Technische Universität Dresden
Institute for Theoretical Computer Science
Chair for Automata Theory

LTCS–Report

Inconsistency- and Error-Tolerant Reasoning w.r.t. Optimal Repairs of $\mathcal{EL}^{\bot}$ Ontologies (Extended Version)

Franz Baader, Francesco Kriegel, Adrian Nuradiansyah

LTCS-Report 24-02

This is an extended version of an article accepted at FoIKS 2024.

Postal Address:
Lehrstuhl für Automatentheorie
Institut für Theoretische Informatik
TU Dresden
01062 Dresden

http://lat.inf.tu-dresden.de

Visiting Address:
Nöthnitzer Str. 46
Dresden

Inconsistency- and Error-Tolerant Reasoning w.r.t. Optimal Repairs of $\mathcal{EL}^\bot$ Ontologies (Extended Version)

Franz Baader[1,2], Francesco Kriegel[1,2], and Adrian Nuradiansyah[1]

[1]Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany
[2]Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI)
firstname.lastname@tu-dresden.de

Abstract. Errors in knowledge bases (KBs) written in a Description Logic (DL) are usually detected when reasoning derives an inconsistency or a consequence that does not hold in the application domain modelled by the KB. Whereas classical repair approaches produce maximal subsets of the KB not implying the inconsistency or unwanted consequence, optimal repairs maximize the consequence sets. In this paper, we extend previous results on how to compute optimal repairs from the DL $\mathcal{EL}$ to its extension $\mathcal{EL}^\bot$, which in contrast to $\mathcal{EL}$ can express inconsistency. The problem of how to deal with inconsistency in the context of optimal repairs was addressed previously, but in a setting where the (fixed) terminological part of the KB must satisfy a restriction on cyclic dependencies. Here, we consider a setting where this restriction is not required. We also show how the notion of optimal repairs obtained this way can be used in inconsistency- and error-tolerant reasoning.

1 Introduction

Description Logics (DLs) [2,3] are a well-investigated family of logic-based knowledge representation formalisms, which have gained particular prominence by the fact that they are the formal basis for the Web ontology language OWL,[1] and are thus employed in many application domains (e.g., biology and medicine [20]). In particular, in the setting of ontology-mediated query answering (OMQA) [15,27], concepts defined in the terminological part (TBox) of a DL knowledge base (KB) can be used as queries or within more complex queries on data, represented in the assertional part (ABox) of the KB.

For example, assume that the TBox $\mathcal{T}$ consists of the concept inclusions (CIs) $\mathit{Man} \sqsubseteq \mathit{Human}$ and $\mathit{Human} \sqcap \exists \mathit{loves}.\mathit{Human} \sqsubseteq \mathit{Caring}$, which say that men are humans and that a human loving some human is caring; and that the ABox $\mathcal{A}$ consists of the assertions $\mathit{Man}(n)$, $\mathit{loves}(n, n)$, and $\mathit{Egoistic}(n)$, which say that Narcissus (represented by the individual name n) is an egoistic man who loves himself. Given this KB, the instance query (IQ) $(\mathit{Caring} \sqcap \mathit{Man})(n)$, which checks whether the individual n is an instance of the concept $\mathit{Caring} \sqcap \mathit{Man}$, returns

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

`true` as an answer, and the conjunctive query (CQ) $\mathit{Human}(x) \wedge \mathit{loves}(x, x)$, which looks for self-loving humans, returns n. Note that, for both queries, we would not obtain this answer if we considered only the data (i.e., the ABox without the TBox). Also note that, in most DLs, this CQ cannot be expressed by instance queries due to the fact that it looks for a cycle in the data.

The DL community has spent considerable effort on designing sound, complete, and terminating inference algorithms for DLs of various degrees of expressiveness, not just for query answering, but also for other inference problems, such as the consistency and the subsumption problem [2,3]. However, even a sound inference procedure can produce consequences that are plainly wrong in the application domain that is modeled by a KB in case this KB contains errors. Inconsistency of a KB is a sure sign that there is something wrong with it, but errors can also be detected if consequences are produced that are not supposed to hold in the application domain. In our example, one may wonder whether it really makes sense to have a KB implying that Narcissus is both egoistic and caring. To correct this, one may thus try to construct a repair of it, i.e., a new KB from which the unwanted consequence $(\mathit{Caring} \sqcap \mathit{Egoistic})(n)$ no longer follows. As pointed out in [4], it is not reasonable to use as a repair an arbitrary KB that does not have the unwanted consequences. Additionally, the repaired KB should (a) not introduce new knowledge and (b) be as close as possible to the original KB. More formally, (a) can be reformulated as saying that every repair must be entailed by the original KB, and the optimality condition (b) chooses repairs that are not strictly entailed by another repair. This still leaves different possibilities for how to formalize the notion of an (optimal) repair, depending on which entailment relation is employed.

Classical repair approaches [24,26,14] read (a) as talking about the explicitly represented knowledge, and thus use the superset relation as the "entailment" relation. Thus, a *classical repair* is a subset of the original KB that does not have the unwanted consequences, and it is *optimal* if it is a maximal such subset. In our example, if we assume that the TBox is correct (and thus should not be changed) there are three optimal classical repairs, obtained by removing one of the three assertions from the ABox. As pointed out in [12], this classical approach has the disadvantage that it is syntax-dependent. If, in our example, we had used the single assertion $(\mathit{Man} \sqcap \mathit{Egoistic})(n)$, which is equivalent to the two assertions $\mathit{Man}(n)$ and $\mathit{Egoistic}(n)$, then a classical repair would need to remove the information that Narcissus is an egoistic man as a whole, thus leading to an optimal classical repair that is weaker than the one obtained when only the assertion $\mathit{Egoistic}(n)$ is removed. Even for the original version of our example, the optimal classical repair obtained by removing the assertion $\mathit{loves}(n, n)$ loses more consequences than necessary. In fact, this repair retains no information at all about love-relationships, though one could actually have kept the information that Narcissus loves something not known to be human (this could, e.g., be a pet), and even more (see below).

To be able to retain such entailed knowledge, we use in [13,5] as entailment relation the usual logical entailment induced by the semantics of TBoxes and

ABoxes, and extend ABoxes to quantified ABoxes (qABoxes), which may contain anonymous individuals. For example, the qABox $\exists\{x,y\}.\mathcal{B}$ with $\mathcal{B}$ equal to

$$\{Man(n), Egoistic(n), loves(n,x), loves(x,y), loves(y,y), Man(y), Egoistic(y)\}$$

is a (still not optimal) repair, which says that Narcissus is an egoistic man who loves something that in turn loves an egoistic man loving himself, where the existentially quantified variables x, y represent anonymous individuals whose names are not known. Using logical entailment is appropriate if we are interested in the answers to CQs that a KB yields. For this reason, this entailment relation is also called CQ-entailment in [5]. CQ-entailment is too strong if we are only interested in instance queries. In this case, it is sufficient to employ IQ-entailment, which looks at what concept assertions a given KB entails, i.e., the KB $\mathcal{K}_1$ IQ-entail the KB $\mathcal{K}_2$ if every concept assertion entailed by the latter is also entailed by the former.

In [5], we investigate the repair problem for the DL $\mathcal{EL}$, which has the top concept ($\top$), conjunction ($C \sqcap D$), and existential restrictions ($\exists r.C$) as concept constructors. More precisely, we consider KBs consisting of an $\mathcal{EL}$ TBox and a qABox, where the TBox is static (i.e., cannot be changed), and repair requests consisting of $\mathcal{EL}$ concept assertions. In the IQ-case, the set of optimal IQ-repairs can always be computed in exponential time, and this set covers all repairs in the sense that every IQ-repair is IQ-entailed by an optimal repair. In the CQ-case, this covering property need no longer hold; in particular, there are repair problems that have a CQ-repair, but no optimal one. To regain this important property, which implies that one does not lose repair options when concentrating on optimal repairs, one must restrict the TBox to being cycle-restricted.[2] With respect to cycle-restricted TBoxes, the set of optimal CQ-repairs can be computed in exponential time, but this computation requires the use of an NP-oracle.[3] These results show that it makes sense to tailor the employed entailment relation to the repair problem at hand. It must be strong enough to take the queries one is interested in and the ones used to specify the repair request into account, but also should not be stronger than that to avoid unnecessary complications such as non-coverage or higher computational complexity.

In this paper, we consider repairs in $\mathcal{EL}^\bot$, which extends $\mathcal{EL}$ with the bottom concept $\bot$, and in which inconsistency can be expressed. In our example, we could then use the CI $Caring \sqcap Egoistic \sqsubseteq \bot$ to say that no one can both be caring and egoistic. With this additional CI, the KB of our example becomes inconsistent. To repair the inconsistency, it is no longer enough to request that Narcissus should not be both caring and egoistic, one must also forbid this for all other individuals, also anonymous ones. For example, the qABox $\exists\{x,y\}.\mathcal{B}$ introduced above would still be inconsistent w.r.t. the extended TBox. Since anything follows from an inconsistent KB, condition (a) can no longer enforce

[2] For example, the CI $Human \sqsubseteq \exists loves.Human$ destroys cycle-restrictedness, whereas the CI $\exists loves.Human \sqsubseteq Human$ does not.

[3] In the CQ-case, using conjunctive instead of instance queries as repair requests may appear to be more appropriate, but this may destroy the covering property [9].

that the repair is related in a reasonable way to the original KB. This problem is solved in [9] for a more expressive DL, according to which any $\mathcal{EL}^\bot$ TBox can be transformed into a normal form that is the union of a positive part $\mathcal{T}_+$ not containing $\bot$ and a bottom part $\mathcal{T}_\bot$ consisting of CIs of the form $C \sqsubseteq \bot$ for $\mathcal{EL}$ concepts C. As entailment relation, one then uses entailment w.r.t. $\mathcal{T}_+$, and the repair request must be extended by global requests $\exists\{x\}.\{C(x)\}$ for all CIs $C \sqsubseteq \bot$ in $\mathcal{T}_\bot$, expressing that C must not be populated by any individual, be it named or anonymous. In our extended example, $\mathcal{T}_\bot$ consists of the CI $\mathit{Caring} \sqcap \mathit{Egoistic} \sqsubseteq \bot$, and $\mathcal{T}_+$ of all other CIs. If we want to repair the inconsistency, then this boils down to repairing the original ABox $\mathcal{A}$ for the global repair request $\exists\{x\}.\{(\mathit{Caring} \sqcap \mathit{Egoistic})(x)\}$ w.r.t. the TBox $\mathcal{T}_+$. In [9], such repair problems were investigated using CQ-entailment, which takes care of global repair requests since they can be represented as CQs. In the present paper, we investigate the IQ-case, which has the advantage that TBoxes need not be cycle-restricted. However, in addition to instance queries, the entailment relation now also needs to take global requests of the form $\exists\{x\}.\{C(x)\}$ into account. Following [11], we additionally allow role assertions to be contained in the repair request, which means that we have to consider gloIRQ-entailment, which takes these three types of queries (instance and roles assertions as well as global requests) into account. Dealing with such extended repair requests and the stronger entailment relation requires non-trivial changes to our repair approach in [5], but in the end we can again show that the set of optimal gloIRQ-repairs can be computed in exponential time, and this set covers all gloIRQ-repairs.

