Signature-Based Abduction with Fresh Individuals and Complex Concepts for Description Logics (Extended Abstract) ⋆

Patrick Koopmann

Institute for Theoretical Computer Science, Technische Universität Dresden, Germany

In abduction, we are given a KB as background knowledge, in combination with a set of facts (the observation) that cannot be deduced from the background knowledge. We are then looking for the missing piece in the background knowledge (the hypothesis) that is needed to make the observation logically entailed [12]. This form of reasoning has many applications: 1) it can be used to *explain* why something cannot be deduced [6,7], to supplement services explaining positive entailments such as justifications [22,4,15] and proofs [1,2], 2) it can be used for *diagnosis* tasks, giving the hypothesis as possible explanation for an unexpected observation [19], and 3) it can be used in *KB repair* to give hints on how to fix missing entailments [23].

There is a variety of research on abduction with description logics. Based on the shape of the hypothesis, one distinguishes between concept abduction [5], TBox abduction [10,23], ABox abduction [8,7,21,20,6,11,9,14,16] and KB abduction [18,12]. We focus on a variant called *signature-based ABox abduction* defined as follows, where by *flat ABox*, we refer to an ABox that does not use complex concepts.

Definition 1. *Let $\mathcal{L}$ be a DL, and denote for an ABox $\mathcal{A}$ by $\mathsf{sig}(\mathcal{A})$ the concept and role names in $\mathcal{A}$, and by $\mathsf{size}(\mathcal{A})$ its size. An $\mathcal{L}$* abduction problem *is then given by a triple $\mathfrak{A} = \langle \mathcal{K}, \Phi, \Sigma \rangle$ with $\mathcal{K}$ an $\mathcal{L}$ KB of* background knowledge*, Φ an $\mathcal{L}$ ABox called the* observation*, and $\Sigma \subseteq \mathsf{N_C} \cup \mathsf{N_R}$ a signature of* abducibles*; and asks whether there exists a* hypothesis for $\mathfrak{A}$*, i.e. an $\mathcal{L}$ ABox $\mathcal{H}$ satisfying*

$$\textbf{\textit{A1.}}\ \mathcal{K} \cup \mathcal{H} \not\models \bot, \quad \textbf{\textit{A2.}}\ \mathcal{K} \cup \mathcal{H} \models \Phi, \text{ and } \quad \textbf{\textit{A3.}}\ \mathsf{sig}(\mathcal{H}) \subseteq \Sigma.$$

If we require $\mathcal{H}$ additionally to be flat, we speak of a flat abduction problem*. A* size-restricted (flat) $\mathcal{L}$ abduction problem *is a tuple $\mathfrak{A} = \langle \mathcal{K}, \Phi, \Sigma, n \rangle$ s.t. $\mathfrak{A}' = \langle \mathcal{K}, \Phi, \Sigma \rangle$ is a (flat) $\mathcal{L}$ abduction problem and n is a number encoded in binary. A* hypothesis for $\mathfrak{A}$ *is then an $\mathcal{L}$ ABox $\mathcal{H}$ which is a hypothesis for $\mathfrak{A}'$ and additionally satisfies $\mathsf{size}(\mathcal{H}) \leq n$.*

As a simplified application example from the geology domain, assume we have observed that in an area near a canal, holes appeared in the street as a result of subsidence due to an unstable ground. A possible explanation could

⋆ This work was supported by the DFG in grant 389792660 as part of TRR 248.

involve the presence of a formation of so-called *evaporite* below the street, which dissolves when in contact with water [13]. Our background knowledge consists of a geology ontology together with data about the area. Among others, it contains the following abbreviated axioms:

1. $\mathrm{EvaFor} \sqcap \exists \mathrm{bord}.(\mathrm{Wat} \sqcap \neg \exists \mathrm{lin}.\mathrm{WatPro}) \sqsubseteq \exists \mathrm{aff}.\mathrm{Dis}$
2. $\mathrm{EvaFor} \sqcap \exists \mathrm{aff}.\mathrm{Dis} \sqsubseteq \forall \mathrm{abov}.\mathrm{Unst}$
3. $(\mathrm{Wat} \sqcup \mathrm{Str}) \sqcap \mathrm{EvaFo} \sqsubseteq \bot$ 4. $\mathrm{Wat}(\mathit{can})$ 5. $\mathrm{Str}(\mathit{str})$,

which state that 1. an **Eva**porite **For**mation which **bo**rders to a **Wat**erway without **Wat**er-**Pro**of **lin**ing will be **aff**ected by **Dis**solution; 2. all ground **abov**e an evaporite formation affected by dissolution is **Uns**table; 3. waterways and **Str**eets are not evaporite formations; 4. *can* is a waterway; 5. *str* is a street. Our observation would be that the street is unstable: $\{\ \mathrm{Unst}(\mathit{str})\ \}$, and we are looking for a hypothesis that uses sufficiently precise vocabulary, and only refers to aspects we have incomplete knowledge about and that can later be verified by a team of geologists: $\Sigma = \{\mathrm{EvaFor}, \mathrm{abov}, \mathrm{bord}, \mathrm{lin}, \ldots\}$. A hypothesis for the resulting abduction problem would then be

