Computing Optimal Repairs of Quantified ABoxes w.r.t. Static $\mathcal{EL}$ TBoxes (Extended Abstract)

Franz Baader, **Patrick Koopmann**, **Francesco Kriegel**, **Adrian Nuradiansyah**
Theoretical Computer Science, Technische Universität Dresden
firstname.lastname@tu-dresden.de

Description Logics (DLs) are a family of logic-based knowledge representation languages (Baader et al. 2017), which are frequently used to formalize ontologies for application domains such as biology and medicine (Hoehndorf, Schofield, and Gkoutos 2015). Such ontologies consist of terminological axioms, which specify the important notions of the application domain and are collected in the TBox, and of assertional axioms, which describe properties of specific individuals and objects and are collected in the ABox. As the size of ontologies grows, the likelihood of them containing errors increases as well. This is particularly problematic if the data stored in the ABox are automatically extracted from text or other sources using natural language processing or machine learning. The reasoning services of DL systems (Kazakov, Krötzsch, and Simancik 2014; Glimm et al. 2014; Steigmiller, Liebig, and Glimm 2014; Haarslev et al. 2012), which derive implicit consequences from the explicitly represented knowledge, are not only useful once an ontology is deployed, but can also support debugging by exhibiting consequences that are not supposed to hold in the application domain or should be hidden for other reasons (Grau and Kostylev 2019; Baader, Kriegel, and Nuradiansyah 2019).

Ontology repair Once such unwanted consequences are detected, it is often not easy to see how to repair the ontology in order to get rid of them without losing too many other consequences. Thus, it is important to develop automated tools that support the repair process. Classical repair approaches based on axiom pinpointing (Schlobach and Cornet 2003; Parsia, Sirin, and Kalyanpur 2005; Meyer et al. 2006; Schlobach et al. 2007; Kalyanpur et al. 2007; Baader and Suntisrivaraporn 2008) compute maximal subsets of the ontology that do not have the consequence. The obtained results thus strongly depend on the syntactic form of the axioms, and may remove more knowledge than is actually necessary. To alleviate this problem, novel repair approaches have been developed that replace axioms by weaker ones (in the sense that they have less consequences) instead of removing them completely (Horridge, Parsia, and Sattler 2008; Lam et al. 2008; Troquard et al. 2018; Baader et al. 2018). However, these approaches still depend on the syntactic representation of the ontology, and in general do not yield optimal repairs.

Optimal repairs Intuitively, an optimal repair of an ontology retains as many of its consequences as possible, while no longer having the unwanted ones. More formally, a *repair* of an ontology $\mathcal{O}$ is an ontology $\mathcal{O}'$ that is entailed by $\mathcal{O}$ and does not entail any of the unwanted consequences. It is *optimal* if any repair that lies between $\mathcal{O}$ and $\mathcal{O}'$ w.r.t. entailment is equivalent to $\mathcal{O}'$. Since this notion is purely based on entailment, it is clearly syntax-independent.

Unfortunately, as shown in (Baader et al. 2018), optimal repairs need not exist even for ontologies written using the inexpressive DL $\mathcal{EL}$. In a recent series of articles, we have been trying to identify settings where optimal repairs always exist and can effectively be computed.

First, we considered in (Baader, Kriegel, and Nuradiansyah 2019) the quite restricted case where there is no TBox, where the ABox is an $\mathcal{EL}$ instance store (Horrocks et al. 2004), i.e., an ABox without role assertions, and where each unwanted consequence is an $\mathcal{EL}$ concept assertion. In this setting, optimal repairs always exist and can be computed in exponential time, which is optimal since there may be exponentially many optimal repairs of exponential size.

Second, we expanded these results in (Baader et al. 2020) to the setting of ABoxes with role assertions, but still without any TBox. More precisely, we considered *quantified* ABoxes, in which some individuals are anonymized by viewing them as existentially quantified variables. The possibility of anonymizing individuals allows us to retain more consequences. In this context, optimal repairs still always exist. The set of all optimal repairs can be computed by an exponential-time algorithm with access to an NP-oracle. More precisely, the paper introduces the notion of canonical repairs and shows that the set of canonical repairs contains the set of optimal repairs. The NP-oracle is needed to remove the non-optimal repairs from this set. We also considered a modified version of entailment (called IQ-entailment) in (Baader et al. 2020), where quantified ABoxes are compared w.r.t. which $\mathcal{EL}$ instance relationships they imply. This notion is sufficient if one only wants to pose instance queries. No NP-oracle is needed for computing the set of all IQ-optimal repairs in exponential time.

The new results Our latest article in this direction (Baader et al. 2021) extends the previous results reported above in three respects, which greatly improves their applicability.

1. While the data is expressed as a quantified ABox as before, we now support a static $\mathcal{EL}$ TBox, i.e., its axioms are assumed to be correct and thus cannot be changed. To ensure existence of optimal repairs in the case of classical entailment, this TBox needs to be cycle-restricted in the sense introduced in (Baader, Borgwardt, and Morawska 2012). Again, we define canonical repairs and show that the set of canonical repairs contains the set of all optimal repairs, but the construction needs to be changed in several non-trivial ways to take the TBox into account. Nevertheless, we obtain the same complexity as for the case without TBox. For IQ-entailment, no restriction on the TBox is necessary and we can again dispense with the NP-oracle.
2. The construction of the canonical repairs described in (Baader et al. 2020) and extended in (Baader et al. 2021) such that it can deal with static $\mathcal{EL}$ TBoxes, is best case exponential. The second important contribution of (Baader et al. 2021) is the design of a new construction, both for classical and IQ-entailment, that is exponential only in the worst case. Basically, the unoptimized construction pre-emptively introduces an exponential number of copies for all objects, whereas the optimized one introduces such copies only if needed, and thus may get by with considerably less copies.
3. To find out whether the introduced repair approaches are in principle viable for non-trivial ontologies, we implemented them[1] and performed experiments on ontologies taken from the OWL Reasoner Competition for the track OWL EL Realisation. In addition to checking how often the implementation was able to compute a repair within a certain timeout, we also compared the sizes of the computed repairs with those of the unoptimized canonical repairs. The experiments show that the optimizations indeed makes our repair approach viable also for fairly large ontologies, at least for the IQ-case.

