Error-Tolerant Reasoning in $\mathcal{EL}$ w.r.t. Optimal ABox Repairs (Extended Abstract)

Franz Baader[1,2], Francesco Kriegel[1], and Adrian Nuradiansyah[1]

[1]Institute of Theoretical Computer Science, Technische Universität Dresden, 01062 Dresden, Germany
[2]Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI) Dresden/Leipzig, Germany

Keywords
Ontology repair, Description logic EL, Error-tolerant reasoning, Computational complexity

Errors in Description Logic (DL) ontologies are often detected when reasoning yields unintuitive consequences. Consider we have data represented in an ABox using role and concept assertions, such as $\mathit{has\_parent}(\mathit{SOUTH}, \mathit{KIM})$ and $\mathit{Famous}(\mathit{KIM})$, which say that South has the famous parent Kim. Additionally, assume there is background information represented in a TBox using concept inclusions such as $\exists \mathit{has\_parent}.\mathit{Famous} \sqsubseteq \mathit{Rich}$ saying that children of famous parents are rich. Suppose that the user realizes that Kim actually is not a parent of South, i.e., the consequence $\mathit{has\_parent}(\mathit{SOUTH}, \mathit{KIM})$ is incorrect, and requests that it has to be removed.

The classical approach for satisfying the repair request is to construct a maximal subset of the ABox that has none of the unwanted consequences w.r.t. the given TBox. In our example, $\mathit{has\_parent}(\mathit{SOUTH}, \mathit{KIM})$ must be removed from the ABox, but this also removes consequences $\exists\, \mathit{has\_parent}.\, \mathit{Famous}(\mathit{SOUTH})$ and $\mathit{Rich}(\mathit{SOUTH})$ that are not problematic for the user. It might be that the user knows that South has a famous parent, but made an error when naming this parent. Thus, they want to get rid of the erroneous role assertion, but not of the concept instance relationships for South that the ABox entails.

Optimal Repairs

From the above concern, we are interested in repairing the ontology in an optimal way such that the unwanted consequences are removed, but instead of preserving a maximal subset of the input ABox, we intend to keep a maximal set of the unobjected consequences. The problem of computing such optimal repairs was already addressed in our previous work [1, 2, 3, 4]. In particular, in [2] we consider a setting, where the TBoxes are static and formulated in $\mathcal{EL}$, while the ABoxes are generalized into so-called *quantified ABoxes (qABoxes)*, which may contain anonymous individuals represented as existentially quantified variables. Moreover, the setting also considers instance repair requests that consist of $\mathcal{EL}$ concept assertions and the optimality of the repairs is defined w.r.t. *IQ-entailment*, where qABoxes are compared w.r.t. which $\mathcal{EL}$

DL 2023: 36th International Workshop on Description Logics, September 2–4, 2023, Rhodes, Greece
franz.baader@tu-dresden.de (F. Baader); francesco.kriegel@tu-dresden.de (F. Kriegel); adrian.nuradiansyah@tu-dresden.de (A. Nuradiansyah)
0000-0002-4049-221X (F. Baader); 0000-0003-0219-0330 (F. Kriegel); 0000-0002-9047-7624 (A. Nuradiansyah)

instance relationships they imply. It is shown in [2] that how a set of IQ-repairs, called *canonical IQ-repairs*, can be computed in exponential time such that it covers the set of all IQ-repairs w.r.t. IQ-entailment. The optimal ones can then be obtained from this set by removing the ones that are strictly IQ-entailed by another one. Since IQ-entailment between qABoxes can be decided in polynomial time, the set of all optimal IQ-repairs can be computed in exponential time [2]. This complexity cannot be improved since there are qABoxes that have exponentially many optimal repairs or for which a single optimal repair may have exponential size.

As the first contribution of our SAC '23 paper [5], we extend the results of [2] to a setting, where the repair request may also contain role assertions, which we call ABox repair requests. First, we consider role repair requests containing only role assertions, and show that in this case we can compute in polynomial time a single optimal repair covering all repairs. In our example, the qABox $\exists\{x\}.\{\mathit{has\_parent}(\mathit{SOUTH}, x), \mathit{Famous}(x), \mathit{Famous}(\mathit{KIM})\}$ is this optimal repair that allows us to retain the consequences $\mathit{Rich}(\mathit{SOUTH})$ and $\exists\, \mathit{has\_parent}.\mathit{Famous}(\mathit{SOUTH})$, although there is no individual in the qABox that is a parent of *SOUTH*.

