TECHNISCHE UNIVERSITÄT DRESDEN

Faculty of Computer Science • Institute of Theoretical Computer Science • Chair of Automata Theory

Explaining and Repairing Description Logic Ontologies

Franz Baader, Patrick Koopmann, Francesco Kriegel

Funded by DFG in Project 430150274 (Repairing Description Logic Ontologies) and partially in Project 389792660 (TRR 248: Foundations of Perspicuous Software Systems).

ReDLO Repairing Description Logic Ontologies

CPEC CENTER FOR PERSPICUOUS COMPUTING

1st European Summer School on Artificial Intelligence (ESSAI 2023), 28 July 2023

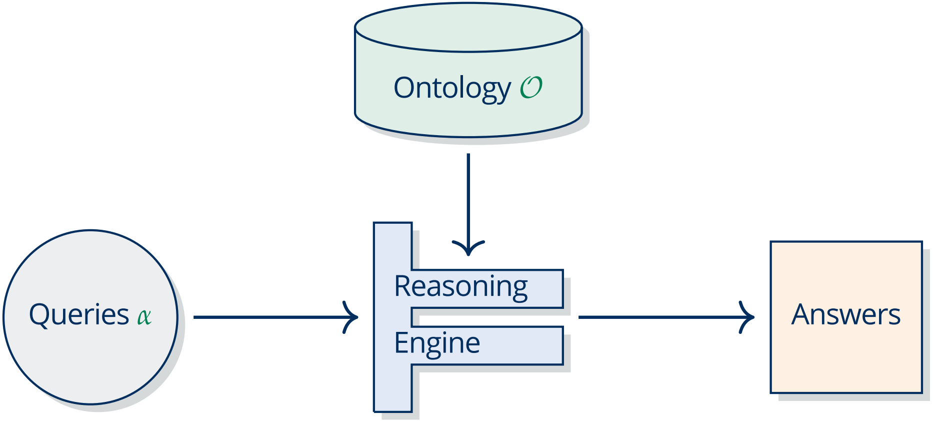
Course Outline

From Classical Repairs to Optimal Repairs

DL Ontologies in Knowledge-based Systems

- DL ontologies can represent knowledge about complex domains, particularly where precision and explainability are essential.
- DL ontologies are used in knowledge-based systems that can perform automated reasoning to answer queries and thereby assist in making decisions based on the encoded knowledge.
- Sometimes faulty query answers are detected, and then the ontology needs to be repaired appropriately.
- Ontology repair thus is a crucial task for ontology maintenance.

Query Answering in Knowledge-based Systems



Query Answering in Knowledge-based Systems

Typical Boolean queries on DL ontologies are:

- Is the individual name a an instance of the concept description C? (Concept assertion $a : C$)
- Is the individual name a related to the individual name b by the role name r? (Role assertion $(a,b) : r$)
- Are all instances of the concept description C also instances of the concept description D? (Concept inclusion $C \sqsubseteq D$)

We focus on the case where the ontology $\mathcal{O}$ entails one or more Boolean queries α (i.e., the answers to these queries are “Yes”), although these are wrong in the underlying domain of interest.

See Session 4 how repairs can be constructed for the case of non-entailed Boolean queries (with answer “No”) that should be entailed.

Classical Repairs

To specify what must be repaired, all unwanted consequences are collected in a repair request.

Definition. A repair request $\mathcal{P}$ is a finite set of Boolean queries.

In the classical way, a repair is then obtained from the input ontology $\mathcal{O}$ by removing axioms such that the remaining ones do not entail the unwanted consequences in $\mathcal{P}$ anymore.

Definition. A classical repair of an ontology $\mathcal{O}$ for $\mathcal{P}$ is an ontology $\mathcal{O}'$ such that

1. $\mathcal{O}' \subseteq \mathcal{O}$
2. $\mathcal{O}' \not\models \alpha$ for each $\alpha \in \mathcal{P}$.

Recall from Session 2 how classical repairs can be computed by means of axiom pinpointing (remove a hitting set of all justifications).

An Example in the Domain of Bicycles

$\mathcal{EL}$ Ontology $\mathcal{O} \coloneqq \mathcal{A} \cup \mathcal{T}$ with components:

- ABox $\mathcal{A} \coloneqq \{ (\mathsf{francesco}, \mathsf{fbike}) : \mathsf{rides},\ \mathsf{fbike} : \mathsf{Mountain\_Bike} \sqcap \exists \mathsf{has\_part}.\mathsf{Full\_Suspension} \}$
- TBox $\mathcal{T} \coloneqq \{ \exists \mathsf{has\_part}.\mathsf{Full\_Suspension} \sqsubseteq \mathsf{Mountain\_Bike},\ \ \mathsf{Mountain\_Bike} \sqsubseteq \mathsf{Bike} \}$

We consider the TBox as static.

The Repair request is $\mathcal{P} \coloneqq \{ \mathsf{francesco} : \exists \mathsf{rides}.\mathsf{Mountain\_Bike} \}$.