Both in $\mathcal{EL}$ and in $\mathcal{EL}^\bot$, a given repair problem may have exponentially many repairs, in the classical as well as in the optimal sense, and it is often hard to decide which one to use. Error-tolerant reasoning does not commit to a single repair, but rather reasons w.r.t. all of them: cautious reasoning returns the answers that follow from all repairs whereas brave reasoning returns the answers that follow from some repair. For classical repairs of $\mathcal{EL}$ TBoxes, it was investigated in [22,25], but for more expressive DLs that can create inconsistencies, error-tolerant reasoning w.r.t. classical repairs had been considered before, for the case where the error is an inconsistency, under the name of inconsistency-tolerant reasoning [17,16,21]. In [10], we investigate error-tolerant reasoning in $\mathcal{EL}$ w.r.t. the optimal repairs introduced in [5], and in [11], this work is extended to take also role assertions in the repair request into account. Here, we make the further extension from $\mathcal{EL}$ to $\mathcal{EL}^\bot$, which then also allows us to do inconsistency-tolerant reasoning. Again, the stronger entailment relation requires non-trivial changes of the approaches developed in [5,11].

This extended version contains all proofs of technical results that needed to be removed from the conference article for space restrictions.

2 Preliminaries

First, we briefly recall syntax and semantics of the DL $\mathcal{EL}$ and of quantified ABoxes, but refer the reader to standard texts on DLs [3] and to [13,5] for more detailed expositions. Then, we introduce the relevant entailment relations and recall some useful results regarding them from [5,7,11]. Finally, we define the notion of optimal repairs, and recall our previously obtained results for them [13,5,7,9].

2.1 The Description Logic $\mathcal{EL}$ and Quantified ABoxes

The *syntax* of $\mathcal{EL}$ concepts is defined inductively as follows. Starting with disjoint sets Σ_C of *concept names* and Σ_R of *role names*, $\mathcal{EL}$ *concepts* are built using the constructors top concept ($\top$), conjunction ($C \sqcap D$), and existential restriction ($\exists r.C$). An $\mathcal{EL}$ *atom* is either a concept name or an existential restriction. We use $\mathsf{Conj}(C)$ to denote the set of all $\mathcal{EL}$ atoms that occur as a top-level conjunct of C. Such concepts can be used to define both terminological and assertional knowledge. An $\mathcal{EL}$ *concept inclusion (CI)* is of the form $C \sqsubseteq D$ for $\mathcal{EL}$ concepts C, D. An $\mathcal{EL}$ *TBox* $\mathcal{T}$ is a finite set of such CIs. Given an additional set Σ_I of *individual names*, disjoint with Σ_C and Σ_R, an $\mathcal{EL}$ *concept assertion* is of the form $C(a)$, where C is an $\mathcal{EL}$ concept and $a \in \Sigma_\mathsf{I}$, and a *role assertion* is of the form $r(a, b)$, where $r \in \Sigma_\mathsf{R}$ and $a, b \in \Sigma_\mathsf{I}$. An $\mathcal{EL}$ *ABox* $\mathcal{A}$ is a finite set of $\mathcal{EL}$ concept assertions and role assertions.

The *semantics* of $\mathcal{EL}$ concepts, TBoxes, and ABoxes is defined in a model-theoretic way. An *interpretation* $\mathcal{I}$ is a pair $(\Delta^\mathcal{I}, \cdot^\mathcal{I})$, where the *domain* $\Delta^\mathcal{I}$ is a non-empty set, and the *interpretation function* $\cdot^\mathcal{I}$ maps each $a \in \Sigma_\mathsf{I}$ to an element $a^\mathcal{I}$ of $\Delta^\mathcal{I}$, each concept name $A \in \Sigma_\mathsf{C}$ to $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$, and each role name $r \in \Sigma_\mathsf{R}$ to a binary relation $r^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$. The interpretation $C^\mathcal{I}$ of an $\mathcal{EL}$ concept C is defined inductively as follows: $\top^\mathcal{I} := \Delta^\mathcal{I}$, $(C \sqcap D)^\mathcal{I} := C^\mathcal{I} \cap D^\mathcal{I}$, and $(\exists r.C)^\mathcal{I} := \{d \in \Delta^\mathcal{I} \mid \exists e \in \Delta^\mathcal{I}$ such that $(d, e) \in r^\mathcal{I}$ and $C^\mathcal{I}\}$. The interpretation $\mathcal{I}$ *satisfies* the CI $C \sqsubseteq D$ (denoted by $\mathcal{I} \models C \sqsubseteq D$) if $C^\mathcal{I} \subseteq D^\mathcal{I}$, the concept assertion $C(a)$ ($\mathcal{I} \models C(a)$) if $a^\mathcal{I} \in C^\mathcal{I}$, and the role assertion $r(a, b)$ ($\mathcal{I} \models r(a, b)$) if $(a^\mathcal{I}, b^\mathcal{I}) \in r^\mathcal{I}$. We further say that $\mathcal{I}$ is a *model* of the $\mathcal{EL}$ TBox $\mathcal{T}$ (ABox $\mathcal{A}$) if each CI in $\mathcal{T}$ (assertion in $\mathcal{A}$) is satisfied by $\mathcal{I}$.

A *quantified ABox (qABox)* $\exists X.\mathcal{A}$ consists of a finite set X of *variables*, which is disjoint with $\Sigma = \Sigma_\mathsf{I} \cup \Sigma_\mathsf{C} \cup \Sigma_\mathsf{R}$, and a *matrix* $\mathcal{A}$, which is a finite set of concept assertions $A(u)$ and role assertions $r(u, v)$, where $A \in \Sigma_\mathsf{C}$, $r \in \Sigma_\mathsf{R}$ and $u, v \in \Sigma_\mathsf{I} \cup X$. Thus, the matrix is an ABox built over the extended signature $\Sigma \cup X$, but cannot contain complex concept descriptions. An *object* of $\exists X.\mathcal{A}$ is either an individual name in Σ_I or a variable in X. We denote the set of objects of $\exists X.\mathcal{A}$ with $\mathsf{Obj}(\exists X.\mathcal{A})$. The interpretation $\mathcal{I}$ is a *model of a qABox* $\exists X.\mathcal{A}$ ($\mathcal{I} \models \exists X.\mathcal{A}$) if there is a variable assignment $\mathcal{Z} : X \to \Delta^\mathcal{I}$ such that the augmented interpretation $\mathcal{I}[\mathcal{Z}]$ that additionally maps each variable x to $\mathcal{Z}(x)$ is a model of the matrix $\mathcal{A}$, i.e., $u^{\mathcal{I}[\mathcal{Z}]} \in A^\mathcal{I}$ for each $A(u) \in \mathcal{A}$ and $(u^{\mathcal{I}[\mathcal{Z}]}, v^{\mathcal{I}[\mathcal{Z}]}) \in r^\mathcal{I}$ for each $r(u, v) \in \mathcal{A}$. As pointed out in [13], qABoxes are syntactic variants of Boolean conjunctive queries [18], i.e., CQs with an empty tuple of answer variables. In addition, $\mathcal{EL}$ ABoxes can be expressed by qABoxes.

2.2 Queries and Entailment Relations

Let α, β be any of the syntactical objects (CI, assertion, ABox, qABox) introduced above, and $\mathcal{T}$ be an $\mathcal{EL}$ TBox. Then, α *entails* β *w.r.t.* $\mathcal{T}$ ($\alpha \models^{\mathcal{T}} \beta$) if each model of α and $\mathcal{T}$ is also a model of β. If $\exists X.\mathcal{A} \models^{\mathcal{T}} C(a)$, then a is called an *instance* of C in $\exists X.\mathcal{A}$ w.r.t. $\mathcal{T}$. In case $\mathcal{T} = \emptyset$, we will sometimes write $\models$ instead of $\models^{\emptyset}$. If $\emptyset \models^{\mathcal{T}} C \sqsubseteq D$, then we also write $C \sqsubseteq^{\mathcal{T}} D$ and say that C *is subsumed by* D *w.r.t.* $\mathcal{T}$; in case $\mathcal{T} = \emptyset$ we simply say that C is subsumed by D. The subsumption and the instance problems in $\mathcal{EL}$ are decidable in polynomial time [1]. With respect to the empty TBox, the instance problem can be characterized as follows:

Lemma 1. [13] *The following statements hold for each qABox $\exists X.\mathcal{A}$:*

1. $\exists X.\mathcal{A} \models C(a)$ *iff* $\mathcal{A} \models C(a)$ *for each concept assertion* $C(a)$.
2. *For each $\mathcal{EL}$ concept C and each object u of $\exists X.\mathcal{A}$, we have $\mathcal{A} \models C(u)$ iff $A(u) \in \mathcal{A}$ for each $A \in \mathsf{Conj}(C)$, and for each existential restriction $\exists r.D \in \mathsf{Conj}(C)$, there is some role assertion $r(u, v) \in \mathcal{A}$ such that $\mathcal{A} \models D(v)$.*

Tractability also holds (w.r.t. an $\mathcal{EL}$ TBox) for entailment between $\mathcal{EL}$ ABoxes and entailment of a concept assertion by a qABox. A role assertion between individuals is entailed by a qABox iff it is contained in its matrix. The entailment problem between qABoxes (with or without $\mathcal{EL}$ TBox) is NP-complete [5,13].

For our purposes, a query language QL is a set of Boolean conjunctive queries. The qABox $\exists X.\mathcal{A}$ QL-entails the qABox $\exists Y.\mathcal{B}$ w.r.t. the TBox $\mathcal{T}$ ($\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{QL}} \exists Y.\mathcal{B}$) if, for each query $\alpha \in \mathsf{QL}$, $\exists Y.\mathcal{B} \models^{\mathcal{T}} \alpha$ implies $\exists X.\mathcal{A} \models^{\mathcal{T}} \alpha$. In our previous work, we have considered three query languages: IQ consists of all $\mathcal{EL}$ concept assertions [13,5,10], IRQ extends IQ by all role assertions between individuals [7,11], and CQ consists of all Boolean conjunctive queries (equivalently: qABoxes) [13,5,9]. As shown in [5], CQ-entailment and model-based entailment coincide, and thus deciding $\models^{\mathcal{T}}_{\mathsf{CQ}}$ is also NP-complete.

In contrast, IQ-entailment between qABoxes can be decided in polynomial time. This is a consequence of the following result from [5]. Given a qABox $\exists X.\mathcal{A}$ and a TBox $\mathcal{T}$, one can compute the IQ-saturation $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A})$ of $\exists X.\mathcal{A}$ w.r.t. $\mathcal{T}$ in polynomial time, and this saturation satisfies $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{IQ}} \exists Y.\mathcal{B}$ iff $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}) \models_{\mathsf{IQ}} \exists Y.\mathcal{B}$. Basically, the IQ-saturation process works as follows: while there is an object u and a CI $C \sqsubseteq D$ in $\mathcal{T}$ such that the matrix of the current qABox entails $C(u)$, but not $D(u)$, the qABox is extended by adding $D(u)$ and then representing this assertion as a qABox. To ensure termination, new variables introduced to express existential restrictions are re-used, i.e., for every concept E occurring in an existential restriction, a single variable x_E is introduced (see [5] for details).