For the ABox repair request, one first applies this approach to the subset consisting of its role assertions, and continue with the approach of [2] to the resulting qABox and the concept assertions in the repair request. However, when removing canonical IQ-repairs that are redundant, instead of using IQ-entailment, one needs to use *IRQ-entailment* that is slightly stronger than IQ-entailment by additionally taking role assertions between named individuals into account (Example 4.10 of [5] explains why IRQ-entailment is needed). It is shown in [3] that the set of all canonical IQ-repairs also covers the set of all IRQ-repairs w.r.t. IRQ-entailment and checking IRQ-entailment between qABoxes can also be done in polynomial time. With these results in place, we can compute the set of all optimal IRQ-repairs for ABox repair requests in exponential time. The repair approach in [4] can also be used to deal with role assertions in the repair request. However, they need to be expressed as concept assertions involving nominals and, together with the unwanted concept assertions, they are removed at once rather than in a two-stage approach as described above.

Error-Tolerant Reasoning

In general, a given repair problem may have exponentially many optimal repairs, both in the classical and in the optimal sense, and thus it is often hard to decide which one to use. Error-tolerant reasoning does not commit to a single repair, but rather reasons w.r.t. all repairs (within the classical or the optimal setting). Brave entailment asks if a given query is entailed by some repair, whereas cautious entailment asks if the query is entailed by all repairs. This was originally considered for classical TBox repairs [6, 7] in which brave entailment is NP-complete and cautious entailment is coNP-complete. For more expressive DLs that create inconsistencies, error-tolerant reasoning was also considered under the name of inconsistent-tolerant reasoning that considers errors as an inconsistency and uses the classical notion of repair [8, 9, 10].

In our RuleML+RR '22 paper [11], we look into error-tolerant reasoning w.r.t. optimal IQ-repairs for instance repair requests, and in our SAC '23 paper [5], we study error-tolerant reasoning w.r.t. optimal IRQ-repairs for ABox repair requests. We show that brave entailment can be reduced to the instance problem in $\mathcal{EL}$ in the setting of [11] and to the $\mathcal{EL}$ ABox entailment problem in the setting of [5]. We particularly use the fact that each repair is entailed by some optimal repair to prove that a given set of ABox assertions is a brave consequence iff it is itself a repair. This shows that brave reasoning is tractable.

Dealing with cautious reasoning is more involved since we need to check what is entailed by all optimal repairs, each of which can have exponential size [12]. To this end, we need to look closer at how optimal repairs are constructed. Since the optimal IQ- and IRQ-repairs can be obtained from the set of canonical IQ-repairs and each canonical IQ-repair is induced by a seed function of polynomial size [2], we showed in [11] and [5] that the cautious entailment problem can be solved by using only seed functions rather than exponentially large induced canonical repairs.

To make this solution feasible, we first showed that the instance problem w.r.t. canonical repairs can be solved in polynomial time in the size of the seed function. Secondly, since not all canonical repairs are optimal, we must be able to decide if a given seed function induces an optimal repair. We were able to show that this problem can also be solved in polynomial time. This yields a coNP-procedure for deciding cautious reasoning: to show that a set of assertions is not a cautious consequence, we guess a seed function and check whether it is minimal and the induced canonical repair does not entail one of the given assertions. If we consider empty TBox and instance queries, then we were able to show that cautious entailment w.r.t. optimal IQ-repairs for instance repair requests can be decided in polynomial time [11].

There is no matching lower bound yet for cautious entailment in both settings. However, if we consider classical repairs [13, 14, 15] rather than the optimal ones, then we can show that cautious reasoning is coNP hard and brave reasoning becomes NP-hard. The NP-hardness of cautious non-entailment is obtained by a reduction from the NP-complete *path via a node* problem [16, 17], while the NP-hardness of the brave entailment via a reduction from the NP-complete *monotone 1-in-3-SAT* problem [18].

