Beyond Optimal: Interactive Identification of Better-than-optimal Repairs

Francesco Kriegel

Theoretical Computer Science, Technische Universität Dresden, Germany
Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Germany

While producing this paper: Center for Perspicuous Computing (CPEC), Germany

40th ACM/SIGAPP Symposium On Applied Computing, 31 March – 4 April 2025

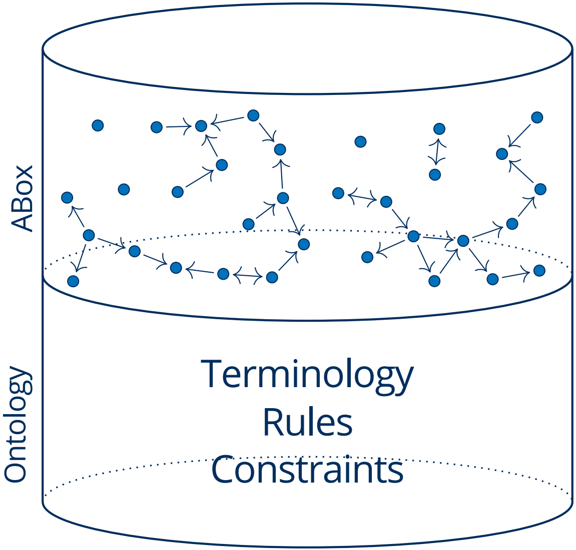
Knowledge-based Systems

Logics, Knowledge Bases, and Models

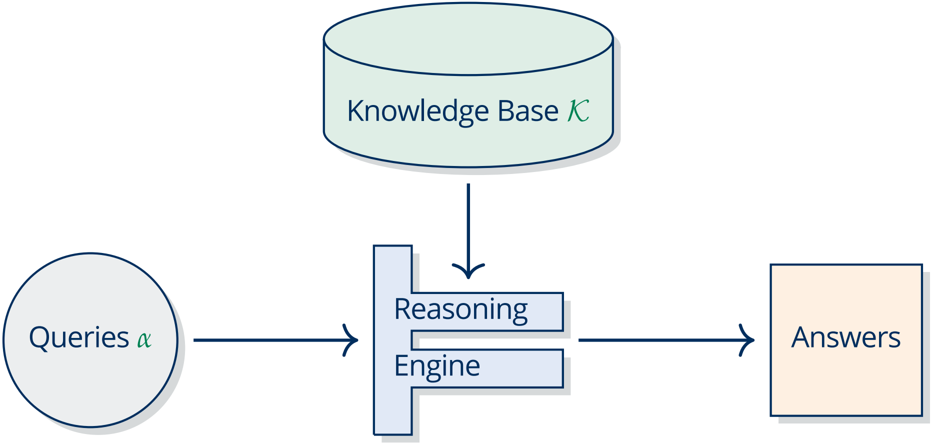
- In general, a logic with model-based semantics consists of
  - a set of all **statements**,
  - a set of all **interpretations**,
  - and a relation $\models$ between them such that $\mathcal{I} \models \alpha$ indicates that the interpretation $\mathcal{I}$ **satisfies** the statement α.
- A **knowledge base (KB)** $\mathcal{K}$ is a finite set of statements.
- $\mathcal{I}$ is a **model** of $\mathcal{K}$ if $\mathcal{I}$ satisfies every statement in $\mathcal{K}$.
- $\mathcal{K}$ is **consistent** if it has a model.
- $\mathcal{K}$ **entails** another KB $\mathcal{K}'$ if every model of $\mathcal{K}$ is also one of $\mathcal{K}'$, written $\mathcal{K} \models \mathcal{K}'$.

ABoxes and Ontologies

- An application domain can be represented by a knowledge base.
- We further assume that the statements are subdivided into **assertions** and **ontological statements**.
- Each KB $\mathcal{K}$ is thus a disjoint union of
  - an **assertion box (ABox)** $\mathcal{A}$ consisting of assertions
  - and an **ontology** $\mathcal{O}$ with ontological statements.



