Signature-Based Abduction with Fresh Individuals and Complex Concepts for Description Logics

Patrick Koopmann
Institute for Theoretical Computer Science, Technische Universität Dresden
patrick.koopmann@tu-dresden.de

Abstract

Given a knowledge base and an observation as a set of facts, ABox abduction aims at computing a *hypothesis* that, when added to the knowledge base, is sufficient to entail the observation. In signature-based ABox abduction, the hypothesis is further required to use only names from a given set. This form of abduction has applications such as diagnosis, KB repair, or explaining missing entailments. It is possible that hypotheses for a given observation only exist if we admit the use of fresh individuals and/or complex concepts built from the given signature, something most approaches so far do not support or only support with restrictions. In this paper, we investigate the computational complexity of this form of abduction—allowing either fresh individuals, complex concepts, or both—for various description logics, and give size bounds on the hypotheses if they exist.

1 Introduction

Description logics (DLs) are a powerful formalism to describe knowledge bases (KBs) containing both general domain knowledge from a DL ontology and a set of facts (the ABox). Using a DL reasoner, we can then infer information that is implicit in the data, and can be logically deduced based on the ontology [Baader *et al.*, 2017]. Sometimes it is useful to not only reason about what logically follows from a DL KB, but to reason also about what does *not* follow. 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 [Elsenbroich *et al.*, 2006]. This form of reasoning has many applications: 1) it can be used to *explain* why something cannot be deduced [Calvanese *et al.*, 2013], 2) it can be used for *diagnosis* tasks, giving the hypothesis as possible explanation for an unexpected observation [Obeid *et al.*, 2019], and 3) it can be used in *KB repair* to give hints on how to fix missing entailments [Wei-Kleiner *et al.*, 2014].

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 involve the presence of a formation of so-called *evaporite* below the street, which dissolves when in contact with water [Fidelibus *et al.*, 2011]. Our background knowledge consists of a geology ontology together with data about the area. Among others, it contains the following abbreviated axioms:

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

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: $\text{Unst}(a_1)$, and a hypothesis based on our background knowledge would be

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

stating that there is an evaporite formation e below the street that borders with the canal, and that the canal has no lining. A team of geologists can then verify the hypothesis by analysing the canal and the ground below the street. We highlight two aspects of this hypothesis: 1) it refers to a previously unknown individual, the evaporite formation, and 2) it uses a complex (composed) DL concept ($\forall \text{lin}.\bot$). We are interested in hypotheses like this for *signature based abduction*, where we are additionally given a signature Σ of *abducibles*—a vocabulary of names to be used within the hypothesis [Koopmann *et al.*, 2020]. The aim of Σ is to restrict to hypotheses that have explanatory character. In the present example, we would exclude *aff* and *Dis* from Σ, as the dissolution alone would be a too shallow explanation, and *Wat*, because we already know the waterways in the area. Furthermore, we are looking at *ABox abduction*, in which observations and hypotheses are ABoxes, in contrast to TBox abduction [Du *et al.*, 2017; Wei-Kleiner *et al.*, 2014], KB abduction [Koopmann *et al.*, 2020; Elsenbroich *et al.*, 2006] or concept abduction [Bienvenu, 2008].

While there are practical approaches to ABox abduction without signature-restriction [Klarman *et al.*, 2011; Halland and Britz, 2012; Pukancová and Homola, 2017], works on signature-based ABox abduction often restrict hypotheses to flat ABoxes with a given set of individuals [Ceylan *et al.*, 2020; Du *et al.*, 2012]—which essentially means that statements in a hypothesis can be picked from a finite set—or they restrict to *rewritable* DLs which have limited expressivity [Du *et al.*, 2014; Calvanese *et al.*, 2013]. As with DLs, we usually have the open-world semantics, in which not all individuals are known, and DLs offer much more expressivity, it is a natural next step to look also at abduction allowing for fresh individuals and complex concepts in the result. This changes the nature of the abduction problem drastically as there is now an unbounded set of axioms that may occur in a hypothesis. Problems such as "*Does axiom α belong to some/every/an optimal solution?*" [Calvanese *et al.*, 2013; Ceylan *et al.*, 2020] become less helpful while new questions become interesting, such as whether we can give bounds on the number of individuals in a hypothesis, or on the overall size of the hypothesis.