Example 2. Starting with the qABox $\exists\emptyset.\{A(a)\}$ and the TBox $\{A \sqsubseteq \exists r.A \sqcap \exists s.(B \sqcap \exists s.A)\}$, we obtain the IQ-saturation $\exists\{x_A, x_{B\sqcap\exists s.A}\}.\{A(a),\ r(a, x_A),\ A(x_A),\ s(a, x_{B\sqcap\exists s.A}),\ B(x_{B\sqcap\exists s.A}),\ s(x_{B\sqcap\exists s.A}, x_A),\ r(x_A, x_A),\ s(x_A, x_{B\sqcap\exists s.A})\}$.

The equivalence $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{IQ}} \exists Y.\mathcal{B}$ iff $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}) \models_{\mathsf{IQ}} \exists Y.\mathcal{B}$ provides us with a polynomial-time reduction of IQ-entailment w.r.t. a TBox to IQ-entailment w.r.t. the empty TBox. The latter in turn corresponds to the existence of a simulation in the other direction [13]. A *simulation* from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$ is a relation $\mathfrak{S} \subseteq \mathsf{Obj}(\exists Y.\mathcal{B}) \times \mathsf{Obj}(\exists X.\mathcal{A})$ such that:

(S1) If a is an individual name, then $(a, a) \in \mathfrak{S}$.
(S2) If $(u, u') \in \mathfrak{S}$ and $A(u) \in \mathcal{B}$, then $A(u') \in \mathcal{A}$.
(S3) If $(u, u') \in \mathfrak{S}$ and $r(u, v) \in \mathcal{B}$, then $(v, v') \in \mathfrak{S}$ and $r(u', v') \in \mathcal{A}$ for some v'.

Since existence of a simulation can be decided in polynomial time [19], IQ-entailment between qABoxes (with or without TBox) is also in P.

This implies that IRQ-entailment is also tractable. In fact, as shown in [11], $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{IRQ}} \exists Y.\mathcal{B}$ iff $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{IQ}} \exists Y.\mathcal{B}$ and $r(a, b) \in \mathcal{B}$ implies $r(a, b) \in \mathcal{A}$ for all $r \in \Sigma_{\mathsf{R}}$ and $a, b \in \Sigma_{\mathsf{I}}$. The second condition can clearly be checked in P.

2.3 Optimal Repairs of Quantified ABoxes w.r.t. Static $\mathcal{EL}$ TBoxes

Assume that we have a qABox (and possibly an $\mathcal{EL}$ TBox) and use the query language QL to extract information from this KB. Usually, one notices that there is something wrong with the given KB if queries are entailed that do not hold in the application domain. Thus, we specify what is to be repaired by a finite set of queries $\mathcal{P}$, which we call a *repair request*, i.e., a repair request is a finite set $\mathcal{P} \subseteq \mathsf{QL}$. As pointed out in the introduction, when defining the notion of an (optimal) repair, it is sufficient to use as entailment relation the one induced by the employed query language.

Definition 3. *Let* QL *be a query language,* $\exists X.\mathcal{A}$ *be a qABox,* $\mathcal{T}$ *be an* $\mathcal{EL}$ *TBox, and* $\mathcal{P}$ *a repair request for* QL*. Then, a* QL-repair *of* $\exists X.\mathcal{A}$ *for* $\mathcal{P}$ *w.r.t.* $\mathcal{T}$ *is a qABox* $\exists Y.\mathcal{B}$ *such that*

(Rep1) $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{QL}} \exists Y.\mathcal{B}$
(Rep2) $\exists Y.\mathcal{B} \not\models^{\mathcal{T}} \alpha$ *for each* $\alpha \in \mathcal{P}$.

This repair is optimal *if it is not strictly* QL*-entailed by another repair. The set* $\mathfrak{R}$ *of* QL*-repairs of* $\exists X.\mathcal{A}$ *for* $\mathcal{P}$ *w.r.t.* $\mathcal{T}$ QL-covers *all* QL*-repairs of* $\exists X.\mathcal{A}$ *for* $\mathcal{P}$ *w.r.t.* $\mathcal{T}$ *if, for each* QL*-repair* $\exists Y.\mathcal{B}$ *of* $\exists X.\mathcal{A}$ *for* $\mathcal{P}$ *w.r.t.* $\mathcal{T}$*, there is* $\exists Z.\mathcal{C} \in \mathfrak{R}$ *such that* $\exists Z.\mathcal{C} \models^{\mathcal{T}}_{\mathsf{QL}} \exists Y.\mathcal{B}$*.*

In our previous work on optimal repairs we have determined situations in which the set of optimal QL-repairs can effectively by computed, and covers all repairs: (1) $\mathsf{QL} = \mathsf{IQ}$ and arbitrary $\mathcal{EL}$ TBox [5]; (2) $\mathsf{QL} = \mathsf{IRQ}$ and arbitrary $\mathcal{EL}$ TBox [7,11]; (3) $\mathsf{QL} = \mathsf{CQ}$, but $\mathcal{P} \subseteq \mathsf{IQ}$ and $\mathcal{T}$ is cycle-restricted [5]; (4) $\mathsf{QL} = \mathsf{CQ}$, but $\mathcal{P} \subseteq \mathsf{gloIRQ}$ and $\mathcal{T}$ is a terminating Horn-$\mathcal{ALCROI}$ TBox [9]. The computation of all optimal QL-repairs can be performed in exponential time for situations (1) and (2), and may additionally require an NP-oracle for (3). For situation (4), the exact complexity of the computation problem has not been determined yet.

3 Global Instance Queries

As mentioned in the introduction, in the presence of CIs of the form $C \sqsubseteq \bot$, we need to express, in the repair request, that concept C must not have any element. This is possible using *global instance queries*, which are of the form $\exists\{x\}.\{C(x)\}$ where C is an $\mathcal{EL}$ concept. Such a query is satisfied in an interpretation $\mathcal{I}$ if $C^{\mathcal{I}} \neq \emptyset$. The query language gloIRQ is obtained from IRQ by adding all global IQs. Since $\mathcal{EL}$ concepts can be expressed by (tree-shaped) conjunctive queries, the inclusions $\mathsf{IRQ} \subseteq \mathsf{gloIRQ} \subseteq \mathsf{CQ}$ hold, which imply the inverse inclusions between the induced entailment relations: $\models^{\mathcal{T}}_{\mathsf{CQ}} \subseteq \models^{\mathcal{T}}_{\mathsf{gloIRQ}} \subseteq \models^{\mathcal{T}}_{\mathsf{IRQ}}$. It is easy to see that these inclusions are strict.

In this section, we give a characterization of gloIRQ-entailment that is based on the existence of certain simulations, and which implies that $\models^{\mathcal{T}}_{\mathsf{gloIRQ}}$ is decidable in polynomial time. We call a simulation $\mathfrak{S}$ from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$ *total* if, for each object $u \in \mathsf{Obj}(\exists Y.\mathcal{B})$, there is an object $v \in \mathsf{Obj}(\exists X.\mathcal{A})$ with $(u, v) \in \mathfrak{S}$.

Lemma 4. *$\exists X.\mathcal{A} \models_{\mathsf{gloIRQ}} \exists Y.\mathcal{B}$ iff there is a total simulation from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$ and $r(a, b) \in \mathcal{B}$ implies $r(a, b) \in \mathcal{A}$ for all $r \in \Sigma_{\mathsf{R}}$ and $a, b \in \Sigma_{\mathsf{I}}$.*

Proof. Recall from [13,5,7] that $\exists X.\mathcal{A} \models_{\mathsf{IRQ}} \exists Y.\mathcal{B}$ iff there is a simulation from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$ and $r(a, b) \in \mathcal{B}$ implies $r(a, b) \in \mathcal{A}$. Thus, it remains to relate the entailed global IQs with totality of the simulation.

For the only-if direction, we assume that $\exists X.\mathcal{A} \models_{\mathsf{gloIRQ}} \exists Y.\mathcal{B}$. The proof of Proposition 23 in [13] shows (under the weaker assumption $\exists X.\mathcal{A} \models_{\mathsf{IQ}} \exists Y.\mathcal{B}$) that $\mathfrak{S} := \{\ (u, v) \mid \mathcal{B} \models C(u) \text{ implies } \mathcal{A} \models C(v) \text{ for each } \mathcal{EL} \text{ concept } C\ \}$ is a simulation. *Assume that $\mathfrak{S}$ is not total.* Then there exists an object $u \in \mathsf{Obj}(\exists Y.\mathcal{B})$ for which there is no object $v \in \mathsf{Obj}(\exists X.\mathcal{A})$ with $(u, v) \in \mathfrak{S}$. Thus, for each $v \in \mathsf{Obj}(\exists X.\mathcal{A})$, there must be an $\mathcal{EL}$ concept C_v with $\mathcal{B} \models C_v(u)$, but $\mathcal{A} \not\models C_v(v)$. Let C be the conjunction of these finitely many concepts. Then $\exists Y.\mathcal{B} \models \exists\{x\}.\{C(x)\}$, and thus $\exists X.\mathcal{A} \models_{\mathsf{gloIRQ}} \exists Y.\mathcal{B}$ yields $\exists X.\mathcal{A} \models \exists\{x\}.\{C(x)\}$. By Proposition 2 in [13], this implies that there is a homomorphism[4] h from $\exists\{x\}.\{C(x)\}$ (expressed as a qABox) to $\exists X.\mathcal{A}$ such that $\mathcal{A} \models C(h(x))$. However, since $C_{h(x)}$ is a conjunct of C, this contradicts the fact that the concepts C_v satisfy $\mathcal{A} \not\models C_v(v)$. Thus, the simulation $\mathfrak{S}$ must be total.

To show the if direction, consider a total simulation $\mathfrak{S}$ from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$, and assume that $\exists Y.\mathcal{B} \models \exists\{x\}.\{C(x)\}$. According to Proposition 2 in [13], there is a homomorphism h from $\exists\{x\}.\{C(x)\}$ (expressed as a qABox) to $\exists Y.\mathcal{B}$. It is easy to construct, by induction on the structure of C, a homomorphism g from $\exists\{x\}.\{C(x)\}$ to $\exists X.\mathcal{A}$: we choose each value $g(u)$ from the non-empty (!) set $\{\ v \mid (h(u), v) \in \mathfrak{S}\ \}$. Another application of Proposition 2 in [13] yields $\exists X.\mathcal{A} \models \exists\{x\}.\{C(x)\}$. □

The proof of the following lemma is similar to the proof of Proposition IV in [6].

Lemma 5. $\exists X.\mathcal{A} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$ *iff* $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}) \models \exists\{x\}.\{C(x)\}$

[4] A homomorphism is a total and functional simulation.

Proof. We can proceed similarly to the proof of Proposition IV in [6]. Regarding the only-if direction, assume that $\exists X.\mathcal{A} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$. The canonical model of the saturation, denoted as $\mathcal{J} := \mathcal{I}_{\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A})}$, is a model of $\exists X.\mathcal{A}$ and $\mathcal{T}$ and thus satisfies the global IQ $\exists\{x\}.\{C(x)\}$. We conclude by induction on C that $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A})$ entails $\exists\{x\}.\{C(x)\}$.