Our approach for checking brave entailment can be used to support computing an optimal repair that retains the consequences that are wanted. By viewing the set of such consequences as a brave consequence, if this set is a repair, then, as shown in [11], we can compute in polynomial time a seed function that induces an optimal repair that entails the given wanted consequences. On the other hand, as argued in [6, 7], since the maintaining organization might need long response time until a repair has published, one can still use cautious entailment to reason w.r.t. the erroneous ontology. Furthermore, if we turn our attention into a privacy-preserving scenario, then one can also use cautious entailment to define a censor [19] that prevents the user to infer sensitive information. The reason is that, in contrast to brave entailment, the set of cautious consequences is closed under (classical) entailment.

The full paper of [11] is published in the proceedings of the 6th International Joint Conference on Rules and Reasoning (RuleML+RR 2022), while the full paper of [5] is published in the proceedings of the 38th Annual ACM Symposium on Applied Computing (SAC 2023).

Acknowledgments

This work was partially supported by the AI competence center ScaDS.AI Dresden/Leipzig and the German Research Foundation (DFG) in Project 430150274 and SFB/TRR 248.

References

[1] F. Baader, F. Kriegel, A. Nuradiansyah, R. Peñaloza, Computing compliant anonymisations of quantified ABoxes w.r.t. $\mathcal{EL}$ policies, in: J. Z. Pan, V. A. M. Tamma, C. d'Amato, K. Janowicz, B. Fu, A. Polleres, O. Seneviratne, L. Kagal (Eds.), The Semantic Web - ISWC 2020 - 19th International Semantic Web Conference, Athens, Greece, November 2-6, 2020, Proceedings, Part I, volume 12506 of *Lecture Notes in Computer Science*, Springer, 2020, pp. 3–20. doi:10.1007/978-3-030-62419-4_1.

[2] F. Baader, P. Koopmann, F. Kriegel, A. Nuradiansyah, Computing optimal repairs of quantified ABoxes w.r.t. static $\mathcal{EL}$ TBoxes, in: A. Platzer, G. Sutcliffe (Eds.), Automated Deduction - CADE 28 - 28th International Conference on Automated Deduction, Virtual Event, July 12-15, 2021, Proceedings, volume 12699 of *Lecture Notes in Computer Science*, Springer, 2021, pp. 309–326. doi:10.1007/978-3-030-79876-5_18.

[3] F. Baader, P. Koopmann, F. Kriegel, A. Nuradiansyah, Optimal abox repair w.r.t. static $\mathcal{EL}$ tboxes: From quantified aboxes back to aboxes, in: P. Groth, M. Vidal, F. M. Suchanek, P. A. Szekely, P. Kapanipathi, C. Pesquita, H. Skaf-Molli, M. Tamper (Eds.), The Semantic Web - 19th International Conference, ESWC 2022, Hersonissos, Crete, Greece, May 29 - June 2, 2022, Proceedings, volume 13261 of *Lecture Notes in Computer Science*, Springer, 2022, pp. 130–146. doi:10.1007/978-3-031-06981-9_8.

[4] F. Baader, F. Kriegel, Pushing optimal abox repair from $\mathcal{EL}$ towards more expressive horn-dls, in: G. Kern-Isberner, G. Lakemeyer, T. Meyer (Eds.), Proceedings of the 19th International Conference on Principles of Knowledge Representation and Reasoning, KR 2022, Haifa, Israel. July 31 - August 5, 2022, 2022. doi:10.24963/kr.2022/3.

[5] F. Baader, F. Kriegel, A. Nuradiansyah, Treating role assertions as first-class citizens in repair and error-tolerant reasoning, in: Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing (SAC '23), March 27–31, 2023, Tallinn, Estonia, Association for Computing Machinery, 2023, pp. 974–982. doi:10.1145/3555776.3577630.

[6] M. Ludwig, R. Peñaloza, Error-tolerant reasoning in the description logic $\mathcal{EL}$, in: E. Fermé, J. Leite (Eds.), Logics in Artificial Intelligence - 14th European Conference, JELIA 2014, Funchal, Madeira, Portugal, September 24-26, 2014. Proceedings, volume 8761 of *Lecture Notes in Computer Science*, Springer, 2014, pp. 107–121. doi:10.1007/978-3-319-11558-0_8.

[7] R. Peñaloza, Error-tolerance and error management in lightweight description logics, Künstliche Intell. 34 (2020) 491–500. doi:10.1007/s13218-020-00684-5.