Without understanding the theoretical properties yet, practical methods for signature-based abduction that admit expressive DL concepts in the hypothesis are presented in [Koopmann *et al.*, 2020; Del-Pinto and Schmidt, 2019]. The authors there consider hypotheses that we would call *complete* in the sense that they cover all hypotheses at the same time. To make this possible, solutions are represented in a very expressive DL using non-classical operators such as fixpoints and axiom disjunction. In this setting, ABox abduction can be reduced to uniform interpolation, which however may produce solutions that are triple exponentially large [Zhao and Schmidt, 2017; Lutz and Wolter, 2011]. A natural question is whether this blow-up is really necessary, or whether we can obtain smaller or simpler hypotheses if we drop the requirement of completeness and look for hypotheses in a classical DL that is sufficient to entail the observation.

Unfortunately, our results indicate that this is not the case: if we only allow for fresh individuals but not for complex concepts, hypotheses may require exponentially many assertions for DLs between $\mathcal{EL}_\bot$ and $\mathcal{ALCI}$, while for $\mathcal{ALCF}$ there does not even exist a general bound. If in addition, we allow for complex concepts, we are able to explain more observations, but the explanations may become *triple exponentially large* in comparison to the input. Motivated by this, we also consider a variant of the abduction problem in which we are additionally given a bound on the size of the hypothesis. To summarize, our contributions are the following.

1. we investigate signature-based ABox abduction for DLs ranging from $\mathcal{EL}$ to $\mathcal{ALCQI}$ where hypotheses may use *fresh individuals*, *complex concepts* or *both*,
2. we give tight bounds on the size of hypotheses if they exist,
3. we analyse the computational complexity of deciding whether a hypothesis exists, and
4. we analyse the complexity of deciding whether a hypothesis of bounded size exists.

Proof details are provided in the technical report [Koopmann, 2021].

2 Description Logics and ABox Abduction

We recall the DLs relevant to this paper [Baader *et al.*, 2017] and provide the formal definition of the abduction problem we consider in this paper.

Let $\mathsf{N_C}$, $\mathsf{N_R}$ and $\mathsf{N_I}$ be three pair-wise disjoint sets of respectively *concept*, *role* and *individual names*. A *role* R is either a role name r or an *inverse role* r^-, where $r \in \mathsf{N_R}$. $\mathcal{EL}$ *concepts* are built according to the following syntax rule, where $A \in \mathsf{N_C}$ and $R \in \mathsf{N_R}$:

$$C ::= \top \mid A \mid C \sqcap C \mid \exists R.C$$

More expressive DLs allow for the following additional concepts, where $n \in \mathbb{N}$, and in brackets, we give the name of the corresponding DL:

$$\bot\ (\mathcal{EL}_\bot)\quad \neg C\ (\mathcal{ALC})\quad \leqslant 1R.\top (\mathcal{ALCF})\quad \leqslant nR.C(\mathcal{ALCQ})$$

In each case, all previous constructs are allowed in the DL as well. Using the letter $\mathcal{I}$ in the DL name we express that in the above, R may also be an inverse role. For example, $(\geq nr^-.C)$ is an $\mathcal{ALCQI}$ concept but not an $\mathcal{ALCQ}$-concept, and $\exists r.\bot$ is an $\mathcal{EL}_\bot$ concept. Additional operators are introduced as abbreviations: $C \sqcup D = \neg(\neg C \sqcap \neg D)$ and $\forall R.C = \neg\exists R.\neg C$.