The best classical repairs are:

1. $\{ \mathsf{fbike} : \mathsf{Mountain\_Bike} \sqcap \exists \mathsf{has\_part}.\mathsf{Full\_Suspension} \} \cup \mathcal{T}$
2. $\{ (\mathsf{francesco}, \mathsf{fbike}) : \mathsf{rides} \} \cup \mathcal{T}$

In both repairs, several other consequences are lost.

Since we can only delete axioms, classical repairs are highly syntax-dependant.

To obtain better repairs, we switch to another objective: instead of removing as few axioms as possible, we rewrite the ontology such that a minimal number of consequences is removed.

A New Definition of Repairs

To avoid dependence on syntax and to allow retaining more consequences, we replace the first condition in the repair definition.

Definition. A repair of an ontology $\mathcal{O}$ for a repair request $\mathcal{P}$ is an ontology $\mathcal{O}'$ such that

1. $\mathcal{O} \models \mathcal{O}'$
2. $\mathcal{O}' \not\models \alpha$ for each $\alpha \in \mathcal{P}$.

Instead of the model-based entailment relation $\models$, we also use weaker entailment relations based on query languages. More details later.

Definition. A repair $\mathcal{O}'$ for $\mathcal{P}$ is optimal if there is no other repair $\mathcal{O}''$ for $\mathcal{P}$ that strictly entails it, i.e., where $\mathcal{O}'' \models \mathcal{O}'$ but $\mathcal{O}' \not\models \mathcal{O}''$.

Since the model-based entailment relation $\models$ is the most general entailment relation, all query languages are taken into account. Thus, in an optimal repair a maximal amount of entailed queries of any type are preserved.

Gentle Repairs

Repairs obtained by Axiom Weakening

Better repairs than the classical ones can be obtained by weakening axioms instead of removing them completely.

Definition. Given axioms α, β, we say that β is weaker than α if α strictly entails β, i.e., $\alpha \models \beta$ but $\beta \not\models \alpha$.

The simplest algorithm to repair the input ontology $\mathcal{O}$ for the repair request $\mathcal{P}$ is as follows.

While $\mathcal{O}$ entails some query in $\mathcal{P}$, do the following:

1. Choose an axiom $\alpha \in \mathcal{O}$.
2. Choose an axiom β that is weaker than α.
3. Replace α by β, i.e., set $\mathcal{O} := (\mathcal{O} \setminus \{\alpha\}) \cup \{\beta\}$.

Weakening of axioms must be restricted to ensure termination.

A Non-terminating Example

- TBox $\mathcal{T} := \{ A \sqsubseteq \exists r.A \}$
- Repair request $\mathcal{P} := \{ A \sqsubseteq \exists r.\top \}$

An infinite chain of subsequent weakenings of the axiom $A \sqsubseteq \exists r.A$ is

$$A \sqsubseteq \exists r.\exists r.A, \quad A \sqsubseteq \exists r.\exists r.\exists r.\exists r.A, \quad A \sqsubseteq \exists r.\exists r.\exists r.\exists r.\exists r.\exists r.\exists r.\exists r.A, \quad \ldots$$

None of the TBoxes $\{ A \sqsubseteq \exists r. \cdots \exists r.A \}$ is a repair.

As solution, we employ weakening relations.

Definition. A weakening relation is an irreflexive, transitive relation $\succ$ on axioms such that, if $\alpha \succ \beta$, then β is weaker than α.

To ensure that some repair is reachable (viz. the trivial repair, which only entails tautologies), $\succ$ must be complete, i.e., if α is no tautology, then there is a tautology β such that $\alpha \succ \beta$.

Then, termination of the simple approach (see last slide) is guaranteed if $\succ$ is well-founded, i.e., for every axiom α, there is no infinite chain $\alpha \succ \beta_1 \succ \beta_2 \succ \cdots$.

Gentle Repair Algorithm

Termination can also be ensured by integrating axiom weakening into the classical repair approach.

Gentle Repair Algorithm. While $\mathcal{O}$ entails some query in $\mathcal{P}$, do the following:

1. Compute all justifications $\mathcal{J}_1, \dots, \mathcal{J}_k$ for any axiom in $\mathcal{P}$.
2. Compute a minimal hitting set $\mathcal{H}$ of $\mathcal{J}_1, \dots, \mathcal{J}_k$.
3. For every axiom $\alpha \in \mathcal{H}$,
   - choose a weaker axiom β with $\alpha \succ \beta$ such that, for each justification $\mathcal{J}_i$ with $\alpha \in \mathcal{J}_i$, no query in $\mathcal{P}$ is entailed by $(\mathcal{J}_i \setminus \{\alpha\}) \cup \{\beta\}$,
   - and then replace α by β, i.e., set $\mathcal{O} := (\mathcal{O} \setminus \{\alpha\}) \cup \{\beta\}$.

In the most general form, we use the weakening relation $\succ$ where $\alpha \succ \beta$ if $\alpha \models \beta$ but $\beta \not\models \alpha$, which is complete but not well-founded.

F. Baader, F. Kriegel, A. Nuradiansyah, R. Peñaloza: **Making Repairs in Description Logics More Gentle.** KR 2018

Termination of the Gentle Repair Algorithm

Theorem. The Gentle Repair Algorithm terminates after at most exponentially many iterations and returns a repair of $\mathcal{O}$ for $\mathcal{P}$.