To show the if direction, consider a model $\mathcal{I}$ of $\exists X.\mathcal{A}$ and $\mathcal{T}$. In [6], we consider a sequence of qABoxes $\exists X_i.\mathcal{A}_i$ obtained by constructing the saturation $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A})$ step by step, and inductively define simulations $\mathfrak{S}_i$ from the canonical models $\mathcal{I}_{\exists X_i.\mathcal{A}_i}$ to $\mathcal{I}$. The key observation for the present proof is that each of these simulations $\mathfrak{S}_i$ is total, and thus the last one is a total simulation from $\mathcal{J}$ to $\mathcal{I}$. Now assume that $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}) \models \exists\{x\}.\{C(x)\}$. We infer that the canonical model $\mathcal{J}$ satisfies $\exists\{x\}.\{C(x)\}$, i.e., there is an element $u \in C^{\mathcal{J}}$. The total simulation relates u to an element v of $\mathcal{I}$. By induction on C, we can show that this implies $v \in C^{\mathcal{I}}$, i.e., also $\mathcal{I}$ satisfies $\exists\{x\}.\{C(x)\}$. Since this holds for all models $\mathcal{I}$ of $\exists X.\mathcal{A}$ and $\mathcal{T}$, we conclude that $\exists X.\mathcal{A} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$. □

Before we can show the main result of this section, we need one more technical lemma.

Lemma 6. *$\exists X.\mathcal{A} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$ iff there exists a global IQ $\exists\{y\}.\{D(y)\}$ with $\exists X.\mathcal{A} \models \exists\{y\}.\{D(y)\}$ and $\exists\{y\}.\{D(y)\} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$.*

Proof. Since the if direction is trivial, we turn our attention to the only-if direction. First note that each global IQ $\exists\{x\}.\{C(x)\}$ is equivalent to the IQ $(\exists u.C)(a)$, where u denotes the universal role with semantics $u^{\mathcal{I}} := \mathsf{Dom}(\mathcal{I}) \times \mathsf{Dom}(\mathcal{I})$ in every interpretation $\mathcal{I}$ [23], and a is an arbitrary individual name. Thus, the assumption $\exists X.\mathcal{A} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$ yields that $\exists X.\mathcal{A} \models^{\mathcal{T}} (\exists u.C)(a)$. According to Statement 2 of Lemma 22 in [23], there is an $\mathcal{EL}$ concept D such that one of the following two statements holds:

- $\exists X.\mathcal{A} \models^{\mathcal{T}} D(a)$ and $D \sqsubseteq^{\mathcal{T}} \exists u.C$
- $\exists X.\mathcal{A} \models^{\mathcal{T}} (\exists u.D)(a)$ and $\exists u.D \sqsubseteq^{\mathcal{T}} \exists u.C$.

Clearly, the first statement implies the second, and the second statement directly yields the claim. □

Proposition 7. *Let $\exists X.\mathcal{A}$ and $\exists Y.\mathcal{B}$ be qABoxes and $\mathcal{T}$ be an $\mathcal{EL}$ TBox. Then, $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{gloIRQ}} \exists Y.\mathcal{B}$ iff $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}) \models_{\mathsf{gloIRQ}} \exists Y.\mathcal{B}$. In addition, $\models^{\mathcal{T}}_{\mathsf{gloIRQ}}$ can be decided in polynomial time.*

Proof. To see that the equivalence holds, note that IQs were already treated in the proof of the corresponding result (Theorem 3) in [5], and role assertions in the proof of Proposition 2 in [7]. Thus, it remains to deal with global IQs.

First, assume $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{gloIRQ}} \exists Y.\mathcal{B}$, and let $\exists Y.\mathcal{B} \models \exists\{x\}.\{C(x)\}$. The latter implies $\exists Y.\mathcal{B} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$, and therefore $\exists X.\mathcal{A} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$. Lemma 5 yields that $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}) \models \exists\{x\}.\{C(x)\}$. Since this holds for all global IQs $\exists\{x\}.\{C(x)\}$ entailed by $\exists Y.\mathcal{B}$, $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}) \models_{\mathsf{gloIRQ}} \exists Y.\mathcal{B}$ follows.

Second, let $\mathsf{sat}_{\mathsf{IQ}}^{\mathcal{T}}(\exists X.\mathcal{A}) \models_{\mathsf{gloIRQ}} \exists Y.\mathcal{B}$, and consider a global IQ $\exists\{x\}.\{C(x)\}$ that is entailed by $\exists Y.\mathcal{B}$ w.r.t. $\mathcal{T}$. By Lemma 6, there is a global IQ $\exists\{y\}.\{D(y)\}$ with $\exists Y.\mathcal{B} \models \exists\{y\}.\{D(y)\} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$. Since $\mathsf{sat}_{\mathsf{IQ}}^{\mathcal{T}}(\exists X.\mathcal{A}) \models_{\mathsf{gloIRQ}} \exists Y.\mathcal{B}$, we infer $\mathsf{sat}_{\mathsf{IQ}}^{\mathcal{T}}(\exists X.\mathcal{A}) \models \exists\{y\}.\{D(y)\}$. Lemma 5 yields $\exists X.\mathcal{A} \models^{\mathcal{T}} \exists\{y\}.\{D(y)\}$, and thus $\exists X.\mathcal{A} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$. Since this argumentation works for all global IQs $\exists\{x\}.\{C(x)\}$, we obtain $\exists X.\mathcal{A} \models_{\mathsf{gloIRQ}}^{\mathcal{T}} \exists Y.\mathcal{B}$.

To show the complexity result, we use Lemma 4 together with the equivalence we have just shown. First recall that the IQ-saturation $\mathsf{sat}_{\mathsf{IQ}}^{\mathcal{T}}(\exists X.\mathcal{A})$ can be computed in polynomial time [5]. Second, the (unique) maximal simulation from a qABox to another can also be computed in polynomial time [19]. This maximal simulation is total iff there is a total simulation, i.e., it suffices to check whether the maximal simulation from $\exists Y.\mathcal{B}$ to $\mathsf{sat}_{\mathsf{IQ}}^{\mathcal{T}}(\exists X.\mathcal{A})$ is total, which can be done in polynomial time. Finally, checking containment of the role assertions involving only individual names obviously needs only polynomial time. $\square$

4 Optimal Repairs of qABoxes w.r.t. $\mathcal{EL}^{\perp}$ TBoxes

The description logic $\mathcal{EL}^{\perp}$ extends $\mathcal{EL}$ with the bottom concept $\perp$, which has the semantics $\perp^{\mathcal{I}} := \emptyset$ in every interpretation $\mathcal{I}$. In contrast to $\mathcal{EL}$, a quantified ABox $\exists X.\mathcal{A}$ can be *inconsistent* w.r.t. an $\mathcal{EL}^{\perp}$ TBox $\mathcal{T}$, which means that there is no model of the qABox and the TBox. Since any query is entailed by such an inconsistent qABox, repairing it for unwanted consequences also encompasses resolving the inconsistency. This problem was tackled in [9] for the more expressive DL $\mathcal{ELROI}(\perp)$ and the query language CQ. Here, we restrict the attention to $\mathcal{EL}^{\perp}$, but consider a smaller query language, which has the advantage the TBoxes need not be restricted to being cycle-restricted and repairs can be computed more efficiently.

4.1 Repairing the Inconsistency

For a qABox $\exists X.\mathcal{A}$ that is inconsistent w.r.t. the $\mathcal{EL}^{\perp}$ TBox $\mathcal{T}$, Condition (Rep1) in the definition of repairs is vacuously true, and thus does not enforce the repair to be related in any way to $\exists X.\mathcal{A}$. In [9], this problem is addressed for $\mathcal{ELROI}(\perp)$ TBoxes, according to which any $\mathcal{EL}^{\perp}$ TBox $\mathcal{T}$ can be normalized such that it is the union of a *positive part* $\mathcal{T}_{+}$ consisting of $\mathcal{EL}$ CIs and an *unsatisfiable part* $\mathcal{T}_{\perp}$ consisting of CIs of the form $C \sqsubseteq \perp$ for $\mathcal{EL}$ concepts C. This separation allows us to characterize inconsistency of qABoxes w.r.t. $\mathcal{EL}^{\perp}$ TBoxes as follows.

Lemma 8. [9, Proposition 17] *Let $\exists X.\mathcal{A}$ be a qABox and $\mathcal{T}$ be an $\mathcal{EL}^{\perp}$ TBox.*

1. *$\exists X.\mathcal{A}$ is inconsistent w.r.t. $\mathcal{T}$ iff there is a CI $C \sqsubseteq \perp \in \mathcal{T}_{\perp}$ with $\exists X.\mathcal{A} \models^{\mathcal{T}_{+}} \exists\{x\}.\{C(x)\}$.*
2. *If $\exists X.\mathcal{A}$ is consistent w.r.t. $\mathcal{T}$, then $\exists X.\mathcal{A} \models^{\mathcal{T}} \exists Y.\mathcal{B}$ iff $\exists X.\mathcal{A} \models^{\mathcal{T}_{+}} \exists Y.\mathcal{B}$.*

Specifically, Statement 1 tells us that an inconsistency can be resolved by ensuring that the repair does not entail any global IQ $\exists\{x\}.\{C(x)\}$ for which the unsatisfiable part $\mathcal{T}_\bot$ contains $C \sqsubseteq \bot$. Statement 2 in turn states that the unsatisfiable part $\mathcal{T}_\bot$ can be ignored when working with a consistent qABox. Thus, to regain a meaningful connection between the original qABox and the repair, $\mathcal{T}_+$ rather than $\mathcal{T}$ is used in Condition (IRep1) below.

Definition 9. [9, Definition 18] *Let* QL *be a query language,* $\exists X.\mathcal{A}$ *be a qABox,* $\mathcal{T}$ *be an* $\mathcal{EL}^\bot$ *TBox, and* $\mathcal{P} \subseteq \mathsf{QL}$ *a repair request. An* inconsistency QL-repair *of* $\exists X.\mathcal{A}$ *for* $\mathcal{P}$ *w.r.t.* $\mathcal{T}$ *is a qABox* $\exists Y.\mathcal{B}$ *such that*

(IRep1) $\exists X.\mathcal{A} \models^{\mathcal{T}_+}_{\mathsf{QL}} \exists Y.\mathcal{B}$,
(IRep2) $\exists Y.\mathcal{B}$ *is consistent w.r.t.* $\mathcal{T}$,
(IRep3) $\exists Y.\mathcal{B} \not\models^{\mathcal{T}} \alpha$ *for each* $\alpha \in \mathcal{P}$.

We say that $\exists Y.\mathcal{B}$ *is* optimal *if it is not strictly* QL*-entailed by another repair.*

In [9], such repairs were investigated for the query language CQ. Here, we are mainly interested in IRQ-repairs, but since resolving the inconsistency requires us to consider global instance queries as well, we use gloIRQ-entailment and also allow the user to include global instance queries in the repair request. The following proposition shows that inconsistency gloIRQ-repairs for $\mathcal{P}$ w.r.t. $\mathcal{T}$ correspond to gloIRQ-repairs for $\mathcal{P}^{\mathcal{T}_\bot}$ w.r.t. $\mathcal{T}_+$, where $\mathcal{P}^{\mathcal{T}_\bot}$ extends $\mathcal{P}$ with the global IQs $\exists\{x\}.\{C(x)\}$ for all CIs $C \sqsubseteq \bot$ in $\mathcal{T}_\bot$. It can be proved by adapting the proof of Theorem 19 in [9].

