and it entails an FO sentence ϕ, denoted $\mathcal{K} \models \phi$, if $\phi^\mathcal{I}$ is true in every $\mathcal{I} \in Mod(\mathcal{K})$.

A Boolean conjunctive query (BCQ) q is an FO sentence of the form $\exists \vec{x}.conj(\vec{x})$, where $conj(\vec{x}) = \alpha_1(\vec{x}) \wedge \ldots \wedge \alpha_n(\vec{x})$, $\vec{x}$ is a sequence of variables, and each $\alpha_i(\vec{x})$ is an atom (possibly with constants) with predicate α_i and variables in $\vec{x}$. The *length* of q is the number of its atoms, denoted by $length(q)$.

In the following, $\mathbf{CQ}$ denotes the language of BCQs over Σ and $\mathcal{V}$, $\mathbf{CQ}_k$ the language of BCQs from $\mathbf{CQ}$ whose maximum length is k, and $\mathbf{GA}$ the language of single atom queries with no variables, i.e., ground atoms or facts. Verifying whether $\mathcal{K} \models \alpha$ for $\mathcal{K} \subseteq \mathbf{FO}$ and $\alpha \in \mathbf{GA}$ is also called *instance checking*.

Description Logics (DLs) are decidable FO languages using only unary and binary predicates, called concepts and roles (for more details on DLs and their relationship with FO we refer the reader to [Baader *et al.*, 2007]). A DL ontology $\mathcal{O}$ is a set $\mathcal{T} \cup \mathcal{A}$, where $\mathcal{T}$ is the *TBox* and $\mathcal{A}$ is the *ABox*, providing intensional and extensional knowledge, respectively. We assume that an ABox is always a set of ground atoms.

In this paper, we consider DL ontologies expressed in *DL-Lite*$_R$ [Calvanese *et al.*, 2007] and $\mathcal{EL}_\bot$, which extends $\mathcal{EL}$ [Baader *et al.*, 2005] with the empty concept $\bot$.

A *DL-Lite*$_R$ TBox is a finite set of assertions of the form $B_1 \sqsubseteq B_2$, $B_1 \sqsubseteq \neg B_2$, $R_1 \sqsubseteq R_2$, $R_1 \sqsubseteq \neg R_2$, where: each R_i, with $i \in \{1, 2\}$, is an *atomic role* $Q \in \Sigma$, or its inverse Q^-; each B_i, with $i \in \{1, 2\}$, is an *atomic concept* $A \in \Sigma$, or a concept of the form $\exists Q$ or $\exists Q^-$, i.e., unqualified existential restrictions, which denote the set of objects occurring as first or second argument of Q, respectively.

An $\mathcal{EL}_\bot$ TBox is a finite set of assertions of the form $C_1 \sqsubseteq C_2$, where each C_i, with $i \in \{1, 2\}$, is: an atomic concept A; a concept of the form $\exists Q.C$, i.e., qualified existential restriction, which denotes the set of objects that the atomic role Q relates to some instance of C; a concept $C \sqcap C'$, i.e., a conjunction of two concepts; or $\bot$, i.e., the *empty concept*.

Besides *DL-Lite*$_R$ and $\mathcal{EL}_\bot$ assertions, we also consider *denial assertions* (or simply *denials*) over concepts and roles, i.e., sentences of the form $\forall \vec{x}.conj(\vec{x}) \rightarrow \bot$, where $conj(\vec{x})$ is such that $\exists \vec{x}.conj(\vec{x})$ is a BCQ whose atoms use only unary and binary predicates. The *length* of the denial is the length of such query. A denial is satisfied by an ontology $\mathcal{O}$ if $\mathcal{O} \not\models \exists \vec{x}.conj(\vec{x})$. We will use denials to specify the policy in CQE. We will also refer to the DL *DL-Lite*$_{R,den}$, which is an extension of *DL-Lite*$_R$ with denials [Lembo *et al.*, 2015].

Given an ontology $\mathcal{O}$ and a language $\mathcal{L} \subseteq \mathbf{FO}$, we denote by $\mathcal{L}(\mathcal{O})$ the subset of formulas of $\mathcal{L}$ over the predicates and constants occurring in $\mathcal{O}$ and the variables in $\mathcal{V}$.

All the complexity results given in this paper are concerned with *data complexity*, that in our framework is the complexity computed only with respect to the size of the ABox.

[1] https://www.w3.org/TR/owl2-profiles/

3 CQE Framework

Our framework for CQE is adapted from the one presented in [Cuenca Grau *et al.*, 2015].

An $\mathcal{L}$ *CQE instance* $\mathcal{E}$ is a triple $\langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$, where $\mathcal{T}$ is a TBox in the DL $\mathcal{L}$, $\mathcal{A}$ is an ABox such that $\mathcal{T} \cup \mathcal{A}$ is consistent, and $\mathcal{P}$ is the *policy*, i.e., a set of denial assertions over the signature of $\mathcal{T} \cup \mathcal{A}$, such that $\mathcal{T} \cup \mathcal{P}$ is consistent. Intuitively, $\mathcal{T}$ is the schema a user interacts with to pose her queries; $\mathcal{A}$ is the dataset underlying the schema; $\mathcal{P}$ specifies the knowledge that cannot be disclosed for confidentiality reasons, in the sense that the user will never get, through query answers, sufficient knowledge to violate the denials in $\mathcal{P}$[2].

We then define a censor for a CQE instance.

Definition 1 Given a CQE instance $\mathcal{E} = \langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$ and a language $\mathcal{L}_c \subseteq \mathbf{FO}(\mathcal{T} \cup \mathcal{A})$, a *censor* for $\mathcal{E}$ in $\mathcal{L}_c$ is a function $\mathsf{cens}_{\mathcal{L}_c}$ that returns a set $\mathsf{cens}_{\mathcal{L}_c}(\mathcal{E}) \subseteq \mathcal{L}_c$ (called the *theory of the censor*) such that: (i) $\mathcal{T} \cup \mathcal{A} \models \phi$, for each $\phi \in \mathsf{cens}_{\mathcal{L}_c}(\mathcal{E})$, and (ii) $\mathcal{T} \cup \mathcal{P} \cup \mathsf{cens}_{\mathcal{L}_c}(\mathcal{E})$ is consistent.