Proof Sketch. In the beginning there can be at most exponentially many justifications for the queries in $\mathcal{P}$, and in every iteration at least one justification is removed.

By restricting the lengths of chains of weakenings, faster termination can be guaranteed.

Definition. A weakening relation $\succ$ is polynomially bounded if, for each axiom α, every chain $\alpha \succ \beta_1 \succ \beta_2 \succ \cdots$ has at most polynomial length.

Proposition. If a polynomially bounded weakening relation $\succ$ is used, then the Gentle Repair Algorithm terminates after at most polynomially many iterations and returns a repair of $\mathcal{O}$ for $\mathcal{P}$.

An Example in the Domain of Bicycles

- ABox $\mathcal{A} \coloneqq \{\, (\mathsf{francesco}, \mathsf{fbike}) : \mathsf{rides},\ \mathsf{fbike} : \mathsf{Mountain\_Bike} \sqcap \exists \mathsf{has\_part}.\mathsf{Full\_Suspension} \,\}$
- Static TBox $\mathcal{T} \coloneqq \{\, \exists \mathsf{has\_part}.\mathsf{Full\_Suspension} \sqsubseteq \mathsf{Mountain\_Bike},\ \ \mathsf{Mountain\_Bike} \sqsubseteq \mathsf{Bike} \,\}$
- Repair request $\mathcal{P} \coloneqq \{\, \mathsf{francesco} : \exists \mathsf{rides}.\mathsf{Mountain\_Bike} \,\}$

The best gentle repairs are:

1. $\{\, \mathsf{fbike} : \mathsf{Mountain\_Bike} \sqcap \exists \mathsf{has\_part}.\mathsf{Full\_Suspension} \,\} \cup \mathcal{T}$
2. $\{\, (\mathsf{francesco}, \mathsf{fbike}) : \mathsf{rides},\ \ \mathsf{fbike} : \exists \mathsf{has\_part}.\top \,\} \cup \mathcal{T}$

None of them is optimal. The first repair could be extended with $\mathsf{francesco} : \exists \mathsf{rides}.\mathsf{Bike}$, whereas we could add $\mathsf{fbike} : \mathsf{Bike}$ to the second repair.

In general, optimal repairs cannot be produced by axiom weakening.

Instead of replacing, in the ontology $\mathcal{O}$ that is currently repaired, an axiom α by a weaker axiom β, we could also replace α by some axiom β with $\mathcal{O} \models \beta$ and $(\mathcal{O} \setminus \{\alpha\}) \cup \{\beta\} \not\models \alpha$, but this not axiom weakening anymore.

Optimal Repairs of ABoxes

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

In the following, we will focus on the case where the input ontology $\mathcal{O}$ consists of

- an ABox $\mathcal{A}$ formulated in the DL $\mathcal{EL}$
- and a TBox $\mathcal{T}$ also formulated in $\mathcal{EL}$ that is free of errors and will not be changed,

and the repair request $\mathcal{P}$ consists of $\mathcal{EL}$ concept assertions $a : C$ and role assertions $(a, b) : r$.

Definition. Given ontologies $\mathcal{O} := \mathcal{A} \cup \mathcal{T}$ and $\mathcal{O}' := \mathcal{A}' \cup \mathcal{T}$ with the same TBox, we often write $\mathcal{A} \models^{\mathcal{T}} \mathcal{A}'$ instead of $\mathcal{O} \models \mathcal{O}'$ and say that $\mathcal{A}$ entails $\mathcal{A}'$ w.r.t. $\mathcal{T}$.

Our main goals are:

1. Canonical construction of all optimal repairs of $\mathcal{A} \cup \mathcal{T}$ for $\mathcal{P}$.
2. Interactive selection of one optimal repair of $\mathcal{A} \cup \mathcal{T}$ for $\mathcal{P}$.

F. Baader, F. Kriegel, A. Nuradiansyah: **Privacy-Preserving Ontology Publishing for $\mathcal{EL}$ Instance Stores.** JELIA 2019
F. Baader, F. Kriegel, A. Nuradiansyah, R. Peñaloza: **Computing Compliant Anonymisations of Quantified ABoxes w.r.t. $\mathcal{EL}$ Policies.** ISWC 2020
F. Baader, P. Koopmann, F. Kriegel, A. Nuradiansyah: **Computing Optimal Repairs of Quantified ABoxes w.r.t. Static $\mathcal{EL}$ TBoxes.** CADE 2021
F. Baader, P. Koopmann, F. Kriegel, A. Nuradiansyah: **Optimal ABox Repair w.r.t. Static $\mathcal{EL}$ TBoxes: from Quantified ABoxes back to ABoxes.** ESWC 2022
F. Baader, P. Koopmann, F. Kriegel: **Optimal Repairs in the Description Logic $\mathcal{EL}$ Revisited.** JELIA 2023

Optimal Repairs in form of ABoxes need not exist

We need to extend the notion of an ABox to ensure that optimal repairs exist.

Consider the ABox $\mathcal{A} := \{ (\mathsf{alice}, \mathsf{bob}) : \mathsf{has\_friend},\ \mathsf{alice} : \mathsf{Famous} \}$.
It entails the following queries:

- Bob is a friend of Alice.
- Alice is famous.
- Bob is a friend of a someone who is famous.