Proposition 10. *Let* $\exists X.\mathcal{A}$ *be a qABox,* $\mathcal{T}$ *be an* $\mathcal{EL}^\bot$ *TBox, and* $\mathcal{P}$ *be a repair request. If* $\mathcal{T}$ *is inconsistent, then there are no inconsistency* gloIRQ*-repairs of* $\exists X.\mathcal{A}$ *for* $\mathcal{P}$ *w.r.t.* $\mathcal{T}$*. Otherwise, the set of all (optimal) inconsistency* gloIRQ*-repairs of* $\exists X.\mathcal{A}$ *for* $\mathcal{P}$ *w.r.t.* $\mathcal{T}$ *coincides with the set of all (optimal)* gloIRQ*-repairs of* $\exists X.\mathcal{A}$ *for* $\mathcal{P}^{\mathcal{T}_\bot}$ *w.r.t.* $\mathcal{T}_+$*.*

For the case where $\mathcal{T}$ is consistent, we compute optimal inconsistency gloIRQ-repairs w.r.t. $\mathcal{T}$ in three stages. To repair the inconsistency of the qABox, we replace $\mathcal{T}$ by the $\mathcal{EL}$ TBox $\mathcal{T}_+$ and $\mathcal{P}$ by the extended repair request $\mathcal{P}^{\mathcal{T}_\bot}$ as described in Proposition 10. Then, we repair for the unwanted role assertions using Proposition 11 below. The main remaining task is then to repair for the unwanted concept assertions and global IQs using gloIRQ-entailment (see Section 4.3).

4.2 Repairing for Role Assertions

We assume that $\mathcal{T}$ is an $\mathcal{EL}$ TBox and $\mathcal{P} \subseteq \mathsf{gloIRQ}$ is a repair request. We denote the set of the role assertions from $\mathcal{P}$ as $\mathcal{P}_{\mathsf{R}}$ and the set of remaining elements as $\mathcal{P}_{\mathsf{C}}$. Given that a role assertion between individuals follows from a qABox iff it is contained in its matrix, one might think that one can deal with role assertions in the repair request by simply removing them. However, this way one may also lose concept assertions involving existential restrictions that could have been retained (see Example 3.3 in [11]). Instead, before removing the role assertions of $\mathcal{P}_\mathsf{R}$, one needs to add a variable x_a as copy of every individual a to the qABox. This copy belongs to the same concept and role assertions as a (see the construction in the proof of Lemma 3.4 in [11] for details). The following proposition can be shown by adapting the proof of Theorem 3.6 in [11] to deal with gloIRQ-repairs rather than IRQ-repairs.

Proposition 11. *Let $\exists X.\mathcal{A}$ be a qABox, $\mathcal{T}$ an $\mathcal{EL}$ TBox, and $\mathcal{P}$ a repair request. Consider the qABox $\exists Z.\mathcal{C}$ constructed from $\exists X.\mathcal{A}$ by first copying all individual names into fresh variables and then removing all role assertions of $\mathcal{P}_\mathsf{R}$, i.e.,*

$$\begin{aligned} Z &\coloneqq X \cup \{\, x_a \mid a \in \Sigma_\mathsf{I} \,\} \\ \mathcal{C} &\coloneqq (\mathcal{A} \cup \{\, A(x_a) \mid A(a) \in \mathcal{A} \,\} \cup \{\, r(x_a, u) \mid r(a, u) \in \mathcal{A} \,\} \\ &\qquad\qquad\qquad\qquad\qquad\qquad \cup \{\, r(u, x_a) \mid r(u, a) \in \mathcal{A} \,\}) \setminus \mathcal{P}_\mathsf{R}. \end{aligned}$$

The set of all (optimal) gloIRQ-repairs of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$ coincides with the set of all (optimal) gloIRQ-repairs of $\exists Z.\mathcal{C}$ for $\mathcal{P}_\mathsf{C}$ w.r.t. $\mathcal{T}$.

Proof. As shown in [11, Theorem 3.6], the qABox $\exists Z.\mathcal{C}$ is an optimal IRQ-repair of $\exists X.\mathcal{A}$ for $\mathcal{P}_\mathsf{R}$ w.r.t. $\mathcal{T}$, and it IRQ-entails every other IRQ-repair. The same proof also works for gloIRQ-entailment since, in the proof of the crucial Lemma 3.4 in [11], the constructed simulation is total, and thus $\exists Z.\mathcal{C}$ is gloIQ-equivalent to $\exists X.\mathcal{A}$, where gloIQ denotes the query language consisting of all IQs and global IQs.

It is easy to see that any gloIRQ-repair of $\exists Z.\mathcal{C}$ for $\mathcal{P}_\mathsf{C}$ w.r.t. $\mathcal{T}$ is also a gloIRQ-repair of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$. In fact, $\exists Z.\mathcal{C}$ does not entail the role assertions from $\mathcal{P}_\mathsf{R}$, and its gloIRQ-repair for $\mathcal{P}_\mathsf{C}$ keeps this property and no longer entails the the elements of $\mathcal{P}_\mathsf{C}$. Conversely, assume that $\exists V.\mathcal{D}$ is a gloIRQ-repair of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$. Since $\exists V.\mathcal{D}$ is also a gloIRQ-repair for $\mathcal{P}_\mathsf{R}$ and $\exists Z.\mathcal{C}$ covers all these repairs, we obtain $\exists Z.\mathcal{C} \models^{\mathcal{T}}_{\mathsf{gloIRQ}} \exists V.\mathcal{D}$. Furthermore, since $\exists V.\mathcal{D}$ does not entail any element from $\mathcal{P}$ w.r.t. $\mathcal{T}$, we conclude that $\exists V.\mathcal{D}$ is a gloIRQ-repair of $\exists Z.\mathcal{C}$ for $\mathcal{P}_\mathsf{C}$ w.r.t. $\mathcal{T}$.

Since the sets of repairs are equal and the same entailment relation $\models^{\mathcal{T}}_{\mathsf{gloIRQ}}$ is employed, the respective sets of optimal repairs also coincide. □

When going to the last stage, we can thus assume that we are given a qABox, an $\mathcal{EL}$ TBox, and a repair request that consists of unwanted concept assertions and global IQs. However, we still need to use gloIRQ-entailment rather than gloIQ-entailment to avoid reintroducing role assertions that have been removed.

4.3 Repairing for Concept Assertions and Global IQs

The main tool used in [5] to compute all optimal repairs is the construction of canonical repairs. It is shown that the set of canonical repairs covers all repairs, and thus the optimal ones can be obtained by removing elements strictly entailed by others. We recall the construction of canonical repairs from [5] and explain the modifications that are necessary to treat global IQs. We denote by $\mathsf{Sub}(\mathcal{T},\mathcal{P})$ and $\mathsf{Atoms}(\mathcal{T},\mathcal{P})$ the set of subconcepts and atoms occurring in the TBox $\mathcal{T}$ and the repair request $\mathcal{P}$. The objects of the canonical repairs are copies of the objects of the saturation $\exists X'.\mathcal{A}' := \mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A})$ of the input qABox $\exists X.\mathcal{A}$. These copies are of the form $\langle\!\langle u,\mathcal{K}\rangle\!\rangle$, where u is an object of $\exists X'.\mathcal{A}'$ and $\mathcal{K}$ is a repair type (see below). Intuitively, $C \in \mathcal{K}$ says that, in the canonical repair, the object $\langle\!\langle u,\mathcal{K}\rangle\!\rangle$ is not an instance of C. Formally, a *repair type* for an object u of $\exists X'.\mathcal{A}'$ is a set of $\mathcal{K} \subseteq \mathsf{Atoms}(\mathcal{T},\mathcal{P})$ such that

(RT1) The object u is an instance of all atoms in $\mathcal{K}$, i.e., the matrix $\mathcal{A}'$ of the saturation entails $C(u)$ for each atom $C \in \mathcal{K}$.
(RT2) The atoms in $\mathcal{K}$ are pairwise subsumption-incomparable, i.e., $C \not\sqsubseteq^{\emptyset} D$ for distinct atoms $C, D \in \mathcal{K}$.
(RT3) If C is an atom in $\mathcal{K}$ and E is a subconcept in $\mathsf{Sub}(\mathcal{T},\mathcal{P})$ with $E \sqsubseteq^{\mathcal{T}} C$ and $\mathcal{A}' \models E(u)$, then there is an atom D in $\mathcal{K}$ with $E \sqsubseteq^{\emptyset} D$.

(RT1) is motivated by the fact that instance relationships that do not hold need not be removed. (RT2) avoids redundancies in $\mathcal{K}$ since having D in $\mathcal{K}$ for $C \sqsubseteq^{\emptyset} D$ also prevents $\langle\!\langle u,\mathcal{K}\rangle\!\rangle$ from being an instance of C. (RT3) ensures that inference with the TBox cannot restore instance relationships for atoms in $\mathcal{K}$.

Given sets $\mathcal{K}$ and $\mathcal{L}$ of $\mathcal{EL}$ concepts (e.g. repair types), we say that $\mathcal{K}$ is *covered by* $\mathcal{L}$ and write $\mathcal{K} \leq \mathcal{L}$ if, for each $C \in \mathcal{K}$, there is $D \in \mathcal{L}$ with $C \sqsubseteq^{\emptyset} D$. The matrix $\mathcal{B}$ of a canonical repair is defined in a way that ensures that indeed each object $\langle\!\langle u,\mathcal{K}\rangle\!\rangle$ is not an instance of any atom in $\mathcal{K}$ for this repair.

(CR1) $\mathcal{B}$ contains all concept assertions $A(\langle\!\langle u,\mathcal{K}\rangle\!\rangle)$ with $A(u) \in \mathcal{A}'$ and $A \notin \mathcal{K}$.
(CR2) $\mathcal{B}$ contains all role assertions $r(\langle\!\langle u,\mathcal{K}\rangle\!\rangle, \langle\!\langle v,\mathcal{L}\rangle\!\rangle)$ with $r(u,v) \in \mathcal{A}'$ and $\mathsf{Succ}(\mathcal{K},r,u) \leq \mathcal{L}$, where $\mathsf{Succ}(\mathcal{K},r,u) := \{\, C \mid \exists r.C \in \mathcal{K} \text{ and } \mathcal{A}' \models C(v) \,\}$.

Finally, for each individual a we select one of its copies $\langle\!\langle a,\mathcal{K}\rangle\!\rangle$ as representation of a in the repair. However, to obtain a genuine repair, the type $\mathcal{K}$ must satisfy an additional condition. Formally, this selection is made by a *repair seed* $\mathcal{S}$, which maps each individual name a to a repair type $\mathcal{S}_a$ for a such that:

(RS) If $C(a)$ is an IQ in $\mathcal{P}$ with $\mathcal{A}' \models C(a)$, then there is an atom D in $\mathcal{S}_a$ such that $C \sqsubseteq^{\emptyset} D$.