[8] M. Bienvenu, R. Rosati, Tractable approximations of consistent query answering for robust ontology-based data access, in: F. Rossi (Ed.), IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013, IJCAI/AAAI, 2013, pp. 775–781. URL: https://www.ijcai.org/Proceedings/13/Papers/121.pdf.

[9] A. Calì, D. Lembo, R. Rosati, On the decidability and complexity of query answering over inconsistent and incomplete databases, in: F. Neven, C. Beeri, T. Milo (Eds.), Proceedings of the Twenty-Second ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, June 9-12, 2003, San Diego, CA, USA, ACM, 2003, pp. 260–271. doi:10.1145/773153.773179.

[10] D. Lembo, M. Lenzerini, R. Rosati, M. Ruzzi, D. F. Savo, Inconsistency-tolerant query

answering in ontology-based data access, J. Web Semant. 33 (2015) 3–29. doi:10.1016/j.websem.2015.04.002.

[11] F. Baader, F. Kriegel, A. Nuradiansyah, Error-tolerant reasoning in the description logic $\mathcal{EL}$ based on optimal repairs, in: G. Governatori, A. Turhan (Eds.), Rules and Reasoning - 6th International Joint Conference on Rules and Reasoning, RuleML+RR 2022, Berlin, Germany, September 26-28, 2022, Proceedings, volume 13752 of *Lecture Notes in Computer Science*, Springer, 2022, pp. 227–243. doi:10.1007/978-3-031-21541-4_15.

[12] F. Baader, F. Kriegel, A. Nuradiansyah, Privacy-preserving ontology publishing for $\mathcal{EL}$ instance stores, in: F. Calimeri, N. Leone, M. Manna (Eds.), Logics in Artificial Intelligence - 16th European Conference, JELIA 2019, Rende, Italy, May 7-11, 2019, Proceedings, volume 11468 of *Lecture Notes in Computer Science*, Springer, 2019, pp. 323–338. doi:10.1007/978-3-030-19570-0_21.

[13] A. Kalyanpur, B. Parsia, M. Horridge, E. Sirin, Finding all justifications of OWL DL entailments, in: Proc. of ISWC'07, volume 4825 of *Lecture Notes in Computer Science*, Springer-Verlag, 2007, pp. 267–280. doi:10.1007/11762256_15.

[14] S. Schlobach, Debugging and semantic clarification by pinpointing, in: A. Gómez-Pérez, J. Euzenat (Eds.), The Semantic Web: Research and Applications, Second European Semantic Web Conference, ESWC 2005, Heraklion, Crete, Greece, May 29 - June 1, 2005, Proceedings, volume 3532 of *Lecture Notes in Computer Science*, Springer, 2005, pp. 226–240. doi:10.1007/11431053_16.

[15] S. Schlobach, Diagnosing terminologies, in: M. M. Veloso, S. Kambhampati (Eds.), Proceedings, The Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference, July 9-13, 2005, Pittsburgh, Pennsylvania, USA, AAAI Press / The MIT Press, 2005, pp. 670–675. URL: https://www.aaai.org/Papers/AAAI/2005/AAAI05-105.pdf.

[16] A. S. LaPaugh, C. H. Papadimitriou, The even-path problem for graphs and digraphs, Networks 14 (1984) 507–513. doi:10.1002/net.3230140403.

[17] S. Fortune, J. E. Hopcroft, J. Wyllie, The directed subgraph homeomorphism problem, Theor. Comput. Sci. 10 (1980) 111–121. doi:10.1016/0304-3975(80)90009-2.

[18] T. J. Schaefer, The complexity of satisfiability problems, in: R. J. Lipton, W. A. Burkhard, W. J. Savitch, E. P. Friedman, A. V. Aho (Eds.), Proceedings of the 10th Annual ACM Symposium on Theory of Computing, May 1-3, 1978, San Diego, California, USA, ACM, 1978, pp. 216–226. doi:10.1145/800133.804350.

[19] G. Cima, D. Lembo, R. Rosati, D. F. Savo, Controlled query evaluation in description logics through instance indistinguishability, in: C. Bessiere (Ed.), Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020, ijcai.org, 2020, pp. 1791–1797. doi:10.24963/ijcai.2020/248.