Now for the repair request $\mathcal{P} := \{ \mathsf{alice} : \mathsf{Famous} \}$, we must remove the assertion that Alice is famous, but we cannot retain the consequence that Bob is a friend of someone famous.

To overcome this issue we extend ABoxes with anonymous individuals (variables).

Quantified ABoxes

Definition. A quantified ABox (qABox) $\exists X.\mathcal{A}$ consists of

1. a finite set X of variables that is disjoint with the signature, and
2. a matrix $\mathcal{A}$ that is a finite set of concept assertions $u : A$ and role assertions $(u, v) : r$ where u, v are individual names or variables, A is a concept name, and r is a role name.

As DLs have model-based semantics, we need to define what the models of a qABox are.

Definition. An interpretation $\mathcal{I}$ is a model of a qABox $\exists X.\mathcal{A}$ if there is a variable assignment $\mathcal{Z} : X \to \mathsf{Dom}(\mathcal{I})$ such that

1. $u^{\mathcal{I}[\mathcal{Z}]} \in A^{\mathcal{I}}$ for each concept assertion $u : A \in \mathcal{A}$, and
2. $(u^{\mathcal{I}[\mathcal{Z}]}, v^{\mathcal{I}[\mathcal{Z}]}) \in r^{\mathcal{I}}$ for each role assertion $(u, v) : r \in \mathcal{A}$,

where $u^{\mathcal{I}[\mathcal{Z}]} := u^{\mathcal{I}}$ if u is an individual name and $u^{\mathcal{I}[\mathcal{Z}]} := u^{\mathcal{Z}}$ otherwise.

Quantified ABoxes

Each $\mathcal{EL}$ ABox has an equivalent qABox, but the converse is not true.

The ABox $\{\, \text{alice} : \text{Cat} \sqcap \exists \text{has\_mother}.\text{Cat} \,\}$
is equivalent to the qABox $\exists \{x\}.\{\, \text{alice} : \text{Cat},\ (\text{alice}, x) : \text{has\_mother},\ x : \text{Cat} \,\}$.

Returning to our example, we can now find an optimal repair in form of a qABox.

- ABox $\mathcal{A} \coloneqq \{\, (\text{carol}, \text{bob}) : \text{has\_friend},\ \text{carol} : \text{Famous} \,\}$
- Repair request $\mathcal{P} \coloneqq \{\, \text{carol} : \text{Famous} \,\}$
- Optimal repair $\exists \{x\}.\{\, (\text{carol}, \text{bob}) : \text{has\_friend},\ (x, \text{bob}) : \text{has\_friend},\ x : \text{Famous} \,\}$

Consequence-based Entailment Relations

Depending on the application, it often suffices to restrict attention to those queries that are typically used to access the ontology.

Definition. Given qABoxes $\exists X.\mathcal{A}$, $\exists Y.\mathcal{B}$, a static TBox $\mathcal{T}$, and a query language QL, we say that $\exists X.\mathcal{A}$ QL-entails $\exists Y.\mathcal{B}$ w.r.t. $\mathcal{T}$ and write $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{QL}} \exists Y.\mathcal{B}$ if $\exists Y.\mathcal{B} \models^{\mathcal{T}} \alpha$ implies $\exists X.\mathcal{A} \models^{\mathcal{T}} \alpha$ for each query $\alpha \in \mathsf{QL}$.

- The query language IQ consists of all concept assertions $a : C$.
- The query language IRQ consists of all concept assertions $a : C$ and all role assertions $(a, b) : r$.
- The query language CQ consists of all Boolean conjunctive queries, which are the same as all qABoxes $\exists X.\mathcal{A}$.

Since the TBox is static, it makes no sense to allow for terminological queries in QL (i.e., no concept inclusions).

Lemma. $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{CQ}} \exists Y.\mathcal{B}$ iff $\exists X.\mathcal{A} \models^{\mathcal{T}} \exists Y.\mathcal{B}$.

Repairs w.r.t. Consequence-based Entailment Relations

Repairs w.r.t. such a consequence-based entailment relation $\models_{\mathsf{QL}}$ are then as follows.

Definition. A QL-repair of a qABox $\exists X.\mathcal{A}$ for a repair request $\mathcal{P}$ w.r.t. a static TBox $\mathcal{T}$ is a qABox $\exists Y.\mathcal{B}$ with

1. $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{QL}} \exists Y.\mathcal{B}$
2. $\exists Y.\mathcal{B} \not\models^{\mathcal{T}} \alpha$ for each $\alpha \in \mathcal{P}$.

A QL-repair $\exists Y.\mathcal{B}$ for $\mathcal{P}$ is optimal if there is no other QL-repair $\exists Z.\mathcal{C}$ that strictly QL-entails it, i.e., where $\exists Z.\mathcal{C} \models^{\mathcal{T}}_{\mathsf{QL}} \exists Y.\mathcal{B}$ but $\exists Y.\mathcal{B} \not\models^{\mathcal{T}}_{\mathsf{QL}} \exists Z.\mathcal{C}$.

Of course, we should require that $\mathcal{P} \subseteq \mathsf{QL}$.