$$\mathcal{H} = \{\quad \mathrm{EvaFor}(e),\ \mathrm{abov}(e, \mathit{str}),\ \mathrm{bord}(e, \mathit{can}),\ \forall \mathrm{lin}.\bot(\mathit{can}) \quad\}$$

stating that there is an evaporite formation e below the street that borders with the canal, and that the canal has no lining. Note that this hypothesis uses a fresh individual name e, as well as a complex concept $\forall\mathrm{lin}.\bot$. The aim of Σ is to restrict to hypotheses that have explanatory character. In the present example, we would for instance also exclude *aff* and *Dis* from Σ, as the dissolution alone would be a too shallow explanation.

Works on signature-based ABox abduction often restrict hypotheses to flat ABoxes with a given set of individuals [7,21,9]—which means that statements in a hypothesis can be picked from a finite set—or they restrict to *rewritable* DLs [11,6]. As with DLs, we usually have the open-world semantics, in which not all individuals are known, and DLs offer much more expressivity, abduction admitting both fresh individuals and complex concepts in the result is well-motivated. Techniques for practical signature-based ABox and KB abduction with complex concepts are presented in [18,8], for a stricter variant where solutions are required to cover all possible solutions, and may use operators from a more expressive DL, however without a theoretical analysis of the problem in terms of complexity. We fill this gap by answering two questions: 1) what is the complexity of deciding whether a solution to the abduction problem exists, and 2) what is the size of the smallest hypothesis in the worst case. Our results are:

1. Both flat and non-flat ABox abduction for $\mathcal{EL}$ always admit polynomially sized hypotheses, whose existence can be decided in polynomial time.
2. Flat ABox abduction is closely related to the query-emptiness problem [3], and one obtains similar complexity bounds. Here, the size of a hypothesis may become exponential already for $\mathcal{EL}_\bot$, it is exponentially bounded for $\mathcal{ALCI}$, and a bound is not computable in general for $\mathcal{ALCF}$. Deciding the

flat ABox abduction problem is ExpTime-complete for $\mathcal{EL}_\bot$, coNExpTime-complete for $\mathcal{ALC}$ and $\mathcal{ALCI}$, and undecidable for $\mathcal{ALCF}$.
3. For $\mathcal{EL}_\bot$, admitting complex concepts is only interesting if we additionally forbid fresh individuals in the hypothesis. Then, they can become double exponential in size, while their existence can still be decided in ExpTime.
4. The most challenging problem turned out to be the case of general ABox abduction in more expressive DLs. For $\mathcal{ALC}$, we found a tight bound on the size of hypotheses which is triple exponential in the input. For deciding their existence, we showed an N2ExpTime$^{\mathrm{NP}}$ upper bound.
5. Finally, the size-bounded abduction problem is NP-complete for $\mathcal{EL}$, it is NExpTime$^{\mathrm{NP}}$-complete for the flat variant in $\mathcal{ALC}$, and in 2ExpTime for $\mathcal{ALCQI}$.

This is an extended abstract of a paper accepted at IJCAI 2021 [17].

References

1. Alrabbaa, C., Baader, F., Borgwardt, S., Koopmann, P., Kovtunova, A.: Finding small proofs for description logic entailments: Theory and practice. In: Albert, E., Kovacs, L. (eds.) LPAR-23: 23rd International Conference on Logic for Programming, Artificial Intelligence and Reasoning. EPiC Series in Computing, vol. 73, pp. 32–67. EasyChair (2020). https://doi.org/https://dx.doi.org/10.29007/nhpp
2. Alrabbaa, C., Baader, F., Borgwardt, S., Koopmann, P., Kovtunova, A.: Finding good proofs for description logic entailments using recursive quality measures. In: Proceedings of the 28th International Conference on Automated Deduction (CADE-28), July 11–16, 2021, Virtual Event, United States (2021), to appear.
3. Baader, F., Bienvenu, M., Lutz, C., Wolter, F.: Query and predicate emptiness in ontology-based data access. J. Artif. Intell. Res. **56**, 1–59 (2016). https://doi.org/10.1613/jair.4866, https://doi.org/10.1613/jair.4866
4. Baader, F., Peñaloza, R., Suntisrivaraporn, B.: Pinpointing in the description logic $\mathcal{EL}^+$. In: KI 2007: Advances in Artificial Intelligence, 30th Annual German Conference on AI, KI 2007, Osnabrück, Germany, September 10-13, 2007, Proceedings. pp. 52–67 (2007)
5. Bienvenu, M.: Complexity of abduction in the $\mathcal{EL}$ family of lightweight description logics. In: Proceedings of KR 2008. pp. 220–230. AAAI Press (2008), http://www.aaai.org/Library/KR/2008/kr08-022.php
6. Calvanese, D., Ortiz, M., Simkus, M., Stefanoni, G.: Reasoning about explanations for negative query answers in DL-Lite. J. Artif. Intell. Res. **48**, 635–669 (2013). https://doi.org/10.1613/jair.3870, https://doi.org/10.1613/jair.3870
7. Ceylan, İ.İ., Lukasiewicz, T., Malizia, E., Molinaro, C., Vaicenavicius, A.: Explanations for negative query answers under existential rules. In: Calvanese, D., Erdem, E., Thielscher, M. (eds.) Proceedings of KR 2020. pp. 223–232. AAAI Press (2020). https://doi.org/10.24963/kr.2020/23, https://doi.org/10.24963/kr.2020/23
8. Del-Pinto, W., Schmidt, R.A.: ABox abduction via forgetting in $\mathcal{ALC}$. In: The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019. pp. 2768–2775. AAAI Press (2019). https://doi.org/10.1609/aaai.v33i01.33012768, https://doi.org/10.1609/aaai.v33i01.33012768

