TECHNISCHE UNIVERSITÄT DRESDEN

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

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)
  - $\mathcal{SROIQ}$ (OWL 2 DL)
  - $\mathcal{ALC}$

F. Baader, I. Horrocks, C. Lutz, U. Sattler: **An Introduction to Description Logic.** Cambridge University Press (2017)
F. Baader, S. Brandt, C. Lutz: **Pushing the $\mathcal{EL}$ Envelope.** IJCAI 2005
D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, R. Rosati: ***DL-Lite:* tractable description logics for ontologies.** AAAI 2005
I. Horrocks, O. Kutz, U. Sattler: **The even more irresistible $\mathcal{SROIQ}$.** KR 2006
Y. Kazakov: **$\mathcal{RIQ}$ and $\mathcal{SROIQ}$ are harder than $\mathcal{SHOIQ}$.** KR 2008
P. Hitzler, M. Krötzsch, B. Parsia, P. F. Patel-Schneider, S. Rudolph: **OWL 2 Web Ontology Language Primer.** W3C Recommendation (2012)

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 $\mathcal{EL}$):

$$\begin{aligned}\mathcal{T} \coloneqq \{ \quad \mathsf{MountainBike} &\sqsubseteq \mathsf{Bike}, \\ \mathsf{Bike} &\sqsubseteq \exists \mathsf{hasPart}.\mathsf{SuspensionFork} \sqcap \exists \mathsf{isSuitableFor}.\mathsf{OffRoadCycling}, \\ \mathsf{SuspensionFork} &\sqsubseteq \mathsf{Fork}, \\ \mathsf{OffRoadCycling} &\sqsubseteq \mathsf{Cycling} \quad \}\end{aligned}$$

$$\mathcal{A} \coloneqq \{ \mathsf{rides}(\mathsf{Francesco}, x), \ \mathsf{Bike}(x) \}$$

- 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 $\mathcal{O}$ entails an axiom α, written $\mathcal{O} \models \alpha$, if each model of $\mathcal{O}$ is a model of α.
- 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*

Y. Kazakov, M. Krötzsch, F. Simancik: **The incredible ELK - from polynomial procedures to efficient reasoning with $\mathcal{EL}$ ontologies.** J. Autom. Reason. (2014)
A. Steigmiller, T. Liebig, B. Glimm: **Konclude: system description.** J. Web Semant. (2014)
D. Tena Cucala, B. Cuenca Grau, I. Horrocks: **Pay-as-you-go consequence-based reasoning for the description logic $\mathcal{SROIQ}$.** Artif. Intell. (2021)
B. Parsia, N. Matentzoglu, R. S. Gonçalves, B. Glimm, A. Steigmiller: **The OWL Reasoner Evaluation (ORE) 2015 Competition Report.** J. Autom. Reason. (2017)

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 $\mathcal{T}$ entails two faulty consequences
  1. $\mathsf{Bike} \sqsubseteq \exists \mathsf{hasPart}.\mathsf{SuspensionFork}$
  2. $\mathsf{Bike} \sqsubseteq \exists \mathsf{isSuitableFor}.\mathsf{OffRoadCycling}$

  } repair request $\mathcal{P}$

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 $\mathcal{P}$.

R. Reiter: **A theory of diagnosis from first principles.** Artif. Intell. (1987)
R. Greiner, B. A. Smith, R. W. Wilkerson: **A correction to the algorithm in Reiter's theory of diagnosis.** Artif. Intell. (1989)
F. Baader, R. Peñaloza, B. Suntisrivaraporn: **Pinpointing in the description logic $\mathcal{EL}^+$.** KI 2007
F. Baader, R. Peñaloza: **Automata-based axiom pinpointing.** J. Autom. Reason. (2010)
F. Baader, R. Peñaloza: **Axiom pinpointing in general tableaux.** J. Log. Comput. (2010)

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

$$\{ \quad \text{MountainBike} \sqsubseteq \text{Bike},$$
$$\sout{\text{Bike} \sqsubseteq \exists \text{hasPart}.\text{SuspensionFork} \sqcap \exists \text{isSuitableFor}.\text{OffRoadCycling}},$$
$$\text{SuspensionFork} \sqsubseteq \text{Fork},$$
$$\text{OffRoadCycling} \sqsubseteq \text{Cycling} \quad \}$$

R. Reiter: **A theory of diagnosis from first principles.** Artif. Intell. (1987)
R. Greiner, B. A. Smith, R. W. Wilkerson: **A correction to the algorithm in Reiter's theory of diagnosis.** Artif. Intell. (1989)
F. Baader, R. Peñaloza, B. Suntisrivaraporn: **Pinpointing in the description logic $\mathcal{EL}^+$.** KI 2007
F. Baader, R. Peñaloza: **Automata-based axiom pinpointing.** J. Autom. Reason. (2010)
F. Baader, R. Peñaloza: **Axiom pinpointing in general tableaux.** J. Log. Comput. (2010)

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 $\succ$.