A *KB* is a set of *axioms*, that is, *concept inclusions* (CIs) $C \sqsubseteq D$, *concept assertions* $C(a)$ and *role assertions* $r(a, b)$, where C, D are concepts, $a, b \in \mathsf{N_I}$ and $r \in \mathsf{N_R}$. If a KB contains only concept and role assertions, it is called *ABox*, and if every concept assertion is of the form $A(a)$, where $A \in \mathsf{N_C}$, *flat ABox*. Given a concept/axiom/KB/ABox E, we denote by $\mathsf{sub}(E)$ the set of (sub-)concepts occurring in E, by $\mathsf{sig}(E)$ the set of concept and role names occuring in E, and by $\mathsf{ind}(E)$ the set of individual names in E. By $\mathsf{size}(E)$, we denote the number of symbols required to write E down, where operators, as well as concept, role and individual names count as one, numbers are encoded in binary and the introduced abbreviations can be used.

The semantics of DLs is defined based on *interpretations*, which are tuples $\mathcal{I} = \langle \Delta^\mathcal{I}, \cdot^\mathcal{I} \rangle$ of a set $\Delta^\mathcal{I}$ of *domain elements* and an *interpretation function* $\cdot^\mathcal{I}$ which maps every $a \in \mathsf{N_I}$ to some $a^\mathcal{I} \in \Delta^\mathcal{I}$, every $A \in \mathsf{N_C}$ to some $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$, every $r \in \mathsf{N_R}$ to some $r^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$, satisfies $(r^-)^\mathcal{I} = (r^\mathcal{I})^-$, and is extended to concepts as follows, where $\#S$ denotes the cardinality of the set S:

$$\bot^\mathcal{I} = \emptyset \quad (C \sqcap D)^\mathcal{I} = C^\mathcal{I} \cap D^\mathcal{I} \quad (\neg C)^\mathcal{I} = \Delta^\mathcal{I} \setminus C^\mathcal{I}$$
$$(\exists R.C)^\mathcal{I} = \{d \in \Delta^\mathcal{I} \mid \langle d, e\rangle \in R^\mathcal{I},\ e \in C^\mathcal{I}\}$$
$$(\leqslant nR.C)^\mathcal{I} = \{d \in \Delta^\mathcal{I} \mid \#\{\langle d, e\rangle \in R^\mathcal{I} \mid e \in C^\mathcal{I}\} \leq n\}$$

$\mathcal{I}$ *satisfies* an axiom α, in symbols $\mathcal{I} \models \alpha$, if $\alpha = C \sqsubseteq D$ and $C^\mathcal{I} \subseteq D^\mathcal{I}$; $\alpha = C(a)$ and $a^\mathcal{I} \in C^\mathcal{I}$; or $\alpha = r(a, b)$ and $\langle a^\mathcal{I}, b^\mathcal{I}\rangle \in r^\mathcal{I}$. If $\mathcal{I}$ satisfies all axioms in a KB $\mathcal{K}$, we write $\mathcal{I} \models \mathcal{K}$ and call $\mathcal{I}$ a *model of* $\mathcal{K}$. A KB *entails* an axiom/KB E, in symbols $\mathcal{K} \models E$, if $\mathcal{I} \models E$ for every model $\mathcal{I}$ of $\mathcal{K}$. If $\mathcal{K}$ has no model, it is *inconsistent* and we write $\mathcal{K} \models \bot$.

We can now define the main reasoning problem we are concerned with in this paper. We consider *signature-based ABox abduction problems*, which for convenience, we just call abduction problems from here on.

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

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

If we require $\mathcal{H}$ additionally to be flat, we speak of a flat abduction problem.

Condition **A1** is required to avoid trivial solutions, since everything follows from a basic inconsistency. Condition **A2** ensures the hypothesis is indeed effective, and Condition **A3** is what makes this a *signature-based* abduction problem.

In addition, different minimality criteria can be put on the hypothesis, such as *size-*, *subset-*, *semantic minimality* [Calvanese *et al.*, 2013], or *completeness* [Koopmann *et al.*, 2020]. In this paper, we consider size minimality, for which the corresponding decision problem is the following.