Intuitively, the censor establishes which are the sentences in $\mathcal{L}_c$ (called the censor language) implied by $\mathcal{T} \cup \mathcal{A}$ that can be divulged to the user while preserving the policy. A censor $\mathsf{cens}_{\mathcal{L}_c}$ for $\mathcal{E}$ in $\mathcal{L}_c$ is *optimal* if there does not exist a censor $\mathsf{cens'}_{\mathcal{L}_c}$ for $\mathcal{E}$ in $\mathcal{L}_c$ such that $\mathsf{cens}_{\mathcal{L}_c}(\mathcal{E}) \subset \mathsf{cens'}_{\mathcal{L}_c}(\mathcal{E})$. The set of theories of all the optimal censors in $\mathcal{L}_c$ for a CQE instance $\mathcal{E}$ is denoted with $\mathit{Th}^{oc}_{\mathcal{L}_c}(\mathcal{E})$.

Hereinafter, to simplify the notation, we will sometimes omit to specify that a censor language is limited to the signature of $\mathcal{T} \cup \mathcal{A}$ (e.g., we will use $\mathbf{CQ}$ instead of $\mathbf{CQ}(\mathcal{T} \cup \mathcal{A})$).

Example 1 To regulate access to information about customers and the medicines they buy, a CQE instance $\mathcal{E} = (\mathcal{T}, \mathcal{A}, \mathcal{P})$ is used in a pharmacy, where $\mathcal{T}$ is an empty TBox, i.e., without assertions, $\mathcal{A} = \{\mathsf{Buy}(c_1, m_A), \mathsf{Buy}(c_1, m_B), \mathsf{Buy}(c_2, m_A)\}$, and $\mathcal{P} = \{\forall x.\mathsf{Buy}(x, m_A) \land \mathsf{Buy}(x, m_B) \rightarrow \bot\}$. The policy specifies as confidential the fact that a customer buys both medicine A and medicine B (this may reveal an embarrassing disease). The optimal censors for $\mathcal{E}$ in $\mathbf{CQ}$ are only $\mathsf{cens}^1_{\mathbf{CQ}}$ and $\mathsf{cens}^2_{\mathbf{CQ}}$, where $\mathsf{cens}^1_{\mathbf{CQ}}(\mathcal{E})$ contains the queries $\exists x.\mathsf{Buy}(x, m_B)$, $\mathsf{Buy}(c_1, m_A)$, $\mathsf{Buy}(c_2, m_A)$, and all the queries in $\mathbf{CQ}$ inferred by them, and $\mathsf{cens}^2_{\mathbf{CQ}}(\mathcal{E})$ contains the queries $\mathsf{Buy}(c_1, m_B)$ and $\mathsf{Buy}(c_2, m_A)$, and all the queries in $\mathbf{CQ}$ inferred by them.

If we instead restrict the censor language to $\mathcal{A}$ (i.e., censor theories can only contain facts of the ABox), we still have only two optimal censors, i.e., $\mathsf{cens}^1_{\mathcal{A}}(\mathcal{E}) = \{\mathsf{Buy}(c_1, m_A), \mathsf{Buy}(c_2, m_A)\}$ and $\mathsf{cens}^2_{\mathcal{A}}(\mathcal{E}) = \{\mathsf{Buy}(c_1, m_B), \mathsf{Buy}(c_2, m_A)\}$.

Below we provide the definition of entailment in CQE. More precisely, we give a definition for each type of censor language considered in this paper.

Definition 2 Given a CQE instance $\mathcal{E} = \langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$ and an FO sentence ϕ, we define the following three decision problems:

(CQ-Cens-Entailment): decide whether $\mathsf{Th} \models \phi$ for every $\mathsf{Th} \in \mathit{Th}^{oc}_{\mathbf{CQ}}(\mathcal{E})$. If this is the case, we write $\mathcal{E} \models^{cqe}_{CQ} \phi$.

(GA-Cens-Entailment): decide whether $\mathsf{Th} \models \phi$ for every $\mathsf{Th} \in \mathit{Th}^{oc}_{\mathbf{GA}}(\mathcal{E})$. If this is the case, we write $\mathcal{E} \models^{cqe}_{GA} \phi$.

(ABox-Cens-Entailment): decide whether $\mathsf{Th} \models \phi$ for every $\mathsf{Th} \in \mathit{Th}^{oc}_{\mathcal{A}}(\mathcal{E})$. If this is the case, we write $\mathcal{E} \models^{cqe}_{ABox} \phi$.

When ϕ is a ground atom, the above entailment problems are called CQ-Cens-Instance-Checking, GA-Cens-Instance-Checking and ABox-Cens-Instance-Checking, respectively.

Example 2 Let $\mathcal{E}$ be the CQE instance of Example 1. We have, for instance, that $\mathcal{E} \models^{cqe}_{ABox} \exists x.\mathsf{Buy}(c_1, x)$, but $\mathcal{E} \not\models^{cqe}_{ABox} \exists x.\mathsf{Buy}(x, m_B)$. Furthermore, we have that $\mathcal{E} \models^{cqe}_{CQ} \exists x.\mathsf{Buy}(c_1, x)$ and $\mathcal{E} \models^{cqe}_{CQ} \exists x.\mathsf{Buy}(x, m_B)$, but $\mathcal{E} \not\models^{cqe}_{CQ} \mathsf{Buy}(c_1, m_B)$.

4 Relationship between CQE and CQA

In this section we discuss the relationship between the CQE framework we have just defined and CQA. To this aim, we first provide a general definition for CQA.

An $\mathcal{L}$ *CQA instance* $\mathcal{J}$ is a pair $\langle \mathcal{T}, \mathcal{A} \rangle$ where $\mathcal{T}$ is a consistent TBox in the DL $\mathcal{L}$, and $\mathcal{A}$ is a DL ABox, where $\mathcal{T} \cup \mathcal{A}$ is a possibly inconsistent ontology. We then give the notion of the *consistent entailment set in a language* $\mathcal{L}$ for an FO theory Ψ and an ABox $\mathcal{A}$, denoted by $\mathit{CES}_{\mathcal{L}}(\Psi, \mathcal{A})$, which is the set $\{\phi \mid \phi \in \mathcal{L}$ and there exists a $\mathcal{A}' \subseteq \mathcal{A}$ such that $\Psi \cup \mathcal{A}'$ is consistent and $\Psi \cup \mathcal{A}' \models \phi\}$.