An Example for the Query Language IRQ

- ABox $\mathcal{A} \coloneqq \{\, \mathsf{alice} : \mathsf{Cat} \,\}$
- TBox $\mathcal{T} \coloneqq \{\, \mathsf{Cat} \sqsubseteq \exists \mathsf{has\_mother}.\mathsf{Cat},\ \exists \mathsf{has\_mother}.\mathsf{Cat} \sqsubseteq \mathsf{Cat} \,\}$
- Repair request $\mathcal{P} \coloneqq \{\, \mathsf{alice} : \mathsf{Cat} \,\}$

The optimal IRQ-repair is $\exists\{x\}.\{\, (\mathsf{alice}, x) : \mathsf{has\_mother},\ (x, x) : \mathsf{has\_mother} \,\}$.
It entails the following queries in IRQ:

- $\mathsf{alice} : \exists \mathsf{has\_mother}.\top$
- $\mathsf{alice} : \exists \mathsf{has\_mother}.\exists \mathsf{has\_mother}.\top$
- $\mathsf{alice} : \exists \mathsf{has\_mother}.\exists \mathsf{has\_mother}.\exists \mathsf{has\_mother}.\top$
- ...

However, this IRQ-repair does not entail that the mother of Alice is the mother of herself, since this is not expressible by a query in IRQ. It would be expressible by a query in CQ, and thus this IRQ-repair is no CQ-repair.

Characterizations of the Consequence-based Entailment Relations

Definition. A homomorphism from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$ is a function h that sends every object of $\exists Y.\mathcal{B}$ to an object of $\exists X.\mathcal{A}$ such that:

1. $h(a) = a$ for each individual name a.
2. If $\mathcal{B}$ contains $u : A$, then $\mathcal{A}$ contains $h(u) : A$.
3. If $\mathcal{B}$ contains $(u, v) : r$, then $\mathcal{A}$ contains $(h(u), h(v)) : r$.

Lemma. $\exists X.\mathcal{A} \models_{\mathsf{CQ}} \exists Y.\mathcal{B}$ iff there is a homomorphism from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$.

Simulations are similar to homomorphisms but allow to relate every object of $\exists Y.\mathcal{B}$ to several objects of $\exists X.\mathcal{A}$.

Lemma. $\exists X.\mathcal{A} \models_{\mathsf{IQ}} \exists Y.\mathcal{B}$ iff there is a simulation from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$.

Lemma. $\exists X.\mathcal{A} \models_{\mathsf{IRQ}} \exists Y.\mathcal{B}$ iff there is a simulation from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$ and $\mathcal{A}$ contains all role assertions $(a, b) : r$ in $\mathcal{B}$ that involve only individual names.

Characterizations of the Consequence-based Entailment Relations

From the homomorphism characterization, we can derive the following.

Lemma. $\exists X.\mathcal{A} \models_{\mathsf{CQ}} \exists Y.\mathcal{B}$ iff there is a finite sequence of applications of the Copy Rule and the Delete Rule that starts with $\exists X.\mathcal{A}$ and ends with $\exists Y.\mathcal{B}$.

- Copy Rule: Choose an object u of $\exists X.\mathcal{A}$ as well as a fresh variable y not occurring in $\exists X.\mathcal{A}$, and return $\exists (X \cup \{y\}).(\mathcal{A} \cup \{\, y : A \mid u : A \in \mathcal{A} \,\} \cup \{\, (t,y) : r \mid (t,u) : r \in \mathcal{A} \,\}$
$\cup \{\, (y,y) : r \mid (u,u) : r \in \mathcal{A} \,\} \cup \{\, (y,v) : r \mid (u,v) : r \in \mathcal{A} \,\})$.
- Delete Rule: Choose an assertion α in $\mathcal{A}$ and return the qABox $\exists X.(\mathcal{A} \setminus \{\alpha\})$, or choose a variable $x \in X$ that does not occur in $\mathcal{A}$ and return the qABox $\exists (X \setminus \{x\}).\mathcal{A}$.

For $\models_{\mathsf{IQ}}$ and $\models_{\mathsf{IRQ}}$ a similar characterization holds but that needs two more rules.

Thus, also repairs can be constructed by copying and deleting.

Constructing Repairs

Assume that the input qABox is $\exists X.\mathcal{A}$, and that there is no TBox for now.

1. We first repair for the role assertions in $\mathcal{P}$.

   For each role assertion $(a,b):r$ in $\mathcal{P}$, we copy a and b into fresh variables and then delete $(a,b):r$ from the matrix $\mathcal{A}$.

2. Afterwards repairing for the concept assertions $a:C$ in $\mathcal{P}$ is more complex, due to the structure of the concept description C.

   We therefore introduce copies $\langle\!\langle u,\mathcal{K}\rangle\!\rangle$ for all objects u in $\exists X.\mathcal{A}$. The second component $\mathcal{K}$ is a repair type that specifies what is deleted for this copy.

   Definition. A repair type for u is a set of atoms (concept names or existential restrictions) occurring in $\mathcal{P}$ such that
   - $\mathcal{A} \models u:C$ for each atom $C \in \mathcal{K}$
   - $C \not\sqsubseteq^{\emptyset} D$ for each two atoms $C,D \in \mathcal{K}$.