F. Baader, F. Kriegel, A. Nuradiansyah, R. Peñaloza: **Making repairs in description logics more gentle.** KR 2018
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 $\succ^{\mathsf{sub}}$ for $\mathcal{EL}$ concept inclusions:
  $C \sqsubseteq D \succ^{\mathsf{sub}} C' \sqsubseteq D'$ if $C = C'$, and $\emptyset \models D \sqsubseteq D'$, and $C' \sqsubseteq D' \not\models C \sqsubseteq D$.
- Problems:
  1. Efficient computation of maximally strong $\succ^{\mathsf{sub}}$-weakenings
  2. Efficient computation of one or all optimal repairs

F. Baader, F. Kriegel, A. Nuradiansyah, R. Peñaloza: **Making repairs in description logics more gentle.** KR 2018
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.

F. Kriegel: **Constructing and extending description logic ontologies using methods of formal concept analysis.** Doctoral thesis (2019)

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.
- 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).

F. Kriegel: **Constructing and extending description logic ontologies using methods of formal concept analysis.** Doctoral thesis (2019)

Generalized-Conclusion Repairs

- Inspired by the gentle repairs w.r.t. $\succ^{\mathsf{sub}}$ 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 $\emptyset \models D \sqsubseteq D'$.

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** $\mathcal{S}$.
  2. Construct the **induced countermodel** $\mathcal{J}_{\mathcal{S}}$.
  3. Replace each concept inclusion $C \sqsubseteq D$ with $C \sqsubseteq D \vee C^{\mathcal{J}_{\mathcal{S}}\mathcal{J}_{\mathcal{S}}}$.
- 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 $\mathcal{T}$ is

$$\{\ \mathsf{MountainBike} \sqsubseteq \mathsf{Bike},$$
$$\mathsf{Bike} \sqsubseteq \sout{\exists \mathsf{hasPart}.\mathsf{SuspensionFork} \sqcap \exists \mathsf{isSuitableFor}.\mathsf{OffRoadCycling}},$$
$$\exists \mathsf{hasPart}.\top \sqcap \exists \mathsf{isSuitableFor}.\top,$$
$$\mathsf{SuspensionFork} \sqsubseteq \mathsf{Fork},$$
$$\mathsf{OffRoadCycling} \sqsubseteq \mathsf{Cycling}\ \}$$

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 $C = C'$.

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!)

Fixed-Premise Repairs

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

$$\{ \quad \mathsf{MountainBike} \sqsubseteq \sout{\mathsf{Bike},}$$

$$\mathsf{Bike} \sqcap \exists \mathsf{hasPart}.\mathsf{SuspensionFork} \sqcap \exists \mathsf{isSuitableFor}.\mathsf{OffRoadCycling},$$

$$\mathsf{Bike} \sqsubseteq \sout{\exists \mathsf{hasPart}.\mathsf{SuspensionFork} \sqcap \exists \mathsf{isSuitableFor}.\mathsf{OffRoadCycling},}$$

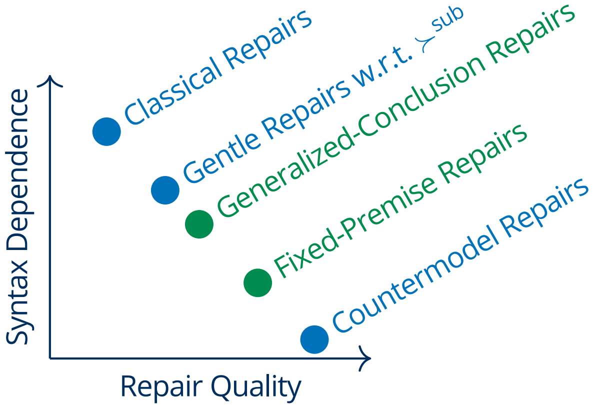
$$\exists \mathsf{hasPart}.\mathsf{Fork} \sqcap \exists \mathsf{isSuitableFor}.\mathsf{Cycling},$$

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

$$\mathsf{OffRoadCycling} \sqsubseteq \mathsf{Cycling} \quad \}$$

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 $R_1 \circ \cdots \circ R_n \sqsubseteq S$
- 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?