A repair for a CQA instance is defined as follows.

Definition 3 A *repair* for a CQA instance $\mathcal{J} = \langle \mathcal{T}, \mathcal{A} \rangle$ in a language $\mathcal{L}_r \subseteq \mathbf{FO}(\mathcal{T} \cup \mathcal{A})$ (called repair language) is a subset $\mathcal{R}$ of $\mathcal{L}_r$ such that: (i) $\mathcal{R} \subseteq \mathit{CES}_{\mathcal{L}_r}(\mathcal{T}, \mathcal{A})$; (ii) $\mathcal{T} \cup \mathcal{R}$ is consistent; (iii) there does not exist any $\mathcal{R}'$ such that $\mathcal{R} \subset \mathcal{R}' \subseteq \mathit{CES}_{\mathcal{L}_r}(\mathcal{T}, \mathcal{A})$ and $\mathcal{T} \cup \mathcal{R}'$ is consistent. We denote by $\mathsf{RepSet}_{\mathcal{L}_r}(\mathcal{J})$ the set of repairs of $\mathcal{J}$.

Definition 3 captures some definitions of repair proposed in the literature, such as the repair at the basis of the prototypical *AR*-semantics or the repair adopted by the *CAR*-semantics [Lembo *et al.*, 2015; Rosati, 2011].

Indeed, given an ontology $\mathcal{O} = \mathcal{T} \cup \mathcal{A}$, repairs in the *AR*-semantics aim to preserve as many facts as possible of those belonging to $\mathcal{A}$. This means that, in a CQA instance adopting the *AR*-semantics, the language $\mathcal{L}_r$ has to be set to $\mathcal{A}$. Differently, the *CAR*-semantics aims to preserve as many facts as possible of those implied by $\mathcal{T}$ and each subset of $\mathcal{A}$ consistent with $\mathcal{T}$. Therefore, to encode such semantics $\mathcal{L}_r$ has to coincide with the set of ground atoms $\mathbf{GA}(\mathcal{O})$.

We now provide some conditions on CQE and CQA instances that allow to establish correspondences between (theories of) censors and repairs. We first analyze CQE instances.

Definition 4 A CQE instance $\mathcal{E} = \langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$ is *CQA-reducible* w.r.t. a language $\mathcal{L}_c \subseteq \mathbf{FO}(\mathcal{T} \cup \mathcal{A})$ if:

(i) for every $\phi \in \mathcal{L}_c$ such that $\mathcal{T} \cup \mathcal{A} \models \phi$ and $\{\phi\} \cup \mathcal{T} \cup \mathcal{P}$ is consistent, there exists $\mathcal{A}' \subseteq \mathcal{A}$ such that $\mathcal{T} \cup \mathcal{A}' \cup \mathcal{P}$ is consistent and $\mathcal{T} \cup \mathcal{A}' \models \phi$;

(ii) for every $\phi \in \mathcal{L}_c$ and every $\mathcal{A}' \subseteq \mathcal{A}$ such that $\mathcal{T} \cup \mathcal{A}' \cup \mathcal{P}$ is consistent, if $\mathcal{T} \cup \mathcal{A}' \cup \mathcal{P} \models \phi$ then $\mathcal{T} \cup \mathcal{A}' \models \phi$.

[2]Our notion of policy generalizes the one given in [Cuenca Grau *et al.*, 2015], where $\mathcal{P}$ is a single CQ.

If $\mathcal{L}_c$ is $\mathbf{CQ}$ (or $\mathbf{GA}$, or $\mathcal{A}$) we say that the instance $\mathcal{E}$ is CQ-Cens-CQA-Reducible (resp. GA-Cens-CQA-Reducible, or ABox-Cens-CQA-Reducible).

In words, condition (i) imposes that every logical consequence of $\mathcal{T} \cup \mathcal{A}$ that is consistent with the policy and the TBox belongs to $CES_{\mathcal{L}_c}(\mathcal{T} \cup \mathcal{P}, \mathcal{A})$, i.e., the consistent entailment set in $\mathcal{L}_c$ for an ontology obtained by putting together the TBox, the ABox, and the policy (which indeed might result inconsistent). Condition (ii) instead says that in a CQE insance that is CQA-reducible w.r.t $\mathcal{L}_c$ the sentences in the policy act as constraints on top of $\mathcal{T} \cup \mathcal{A}$, since they never contribute to infer new formulas from $\mathcal{L}_c$ if added to $\mathcal{T} \cup \mathcal{A}$ (notice however that $\mathcal{T} \cup \mathcal{A}$ can contradict denials in $\mathcal{P}$, and thus in Definition 4 we consider subsets of $\mathcal{A}$ that are consistent with $\mathcal{T} \cup \mathcal{P}$).

Example 3 The instance $\mathcal{E} = \langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$ with $\mathcal{T} = \{A \sqsubseteq B\}$, $\mathcal{A} = \{A(d)\}$, $\mathcal{P} = \{\forall x. A(x) \rightarrow \bot\}$ is not GA-Cens-CQA-Reducible, since it does not respect condition (i), even though it satisfies condition (ii) (in a trivial way). Instead, $\mathcal{E}' = \langle \mathcal{T}, \mathcal{A}', \mathcal{P}' \rangle$ with $\mathcal{A}' = \{A(d), B(d)\}$, $\mathcal{P} = \{\forall x. A(x) \wedge B(x) \rightarrow \bot\}$ and $\mathcal{T}$ as before, is GA-Cens-CQA-Reducible.

For CQA-reducible instances the following result holds.