Canonical Repairs

Assume that $\exists X.\mathcal{A}$ is already repaired for all role assertions in $\mathcal{P}$. Then, the canonical repairs of $\exists X.\mathcal{A}$ for the concept assertions in $\mathcal{P}$ have all copies $\langle\!\langle u, \mathcal{K} \rangle\!\rangle$ as objects and their matrix $\mathcal{B}$ consists of the following assertions:

- $\mathcal{B}$ contains $\langle\!\langle u, \mathcal{K} \rangle\!\rangle : A$ if $\mathcal{A}$ contains $u : A$ and $A \notin \mathcal{K}$
- $\mathcal{B}$ contains $(\langle\!\langle u, \mathcal{K} \rangle\!\rangle, \langle\!\langle v, \mathcal{L} \rangle\!\rangle) : r$ if $\mathcal{A}$ contains $(u, v) : r$ and, for each atom $\exists r.C \in \mathcal{K}$ with $\mathcal{A} \models v : C$, there is an atom $D \in \mathcal{L}$ with $C \sqsubseteq^{\emptyset} D$.

The canonical repairs differ in the selection which of the copies represent the individual names.

Definition. A repair seed $\mathcal{S}$ maps each individual name a to a repair type $\mathcal{S}(a)$ for a with

- if $a : C \in \mathcal{P}$, then there is an atom $D \in \mathcal{S}(a)$ with $C \sqsubseteq^{\emptyset} D$.

The canonical repair $\mathsf{rep}(\exists X.\mathcal{A}, \mathcal{S})$ induced by $\mathcal{S}$ is the qABox $\exists Y.\mathcal{B}$ (with the above matrix $\mathcal{B}$) where individual names a and their copies $\langle\!\langle a, \mathcal{S}(a) \rangle\!\rangle$ are used as synonyms and the other copies $\langle\!\langle u, \mathcal{K} \rangle\!\rangle$ are the variables in Y.

Canonical IQ- and IRQ-repairs w.r.t. a TBox

The TBox $\mathcal{T}$ is taken into account in two ways:

1. Before we construct the canonical repair, we saturate the input qABox $\exists X.\mathcal{A}$ w.r.t. $\mathcal{T}$.

   IQ-Saturation Rule: Choose an object u of $\exists X.\mathcal{A}$ as well as a CI $C \sqsubseteq D$ in $\mathcal{T}$ with $\mathcal{A} \models u:C$ but $\mathcal{A} \not\models u:D$, and return the qABox obtained from $\exists X.\mathcal{A}$ by IQ-unfolding D at u, where "IQ-unfolding E at v" is a recursive operation that does the following:
   - For each concept name $A \in \mathsf{Conj}(E)$, add the assertion $v:A$ to $\mathcal{A}$.
   - For each existential restriction $\exists r.F \in \mathsf{Conj}(E)$, add the assertion $(v, x_F):r$ to $\mathcal{A}$ and, if $x_F \notin X$, then add the variable x_F to X and IQ-unfold F at x_F.

2. During the construction of the canonical repair, we must ensure that removed consequences cannot be restored with the TBox.

   We therefore require that each repair type $\mathcal{K}$ for u additionally satisfies:
   - For each axiom $C \sqsubseteq D \in \mathcal{T}$ with $\mathcal{A} \models u:C$ and for each atom $E \in \mathcal{K}$ with $D \sqsubseteq^{\emptyset} E$, there is an atom $F \in \mathcal{K}$ with $C \sqsubseteq^{\emptyset} F$.

Canonical CQ-repairs w.r.t. a TBox

To construct CQ-repairs, we need to modify the Saturation Rule such that it always adds a fresh variable (instead of trying to reuse x_F). Thus, saturation need not terminate.

- The majority of real-world TBoxes is cycle-restricted and then saturation terminates.
- For the others, we can compute a finite representation of the infinite CQ-saturation as well as of every infinite CQ-repair constructed from it.

We still require $\mathcal{P} \subseteq \mathsf{IRQ}$, as otherwise optimal repairs need not exist.
Counterexample: $\mathcal{A} \coloneqq \{(a,a):r\}$ and $\mathcal{P} \coloneqq \{\exists\{x\}.\{(x,x):r\}\}$

- ABox $\mathcal{A} \coloneqq \{\, \mathsf{alice} : \mathsf{Cat} \,\}$
- TBox $\mathcal{T} \coloneqq \{\, \mathsf{Cat} \sqsubseteq \exists \mathsf{has\_mother}.\mathsf{Cat},\ \exists \mathsf{has\_mother}.\mathsf{Cat} \sqsubseteq \mathsf{Cat} \,\}$
- Repair request $\mathcal{P} \coloneqq \{\, \mathsf{alice} : \mathsf{Cat} \,\}$
- Optimal IQ- and IRQ-repair $\exists\{x\}.\{\, (\mathsf{alice}, x) : \mathsf{has\_mother},\ (x,x) : \mathsf{has\_mother} \,\}$
- Optimal CQ-repair $\exists\{x_1, x_2, \ldots\}.\{\, (\mathsf{alice}, x_1) : \mathsf{has\_mother},\ (x_1, x_2) : \mathsf{has\_mother},\ \ldots \,\}$

Main Results on Canonical Repairs

Let QL be one of the query languages IQ, IRQ, and CQ.