Definition 2. *A* size-restricted (flat) $\mathcal{L}$ abduction problem *is a tuple $\mathfrak{A} = \langle \mathcal{K}, \Phi, \Sigma, n \rangle$, where $\mathfrak{A}' = \langle \mathcal{K}, \Phi, \Sigma \rangle$ is a (flat) $\mathcal{L}$ abduction problem and n is a number encoded in binary. A* hypothesis for $\mathfrak{A}$ *is an $\mathcal{L}$-ABox $\mathcal{H}$ which is a hypothesis for $\mathfrak{A}'$ and additionally satisfies $\mathsf{size}(\mathcal{H}) \leq n$.*

3 Flat ABox Abduction

We first consider flat ABox abduction: the size of hypotheses, and the complexity of deciding their existence. This problem is similar to that of *query emptiness* [Baader *et al.*, 2016] which asks whether for a given set $\mathcal{T}$ of CIs, Boolean query q and signature Σ, there exists a flat ABox $\mathcal{A}$ with $\mathsf{sig}(\mathcal{A}) \subseteq \Sigma$, s.t. $\mathcal{T} \cup \mathcal{A}$ entails that query. Query emptiness for *instance queries* is essentially the special case of flat ABox abduction where the observation is of the form $A(a)$ and the background knowledge contains only CIs. This immediately gives us lower bounds for the decision problem of flat ABox abduction. The other results in this section do not follow, but can be shown similarly as in [Baader *et al.*, 2016]. Similar to query emptiness, flat ABox abduction only becomes interesting if the DL is powerful enough to create inconsistencies. Otherwise, we can construct a trivial hypothesis candidate as

$$\begin{aligned}\mathcal{H} =& \{A(a) \mid A \in (\Sigma \cap \mathsf{N_C}), a \in \mathsf{ind}(\mathcal{K} \cup \Phi)\} \\ & \cup \{r(a,b) \mid r \in (\Sigma \cap \mathsf{N_R}), a, b \in \mathsf{ind}(\mathcal{K} \cup \Phi)\}.\end{aligned}$$

Clearly, $\mathcal{H}$ satisfies **A1** and **A3**. If it does not satisfy **A2**, then neither would any other ABox. Consequently, if there is a solution to the abduction problem, then $\mathcal{H}$ must be such a solution. This means that for $\mathcal{EL}$, flat ABox abduction can always be performed in polynomial time. On the other hand, already for $\mathcal{EL}_\bot$, solutions may require exponentially many fresh individual names.

Theorem 1. *There exists a family $\langle \mathcal{K}_n, \Phi, \Sigma \rangle_{n>0}$ of flat $\mathcal{EL}_\bot$ abduction problems s.t. every hypothesis is of size exponential in the size of $\mathcal{K}_n$.*

Proof. We set $\Phi = A(a)$ and $\Sigma = \{B, r\}$, and let $\mathcal{K}_n$ use concept names $X_1, \overline{X}_1, \ldots, X_n, \overline{X}_n$ to encode an exponential counter. $\mathcal{K}_n$ uses this to ensure that an r-chain of distinct 2^n elements from a to an instance of B entails $A(a)$.

$$\begin{aligned} B &\sqsubseteq \overline{X}_1 \sqcap \ldots \sqcap \overline{X}_n \\ \exists r.(\overline{X}_i \sqcap X_{i-1} \sqcap \ldots X_1) &\sqsubseteq X_i \quad \text{for } i \in [\![1,n]\!] \\ \exists r.(X_i \sqcap X_{i-1} \sqcap \ldots \sqcap X_1) &\sqsubseteq \overline{X}_i \quad \text{for } i \in [\![1,n]\!] \\ \exists r.(\overline{X_i} \sqcap \overline{X}_j) &\sqsubseteq \overline{X}_i \quad \text{for } i,j \in [\![1,n]\!], j < i \\ \exists r.(X_i \sqcap \overline{X}_j) &\sqsubseteq X_i \quad \text{for } i,j \in [\![1,n]\!], j < i \\ X_i \sqcap \overline{X_i} &\sqsubseteq \bot \quad \text{for } i \in [\![1,n]\!] \\ X_1 \sqcap \ldots \sqcap X_n &\sqsubseteq A \end{aligned}$$

The only way to produce this chain of 2^n elements is using $2^n - 1$ role assertions, which establishes our lower bound. □

This bound remains tight if we add expressivity up to $\mathcal{ALCI}$, while we lose any bound on the size once we additionally allow concepts of the form $\leqslant 1 r.\top$.