Reasoning


Knowledge Base $\mathcal{K}$
Queries α
Reasoning
Engine
Answers

Optimal Repairs

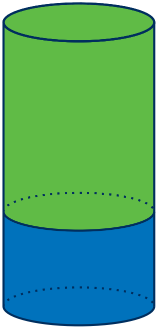
Repairs

- A **repair request** is an assertion set $\mathcal{P} \coloneqq \mathcal{P}_+ \uplus \mathcal{P}_-$ partitioned into
  - an **addition part** $\mathcal{P}_+$
  - and a **removal part** $\mathcal{P}_-$.
- Of a consistent KB $\mathcal{K} \coloneqq \mathcal{A} \uplus \mathcal{O}$, an **ABox repair** for $\mathcal{P}$ is an ABox $\mathcal{B}$ such that
  - $\mathcal{B} \cup \mathcal{O}$ is consistent,
  - $\mathcal{B} \cup \mathcal{O} \models \alpha$ for each $\alpha \in \mathcal{P}_+$, and
  - $\mathcal{B} \cup \mathcal{O} \not\models \beta$ for each $\beta \in \mathcal{P}_-$.

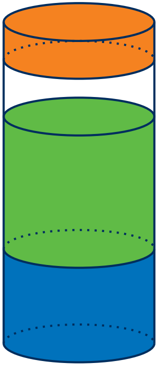
Repairs

- A **repair request** is an assertion set $\mathcal{P} \coloneqq \mathcal{P}_+ \uplus \mathcal{P}_-$ partitioned into
  - an **addition part** $\mathcal{P}_+$
  - and a **removal part** $\mathcal{P}_-$.
- Of a consistent KB $\mathcal{K} \coloneqq \mathcal{A} \uplus \mathcal{O}$, an **ABox repair** for $\mathcal{P}$ is an ABox $\mathcal{B}$ such that
  - $\mathcal{B} \cup \mathcal{O}$ is consistent,
  - $\mathcal{B} \cup \mathcal{O} \models \alpha$ for each $\alpha \in \mathcal{P}_+$, and
  - $\mathcal{B} \cup \mathcal{O} \not\models \beta$ for each $\beta \in \mathcal{P}_-$.
- We write $\mathcal{B} \geq \mathcal{C}$ and say that $\mathcal{B}$ is **at least as good as** another repair $\mathcal{C}$ if
  - $\mathcal{B} \cup \mathcal{O} \models \gamma$ for each $\gamma \in \mathcal{C}$ with $\mathcal{K} \models \gamma$ (i.e. $\mathcal{B}$ entails all retained knowledge in $\mathcal{C}$),
  - and $\mathcal{C} \cup \mathcal{O} \models \gamma$ for each $\gamma \in \mathcal{B}$ with $\mathcal{K} \not\models \gamma$ (i.e. $\mathcal{C}$ entails all additional knowledge in $\mathcal{B}$).
- Moreover, we write $\mathcal{B} > \mathcal{C}$ and say that $\mathcal{B}$ is **better than** $\mathcal{C}$ if $\mathcal{B} \geq \mathcal{C}$ but $\mathcal{C} \not\geq \mathcal{B}$, i.e. either less knowledge is added or less knowledge is removed.