9. Du, J., Qi, G., Shen, Y., Pan, J.Z.: Towards practical ABox abduction in large description logic ontologies. Int. J. Semantic Web Inf. Syst. **8**(2), 1–33 (2012). https://doi.org/10.4018/jswis.2012040101, https://doi.org/10.4018/jswis.2012040101
10. Du, J., Wan, H., Ma, H.: Practical TBox abduction based on justification patterns. In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. pp. 1100–1106 (2017), http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14402
11. Du, J., Wang, K., Shen, Y.: A tractable approach to ABox abduction over description logic ontologies. In: Brodley, C.E., Stone, P. (eds.) Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence. pp. 1034–1040. AAAI Press (2014), http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8191
12. Elsenbroich, C., Kutz, O., Sattler, U.: A case for abductive reasoning over ontologies. In: Proceedings of the OWLED'06 Workshop on OWL: Experiences and Directions (2006), http://ceur-ws.org/Vol-216/submission_25.pdf
13. Fidelibus, M.D., Gutiérrez, F., Spilotro, G.: Human-induced hydrogeological changes and sinkholes in the coastal gypsum karst of Lesina Marina area (Foggia Province, Italy). Engineering Geology **118**(1-2), 1–19 (2011)
14. Halland, K., Britz, K.: ABox abduction in $\mathcal{ALC}$ using a DL tableau. In: 2012 South African Institute of Computer Scientists and Information Technologists Conference, SAICSIT '12. pp. 51–58 (2012). https://doi.org/10.1145/2389836.2389843, https://doi.org/10.1145/2389836.2389843
15. Horridge, M.: Justification Based Explanation in Ontologies. Ph.D. thesis, University of Manchester, UK (2011), https://www.research.manchester.ac.uk/portal/files/54511395/FULL_TEXT.PDF
16. Klarman, S., Endriss, U., Schlobach, S.: ABox abduction in the description logic $\mathcal{ALC}$. Journal of Automated Reasoning **46**(1), 43–80 (2011). https://doi.org/10.1007/s10817-010-9168-z, https://doi.org/10.1007/s10817-010-9168-z
17. Koopmann, P.: Signature-based abduction with fresh individuals and complex concepts for description logics. In: Zhou, Z. (ed.) Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI 2021, Virtual Event / Montreal, Canada, 19-27 August 2021. pp. 1929–1935. ijcai.org (2021). https://doi.org/10.24963/ijcai.2021/266
18. Koopmann, P., Del-Pinto, W., Tourret, S., Schmidt, R.A.: Signature-based abduction for expressive description logics. In: Calvanese, D., Erdem, E., Thielscher, M. (eds.) Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020. pp. 592–602. AAAI Press (2020). https://doi.org/10.24963/kr.2020/59
19. Obeid, M., Obeid, Z., Moubaiddin, A., Obeid, N.: Using description logic and Abox abduction to capture medical diagnosis. In: Wotawa, F., Friedrich, G., Pill, I., Koitz-Hristov, R., Ali, M. (eds.) Advances and Trends in Artificial Intelligence. From Theory to Practice. pp. 376–388. Springer International Publishing, Cham (2019)
20. Pukancová, J., Homola, M.: Tableau-based ABox abduction for the $\mathcal{ALCHO}$ description logic. In: Proceedings of the 30th International Workshop on Description Logics (2017), http://ceur-ws.org/Vol-1879/paper11.pdf
21. Pukancová, J., Homola, M.: The AAA ABox abduction solver. Künstliche Intell. **34**(4), 517–522 (2020). https://doi.org/10.1007/s13218-020-00685-4, https://doi.org/10.1007/s13218-020-00685-4

22. Schlobach, S., Cornet, R.: Non-standard reasoning services for the debugging of description logic terminologies. In: Gottlob, G., Walsh, T. (eds.) Proc. of the 18th Int. Joint Conf. on Artificial Intelligence (IJCAI 2003). pp. 355–362. Morgan Kaufmann, Acapulco, Mexico (2003), http://ijcai.org/Proceedings/03/Papers/053.pdf
23. Wei-Kleiner, F., Dragisic, Z., Lambrix, P.: Abduction framework for repairing incomplete $\mathcal{EL}$ ontologies: Complexity results and algorithms. In: Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence. pp. 1120–1127. AAAI Press (2014), http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8239