Theorem 2. *If there exists a hypothesis for a flat $\mathcal{L}$ abduction problem $\mathfrak{A}$, then there exists one of size*

1. *polynomial in the size of $\mathfrak{A}$ if $\mathcal{L} = \mathcal{EL}$,*
2. *exponential in the size of $\mathfrak{A}$ if $\mathcal{L} = \mathcal{ALCI}$, and*
3. *if $\mathcal{L} = \mathcal{ALCF}$, no general upper bound based on $\mathfrak{A}$ can be given.*

Proof sketch. We already established the bound for $\mathcal{L} = \mathcal{EL}$. For $\mathcal{ALCI}$, we assume there exists some hypothesis $\mathcal{H}_0$, based on which we build one of bounded size. For this, we pick any model $\mathcal{I}$ of $\mathcal{H}_0 \cup \mathcal{K}$, which allows us to identify individual names $a \in \mathsf{ind}(\mathcal{H}_0)$ using at most exponentially many *types* $\mathrm{tp}(a) = \mathrm{tp}(a^\mathcal{I})_\mathcal{I}$, where

$$\mathrm{tp}(d)_\mathcal{I} = \{C \in \mathsf{sub}(\mathcal{K} \cup \Phi) \mid d \in C^\mathcal{I}\}.$$

We associate to every type $\mathrm{tp}(a)$ an individual name $b_{\mathrm{tp}(a)}$, and define $h : \mathsf{ind}(\mathcal{H}_0) \to \mathsf{N_I}$ by $h(a) = a$ if $a \in \mathsf{ind}(\mathcal{K} \cup \Phi)$ and $h(a) = b_{\mathrm{tp}(a)}$ otherwise. The hypothesis $\mathcal{H}$ is then:

$$\{A(h(a)) \mid A(a) \in \mathcal{H}_0\} \cup \{r(h(a), h(b)) \mid r(a,b) \in \mathcal{H}_0\}.$$

Based on $\mathcal{I}$ and h, one can construct a model for $\mathcal{K} \cup \mathcal{H}$, so that $\mathcal{H}$ satisfies **A1**. Because $\mathcal{H}_0$ satisfies **A3**, by construction, so does $\mathcal{H}$. Finally, using the fact that h is a homomorphism from $\mathcal{K} \cup \mathcal{H}_0$ into $\mathcal{K} \cup \mathcal{H}$ s.t. for every $a \in \mathsf{ind}(\mathcal{K} \cup \Phi)$ $h(a) = a$, we can show that $\mathcal{K} \cup \mathcal{H} \models \Phi$, and thus **A2**.

For $\mathcal{L} = \mathcal{ALCF}$, we note that if there was a bound on the size of hypotheses, we could decide the instance query emptiness problem for $\mathcal{ALCF}$ by iterating over all candidates, contradicting that this problem is undecidable for $\mathcal{ALCF}$ [Baader *et al.*, 2016]. □

The proof of Theorem 2 indicates how types can be used to perform abduction, which is used in the following theorem.

Theorem 3. *Flat $\mathcal{L}$ ABox abduction is*

- P-*complete for $\mathcal{L} = \mathcal{EL}$,*
- ExpTime-*complete for $\mathcal{L} = \mathcal{EL}_\bot$,*
- coNExpTime-*complete for $\mathcal{L} = \mathcal{ALCI}$,*
- *undecidable for $\mathcal{L} = \mathcal{ALCF}$.*

