Optimal Fixed-Premise Repairs of $\mathcal{EL}$ TBoxes

Francesco Kriegel
Funded by DFG in Project 430150274.

ReDLO
Repairing Description Logic Ontologies

45th German Conference on Artificial Intelligence (KI 2022),
23 September 2022
Description Logics (DLs)

- DLs are designed for **Knowledge Representation and Reasoning (KRR).**
- Trade-off between representation power and reasoning costs.
- DLs provide the logical underpinning of the **OWL 2 Web Ontology Language,** which is a recommendation of the World Wide Web Consortium (W3C).

**Examples:**
- \(\mathcal{EL}^{++}\) (OWL 2 EL)
- **DL-Lite** (OWL 2 QL)
- **SROIQ** (OWL 2 DL)
- **ALC**

F. Baader, S. Brandt, C. Lutz: *Pushing the \(\mathcal{EL}\) Envelope.* IJCAI 2005
Y. Kazakov: *RIQ and SROIQ are harder than SHOIQ.* KR 2008
Ontologies

- Knowledge on a particular domain can be represented as an ontology.
- Each DL **ontology** $\mathcal{O}$ consists of axioms and is divided into
  1. a **TBox** $\mathcal{T}$ (terminology, global knowledge)
  2. and an **ABox** $\mathcal{A}$ (the data, local knowledge).
Ontologies

Knowledge on a particular domain can be represented as an ontology.

Each DL ontology $\mathcal{O}$ consists of axioms and is divided into:

1. a TBox $\mathcal{T}$ (terminology, global knowledge)
2. and an ABox $\mathcal{A}$ (the data, local knowledge).

Example (formulated in $\mathcal{EL}$):

\[
\mathcal{T} := \{ \text{MountainBike} \sqsubseteq \text{Bike}, \\
\quad \text{Bike} \sqsubseteq \exists \text{hasPart}.\text{SuspensionFork} \sqcap \exists \text{isSuitableFor}.\text{OffRoadCycling}, \\
\quad \text{SuspensionFork} \sqsubseteq \text{Fork}, \\
\quad \text{OffRoadCycling} \sqsubseteq \text{Cycling} \} 
\]

\[
\mathcal{A} := \{ \text{rides(Francesco, x)}, \text{Bike}(x) \} 
\]

The TBox $\mathcal{T}$ will be the running example.
**Reasoning**

- **Reasoning** is the task of deriving implicit consequences from the explicit axioms in an ontology.
- DLs have a **model-theoretic semantics** under open-world assumption.
- Standard reasoning task: Deciding **entailment** \( \models \)
- An ontology \( O \) entails an axiom \( \alpha \), written \( O \models \alpha \), if each model of \( O \) is a model of \( \alpha \).
- Decision procedures for entailment are implemented in **reasoners**.
  - \( \mathcal{EL}^{++} \) (OWL 2 EL): \( CEL/jcel, ELepHant, ELK \)
  - \( \mathcal{SROIQ} \) (OWL 2 DL): \( Chainsaw, FaCT++/Jfact, HermiT, Konclude, MORe, PAGOdA, Pellet, Racer, Sequoia, TrOWL \)

Repairs

- An **ontology can contain axioms that are incorrect** in the underlying domain, especially if
  - it was constructed from incomplete data
  - or using inexact methods based on machine learning.
- Such errors are detected when a **reasoner generates faulty consequences**.
- **Goal**: **Repair the ontology** for these unwanted consequences.
Repairs

- An **ontology can contain axioms that are incorrect** in the underlying domain, especially if
  - it was constructed from incomplete data
  - or using inexact methods based on machine learning.
- Such errors are detected when a **reasoner generates faulty consequences**.
- Goal: **Repair the ontology** for these unwanted consequences.
- My paper focuses on repairing $\mathcal{EL}$ TBoxes.

**Running example:** The TBox $T$ entails two faulty consequences

1. $\text{Bike} \sqsubseteq \exists \text{hasPart. SuspensionFork}$
2. $\text{Bike} \sqsubseteq \exists \text{isSuitableFor. OffRoadCycling}$

} repair request $\mathcal{P}$
Related Work: Classical Repairs

- **Classical Repair Approach:** Delete axioms.
- Each classical repair is obtained by deleting from $T$ all axioms in a hitting set of all justifications for $P$.