Proposition. For each repair seed $\mathcal{S}$, the induced canonical QL-repair $\mathsf{rep}^{\mathcal{T}}_{\mathsf{QL}}(\exists X.\mathcal{A}, \mathcal{S})$ is a QL-repair of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$.

Proof Sketch. We consider the case $\mathsf{QL} = \mathsf{CQ}$, the others are similar. Denote $\mathsf{rep}^{\mathcal{T}}_{\mathsf{QL}}(\exists X.\mathcal{A}, \mathcal{S})$ as $\exists Y.\mathcal{B}$.

- The mapping h with $h(\langle\!\langle u, \mathcal{K} \rangle\!\rangle) := u$ is a homomorphism from $\exists Y.\mathcal{B}$ to the saturation of $\exists X.\mathcal{A}$ w.r.t. $\mathcal{T}$. It follows that $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{CQ}} \exists Y.\mathcal{B}$.
- Recall that we have initially repaired for all role assertions in $\mathcal{P}$. Furthermore, we can show that $\mathcal{B} \not\models \langle\!\langle u, \mathcal{K} \rangle\!\rangle : C$ for each atom $C \in \mathcal{K}$ (by induction on C). It follows that $\exists Y.\mathcal{B}$ does not entail any concept assertion in $\mathcal{P}$.

Main Results on Canonical Repairs

Let QL be one of the query languages IQ, IRQ, and CQ.

Proposition. If $\exists Y.\mathcal{B}$ is a QL-repair of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$, then there is a repair seed $\mathcal{S}$ such that $\mathsf{rep}^{\mathcal{T}}_{\mathsf{QL}}(\exists X.\mathcal{A}, \mathcal{S})$ QL-entails $\exists Y.\mathcal{B}$ w.r.t. $\mathcal{T}$.

Proof Sketch. We consider the case $\mathsf{QL} = \mathsf{CQ}$, the others are similar. Denote the saturation of $\exists X.\mathcal{A}$ w.r.t. $\mathcal{T}$ as $\exists X'.\mathcal{A}'$. Since $\exists Y.\mathcal{B}$ is a repair, there is a homomorphism h from $\exists Y.\mathcal{B}$ to $\exists X'.\mathcal{A}'$. For each object u of $\exists Y.\mathcal{B}$, a repair type for u is

$$\mathcal{F}(u) \coloneqq \{\, C \mid C \text{ is an atom in } \mathcal{P} \text{ or } \mathcal{T} \text{ where } \mathcal{B} \not\models^{\mathcal{T}} u : C \text{ and } \mathcal{A}' \models h(u) : C \,\}.$$

The mapping k with $k(u) \coloneqq \langle\!\langle h(u), \mathcal{F}(u) \rangle\!\rangle$ is a homomorphism from $\exists Y.\mathcal{B}$ to the canonical repair induced by repair seed $\mathcal{S}$ with $\mathcal{S}(a) \coloneqq \mathcal{F}(a)$. It follows that $\mathsf{rep}^{\mathcal{T}}_{\mathsf{CQ}}(\exists X.\mathcal{A}, \mathcal{S}) \models_{\mathsf{CQ}} \exists Y.\mathcal{B}$.

Thus, every optimal repair is equivalent to a canonical repair.

Main Results on Canonical Repairs

We obtain the following main results.

Theorem.

1. The set of all optimal IQ-repairs of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$ can be computed in exponential time.
2. The set of all optimal IRQ-repairs of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$ can be computed in exponential time.
3. A set of finite representations of all optimal CQ-repairs of $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$ can be computed in exponential time with access to an NP-oracle.

Interactive Selection of a Repair Seed

Although there exponentially many canonical repairs, we can interactively select one sensible repair seed $\mathcal{S}$ in polynomial time.

We initialize $\mathcal{S}$ by adding the concept description C to $\mathcal{S}(a)$ for each unwanted consequence $a:C$ in $\mathcal{P}$ with $\exists X.\mathcal{A} \models^{\mathcal{T}} a:C$. Then we exhaustively apply the following three rules to $\mathcal{S}$.

1. If $\mathcal{S}(a)$ contains a conjunction $C_1 \sqcap \cdots \sqcap C_n$, then choose an atom C_i and add C_i to $\mathcal{S}(a)$.
2. If $\mathcal{S}(a)$ contains an existential restriction $\exists r.D$ and $\exists X.\mathcal{A}$ contains a role assertion $(a,b):r$ where $\exists X.\mathcal{A} \models^{\mathcal{T}} b:D$, then
   - either add D to $\mathcal{S}(b)$ (in which case the repair will entail $(a,b):r$ but not $b:D$ anymore)
   - or do nothing (then the repair will entail $b:D$ but not $(a,b):r$ anymore).
3. If $\mathcal{S}(a)$ contains an atom E and the TBox $\mathcal{T}$ contains a concept inclusion $C \sqsubseteq D$ with $D \sqsubseteq^{\emptyset} E$ and $\exists X.\mathcal{A} \models^{\mathcal{T}} a:C$, then we add C to $\mathcal{S}(a)$.