Theorem 1 *Let $\mathcal{E} = \langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$ be a CQE instance and let $\mathcal{L}_c \subseteq \mathbf{FO}(\mathcal{T} \cup \mathcal{A})$, such that $\mathcal{E}$ is CQA-reducible w.r.t. $\mathcal{L}_c$. Then $Th^{oc}_{\mathcal{L}_c}(\mathcal{E}) = \mathsf{RepSet}_{\mathcal{L}_c}(\langle \mathcal{T} \cup \mathcal{P}, \mathcal{A} \rangle)$.*

Below we consider reducibility of CQA instances into CQE ones, and provide a notion analogous to Definition 4.

Definition 5 A CQA instance $\mathcal{J} = \langle \mathcal{T}, \mathcal{A} \rangle$ is *CQE-reducible* w.r.t. a language $\mathcal{L}_r$ if there exists a partition $\mathcal{T}_P \cup \mathcal{T}_N$ of $\mathcal{T}$ such that $\mathcal{T}_P \cup \mathcal{A}$ is consistent, $\mathcal{T}_N$ is equivalent to a set of denials, and:

(i) for every $\phi \in \mathcal{L}_r$, such that $\mathcal{T}_P \cup \mathcal{A} \models \phi$ and $\{\phi\} \cup \mathcal{T}$ is consistent, there exists $\mathcal{A}' \subseteq \mathcal{A}$ such that $\mathcal{T} \cup \mathcal{A}'$ is consistent and $\mathcal{T} \cup \mathcal{A}' \models \phi$;

(ii) for every $\phi \in \mathcal{L}_r$ and every $\mathcal{A}' \subseteq \mathcal{A}$ such that $\mathcal{T} \cup \mathcal{A}'$ is consistent, if $\mathcal{T} \cup \mathcal{A}' \models \phi$ then $\mathcal{T}_P \cup \mathcal{A}' \models \phi$.

If $\mathcal{L}_r$ is $\mathbf{CQ}$ (or $\mathbf{GA}$, or $\mathcal{A}$) we say that the instance $\mathcal{J}$ is CQ-Rep-CQE-Reducible (resp. GA-Rep-CQE-Reducible, or ABox-Rep-CQE-Reducible).

Intuitively, the above definition says that in a CQE-reducible instance we can identify a portion $\mathcal{T}_N$ of $\mathcal{T}$ such that its assertions act as constraints on the ontology $\mathcal{T}_P \cup \mathcal{A}$ (cond. (ii)), thus $\mathcal{T}_N$ behaves as a policy in a CQE instance. At the same time, each logical consequence in $\mathcal{L}_r$ of $\mathcal{T}_P \cup \mathcal{A}$ consistent with $\mathcal{T}$ must belong to $CES_{\mathcal{L}_r}(\mathcal{T}, \mathcal{A})$ (cond. (i)).

CQE-reducible instances have the following property.

Theorem 2 *Let $\mathcal{J} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a CQA instance, such that $\mathcal{J}$ is CQE-reducible w.r.t $\mathcal{L}_r$ and $\mathcal{T} = \mathcal{T}_P \cup \mathcal{T}_N$. Then $\mathsf{RepSet}_{\mathcal{L}_r}(\mathcal{J}) = Th^{oc}_{\mathcal{L}_r}(\langle \mathcal{T}_P, \mathcal{A}, \mathcal{T}_N \rangle)$.*

We now rephrase entailment in CQA [Lembo *et al.*, 2015]. As done for CQE, we define three entailment problems. That is, given a CQA instance $\mathcal{J} = \langle \mathcal{T}, \mathcal{A} \rangle$ and an FO sentence ϕ, we define: *(CQ-Rep-Entailment)*, i.e., decide whether $\mathcal{T} \cup \mathcal{R} \models \phi$ for every $\mathcal{R} \in \mathsf{RepSet}_{\mathbf{CQ}}(\mathcal{J})$, denoted by $\mathcal{J} \models^{cqa}_{CQ} \phi$; *(GA-Rep-Entailment)*, i.e., decide whether $\mathcal{T} \cup \mathcal{R} \models \phi$ for every $\mathcal{R} \in \mathsf{RepSet}_{\mathbf{GA}}(\mathcal{J})$, denoted by $\mathcal{J} \models^{cqa}_{GA} \phi$; *(ABox-Rep-Entailment)*, i.e., decide whether $\mathcal{T} \cup \mathcal{R} \models \phi$ for every $\mathcal{R} \in \mathsf{RepSet}_{\mathcal{A}}(\mathcal{J})$, denoted by $\mathcal{J} \models^{cqa}_{ABox} \phi$. Notice that the last two types of entailment coincide with entailment under CAR- and AR-semantics, respectively.

The following result follows immediately from Definition 2, the definition of entailment in CQA, and Theorem 1.

Corollary 1 *Let $X \in \{CQ, GA, ABox\}$, and let $\mathcal{E} = \langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$ be a CQE instance, such that $\mathcal{E}$ is X-Cens-CQA-reducible, and ϕ an FO sentence. Then, $\mathcal{E} \models^{cqe}_{X} \phi$ iff $\mathcal{J} \models^{cqa}_{X} \phi$, where $\mathcal{J} = \langle \mathcal{T} \cup \mathcal{P}, \mathcal{A} \rangle$.*

Analogously, the following result follows from Definition 2, the definition of entailment in CQA, and Theorem 2.

Corollary 2 *Let $X \in \{CQ, GA, ABox\}$, and let $\mathcal{J} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a CQA instance with $\mathcal{T} = \mathcal{T}_P \cup \mathcal{T}_N$, such that $\mathcal{J}$ is X-Rep-CQE-reducible, and ϕ an FO sentence. Then, $\mathcal{J} \models^{cqa}_{X} \phi$ iff $\mathcal{E} \models^{cqe}_{X} \phi$, where $\mathcal{E} = \langle \mathcal{T}_P, \mathcal{A}, \mathcal{T}_N \rangle$.*

5 CQE under Restricted Censor Languages

In this section we establish data complexity of CQE instance checking and entailment of BCQs for both *DL-Lite*$_R$ and $\mathcal{EL}_\bot$ CQE instances when the censor language is either the ABox of the instance or $\mathbf{GA}$. For the former case, we establish our complexity results by exploiting a mutual reduction between entailment in CQE and CQA. For the latter case, the two frameworks behave in a slightly different way, and thus we also need to use techniques tailored to the CQE setting. The results showed in this section allow us to clarify the computational properties of query answering in CQE when we adopt a restricted censor language, i.e., which can be less expressive than the query language, as in the case of GA-Cens-Entailment and ABox-Cens-Entailment of BCQs.