Comparing Repairs

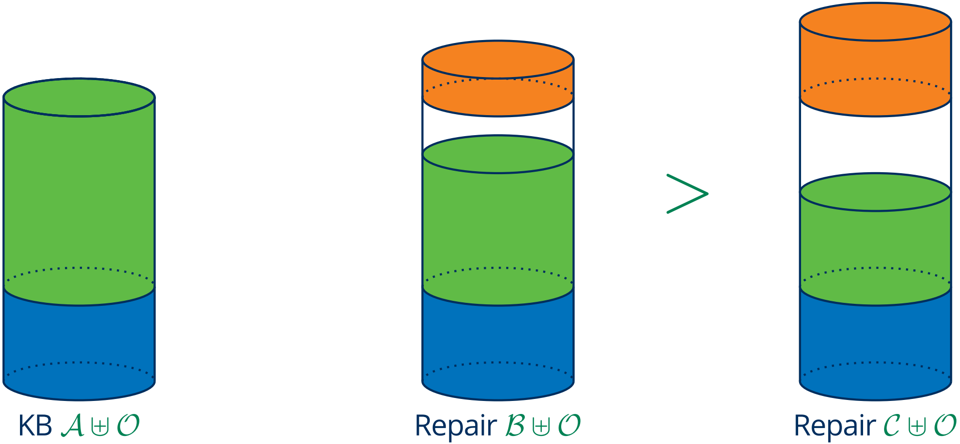


KB $\mathcal{A} \uplus \mathcal{O}$



Repair $\mathcal{B} \uplus \mathcal{O}$

Comparing Repairs

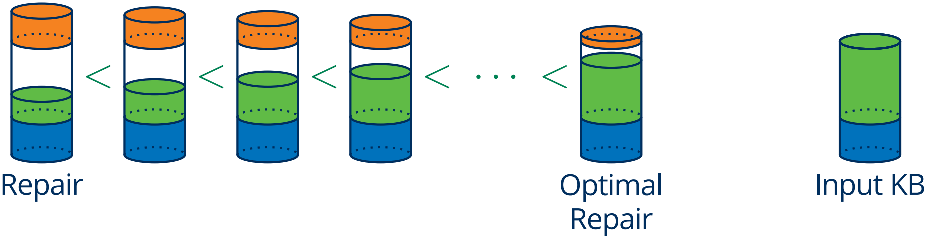


Optimality

- A repair $\mathcal{B}$ **optimal** if there is no repair better than $\mathcal{B}$.
- $\mathcal{P}$ is **optimally coverable** w.r.t. $\mathcal{K}$ if every repair of $\mathcal{K}$ for $\mathcal{P}$ is at most as good as some optimal one.

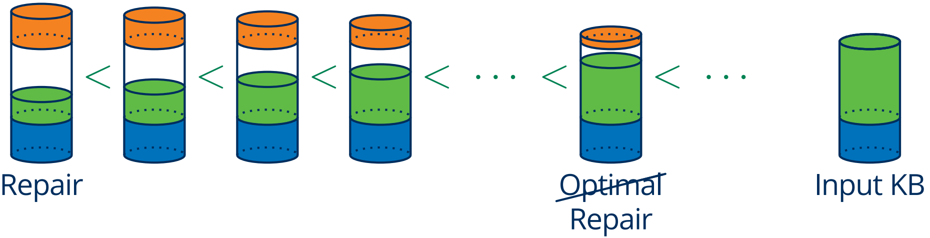
Optimality

- A repair $\mathcal{B}$ **optimal** if there is no repair better than $\mathcal{B}$.
- $\mathcal{P}$ is **optimally coverable** w.r.t. $\mathcal{K}$ if every repair of $\mathcal{K}$ for $\mathcal{P}$ is at most as good as some optimal one.



Optimality

- A repair $\mathcal{B}$ **optimal** if there is no repair better than $\mathcal{B}$.
- $\mathcal{P}$ is **optimally coverable** w.r.t. $\mathcal{K}$ if every repair of $\mathcal{K}$ for $\mathcal{P}$ is at most as good as some optimal one.



Recap: Results on Computing Optimal Repairs

- Quantified ABoxes $\exists X.\mathcal{A}$ consisting of assertions $u : A$ and $(u,v) : r$.
- $\mathcal{EL}$ ontologies consisting of inclusions $C \sqsubseteq D$ where $C ::= \top \mid A \mid C \sqcap C \mid \exists r.C$
- Horn-$\mathcal{ALCROI}$ ontologies consisting of inclusions $C \sqsubseteq D$ and $R_1 \circ \cdots \circ R_n \sqsubseteq S$ where
  $C ::= \bot \mid \top \mid \{a\} \mid A \mid C \sqcap C \mid \exists R.C \mid \exists \rho.C \mid C \sqcup C$
  $D ::= \bot \mid \top \mid \{a\} \mid A \mid C \sqcap C \mid \exists R.C \mid \forall \rho.C \mid \neg C \mid \neg C \sqcup D$
  $R ::= r \mid r^-$
  $\rho ::= r \mid r^- \mid \rho \circ \rho \mid \rho + \rho \mid \rho^*$
- In these settings, every repair request $\mathcal{P}$ consisting of concept assertions $a : C$ and role assertions $(a,b) : r$ is optimally coverable, and there is a canonical form of these optimal repairs that can be computed in exponential time.

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, F. Kriegel: **Pushing Optimal ABox Repair from $\mathcal{EL}$ Towards More Expressive Horn-DLs.** KR 2022

Recap: Results on Computing Optimal Repairs

What's more:

- To ensure that repairs in form of finite quantified ABoxes exist, an acyclicity/termination condition on the ontology must be imposed.
- We also investigated how finite representations of infinite quantified ABoxes as repairs can be computed when these conditions are not fulfilled.
- The ontology is considered as static, i.e. the errors are only in the ABox. If this is not the case, then also $\mathcal{EL}$ ontologies/TBoxes can be optimally repaired (with fixed premises).