Proof sketch. For $\mathcal{EL}$, we already described a method to compute and verify hypotheses in P. For the other DLs, we use type elimination to compute in exponential time the set $T_{\mathcal{K}\cup\Phi}$ of types $t \subseteq \mathsf{sub}(\mathcal{K}\cup\Phi)$ that can occur in models of $\mathcal{K}$. For a type t and a set T of types, we denote by $\mathrm{S}^r_T(t) \subseteq T$ the types an r-succesor of a t-individual can have. We iterate over all assignments $s : \mathsf{ind}(\mathcal{K}\cup\Phi) \to T_{\mathcal{K}\cup\Psi}$ for which we set

$$\begin{aligned}\mathcal{H}_s =&\{A(a) \mid a \in I, A \in s(a) \cap \Sigma\}\\ &\cup \{r(a, b) \mid a, b \in I, r \in \Sigma, s(b) \in \mathrm{S}^r_{T_{\mathcal{K}\cup\Phi}}(s(a))\},\end{aligned}$$

where I contains an additional individual name for every type and s is extended to I. If there exists a hypothesis, then the hypothesis constructed in the proof for Theorem 2 is contained in some such $\mathcal{H}_s$ s.t. $\mathcal{K} \cup \mathcal{H}_s \not\models \bot$, so that we can directly check **A1**–**A3** from Def. 2 on the generated $\mathcal{H}_s$. **A1** can be verified syntactically using the types. For $\mathcal{EL}_\bot$, **A2** is verified in time polynomial in the size of $\mathcal{H}_s$ and thus in EXPTIME. For $\mathcal{ALCI}$, we guess a model of $\mathcal{K} \cup \mathcal{H}_s$ that does not entail Φ by assigning a type to every individual and possibly adding a new element per type. If this fails, $\mathcal{K}\cup\mathcal{H}_s \models \Phi$. ☐

4 Abduction with Complex Concepts

As illustrated in the introduction, abduction may only be successful if we also admit complex concepts in the hypothesis. Determining such hypotheses turns out to be more challenging than for flat hypotheses, and we cannot find a correspondence to a known problem as for flat abduction. Indeed, one might assume such a relation to uniform interpolation: given a KB $\mathcal{K}$ and a signature Σ, the uniform interpolant of $\mathcal{K}$ for Σ is a Σ ontology that captures all entailments of $\mathcal{K}$ within Σ [Koopmann and Schmidt, 2015]. By negating the observation, this can be used to perform *complete* abduction [Koopmann *et al.*, 2020; Del-Pinto and Schmidt, 2019], that is, to compute a hypothesis that would be entailed by any other hypothesis. However, if we are interested just in computing any hypothesis rather than a complete one, this correspondence falls short, as uniform interpolants have stronger requirements than hypotheses, and the reasons for non-existence are different: namely, capturing *all* entailments of $\mathcal{K}$ in Σ in the uniform interpolant, using only names from Σ, may require infinitely many axioms in case of cyclic axioms. In contrast, for abduction, non-existence is always due to Condition **A1**.

We consider abduction for $\mathcal{ALC}$, $\mathcal{EL}$, and $\mathcal{EL}_\bot$, starting with the latter. In $\mathcal{EL}$ and $\mathcal{EL}_\bot$, complex concepts do not bring much benefit compared to fresh individuals: an $\mathcal{EL}_\bot$ concept can only state the existence of role successors, which we can also do in flat ABoxes. In fact, for $\mathcal{EL}_\bot$, if we allow complex concepts *instead* of fresh individuals , hypotheses even get more complex.

Theorem 4. *There exists a family of $\mathcal{EL}_\bot$ abduction problems for which every hypothesis without fresh individuals is at least of double exponential size. If there exists such a hypothesis, there always exists one of at most double exponential size, whose existence can be decided in exponential time.*

Proof sketch. The family of abduction problems is obtained similarly as in the proof for Theorem 1, only that we now use two roles r and s To get the corresponding upper bound, we first flatten an existing hypothesis $\mathcal{H}_0$ and again simplify the ABox based on the types in some model, however this time making sure the resulting flat ABox can be translated back into a complex one without fresh individuals.