Given a repair seed $\mathcal{S}$, the variable set Y consists of all copies $\langle\!\langle u,\mathcal{K}\rangle\!\rangle$ except those of the form $\langle\!\langle a,\mathcal{S}_a\rangle\!\rangle$, which are treated as synonyms of the individual names. We then call $\exists Y.\mathcal{B}$ the *canonical repair* induced by $\mathcal{S}$, and denote it by $\mathsf{rep}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A},\mathcal{S})$.

In [9], global IQs in the repair request were treated in the context of a more expressive DL (also containing inverse roles, role inclusions, and nominals) and for CQ-entailment. Here, we present a simpler treatment, which is also easier to handle algorithmically. The following condition, which is introduced in [9], is a variant of (RS) that deals with global IQs by forbidding their concepts for all objects of the repair, and not just for individual names:

(RT4) If $\exists\{x\}.\{C(x)\}$ is a global IQ in $\mathcal{P}$ and $\mathcal{A}' \models C(u)$, then there is an atom D in $\mathcal{K}$ such that $C \sqsubseteq^{\emptyset} D$.

The following example shows that this condition is not sufficient to repair for global IQs since it does not prevent the TBox from restoring global IQs.

Example 12. Consider the TBox $\{A \sqsubseteq \exists r.B\}$, the ABox $\{A(a)\}$, and the repair request $\{\exists\{x\}.\{B(x)\}\}$. The saturation is $\exists\{y\}.\{A(a), r(a,y), B(y)\}$. With only the above conditions, a repair seed $\mathcal{S}$ could map a to the empty repair type. In the canonical repair induced by this repair seed, the assertion $A(a)$ is not removed, and inference with the TBox will re-introduce an r-successor of a that is an instance of B. Thus, the unwanted global IQ is still entailed.

In [9], this problem is dealt with by introducing a (rather complex) condition on repair seed, called *admissibility*, which also deals with inverse roles, role inclusions, and nominals. Here, we treat it by introducing a new condition on repair types, which is stronger than (RT4) and easier to check than admissibility.

Definition 13. *An $\mathcal{EL}$ concept D is* globally forbidden *w.r.t. $\mathcal{P}$ if it entails a global IQ in $\mathcal{P}$, i.e. if $\exists\{y\}.\{D(y)\} \models^{\mathcal{T}} \exists\{x\}.\{C(x)\}$ for some $\exists\{x\}.\{C(x)\} \in \mathcal{P}$.*

Repairs must not entail $\exists\{y\}.\{D(y)\}$ for any globally forbidden concept D.

(RT5) If C is a subconcept in $\mathsf{Sub}(\mathcal{T},\mathcal{P})$ that is globally forbidden w.r.t. $\mathcal{P}$ and $\mathcal{A}' \models C(u)$, then there is an atom D in $\mathcal{K}$ with $C \sqsubseteq^{\emptyset} D$.

Since $\exists\{x\}.\{C(x)\} \in \mathcal{P}$ obviously implies that C is globally forbidden, (RT5) encompasses (RT4). In the remainder of this section, we will verify that repair types additionally satisfying (RT5) correctly treat the global IQs in $\mathcal{P}$. With respect to this extended definition of repair types, we denote the canonical repair induced by the repair seed $\mathcal{S}$ as $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A},\mathcal{S})$.

Lemma 14. *Let $\mathcal{S}$ be a repair seed and consider a subconcept $C \in \mathsf{Sub}(\mathcal{T},\mathcal{P})$. The matrix of $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A},\mathcal{S})$ entails $C(\langle\langle u,\mathcal{K}\rangle\rangle)$ iff the matrix of $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A})$ entails $C(u)$ and no atom in $\mathcal{K}$ subsumes C.*

Proof. The proof is almost the same as for Lemma XII in [5], except that we must extend the last case in the if direction where $C = \exists r.D$ is an existential restriction.

According to Lemma 1, it follows from the preconditions that there exists some object v such that $\mathcal{A}'$ contains $r(u,v)$ and entails $D(v)$. We infer that $D \not\sqsubseteq^{\mathcal{T}} E$ for each $\exists r.E \in \mathcal{K}$ (otherwise $\exists r.D \sqsubseteq^{\mathcal{T}} \exists r.E$, and by (RT3) $\exists r.D$ would be subsumed by an atom in $\mathcal{K}$, a contradiction). Thus for each $\exists r.E \in \mathcal{K}$, there is some atom $F_E \in \mathsf{Conj}(E)$ such that $D \not\sqsubseteq^{\mathcal{T}} F_E$.

We also need to take care of Conditions (RT4) and (RT5). Therefore let $G \in \mathsf{Sub}(\mathcal{T},\mathcal{P})$ be globally forbidden and $\mathcal{A}' \models G(u)$. Since no atom in $\mathcal{K}$ subsumes $\exists r.D$ and $\mathcal{K}$ satisfies (RT5), it follows that $\exists r.D$ is not globally forbidden. So D is not globally forbidden either and thus $D \not\sqsubseteq^{\mathcal{T}} G$, i.e., there is an atom $F_G \in \mathsf{Conj}(G)$ such that $D \not\sqsubseteq^{\mathcal{T}} F_G$.

According to Lemma XI in [5] there is a repair type $\mathcal{L}$ for v that covers the set $\mathcal{L}_0 := \mathsf{Max}_{\sqsubseteq^\emptyset}(\{\, F_E \mid \exists r.E \in \mathcal{K} \text{ and } \mathcal{A}' \models E(v)\,\} \cup \{\, F_G \mid G \text{ is globally forbidden and } \mathcal{A}' \models G(v)\,\})$ and that does not contain an atom subsuming D. Although Lemma XI only guarantees that $\mathcal{L}$ is a repair type as per the old definition, i.e., satisfies only (RT1), (RT2), and (RT3), our construction of $\mathcal{L}_0$ guarantees that also (RT4) and (RT5) are fulfilled since $\mathcal{L}$ covers $\mathcal{L}_0$.

Since $D \in \mathsf{Sub}(\mathcal{R}, \mathcal{T})$, we can apply the induction hypothesis and obtain that $\mathcal{B} \models D(y_{v,\mathcal{L}})$. By the very construction of $\mathcal{L}$, it follows that the matrix $\mathcal{B}$ contains the role assertion $r(y_{u,\mathcal{K}}, y_{v,\mathcal{L}})$. Thus, we conclude that $\mathcal{B} \models C(y_{u,\mathcal{K}})$. □

Proposition 15. *Let $\exists X.\mathcal{A}$ be a qABox, $\mathcal{T}$ an $\mathcal{EL}$ TBox, and $\mathcal{P} \subseteq$ gloIQ a repair request. For each repair seed $\mathcal{S}$, the induced canonical repair is a* gloIRQ*-repair of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$. Conversely, every* gloIRQ*-repair of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$ is* gloIRQ*-entailed by a canonical repair.*

Proof. The proof of the first claim is the same as for Proposition 8 in [5], but uses Lemma 14 instead of Lemma XII.

We proceed with the second claim and consider a gloIRQ-repair $\exists Y.\mathcal{B}$. Since it fulfills Condition (Rep1), Proposition 7 yields a total simulation $\mathfrak{S}$ from $\exists Y.\mathcal{B}$ to the saturation $\exists X'.\mathcal{A}' := \mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A})$. Like in the proof of Proposition 8 in [5], we define the mapping $\mathcal{F} \colon \mathfrak{S} \to \wp(\mathsf{Atoms}(\mathcal{T}, \mathcal{P}))$ by

$$\mathcal{F}(t, v) := \mathsf{Max}_{\sqsubseteq^\emptyset}\{\, C \mid C \in \mathsf{Atoms}(\mathcal{T}, \mathcal{P}),\ \mathcal{A}' \models C(v), \text{ and } \mathcal{B} \not\models^{\mathcal{T}} C(t)\,\}$$

for all $(t, v) \in \mathfrak{S}$. In [5] we showed that each set $\mathcal{F}(t, v)$ satisfies Conditions (RT1), (RT2), and (RT3). To conclude that $\mathcal{F}(t, v)$ is a repair type for v, we need to verify that Conditions (RT4) and (RT5) also hold.

(RT4) Let $\exists\{x\}.\{C(x)\}$ be a global IQ in $\mathcal{P}$ with $\mathcal{A}' \models C(v)$. Since $\exists Y.\mathcal{B}$ satisfies Condition (Rep2), it follows that $\mathcal{B} \not\models^{\mathcal{T}} C(t)$. Thus there is a top-level conjunct $D \in \mathsf{Conj}(C)$ with $\mathcal{A}' \models D(v)$ and $\mathcal{B} \not\models^{\mathcal{T}} D(t)$. We conclude that either D itself or an atom subsuming D is contained in $\mathcal{F}(t, v)$.

(RT5) Let $C \in \mathsf{Sub}(\mathcal{T}, \mathcal{P})$ be globally forbidden and $\mathcal{A}' \models C(v)$, i.e., there is a global IQ $\exists\{y\}.\{E(y)\}$ in $\mathcal{P}$ with $\exists\{x\}.\{C(x)\} \models^{\mathcal{T}} \exists\{y\}.\{E(y)\}$. Since $\exists Y.\mathcal{B}$ satisfies Condition (Rep2), it follows that $\mathcal{B} \not\models^{\mathcal{T}} \exists\{y\}.\{E(y)\}$ and thus in particular that $\mathcal{B} \not\models^{\mathcal{T}} C(t)$. Thus there is a top-level conjunct $D \in \mathsf{Conj}(C)$ with $\mathcal{A}' \models D(v)$ and $\mathcal{B} \not\models^{\mathcal{T}} D(t)$. We conclude that either D itself or an atom subsuming D is contained in $\mathcal{F}(t, v)$.

Like in [5] we obtain a repair seed $\mathcal{S}$ by defining $\mathcal{S}_a := \mathcal{F}(a, a)$ for each individual name a, and the relation $\mathfrak{T} := \{\, (t, \langle\!\langle v, \mathcal{F}(t, v)\rangle\!\rangle) \mid (t, v) \in \mathfrak{S}\,\}$ is a simulation from the considered repair $\exists Y.\mathcal{B}$ to the induced repair $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S})$. Furthermore, $\mathfrak{T}$ is total since $\mathfrak{S}$ is total and $\mathcal{F}$ is defined for each pair in $\mathfrak{S}$. The role assertions involving only individual names are treated as in the proof of Proposition 21 in [8]. Proposition 7 yields the claim. □

This proposition shows that the set of canonical repairs is a set of gloIRQ-repairs that covers all gloIRQ-repairs. The definition of canonical repairs implies that there are at most exponentially many such repairs of at most exponential size, which can be computed in exponential time. Up to equivalence, the optimal gloIRQ-repairs can be obtained from this set by removing redundant elements, i.e., ones that are strictly gloIRQ-entailed by other elements. Since gloIRQ-entailment is in P (Proposition 7), we obtain the following complexity result for computing the set of all optimal repairs, and clearly this set still covers all repairs.