F. Baader, P. Koopmann, F. Kriegel: **Optimal Repairs in the Description Logic $\mathcal{EL}$ Revisited.** JELIA 2023
F. Kriegel: **Optimal Fixed-Premise Repairs of $\mathcal{EL}$ TBoxes.** KI 2022

Disputable Consequences

Beyond Optimal

- Recall: optimal repairs preserve as much as possible knowledge entailed by the input KB.
- Prima facie, optimality seems to be desired.

Toy example in the medical domain:

- ABox $\{\mathsf{bob} : \exists \mathsf{shows.SignOrSymptom1},\ \mathsf{bob} : \exists \mathsf{shows.SignOrSymptom2}\}$
- TBox $\{\exists \mathsf{shows.SignOrSymptom1} \sqcap \exists \mathsf{shows.SignOrSymptom2} \sqsubseteq \exists \mathsf{has.DiseaseA}\}$
- Repair request $\{\mathsf{bob} : \exists \mathsf{shows.SignOrSymptom1}\}$
- Optimal repair contains/entails $\mathsf{bob} : \exists \mathsf{has.DiseaseA}$

Another toy example:

- Quantified ABox $\exists\{x\}.\{(\mathsf{alice}, x) : \mathsf{ride},\ x : \mathsf{MountainBike}\}$
- TBox $\{\mathsf{MountainBike} \sqsubseteq \mathsf{Bike}\}$
- Repair request $\{\mathsf{alice} : \exists \mathsf{rides.MountainBike}\}$
- Optimal repair contains/entails $\mathsf{alice} : \exists \mathsf{rides.Bike}$

Disputable Consequences

Formally:

Given a consistent KB $\mathcal{K}$ and a feasible repair request $\mathcal{P}$, a **disputable consequence** is an assertion γ s.t.

- $\mathcal{K} \cup \mathcal{P}_+ \models \gamma$,
- there is a repair $\mathcal{B}$ with $\mathcal{B} \cup \mathcal{O} \models \gamma$,
- there is a repair $\mathcal{B}$ with $\mathcal{B} \cup \mathcal{O} \not\models \gamma$, and
- for each repair $\mathcal{B}$, the KB $\mathcal{B} \cup \mathcal{O}$ does not entail any substantiation* of γ.

*A **substantiation** of an assertion γ is an ABox $\mathcal{J}$ s.t. $\mathcal{A} \cup \mathcal{P}_+ \models \mathcal{J}$ and $\mathcal{J} \cup \mathcal{O} \models \gamma$.

Interactive Repair Strategy

Deterministic Repair Requests

- If $\mathcal{P}$ is optimally coverable w.r.t. $\mathcal{K}$ and there is exactly one optimal repair of $\mathcal{K}$ w.r.t. $\mathcal{P}$ up to equivalence w.r.t. $\mathcal{O}$, then $\mathcal{P}$ is **deterministic** w.r.t. $\mathcal{K}$.
- $\mathcal{P}'$ is a **refinement** of $\mathcal{P}$ if $\mathcal{P}_+ \subseteq \mathcal{P}'_+$ and $\mathcal{P}_- \subseteq \mathcal{P}'_-$ and at least one of these set inclusions is strict.
- When given a non-deterministic repair request $\mathcal{P}$, we could either compute an arbitrary optimal repair, or rather refine $\mathcal{P}$ by user/expert interaction to a deterministic request and so identify a useful repair.

$$\mathcal{P} = \mathcal{P}_0 \xrightarrow{\text{refine}} \mathcal{P}_1 \xrightarrow{\text{refine}} \mathcal{P}_2 \xrightarrow{\text{refine}} \dots \xrightarrow{\text{refine}} \mathcal{P}_n \text{ deterministic}$$

- **How to efficiently find a deterministic refinement?**

Interactive Repair Strategy

Results in the paper:

- Considered setting: KB $\mathcal{K}$ consisting of a quantified ABox $\exists X.\mathcal{A}$ and an $\mathcal{EL}$ ontology $\mathcal{O}$.
- The paper presents an interaction strategy with which a deterministic refinement can be identified in polynomially many steps (i.e. the number of questions that need to be answered by the users/experts is polynomial).
- Every (theoretically) optimal repair can be found with the strategy.
- Strategy runs in two phases:
  1. Phase 1 is devoted to identifying the causes of the initially reported errors in $\mathcal{P}$.
  2. Phase 2 first computed all disputable consequences, and then proceeds with these as Phase 1.
- Although deciding disputable consequences in this setting is coNP-complete, they can be computed rather efficiently in practise.

Do you have questions or comments?