The same care has to be taken when we modify the method used for Theorem 3. Specifically, we have to make sure that the hypothesis $\mathcal{H}_s$ that we generate for a given mapping $s : \mathsf{ind}(\mathcal{K} \cup \Phi) \to T_{\mathcal{K}\cup\Phi}$ does not contain cycles between fresh individuals, so that it can be translated into a hypothesis without fresh individuals. Our fresh individuals are now of the form $b_{a,t,i}$, where $a \in \mathsf{ind}(\mathcal{K} \cup \Phi)$, $t \in T = T_{\mathcal{K}\cup\Phi}$, and $i \in [\![1, 2^{|T|}]\!]$. Set $b_{a,s(a),0} = a$. $\mathcal{H}_s$ is then defined as:

$$\begin{aligned}\mathcal{H}_s = &\big\{A(b_{a,t,k}) \mid a \in \mathsf{ind}(\mathfrak{A}), t \in T, k \in [\![0, 2^{|T|}]\!],\\ &\qquad A \in (T \cap \Sigma)\big\}\\ &\cup\ \big\{r(b_{a,t_1,k}, b_{a,t_2,k+1}) \mid a \in \mathsf{ind}(\mathcal{K} \cup \Phi), t_1 \in T, r \in \Sigma,\\ &\qquad t_2 \in \mathrm{S}^r_T(t_1), k \in [\![0, 2^{|T|} - 1]\!]\big\}\\ &\cup\ \big\{r(a, b) \mid s(b) \in \mathrm{S}^r_T(s(a))\big\}\end{aligned}$$

☐

The hardness result requires again $\bot$: for $\mathcal{EL}$, we can always use a flat solution as in the last section. In contrast, with more expressivity, the problem becomes even harder, even if we do admit fresh individuals. The reason is that for concepts of the form $\forall r.C$, fresh individuals cannot come to the rescue anymore, and disjunctions may become necessary. The following theorem is shown by a modification of a construction in [Ghilardi *et al.*, 2006].

Theorem 5. *There is a family of $\mathcal{ALC}$ abduction problems for which the smallest (non-flat) ABox hypotheses are triple exponential in size.*

We can show that this bound is tight.

Theorem 6. *Let $\mathfrak{A}$ be an $\mathcal{ALC}$ abduction problem. Then, there exists a hypothesis for $\mathfrak{A}$ iff there exists a hypothesis of triple exponential size.*

To show this theorem, we use a technique similar as for Theorem 2: we take an arbitrary hypothesis, and transform it into one of triple exponential size. However, this time, a construction based on a single model is not sufficient, and we have to take into account an appropriate abstraction of *several* models of $\mathcal{K} \cup \mathcal{H}_0$. We thus proceed as follows:

1. we abstract the KB $\mathcal{K} \cup \mathcal{H}_0$ into a model abstraction,
2. we reduce the size of this abstraction,
3. based on which we construct a hypothesis $\mathcal{H}$ of triple exponential size.

In the model abstraction, elements are represented as nodes $v \in V$ that are labeled with a set $\lambda(v)$ of types with the intuitive meaning "this element may have one of the types in $\lambda(v)$". Role relations are represented using tuples $\langle v_1, t, r, v_2\rangle$ which are read as: if the node v_1 has type t, then it has an r-successor corresponding to v_2. Roughly, from a model abstraction, we can obtain a model using the following inductive procedure: 1) start with the nodes that represent individuals, 2) assign to each node a type from its label set, 3) if for those types, the node requires successor nodes, add those and continue in 2). To allow for unbounded paths in models for finite acyclic model abstractions, we further have "open" nodes whose role successors are only restricted by the TBox.

Definition 3. *An* interpretation abstraction *for* $\langle \mathcal{K}, \Phi, \Sigma \rangle$ *is a tuple* $\mathfrak{I} = \langle V, \lambda, s, \mathfrak{R}, F \rangle$*, where*

- V *is a set of* nodes*,*
- $\lambda : V \to 2^{T_{\mathcal{K}\cup\Phi}}$ *maps each node to a set of types,*
- *partial function* $s : \mathsf{N_I} \nrightarrow V$ *assigns individuals to nodes*
- $\mathfrak{R} \subseteq \big(V \times T_{\mathcal{K}\cup\Phi} \times (\Sigma \cap \mathsf{N_R}) \times V\big)$ *is the* role assignment*,*
- *and* $F \subseteq V$ *is the set of* open nodes*.*