We start by setting the censor language to the assertions in the ABox.

Theorem 3 *Each DL-Lite$_R$ or $\mathcal{EL}_\bot$ CQE instance is ABox-Cens-CQA-Reducible, and each DL-Lite$_R$ or $\mathcal{EL}_\bot$ CQA instance is ABox-Rep-CQE-Reducible.*

Then, we establish an upper bound for entailment of BCQs.

Theorem 4 *ABox-Cens-Entailment of BCQs is in coNP in data complexity for both DL-Lite$_R$ and $\mathcal{EL}_\bot$ CQE instances.*

Proof (sketch). From Theorem 3 (direction from CQE to CQA) and Corollary 1, it follows that ABox-Cens-Entailment in *DL-Lite*$_R$ is equivalent to ABox-Rep-Entailment in *DL-Lite*$_{R,den}$, i.e., entailment under the *AR*-semantics, and ABox-Cens-Entailment in $\mathcal{EL}_\bot$ is equivalent to ABox-Rep-Entailment in $\mathcal{EL}_\bot$ plus denials. The thesis then follows from the fact that entailment of BCQs is in coNP in data complexity in CQA under the *AR*-semantics, for both *DL-Lite*$_{R,den}$ [Lembo *et al.*, 2015], and $\mathcal{EL}_\bot$ plus denials. This last result is a consequence of an analogous complexity result shown in [Rosati, 2011] for $\mathcal{EL}_\bot$. ■

The following theorem provides matching lower bounds for the results of Theorem 4.

Theorem 5 *ABox-Cens-Instance-Checking is coNP-hard in data complexity for both DL-Lite$_R$ and $\mathcal{EL}_\bot$ CQE instances.*

Proof (sketch). The results follow from Theorem 3 (direction from CQA to CQE), Corollary 2, and from coNP-hardness of instance checking of CQA under the AR-semantics for both *DL-Lite*$_R$ [Lembo *et al.*, 2015] and $\mathcal{EL}_\perp$ [Rosati, 2011]. ∎

Theorem 4 and Theorem 5 actually imply that both ABox-Cens-Instance-Checking and ABox-Cens-Entailment of BCQs are coNP-complete in data complexity for both *DL-Lite*$_R$ and $\mathcal{EL}_\perp$ CQE instances.

We now consider the case in which the censor language coincides with $\mathbf{GA}$. In this case, *DL-Lite*$_R$ and $\mathcal{EL}_\perp$ CQE instances are *not* always CQA-reducible, as shown in Example 3, where the non-reducible instance is both *DL-Lite*$_R$ and $\mathcal{EL}_\perp$. Reducibility in the other way round is also not always possible. However, for *DL-Lite*$_R$ we can show some weaker, but useful, properties.

Proposition 1 *Each DL-Lite*$_R$ *CQE instance* $\langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$*, such that* $\mathcal{T} \cup \mathcal{P} \cup \{\alpha\}$ *is consistent for each* $\alpha \in \mathcal{A}$*, is GA-Cens-CQA-Reducible. Also, each DL-Lite*$_R$ *CQA instance* $\langle \mathcal{T}, \mathcal{A} \rangle$*, such that* $\mathcal{T} \cup \{\alpha\}$ *is consistent for each* $\alpha \in \mathcal{A}$*, is GA-Rep-CQE-Reducible.*

Proposition 1 is used to prove the following theorem, which in fact is stated for general *DL-Lite*$_R$ CQE instances.

Theorem 6 *GA-Cens-Instance-Checking and GA-Cens-Entailment of BCQs are respectively in* AC^0 *and coNP-complete in data complexity for DL-Lite*$_R$ *CQE instances.*

Proof (sketch). For CQE instances satisfying the condition in Proposition 1 (direction from CQE to CQA), the membership results follow from that proposition, Corollary 1, and from the fact that GA-Rep-Entailment, i.e., entailment under CAR-semantics, of ground atoms and of BCQs over *DL-Lite*$_{R,den}$ CQA instances are respectively in AC^0 (which follows from the results in [Lembo *et al.*, 2015; Lembo *et al.*, 2011]) and in coNP (which follows from the results in [Lembo *et al.*, 2010]). The case of general *DL-Lite*$_R$ instances can be proved by adapting the techniques used to prove the mentioned AC^0 and coNP membership for CQA. coNP-hardness for GA-Cens-Entailment of BCQs follows from Proposition 1 (direction from CQA to CQE), Corollary 2, and from coNP-hardness of GA-Rep-Entailment of BCQs in *DL-Lite*$_R$ [Lembo *et al.*, 2010]. ∎

Let us now turn to $\mathcal{EL}_\perp$. While to establish GA-Rep-Entailment of BCQs one needs to check that the consequences follow from subsets of the ABox that are consistent with the denials [Rosati, 2011], this is not needed to establish GA-Cens-Entailment, leading to a lower upper bound, i.e., coNP. coNP-hardness then follows from the coNP-hardness of ABox-Rep-Entailment of BCQs shown in [Bienvenu and Bourgaux, 2016, Theorem 17], which carries over also to CAR-semantics and CQE (the semantic difference between CQE and CQA in this case does not show up).

Theorem 7 *GA-Cens-Entailment of BCQs is coNP-complete in data complexity for* $\mathcal{EL}_\perp$ *CQE instances.*

6 CQE under Full Censor Language

In this section we study entailment of BCQs under our CQE framework for both *DL-Lite*$_R$ and $\mathcal{EL}_\perp$ CQE instances and more expressive censor languages.