Proposition 16. *For each qABox $\exists X.\mathcal{A}$, each $\mathcal{EL}$ TBox $\mathcal{T}$, and each repair request $\mathcal{P} \subseteq \mathsf{gloIQ}$, the set of all optimal gloIRQ-repairs can be computed in exponential time, up to gloIRQ-equivalence, and this set covers all gloIRQ-repairs.*

Putting Propositions 10, 11, and 16 together, yields the main result of this section.

Theorem 17. *Given a qABox $\exists X.\mathcal{A}$, an $\mathcal{EL}^\perp$ TBox $\mathcal{T}$, and a repair request $\mathcal{P} \subseteq \mathsf{gloIRQ}$, the set of all optimal inconsistency gloIRQ-repairs can be computed in exponential time, up to gloIRQ-equivalence, and this set covers all inconsistency gloIRQ-repairs.*

5 Inconsistency- and Error-Tolerant Reasoning

Error-tolerant reasoning does not commit to a single repair, but rather reasons w.r.t. all of them. In [10,11], we have investigated error-tolerant reasoning in a setting where the query language is IRQ and the TBox is written in $\mathcal{EL}$. Here, we extend the obtained results to $\mathcal{EL}^\perp$ TBoxes and the query language gloIRQ. Inconsistency-tolerant reasoning is the special case where the repair request is $\{\perp(a)\}$ for some individual a, which is entailed by a KB iff the KB is inconsistent. We assume in the following that the repair request $\mathcal{P}$ is *solvable*, i.e., has a repair, which is the case if none of the queries in $\mathcal{P}$ is entailed by $\mathcal{T}$ alone.

Definition 18. *Let $\exists X.\mathcal{A}$ be a qABox, $\mathcal{T}$ an $\mathcal{EL}^\perp$ TBox, $\mathcal{P} \subseteq \mathsf{gloIRQ}$ a solvable repair request, and $\alpha \in \mathsf{gloIRQ}$ a query. Then α is* bravely (cautiously) gloIRQ-entailed *by $\exists X.\mathcal{A}$ w.r.t. $\mathcal{P}$ and $\mathcal{T}$ if it is entailed w.r.t. $\mathcal{T}$ by some (all) optimal inconsistency gloIRQ-repair(s) of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$.*

As shown in [10] for a restricted setting, brave entailment can be reduced to classical entailment. To this purpose, let $\exists Z.\mathcal{C}$ be the qABox representation of $\exists\emptyset.\{\alpha\}$ if α is a concept or role assertion, and α itself if it is a global IQ. Since every inconsistency gloIRQ-repair is entailed by an optimal one, α is bravely gloIRQ-entailed by $\exists X.\mathcal{A}$ w.r.t. $\mathcal{T}$ and $\mathcal{P}$ iff $\exists Z.\mathcal{C}$ is an inconsistency gloIRQ-repair of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$. This reduces checking brave gloIRQ-entailment to deciding whether a given qABox satisfies Conditions (IRep1)–(IRep3). Since each condition can verified in polynomial time, we obtain the following result.

Theorem 19. *Brave gloIRQ-entailment is in* P*.*

Dealing with cautious entailment is more involved. Since a given repair problem may have exponentially many optimal repairs of exponential size, the naïve approach to solve cautious entailment, which computes all optimal repairs and checks whether each of them entails the query α, would require exponential time. The approach employed in [10] (for a more restricted setting) to reduce the complexity to coNP proceeds as follows: to check whether α is *not* cautiously entailed, it guesses a mapping $\mathcal{S}$ from individuals a to sets of atoms $\mathcal{S}_a$ and then checks whether

1. $\mathcal{S}$ is a repair seed,
2. the repair seed $\mathcal{S}$ induces an optimal repair,
3. the optimal repair induced by $\mathcal{S}$ does not entail α.

To extend this approach to our setting, we must show that Conditions 1–3 can be checked in polynomial time for the query language gloIRQ and $\mathcal{EL}$ TBoxes. In fact, we have shown in the previous section that the optimal inconsistency gloIRQ-repairs w.r.t. $\mathcal{EL}^\perp$ TBoxes we are interested in can actually be obtained as optimal gloIRQ-repairs w.r.t. $\mathcal{EL}$ TBoxes.

Regarding Condition 1, it is easy to see that (RT1)–(RT3), (RT5), and (RS) can be checked in polynomial time. Note that this shows an advantage of (RT5) over the admissibility condition in [9] since for the latter it is less clear how to test it in P. To deal with Condition 2, we use a pre-order on repair seeds that reflects gloIRQ-entailment between the induced canonical gloIRQ-repairs.

Definition 20. [8,11] *Given repair seeds $\mathcal{S}$ and $\mathcal{R}$, we say that $\mathcal{S}$ is* IRQ-covered *by $\mathcal{R}$ (write $\mathcal{S} \leq_{\mathsf{IRQ}} \mathcal{R}$) if $\mathcal{S}_a \leq \mathcal{R}_a$ for each $a \in \Sigma_\mathsf{I}$, and $\mathsf{Succ}(\mathcal{R}_a, r, b) \leq \mathcal{R}_b$ implies $\mathsf{Succ}(\mathcal{S}_a, r, b) \leq \mathcal{S}_b$ for all $r(a,b) \in \mathcal{A}$ with $a, b \in \Sigma_\mathsf{I}$.*

Lemma 21. *$\mathcal{S} \leq_{\mathsf{IRQ}} \mathcal{R}$ iff $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S}) \models^{\mathcal{T}}_{\mathsf{gloIRQ}} \mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{R})$.*

Proof. With respect to the old definition of repair types (without (RT5)), it was shown in [7] that $\mathcal{S} \leq_{\mathsf{IRQ}} \mathcal{R}$ iff $\mathsf{rep}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}, \mathcal{S}) \models^{\mathcal{T}}_{\mathsf{IRQ}} \mathsf{rep}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}, \mathcal{R})$. The same proof, but using Lemma 14 instead of Lemma XII in [6], shows that $\mathcal{S} \leq_{\mathsf{IRQ}} \mathcal{R}$ iff $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S}) \models^{\mathcal{T}}_{\mathsf{IRQ}} \mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{R})$. Since all can. repairs (for the same input) entail the same global IQs, the latter is actually a gloIRQ-entailment. □

Given any pre-order $\leq$, we write $\alpha < \beta$ if $\alpha \leq \beta$ and $\beta \not\leq \alpha$, and say that α is $\leq$-minimal if there is no β such that $\beta < \alpha$. Lemma 21 implies that, up to gloIRQ-equivalence, the optimal gloIRQ-repairs are exactly the canonical gloIRQ-repairs induced by $\leq_{\mathsf{IRQ}}$-minimal repair seeds. Thus, to decide optimality in polynomial time, it remains to show the following result.

Proposition 22. *$\leq_{\mathsf{IRQ}}$-minimality of repair seeds is in* P.

Proof. With respect to the old definition of repair types, this result follows from Lemma 5.7 in [11], which states that a repair seed $\mathcal{S}$ is not $\leq_{\mathsf{IRQ}}$-minimal iff there exists an individual a and an atom $D \in \mathcal{S}_a$ such that the lowering $\mathsf{low}(\mathcal{S}, D(a))$ is a repair seed. Intuitively, if there is a repair seed $\mathcal{S}'$ that is strictly smaller than

$\mathcal{S}$, then there must by an individual a and an atom $D \in \mathcal{S}_a$ such that $D \notin \mathcal{S}'_a$. However, just removing D from $\mathcal{S}_a$ is not sufficient since such a removal also has an impact on other parts of the repair seed. The (rather intricate) definition of the lowering function (see Definition 5.3 in [11]) takes care of such effects.

The same proof as in [11] also applies w.r.t. the extended definition of repair types, but in the only-if direction we additionally need to verify that each repair type assigned by the lowering $\mathsf{low}(\mathcal{S}, D(a))$ satisfies (RT5).

To see this, assume that $\mathcal{S}$ is not $\leq_{\mathsf{IRQ}}$-minimal, i.e., there is a repair seed $\mathcal{R}$ such that $\mathcal{R} <_{\mathsf{IRQ}} \mathcal{S}$. By Lemma 5.5 in [11], there exists $a \in \Sigma_{\mathsf{I}}$ and $D \in \mathcal{S}_a$ such that $\mathcal{R} \leq_{\mathsf{IQ}} \mathsf{low}(\mathcal{S}, D(a))$. As argued in [11], the lowering is a repair seed w.r.t. the old definition of repair types. Since the repair type assigned to an individual b by $\mathsf{low}(\mathcal{S}, D(a))$ covers the repair type $\mathcal{R}_b$, and the latter satisfies Conditions (RT5), it follows that also the former covers all globally forbidden concepts, and thus satisfies (RT5). $\square$

It remains to show that Condition 3 can be checked in polynomial time.

Proposition 23. *Given a qABox $\exists X.\mathcal{A}$, an $\mathcal{EL}$ TBox $\mathcal{T}$, a repair request $\mathcal{P} \subseteq$ gloIRQ, a query $\alpha \in \mathsf{gloIRQ}$, and a repair seed $\mathcal{S}$, we can decide in polynomial time whether α is entailed w.r.t. $\mathcal{T}$ by the canonical gloIRQ-repair induced by $\mathcal{S}$.*

Proof. We prove three subclaims regarding the different kinds of queries.

1. $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S})$ entails a concept assertion/IQ $C(a)$ iff $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}) \models C(a)$ and no atom in $\mathcal{S}_a$ subsumes C w.r.t. $\mathcal{T}$.
2. $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S})$ entails a global IQ $\exists\{x\}.\{C(x)\}$ iff $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}) \models \exists\{x\}.\{C(x)\}$ and C is not globally forbidden.
3. $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S})$ entails a role assertion $r(a,b)$ iff $r(a,b) \in \mathcal{A}$, $r(a,b) \notin \mathcal{P}$, and $\mathsf{Succ}(\mathcal{S}_a, r, b) \leq \mathcal{S}_b$.

The first claim is mentioned in Lemma 3 in [10], but for the old definition of repair types. It is straightforward to adapt the proof of Lemma 14 such that C can be an arbitrary $\mathcal{EL}$ concept and the if direction requires that no atom in $\mathcal{K}$ subsumes C w.r.t. $\mathcal{T}$. (Then $\exists r.D \not\sqsubseteq^{\mathcal{T}} \exists r.E$ follows from the assumption and not by (RT3).) Regarding the only-if direction, if an atom D in $\mathcal{S}_a$ subsumes C w.r.t. $\mathcal{T}$, i.e., $C \sqsubseteq^{\mathcal{T}} D$, then $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S}) \not\models D(a)$ by Lemma 14, and thus $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S}) \not\models C(a)$ since $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S})$ is saturated (see Lemma XIII in [6]).