F. Baader, R. Peñaloza, B. Suntisrivaraporn: *Pinpointing in the description logic $\mathcal{EL}^+*_. KI 2007
Related Work: Classical Repairs

■ **Classical Repair Approach:** Delete axioms.

■ Each classical repair is obtained by deleting from $\mathcal{T}$ all axioms in a hitting set of all justifications for $\mathcal{P}$.

■ **Running example:** A classical repair of $\mathcal{T}$ is

{ MountainBike ⊑ Bike,  
  Bike ⊑ ∃hasPartSuspensionFork ⊓ ∃isSuitableForOffRoadCycling,  
  SuspensionFork ⊑ Fork,  
  OffRoadCycling ⊑ Cycling }
Related Work: Gentle Repairs

- **Gentle Repair Approach:** Weaken axioms.
- A hitting set of all justifications for $\mathcal{P}$ is still needed to construct a gentle repair, but now all axioms in it are weakened according to a weakening relation $\triangleright$.

F. Kriegel: *Navigating the $\mathcal{EL}$ subsumption hierarchy.* DL 2021
Related Work: Gentle Repairs

- **Gentle Repair Approach:** Weaken axioms.
- A hitting set of all justifications for $\mathcal{P}$ is still needed to construct a gentle repair, but now all axioms in it are weakened according to a weakening relation $\succ$.
- A weakening relation $\succ_{\text{sub}}$ for $\mathcal{EL}$ concept inclusions:
  
  \[
  C \sqsubseteq D \succ_{\text{sub}} C' \sqsubseteq D' \text{ if } C = C', \text{ and } \emptyset \models D \sqsubseteq D', \text{ and } C' \sqsubseteq D' \not\models C \sqsubseteq D.
  \]

- **Problems:**
  1. Efficient computation of maximally strong $\succ_{\text{sub}}$-weakenings
  2. Efficient computation of one or all optimal repairs

---


F. Kriegel: *Navigating the $\mathcal{EL}$ subsumption hierarchy*. DL 2021
Related Work: Countermodel Repairs

- The unwanted consequences in $\mathcal{P}$ are entailed since no counterexamples were known during the construction of the TBox $\mathcal{T}$.
- A model containing such counterexamples can now be obtained from the user or be constructed automatically. The TBox is then rewritten according to the countermodel.
- **Repair-by-Countermodel Approach**: Axiomatize the logical intersection of the TBox and a countermodel.

Related Work: Countermodel Repairs

- The unwanted consequences in $P$ are entailed since no counterexamples were known during the construction of the TBox $\mathcal{T}$.

- A model containing such counterexamples can now be obtained from the user or be constructed automatically. The TBox is then rewritten according to the countermodel.

- **Repair-by-Countermodel Approach:** Axiomatize the logical intersection of the TBox and a countermodel.

- Advantage: Axiomatization method is very precise and thus produces repairs that retain large amounts of knowledge.

- Disadvantage: Repairs are often large (and cannot be made smaller).

---

Generalized-Conclusion Repairs

- Inspired by the gentle repairs w.r.t. $\succ_{\text{sub}}$ as well as by the countermodel repairs, and in order to tackle their problems, a novel type of repairs is introduced.

- A generalized-conclusion repair (GC-repair) $\mathcal{T}'$ of $\mathcal{T}$ is a repair such that additionally: For each $C' \sqsubseteq D' \in \mathcal{T}'$, there is $C \sqsubseteq D \in \mathcal{T}$ such that $C = C'$ and $\emptyset \models D \sqsubseteq D'$. 

Canonical construction of GC-repairs:

1. Choose a polynomial-size repair seed $S$.
2. Construct the induced countermodel $J_S$.
3. Replace each concept inclusion $C \sqsubseteq D$ with $C \sqsubseteq D \vee C_{J_S}$ and $\emptyset \models D \sqsubseteq D'$. 

Main result: For each TBox and each repair request, the set of all optimal GC-repairs can be computed in exponential time, and every GC-repair is entailed by an optimal one.
Generalized-Conclusion Repairs

- Inspired by the gentle repairs w.r.t. $\succ^{sub}$ as well as by the countermodel repairs, and in order to tackle their problems, a novel type of repairs is introduced.

- A **generalized-conclusion repair (GC-repair)** $T'$ of $T$ is a repair such that additionally:
  
  For each $C' \sqsubseteq D' \in T'$, there is $C \sqsubseteq D \in T$ such that $C = C'$ and $\emptyset \models D \sqsubseteq D'$.