Algorithm 1: Algorithm *CQ-Ent-DL-Lite*$_R(\mathcal{E}, q)$

Input: *DL-Lite*$_R$ CQE instance $\mathcal{E} = \langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$, BCQ q
Output: true if $\mathcal{E} \models^{cqe}_{CQ} q$, false otherwise

```
let h be the maximum length of a denial in P;
let k = max(h, length(q));
Φ = CQEntailedSubset(T, A, k);
for i = 1 to k do
    remove from Φ every subset Φ′ such that |Φ′| = i
    and T ∪ P ∪ Φ′ is inconsistent;
if q ∈ Φ then return true else return false
```

We first concentrate on *DL-Lite*$_R$ and study CQ-Cens-Entailment of BCQs. We start by providing the following crucial property, which says that to solve this problem it is possible to resort to CQ$_k$-Cens-Entailment, i.e., CQE entailment defined over theories of censors using $\mathbf{CQ}_k$ as censor language, denoted $\models^{cqe}_{CQ_k}$.

Theorem 8 *Let* $\mathcal{E} = \langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$ *be a DL-Lite*$_R$ *CQE instance, let* q *be a BCQ, and let* $k = \max(h, \mathit{length}(q))$*, where* h *is the maximum length of a denial assertion in* $\mathcal{P}$*. Then,* $\mathcal{E} \models^{cqe}_{CQ} q$ *iff* $\mathcal{E} \models^{cqe}_{CQ_k} q$.

Proof (sketch). We prove the if direction (the other one is trivial). The hypothesis implies that q belongs to every theory of optimal censor for $\mathcal{E}$ in $\mathbf{CQ}_k$, i.e., $q \in \Psi$ for every $\Psi \in \mathit{Th}^{oc}_{\mathbf{CQ}_k}(\mathcal{E})$. Now, suppose $\mathcal{E} \not\models^{cqe}_{CQ} q$. Then, there exists $\Psi \in \mathit{Th}^{oc}_{\mathbf{CQ}}(\mathcal{E})$ such that $q \notin \Psi$. So, $\Psi \cup \{q\} \cup \mathcal{T} \cup \mathcal{P}$ is inconsistent.

Now, the following property can be shown: $\Psi \cup \{q\} \cup \mathcal{T} \cup \mathcal{P}$ is inconsistent iff there exists $\phi \in \mathcal{P}$ such that $\Psi \cup \{q\} \cup \{\phi\}$ is inconsistent. This property follows from the fact that, when $\mathcal{T}$ is a *DL-Lite*$_R$ TBox, the inconsistency of $\mathcal{T} \cup \mathcal{P}$ with respect to a set of BCQs Ψ' implies the existence of a denial assertion that is entailed by $\mathcal{T} \cup \mathcal{P}$ and is violated by Ψ', and from the fact that Ψ is a deductively closed set of BCQs with respect to $\mathcal{T}$. Moreover, since the length of a denial assertion ϕ is not greater than k, it is immediate to verify that if $\Psi \cup \{q\} \cup \{\phi\}$ is inconsistent then $(\Psi \cap \mathbf{CQ}_k) \cup \{q\} \cup \{\phi\}$ is inconsistent.

On the other hand, since $(\Psi \cap \mathbf{CQ}_k) \cup \mathcal{T} \cup \mathcal{P}$ is consistent, there exists $\Psi' \in \mathit{Th}^{oc}_{\mathbf{CQ}_k}(\mathcal{E})$ such that $\Psi \cap \mathbf{CQ}_k \subseteq \Psi'$, but since, by hypothesis, $\mathcal{E} \models^{cqe}_{CQ_k} q$, it follows that $q \in \Psi'$, which implies that $(\Psi \cap \mathbf{CQ}_k) \cup \{q\} \cup \{\phi\}$ is consistent. This leads to a contradiction, and thus the thesis follows. ∎

Hereinafter, without loss of generality, we assume that all formulas of the language $\mathbf{CQ}_k$ and the query q of the entailment problem use the set of $2k$ variables $\{x_1, \ldots, x_{2k}\}$.

For deciding CQ-Cens-Entailment of BCQs for the *DL-Lite*$_R$ case we define Algorithm 1, in which *CQEntailedSubset*$(\mathcal{T}, \mathcal{A}, k)$ is the function returning the set of BCQs from $\mathbf{CQ}_k$ that are entailed by $\mathcal{T} \cup \mathcal{A}$. It is immediate to verify that this function can be computed in polynomial time w.r.t. the size of $\mathcal{A}$.

Informally, the algorithm first computes an integer k, based on the length of the query q and of the denials in $\mathcal{P}$; then, it computes the set Φ that represents the intersection of the theories of the optimal censors for the CQE instance $\mathcal{E}$ in $\mathbf{CQ}_k$:

Algorithm 2: Algorithm CQ_k-*Ent*-$\mathcal{EL}_\perp(\mathcal{E}, q)$

Input: $\mathcal{EL}_\perp$ CQE instance $\mathcal{E} = \langle \mathcal{T}, \mathcal{A}, \mathcal{P} \rangle$, BCQ q
Output: true if $\mathcal{E} \models^{cqe}_{CQ_k} q$, false otherwise
let h be the maximum length of a denial assertion in $\mathcal{P}$;
let $k = \max(h, \mathit{length}(q))$;
let $\Phi = \mathit{CQEntailedSubset}(\mathcal{T}, \mathcal{A}, k)$;
if there exists $\Phi' \subseteq \Phi$ such that
Φ' is a maximal subset of Φ consistent with $\mathcal{T} \cup \mathcal{P}$
and $q \notin \Phi'$
then return false **else** return true

this is done by removing from Φ all formulas that belong to minimal subsets of Φ that are inconsistent with $\mathcal{T} \cup \mathcal{P}$; finally, it checks whether q belongs to Φ.

Theorem 9 *Let $\mathcal{E}$ be a DL-Lite$_R$ CQE instance, and q a BCQ. Then, $\mathcal{E} \models^{cqe}_{CQ} q$ iff CQ-Ent-DL-Lite$_R(\mathcal{E}, q)$ returns true.*

Proof (sketch). It can easily be shown that the set Φ computed by *CQ-Ent-DL-Lite*$_R(\mathcal{E}, q)$ is the set $\bigcap_{\Psi \in Th^{oc}_{\mathbf{CQ}_k}(\mathcal{E})} \Psi$. This property is based on the fact that, in the case of *DL-Lite*$_R$ TBoxes, every minimal subset of the set returned by *CQEntailedSubset*$(\mathcal{T}, \mathcal{A}, k)$ that is inconsistent with $\mathcal{T} \cup \mathcal{P}$ contains at most k formulas. Then, it follows that $\mathcal{E} \models^{cqe}_{CQ_k} q$ iff $q \in \bigcap_{\Psi \in Th^{oc}_{\mathbf{CQ}_k}(\mathcal{E})} \Psi$. Consequently, $\mathcal{E} \models^{cqe}_{CQ_k} q$ iff $q \in \Phi$. Therefore, from Theorem 8 it follows that $\mathcal{E} \models^{cqe}_{CQ} q$. ∎

We are now ready to state a complexity result for the CQ-Cens-Entailment problem in the case of *DL-Lite*$_R$ TBoxes.