Next, we consider the second claim. The relation consisting of all pairs $(\langle\!\langle u, \mathcal{K}\rangle\!\rangle, u)$ is a total simulation from $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S})$ to $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A})$. Thus, if $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S}) \models \exists\{x\}.\{C(x)\}$, then Lemma 4 yields that $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A}) \models \exists\{x\}.\{C(x)\}$. Further recall that no repair entails $\exists\{x\}.\{C(x)\}$ for any globally forbidden concept C. It remains to prove the if direction. By assumption, there is an object u of $\mathsf{sat}^{\mathcal{T}}_{\mathsf{IQ}}(\exists X.\mathcal{A})$ that is an instance of C. Since C is not globally forbidden, we can build a repair type $\mathcal{K}$ for u that does not contain any subsumer of C. It follows that $\langle\!\langle u, \mathcal{K}\rangle\!\rangle$ is an instance of C by Lemma 14, and thus $\mathsf{rep}^{\mathcal{T}}_{\mathsf{gloIRQ}}(\exists X.\mathcal{A}, \mathcal{S}) \models \exists\{x\}.\{C(x)\}$.

Last, the third claim is shown in Lemma 4.5 in [11], and the same proof works here. □

In sum, we have now shown that the Conditions 1–3 introduced above can be checked in polynomial time. This provides us with an NP-procedure for cautious non-entailment, not only for optimal gloIRQ-repairs w.r.t. $\mathcal{EL}$ TBoxes, but also for optimal inconsistency gloIRQ-repairs w.r.t. $\mathcal{EL}^\bot$ TBoxes.

Theorem 24. *Cautious* gloIRQ*-entailment is in* coNP*.*

It is not clear whether this upper bound is tight. For the case of classical repairs, coNP-completeness of cautious entailment is shown in [11], but the hardness proof cannot easily be adapted to the case of optimal repairs.

6 Conclusion

We have shown that our previous work on optimal repairs [5] and error-tolerant reasoning w.r.t. optimal repairs [10,11] can be extended from TBoxes written in the DL $\mathcal{EL}$ to $\mathcal{EL}^\bot$ TBoxes. From a practical point of view, $\bot$ can be used to express disjointness of concepts, which means that certain modelling errors can be detected as inconsistencies, as illustrated in our introductory example. From a theoretical point of view, $\mathcal{EL}^\bot$ is a minimal extension of $\mathcal{EL}$ that can express inconsistency. This allows us to investigate the effect that inconsistency has on repairs (such as the need for considering global instance queries in repair requests) without being distracted by clutter caused by other constructors, as e.g. the ones of $\mathcal{ELROI}(\bot)$. In contrast to the treatment of optimal repairs for the DL $\mathcal{ELROI}(\bot)$ in [9], we use here a different entailment relation (gloIRQ-entailment) in the definition of optimal repairs. This relation has the advantage that TBoxes need not be cycle-restricted, and is appropriate if one does not intend to use general conjunctive queries to access the knowledge base. We conjecture that the approach developed in this paper can be extended to $\mathcal{ELRO}(\bot)$ TBoxes, though this would probably cause a higher complexity for computing optimal repairs and for cautious reasoning. As pointed out in [9], in the presence of inverse roles, finite IQ-saturation cannot always work, and thus the motivation for using gloIRQ-entailment rather than CQ-entailment is no longer there. The complexity of error-tolerant reasoning w.r.t. optimal CQ-repairs [5] still needs to be investigated. Since entailment of a CQ by a qABox is already NP-hard, the best complexity for cautious entailment we can hope for is then on the second level of the polynomial hierarchy.

Acknowledgements This work has been supported by Deutsche Forschungsgemeinschaft (DFG) in projects 430150274 (Repairing Description Logic Ontologies) and 389792660 (TRR 248: Foundations of Perspicuous Software Systems).

References

1. Baader, F., Brandt, S., Lutz, C.: Pushing the $\mathcal{EL}$ envelope. In: Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI). pp. 364–369. Professional Book Center (2005), http://ijcai.org/Proceedings/05/Papers/0372.pdf
2. Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F. (eds.): The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press (2003). https://doi.org/10.1017/cbo9780511711787
3. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: An Introduction to Description Logic. Cambridge University Press (2017). https://doi.org/10.1017/9781139025355
4. Baader, F., Koopmann, P., Kriegel, F.: Optimal repairs in the description logic $\mathcal{EL}$ revisited. In: Proceedings of the 18th European Conference on Logics in Artificial Intelligence (JELIA). Lecture Notes in Computer Science, vol. 14281, pp. 11–34. Springer (2023). https://doi.org/10.1007/978-3-031-43619-2_2
5. Baader, F., Koopmann, P., Kriegel, F., Nuradiansyah, A.: Computing optimal repairs of quantified ABoxes w.r.t. static $\mathcal{EL}$ TBoxes. In: Proceedings of the 28th International Conference on Automated Deduction (CADE). Lecture Notes in Computer Science, vol. 12699, pp. 309–326. Springer (2021). https://doi.org/10.1007/978-3-030-79876-5_18
6. Baader, F., Koopmann, P., Kriegel, F., Nuradiansyah, A.: Computing optimal repairs of quantified ABoxes w.r.t. static $\mathcal{EL}$ TBoxes (extended version). LTCS-Report 21-01, Chair of Automata Theory, Institute of Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany (2021). https://doi.org/10.25368/2022.64
7. Baader, F., Koopmann, P., Kriegel, F., Nuradiansyah, A.: Optimal ABox repair w.r.t. static $\mathcal{EL}$ TBoxes: From quantified ABoxes back to ABoxes. In: Proceedings of the 19th Extended Semantic Web Conference (ESWC). Lecture Notes in Computer Science, vol. 13261, pp. 130–146. Springer (2022). https://doi.org/10.1007/978-3-031-06981-9_8
8. Baader, F., Koopmann, P., Kriegel, F., Nuradiansyah, A.: Optimal ABox repair w.r.t. static $\mathcal{EL}$ TBoxes: from quantified ABoxes back to ABoxes (extended version). LTCS-Report 22-01, Chair of Automata Theory, Institute of Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany (2022). https://doi.org/10.25368/2022.65
9. Baader, F., Kriegel, F.: Pushing optimal ABox repair from $\mathcal{EL}$ towards more expressive Horn-DLs. In: Proceedings of the 19th International Conference on Principles of Knowledge Representation and Reasoning (KR) (2022). https://doi.org/10.24963/kr.2022/3
10. Baader, F., Kriegel, F., Nuradiansyah, A.: Error-tolerant reasoning in the description logic $\mathcal{EL}$ based on optimal repairs. In: Proceedings of the 6th International Joint Conference on Rules and Reasoning (RuleML+RR). Lecture Notes in Computer Science, vol. 13752, pp. 227–243. Springer (2022). https://doi.org/10.1007/978-3-031-21541-4_15
11. Baader, F., Kriegel, F., Nuradiansyah, A.: Treating role assertions as first-class citizens in repair and error-tolerant reasoning. In: Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing (SAC). pp. 974–982. ACM (2023). https://doi.org/10.1145/3555776.3577630

12. Baader, F., Kriegel, F., Nuradiansyah, A., Peñaloza, R.: Making repairs in description logics more gentle. In: Proceedings of the 16th International Conference on Principles of Knowledge Representation and Reasoning (KR). pp. 319–328. AAAI Press (2018), https://cdn.aaai.org/ocs/18056/18056-78653-1-PB.pdf
13. Baader, F., Kriegel, F., Nuradiansyah, A., Peñaloza, R.: Computing compliant anonymisations of quantified ABoxes w.r.t. $\mathcal{EL}$ policies. In: Proceedings of the 19th International Semantic Web Conference (ISWC). Lecture Notes in Computer Science, vol. 12506, pp. 3–20. Springer (2020). https://doi.org/10.1007/978-3-030-62419-4_1
14. Baader, F., Suntisrivaraporn, B.: Debugging SNOMED CT using axiom pinpointing in the description logic $\mathcal{EL}^+$. In: Proceedings of the 3rd International Conference on Knowledge Representation in Medicine (KR-MED). CEUR Workshop Proceedings, vol. 410. CEUR-WS.org (2008), https://ceur-ws.org/Vol-410/Paper01.pdf
15. Bienvenu, M., Ortiz, M.: Ontology-mediated query answering with data-tractable description logics. In: Tutorial Lectures of the 11th Reasoning Web Summer School. Lecture Notes in Computer Science, vol. 9203, pp. 218–307. Springer (2015). https://doi.org/10.1007/978-3-319-21768-0_9
16. Bienvenu, M., Rosati, R.: Tractable approximations of consistent query answering for robust ontology-based data access. In: Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI). pp. 775–781. IJCAI/AAAI (2013), https://www.ijcai.org/Proceedings/13/Papers/121.pdf
17. Calì, A., Lembo, D., Rosati, R.: On the decidability and complexity of query answering over inconsistent and incomplete databases. In: Proceedings of the 22nd ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS). pp. 260–271. ACM (2003). https://doi.org/10.1145/773153.773179
18. Gottlob, G., Leone, N., Scarcello, F.: The complexity of acyclic conjunctive queries. Journal of the ACM **48**(3), 431–498 (2001). https://doi.org/10.1145/382780.382783
19. Henzinger, M.R., Henzinger, T.A., Kopke, P.W.: Computing simulations on finite and infinite graphs. In: Proceedings of the 36th Annual Symposium on Foundations of Computer Science (FOCS). pp. 453–462. IEEE Computer Society (1995). https://doi.org/10.1109/sfcs.1995.492576
20. Hoehndorf, R., Schofield, P.N., Gkoutos, G.V.: The role of ontologies in biological and biomedical research: a functional perspective. Briefings in Bioinformatics **16**(6), 1069–1080 (2015). https://doi.org/10.1093/bib/bbv011
21. Lembo, D., Lenzerini, M., Rosati, R., Ruzzi, M., Savo, D.F.: Inconsistency-tolerant query answering in ontology-based data access. Journal of Web Semantics **33**, 3–29 (2015). https://doi.org/10.1016/j.websem.2015.04.002
22. Ludwig, M., Peñaloza, R.: Error-tolerant reasoning in the description logic $\mathcal{EL}$. In: Proceedings of the 14th European Conference on Logics in Artificial Intelligence (JELIA). Lecture Notes in Computer Science, vol. 8761, pp. 107–121. Springer (2014). https://doi.org/10.1007/978-3-319-11558-0_8
23. Lutz, C., Wolter, F.: Deciding inseparability and conservative extensions in the description logic $\mathcal{EL}$. Journal of Symbolic Computation **45**(2), 194–228 (2010). https://doi.org/10.1016/j.jsc.2008.10.007
24. Parsia, B., Sirin, E., Kalyanpur, A.: Debugging OWL ontologies. In: Proceedings of the 14th International Conference on World Wide Web, (WWW). pp. 633–640. ACM (2005). https://doi.org/10.1145/1060745.1060837

25. Peñaloza, R.: Error-tolerance and error management in lightweight description logics. Künstliche Intelligenz **34**(4), 491–500 (2020). https://doi.org/10.1007/s13218-020-00684-5
26. Schlobach, S., Huang, Z., Cornet, R., van Harmelen, F.: Debugging incoherent terminologies. Journal of Automated Reasoning **39**(3), 317–349 (2007). https://doi.org/10.1007/s10817-007-9076-z
27. Xiao, G., Calvanese, D., Kontchakov, R., Lembo, D., Poggi, A., Rosati, R., Zakharyaschev, M.: Ontology-based data access: A survey. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI). pp. 5511–5519. ijcai.org (2018). https://doi.org/10.24963/ijcai.2018/777