- Canonical construction of GC-repairs:
  
  1. Choose a polynomial-size repair seed $S$.
  2. Construct the induced countermodel $J_S$.
  3. Replace each concept inclusion $C \sqsubseteq D$ with $C \sqsubseteq D \lor C J_S J_S$.

- Main result: For each TBox and each repair request, the set of all optimal GC-repairs can be computed in exponential time, and every GC-repair is entailed by an optimal one.
Generalized-Conclusion Repairs

**Running example:** An optimal GC-repair of $T$ is

\[
\begin{align*}
\{ & \text{MountainBike} \sqsubseteq \text{Bike}, \\
& \text{Bike} \sqsubseteq \exists\text{hasPart. SuspensionFork} \sqsubseteq \exists\text{isSuitableFor. OffRoadCycling}, \\
& \exists\text{hasPart. T} \sqcap \exists\text{isSuitableFor. T}, \\
& \text{SuspensionFork} \sqsubseteq \text{Fork}, \\
& \text{OffRoadCycling} \sqsubseteq \text{Cycling} \}
\end{align*}
\]
Fixed-Premise Repairs

- As seen in the last example, GC-repairs might not be satisfactory. We thus define:
- A **fixed-premise repair (FP-repair)** $\mathcal{T}'$ of $\mathcal{T}$ is a repair that satisfies the following additional condition: For each $C' \sqsubseteq D' \in \mathcal{T}'$, there is $C \sqsubseteq D \in \mathcal{T}$ such that $C = C'$. 

FP-repairs can be computed by a little modification to the framework for GC-repairs. Main result: For each TBox and each repair request, the set of all optimal FP-repairs can be computed in exponential time, and every FP-repair is entailed by an optimal one. Contrary to GC-repairs, optimal FP-repairs might need additional expressivity. (But this is no problem!)
As seen in the last example, GC-repairs might not be satisfactory. We thus define:

A **fixed-premise repair (FP-repair)** $T'$ of $T$ is a repair that satisfies the following additional condition: For each $C' \sqsubseteq D' \in T'$, there is $C \sqsubseteq D \in T$ such that $C = C'$.

FP-repairs can be computed by a little modification to the framework for GC-repairs.

Main result: For each TBox and each repair request, **the set of all optimal FP-repairs can be computed in exponential time**, and **every FP-repair is entailed by an optimal one**.

Contrary to GC-repairs, optimal FP-repairs might need additional expressivity. (But this is no problem!)
Fixed-Premise Repairs

**Running example:** An optimal FP-repair of $\mathcal{T}$ is

$$\{ \text{MountainBike} \sqsubseteq \text{Bike},$$

$$\text{Bike} \sqcap \exists \text{hasPart} \cdot \text{SuspensionFork} \sqcap \exists \text{isSuitableFor} \cdot \text{OffRoadCycling},$$

$$\text{Bike} \sqsubseteq \exists \text{hasPart} \cdot \text{SuspensionFork} \sqcap \exists \text{isSuitableFor} \cdot \text{OffRoadCycling},$$

$$\exists \text{hasPart} \cdot \text{Fork} \sqcap \exists \text{isSuitableFor} \cdot \text{Cycling},$$

$$\text{SuspensionFork} \sqsubseteq \text{Fork},$$

$$\text{OffRoadCycling} \sqsubseteq \text{Cycling} \}$$
Conclusion

- A novel approach to repairing $\mathcal{EL}$ TBoxes for unwanted concept inclusions has been developed.
- Two variants: GC-repairs and FP-repairs
- Each optimal repair is characterized by a polynomial-size repair seed.
- Optimal repairs can be computed in exponential time.
- Prototypical implementation: https://github.com/francesco-kriegel/right-repairs-of-el-tboxes
- Repair seed is obtained by user interaction.
Next Steps

- More expressivity:
  - Nominals \( \{a\} \) (also adds supports for ABox axioms)
  - Bottom concept \( \bot \)
  - Inverse roles \( r^- \)
  - Role inclusions \( R_1 \circ \cdots \circ R_n \subseteq S \)

- Support for a partitioning of the ontology into a static part and a refutable part.

- Improvement of FP-repairs by selective, automatic introduction of new premises (can currently be done manually).

F. Baader, F. Kriegel: **Pushing optimal ABox repair from \( \mathcal{EL} \) towards more expressive Horn-DLs.** KR 2022
Do you have questions or comments?