Finally, we only keep subsumption-maximal concepts in each set $\mathcal{S}(a)$.

Outlook

More Expressive DLs

The canonical repair approach can be extended from $\mathcal{EL}$ to the more expressive DL Horn-$\mathcal{ALCROI}$ with regular RBoxes and regular path expressions:

- Roles $R ::= r \mid r^-$
- Regular path expressions $\rho ::= \varepsilon \mid r \mid r^- \mid \rho + \rho \mid \rho \circ \rho \mid \rho^*$
- TBox $\mathcal{T}$ consisting of concept inclusions $C \sqsubseteq D$ where
  $C ::= \bot \mid \top \mid A \mid \{a\} \mid C \sqcap C \mid \exists \rho.C \mid C \sqcup C$
  $D ::= \bot \mid \top \mid A \mid \{a\} \mid C \sqcap C \mid \exists R.C \mid \forall R.C \mid \neg C \mid \neg C \sqcup D$
- Regular RBox $\mathcal{R}$ consisting of role inclusions $R_1 \circ \cdots \circ R_n \sqsubseteq S$
- Repair request $\mathcal{P}$ consisting of concept assertions $a : C$ and role assertions $(a, b) : r$
- A restricted class of queries in CQ are also allowed in $\mathcal{P}$.

F. Baader, F. Kriegel: **Pushing optimal ABox repair from $\mathcal{EL}$ towards more expressive Horn-DLs.** KR 2022

Error-tolerant Reasoning

If users of an ontology have found an error which they cannot repair themselves, then they need to wait for the repaired version from the publishing organization. Meanwhile, they can employ error-tolerant reasoning.

Definition. Let QL be a query language, $\mathcal{O}$ an ontology, $\mathcal{P}$ a repair request, and α a query.

- α is bravely QL-entailed by $\mathcal{O}$ for $\mathcal{P}$ if some optimal QL-repair of $\mathcal{O}$ for $\mathcal{P}$ entails α.
- α is cautiously QL-entailed by $\mathcal{O}$ for $\mathcal{P}$ if every optimal QL-repair of $\mathcal{O}$ for $\mathcal{P}$ entails α.

Proposition. Given a qABox $\exists X.\mathcal{A}$, $\mathcal{EL}$ TBox $\mathcal{T}$, request $\mathcal{P} \subseteq$ IRQ, and query $\alpha \in$ IRQ.

- It can be decided in polynomial time if α is bravely IRQ-entailed by $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$.
- It can be decided in non-deterministic polynomial time if α is cautiously IRQ-entailed by $\exists X.\mathcal{A}$ for $\mathcal{P}$ w.r.t. $\mathcal{T}$.

F. Baader, F. Kriegel, A. Nuradiansyah: **Error-Tolerant Reasoning in the Description Logic $\mathcal{EL}$ Based On Optimal Repairs.** RuleML+RR 2022
F. Baader, F. Kriegel, A. Nuradiansyah: **Treating Role Assertions as First-class Citizens in Repair and Error-tolerant Reasoning.** SAC 2023

Privacy-preserving Ontology Publishing

Instead of removing errors, it can also be necessary to remove consequences for privacy.

- Repair request $=$ Policy
- Repair $=$ Compliant anonymization

Definition. $\mathcal{O}$ is compliant with $\mathcal{P}$ if $\mathcal{O} \not\models \alpha$ for each $\alpha \in \mathcal{P}$.

For some use cases, compliance is not enough.

Definition. $\mathcal{O}$ is safe for $\mathcal{P}$ if $\mathcal{O} \cup \mathcal{O}'$ is compliant with $\mathcal{P}$ for each compliant $\mathcal{O}'$.

F. Baader, F. Kriegel, A. Nuradiansyah: **Privacy-Preserving Ontology Publishing for $\mathcal{EL}$ Instance Stores.** JELIA 2019
F. Baader, F. Kriegel, A. Nuradiansyah, R. Peñaloza: **Computing Compliant Anonymisations of Quantified ABoxes w.r.t. $\mathcal{EL}$ Policies.** ISWC 2020
F. Baader, F. Kriegel, A. Nuradiansyah, R. Peñaloza: **Safety of Quantified ABoxes w.r.t. Singleton $\mathcal{EL}$ Policies.** SAC 2021
F. Baader, P. Koopmann, F. Kriegel, A. Nuradiansyah, R. Peñaloza: **Privacy-Preserving Ontology Publishing: […].** DL 2021

Optimal Repairs of TBoxes

- TBoxes are usually much smaller than ABoxes.
- Usually, more care is exercised when maintaining TBoxes, while ABoxes are often just collections of observed data.
- But: TBoxes might also contain errors, which can even lead to wrongly answered assertional queries.
- TBox repairs are thus also valuable.

Theorem. With fixed premises, all optimal repairs of an $\mathcal{EL}$ TBox can be constructed in exponential time.

(Same complexity as for optimal repairs of ABoxes.)

F. Kriegel: **Optimal Fixed-Premise Repairs of $\mathcal{EL}$ TBoxes.** KI 2022

Do you have questions or comments?