Theorem 10 *CQ-Cens-Entailment of BCQs is in PTIME in data complexity for DL-Lite$_R$ CQE instances.*

Proof. We prove that the *CQ-Ent-DL-Lite*$_R$ algorithm provides a polynomial-time upper bound in data complexity.

First, as already mentioned, *CQEntailedSubset*$(\mathcal{T}, \mathcal{A}, k)$ can be computed in time polynomial w.r.t. the size of $\mathcal{A}$. Then, it is also easy to verify that checking the consistency of an ontology consisting of a *DL-Lite*$_R$ TBox, a policy and a set of BCQs can be done in polynomial time as well: indeed, to check consistency, the BCQs can be encoded into a standard ABox with fresh constant symbols to represent the existential variables of the BCQs, while the policy corresponds to a set of denial assertions in the DL *DL-Lite*$_{R,den}$, that is, $\mathcal{T} \cup \mathcal{P}$ is a *DL-Lite*$_{R,den}$ TBox. We can then check the consistency of the resulting *DL-Lite*$_{R,den}$ ontology in polynomial time [Lembo *et al.*, 2015]. This implies that the for-loop of the algorithm can be executed in polynomial time with respect to the size of $\mathcal{A}$, since it corresponds to the execution of a polynomial number of inconsistency checks of polynomially-sized *DL-Lite*$_{R,den}$ ontologies of the above form. ∎

We now consider the case of $\mathcal{EL}_\perp$ CQE instances. We restrict our analysis to CQ$_k$-Cens-Entailment of BCQs, for which we provide the nondeterministic Algorithm 2. While Algorithm 1 constructs a finite fragment of the intersection of the theories of all the optimal censors for $\mathcal{E}$ in $\mathbf{CQ}$, Algorithm 2 looks for the existence of a theory of an optimal censor for $\mathcal{E}$ in $\mathbf{CQ}_k$ that does not contain the query q.

The following theorem easily follows from Definition 2.

Theorem 11 *Let $\mathcal{E}$ be a $\mathcal{EL}_\perp$ CQE instance, and q a BCQ. Then, $\mathcal{E} \models^{cqe}_{CQ_k} q$ iff CQ$_k$-Ent-$\mathcal{EL}_\perp(\mathcal{E}, q)$ returns true.*

CQ_k-*Ent*-$\mathcal{EL}_\perp$ allows us to establish a coNP upper bound of the data complexity of CQ$_k$-Cens-Entailment of BCQs for $\mathcal{EL}_\perp$ CQE instances (note that *CQEntailedSubset*$(\mathcal{T}, \mathcal{A}, k)$ can be computed in polynomial time in the size of $\mathcal{A}$ also when $\mathcal{T}$ is an $\mathcal{EL}_\perp$ TBox). It is in fact not difficult to show that the above bound is tight.

Theorem 12 *CQ$_k$-Cens-Entailment of BCQs is coNP-complete in data complexity for $\mathcal{EL}_\perp$ CQE instances.*

7 Discussion and Conclusions

The complexity results for *DL-Lite*$_R$ TBoxes given in this paper show a surprising aspect. In fact, the complexity of entailment of BCQs when the censor language is restricted either to the ABox or to the set of ground atoms is harder than when the censor language is $\mathbf{CQ}$. The explanation of this lies in the fact that, in the latter case, it is possible to establish the entailment by computing the intersection of (a finite and polynomial representation of) all the theories of optimal censors (see Theorem 9), which, as shown in the previous section, can be done in polynomial time in data complexity. This property does not hold for the more restricted censor languages, which require to consider separately all the theories (actually, an exponential number of finite approximations of such theories).

The above property also holds for CQ$_k$-Cens-Entailment of BCQs for $\mathcal{EL}_\perp$ TBoxes. However, in this case it is not possible to compute (a finite representation of) the intersection of all the theories of the optimal censors in PTIME: indeed, differently from *DL-Lite*$_R$, for $\mathcal{EL}_\perp$ the size of the minimal subsets of the set returned by *CQEntailedSubset*$(\mathcal{T}, \mathcal{A}, k)$ that are inconsistent with $\mathcal{T} \cup \mathcal{P}$ is not independent of the size of $\mathcal{A}$. This explains the coNP-hardness in this case.

The present work can be extended in several directions. For $\mathcal{EL}_\perp$ CQE instances we left open the lower bounds of GA-Cens-Instance-Checking and ABox-Cens-Instance-Checking, and the complexity of CQ-Cens-Entailment of BCQs. Also, the PTIME upper bound for CQ-Cens-Entailment of BCQs over *DL-Lite*$_R$ CQE instances should be refined. We believe that an AC^0 bound can be shown in this case: in particular, the *first-order rewritability* of CQE could be proved by adapting and extending query rewriting techniques for CQA in the DL *DL-Lite*$_{R,den}$ [Lembo *et al.*, 2015]. Then, the complexity analysis of CQE could be extended to other DLs, policy and censor languages. Also, based on the complexity analysis of CQE presented in this paper, it would be very important to look for practical techniques allowing for the implementation of CQE extensions of current DL reasoners and Ontology-based Data Access systems [Calvanese *et al.*, 2017; De Giacomo *et al.*, 2012].

Acknowledgements

This work is partially supported by the EC within the H2020 under grant agreement 825333 (MOSAICrOWN). We thank the reviewers for pointing out a potential problem in the proof of Theorem 8 in the submitted version of the paper.