Acknowledgements

This work has been supported by DFG in project number 430150274 and TRR248 (cpec, grant 389792660).

References

Baader, F., and Suntisrivaraporn, B. 2008. Debugging SNOMED CT Using Axiom Pinpointing in the Description Logic $\mathcal{EL}^+$. In *Proc. of KR-MED'08*.

Baader, F.; Horrocks, I.; Lutz, C.; and Sattler, U. 2017. *An Introduction to Description Logic*. Cambridge University Press.

Baader, F.; Kriegel, F.; Nuradiansyah, A.; and Peñaloza, R. 2018. Making Repairs in Description Logics More Gentle. In *Proc. of KR'18*, 319–328.

Baader, F.; Kriegel, F.; Nuradiansyah, A.; and Peñaloza, R. 2020. Computing Compliant Anonymisations of Quantified ABoxes w.r.t. $\mathcal{EL}$ Policies. In *Proc. of ISWC'20*, 3–20.

Baader, F.; Koopmann, P.; Kriegel, F.; and Nuradiansyah, A. 2021. Computing Optimal Repairs of Quantified ABoxes w.r.t. Static $\mathcal{EL}$ TBoxes. In *Proc. of CADE'21*, 309–326.

Baader, F.; Borgwardt, S.; and Morawska, B. 2012. Extending Unification in $\mathcal{EL}$ towards General TBoxes. In *Proc. of KR'12*, 568–572.

Baader, F.; Kriegel, F.; and Nuradiansyah, A. 2019. Privacy-Preserving Ontology Publishing for $\mathcal{EL}$ Instance Stores. In *Proc. of JELIA'19*, 323–338.

Glimm, B.; Horrocks, I.; Motik, B.; Stoilos, G.; and Wang, Z. 2014. Hermit: An OWL 2 reasoner. *J. Autom. Reason.* 53(3):245–269.

Grau, B. C., and Kostylev, E. V. 2019. Logical Foundations of Linked Data Anonymisation. *J. Artif. Intell. Res.* 64:253–314.

Haarslev, V.; Hidde, K.; Möller, R.; and Wessel, M. 2012. The RacerPro Knowledge Representation and Reasoning System. *Semantic Web* 3(3):267–277.

Hoehndorf, R.; Schofield, P. N.; and Gkoutos, G. V. 2015. The Role of Ontologies in Biological and Biomedical Research: A Functional Perspective. *Brief. Bioinform.* 16(6):1069–1080.

Horridge, M.; Parsia, B.; and Sattler, U. 2008. Laconic and Precise Justifications in OWL. In *Proc. of ISWC'08*, 323–338.

Horrocks, I.; Li, L.; Turi, D.; and Bechhofer, S. 2004. The Instance Store: DL Reasoning with Large Numbers of Individuals. In *Proc. of DL'04*.

Kalyanpur, A.; Parsia, B.; Horridge, M.; and Sirin, E. 2007. Finding All Justifications of OWL DL Entailments. In *Proc. of ISWC'07*, 267–280.

Kazakov, Y.; Krötzsch, M.; and Simancik, F. 2014. The Incredible ELK - From Polynomial Procedures to Efficient Reasoning with $\mathcal{EL}$ Ontologies. *Journal of Automed Reasoning* 53(1):1–61.

Lam, J. S. C.; Sleeman, D. H.; Pan, J. Z.; and Vasconcelos, W. W. 2008. A Fine-Grained Approach to Resolving Unsatisfiable Ontologies. *J. Data Semant.* 10:62–95.

Meyer, T.; Lee, K.; Booth, R.; and Pan, J. Z. 2006. Finding Maximally Satisfiable Terminologies for the Description Logic $\mathcal{ALC}$. In *Proc. of AAAI'06*, 269–274.

Parsia, B.; Sirin, E.; and Kalyanpur, A. 2005. Debugging OWL ontologies. In *Proc. of WWW'05*, 633–640.

Schlobach, S., and Cornet, R. 2003. Non-Standard Reasoning Services for the Debugging of Description Logic Terminologies. In *Proc. of IJCAI'03*, 355–362.

Schlobach, S.; Huang, Z.; Cornet, R.; and Harmelen, F. 2007. Debugging Incoherent Terminologies. *J. Automated Reasoning* 39(3):317–349.

Steigmiller, A.; Liebig, T.; and Glimm, B. 2014. Konclude: System description. *J. Web Semant.* 27-28:78–85.

Troquard, N.; Confalonieri, R.; Galliani, P.; Peñaloza, R.; Porello, D.; and Kutz, O. 2018. Repairing Ontologies via Axiom Weakening. In *Proc. of AAAI'18*, 1981–1988.

---

[1] https://github.com/de-tu-dresden-inf-lat/abox-repairs-wrt-static-tbox