References

[Baader *et al.*, 2005] Franz Baader, Sebastian Brandt, and Carsten Lutz. Pushing the $\mathcal{EL}$ envelope. In *Proc. of IJCAI*, pages 364–369, 2005.

[Baader *et al.*, 2007] Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press, 2nd edition, 2007.

[Benedikt *et al.*, 2018] Michael Benedikt, Bernardo Cuenca Grau, and Egor V. Kostylev. Logical foundations of information disclosure in ontology-based data integration. *AIJ*, 262:52–95, 2018.

[Bertossi and Li, 2013] Leopoldo E. Bertossi and Lechen Li. Achieving data privacy through secrecy views and null-based virtual updates. *IEEE Trans. Knowl. Data Eng.*, 25(5):987–1000, 2013.

[Bertossi, 2011] Leopoldo E. Bertossi. *Database Repairing and Consistent Query Answering*. Synthesis Lectures on Data Management. Morgan & Claypool Publishers, 2011.

[Bienvenu and Bourgaux, 2016] Meghyn Bienvenu and Camille Bourgaux. Inconsistency-tolerant querying of description logic knowledge bases. In *RW Tutorial Lectures*, pages 156–202, 2016.

[Biskup and Bonatti, 2004a] Joachim Biskup and Piero A. Bonatti. Controlled query evaluation for enforcing confidentiality in complete information systems. *Int. J. Inf. Sec.*, 3(1):14–27, 2004.

[Biskup and Bonatti, 2004b] Joachim Biskup and Piero A. Bonatti. Controlled query evaluation for known policies by combining lying and refusal. *Ann. Math. Artif. Intell.*, 40(1-2):37–62, 2004.

[Biskup and Weibert, 2008] Joachim Biskup and Torben Weibert. Keeping secrets in incomplete databases. *Int. J. Inf. Sec.*, 7(3):199–217, 2008.

[Bonatti and Sauro, 2013] Piero A. Bonatti and Luigi Sauro. A confidentiality model for ontologies. In *Proc. of ISWC*, pages 17–32, 2013.

[Calvanese *et al.*, 2007] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, and Riccardo Rosati. Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. *J. of Automated Reasoning*, 39(3):385–429, 2007.

[Calvanese *et al.*, 2012] Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, and Riccardo Rosati. View-based query answering in description logics: Semantics and complexity. *JCSS*, 78(1):26–46, 2012.

[Calvanese *et al.*, 2017] Diego Calvanese, Benjamin Cogrel, Sarah Komla-Ebri, Roman Kontchakov, Davide Lanti, Martin Rezk, Mariano Rodriguez-Muro, and Guohui Xiao. Ontop: Answering SPARQL queries over relational databases. *Semantic Web*, 8(3):471–487, 2017.

[Cuenca Grau and Horrocks, 2008] Bernardo Cuenca Grau and Ian Horrocks. Privacy-preserving query answering in logic-based information systems. In *Proc. of ECAI*, pages 40–44, 2008.

[Cuenca Grau *et al.*, 2013] Bernardo Cuenca Grau, Evgeny Kharlamov, Egor V. Kostylev, and Dmitriy Zheleznyakov. Controlled query evaluation over OWL 2 RL ontologies. In *Proc. of ISWC*, pages 49–65, 2013.

[Cuenca Grau *et al.*, 2015] Bernardo Cuenca Grau, Evgeny Kharlamov, Egor V. Kostylev, and Dmitriy Zheleznyakov. Controlled query evaluation for datalog and OWL 2 profile ontologies. In *Proc. of IJCAI*, pages 2883–2889, 2015.

[De Giacomo *et al.*, 2012] Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi, Riccardo Rosati, Marco Ruzzi, and Domenico Fabio Savo. MASTRO: A reasoner for effective Ontology-Based Data Access. In *Proc. of ORE*, 2012.

[Lembo *et al.*, 2010] Domenico Lembo, Maurizio Lenzerini, Riccardo Rosati, Marco Ruzzi, and Domenico Fabio Savo. Inconsistency-tolerant semantics for description logics. In *Proc. of RR*, pages 103–117, 2010.

[Lembo *et al.*, 2011] Domenico Lembo, Maurizio Lenzerini, Riccardo Rosati, Marco Ruzzi, and Domenico Fabio Savo. Query rewriting for inconsistent *DL-Lite* ontologies. In *Proc. of RR*, 2011.

[Lembo *et al.*, 2015] Domenico Lembo, Maurizio Lenzerini, Riccardo Rosati, Marco Ruzzi, and Domenico Fabio Savo. Inconsistency-tolerant query answering in ontology-based data access. *J. of Web Semantics*, 33:3–29, 2015.

[Lembo *et al.*, 2018] Domenico Lembo, Riccardo Rosati, and Domenico Fabio Savo. A comprehensive framework for controlled query evaluation, consistent query answering and KB updates in Description Logics. In *Proc. of KR*, pages 653–654, 2018.

[Rosati, 2011] Riccardo Rosati. On the complexity of dealing with inconsistency in Description Logic ontologies. In *Proc. of IJCAI*, pages 1057–1062, 2011.

[Sicherman *et al.*, 1983] George L. Sicherman, Wiebren de Jonge, and Reind P. van de Riet. Answering queries without revealing secrets. *ACM Trans. Database Syst.*, 8(1):41–59, 1983.

[Stouppa and Studer, 2009] Phiniki Stouppa and Thomas Studer. Data privacy for $\mathcal{ALC}$ knowledge bases. In *Proc. of LFCS 2009*, pages 409–421, 2009.

[Studer and Werner, 2014] Thomas Studer and Johannes Werner. Censors for boolean description logic. *Trans. Data Privacy*, 7(3):223–252, 2014.

[Tao *et al.*, 2010] Jia Tao, Giora Slutzki, and Vasant G. Honavar. Secrecy-preserving query answering for instance checking in $\mathcal{EL}$. In *Proc. of RR*, pages 195–203, 2010.

[Tao *et al.*, 2014] Jia Tao, Giora Slutzki, and Vasant G. Honavar. A conceptual framework for secrecy-preserving reasoning in knowledge bases. *ACM TOCL*, 16(1):3:1–3:32, 2014.