$\mathfrak{I}$ abstracts an interpretation $\mathcal{I}$ *if there is a subset* $\Delta' \subseteq \Delta^{\mathcal{I}}$ *and a function* $h : \Delta' \to V$ *s.t. for every* $d \in \Delta'$ *and* $r \in (\Sigma \cap \mathsf{N_R})$*:*

I1 *for all* $a \in \mathsf{N_I}$*, if* $s(a)$ *is defined, then* $s(a) = h(a^{\mathcal{I}})$

I2 $\mathrm{tp}(d)_{\mathcal{I}} \in \lambda(h(d))$*,*

I3 *if* $h(d) \notin F$*, then for every* $e \in \Delta^{\mathcal{I}}$*,* $\langle d, e \rangle \in r^{\mathcal{I}}$ *iff* $e \in \Delta'$ *and* $\langle h(d), \mathrm{tp}(d)_{\mathcal{I}}, r, h(e) \rangle \in \mathfrak{R}$*.*

We need some additional requirements to make sure an interpretation abstraction can be represented as an $\mathcal{ALC}$ ABox $\mathcal{H}$ s.t. $\mathsf{sig}(\mathcal{H}) \subseteq \Sigma$. We call a node $v \in V$ for which there exists $a \in \mathsf{N_I}$ with $s(a) = v$ *internal node*, and otherwise *outgoing node*. If $v = s(a)$ for $a \in \mathsf{ind}(\mathcal{K} \cup \Phi)$, we call v *named node*. A *path in* $\langle V, \lambda, s, \mathfrak{R}, F \rangle$ is a sequence

$$\pi = v_0, t_0, r_0, v_1, t_1, r_1, \ldots t_{n-1} r_{n-1} v_n$$

s.t. for each $i \in [\![0, n-1]\!]$, $\langle v_i, t_i, r_i, t_{i+1} \rangle \in \mathfrak{R}$. π is cyclic if it contains a node twice, and its *length* is its number of nodes.

Definition 4. $\mathfrak{I} = \langle V, \lambda, s, \mathfrak{R}, F \rangle$ *is called* $\mathcal{ALC}$-conform *if*

D1 *there is no cyclic path between outgoing nodes ,*

D2 *for every internal node* v*, if* $\langle v, t, r, v' \rangle \in \mathfrak{R}$*, then* $\langle v, t', r, v' \rangle \in \mathfrak{R}$ *for every* $t' \in \lambda(v')$*, and*

D3 *for every* $\langle v_1, t, r, v_2 \rangle \in \mathfrak{R}$*, where* v_2 *is internal, there exists* $\langle v_1, t, r, v_2' \rangle \in \mathfrak{R}$ *s.t.* v_2' *is outgoing.*

We say that $\mathfrak{I}$ *is* Σ-complete *if*

D4 *for every* $v \in V$*, and* $t_1 \in \lambda(v)$*,* $\lambda(v)$ *contains every type* $t_2 \in T_{\mathcal{K}\cup\Phi}$ *s.t.* $\lambda(t_1) \cap \Sigma = \lambda(t_2) \cap \Sigma$*, and*

D5 *for every* $\langle v_1, t, r, v_2 \rangle \in \mathfrak{R}$ *and* $t' \in T_{\mathcal{K}\cup\Phi}$ *s.t.* $t \cap \Sigma = t' \cap \Sigma$*, also* $\langle v_1, t', r, v_2 \rangle \in \mathfrak{R}$*.*

D1 ensures that we can represent the outgoing paths from a node $s(a)$ in a single assertion $C(a)$. **D2** expresses that the relations between internal nodes is independent of the type, which allows to represent them using role assertions. **D3** is needed to capture allowed paths using universal role restrictions. **D4** and **D5** ensure that $\mathfrak{I}$ can be captured using only names in Σ.

From here on, we fix an abduction problem $\mathfrak{A} = \langle \mathcal{K}, \Phi, \Sigma \rangle$. We say that $\mathfrak{I}$ *explains* Φ iff some model of $\mathcal{K}$ is abstracted by $\mathfrak{I}$ and every model of $\mathcal{K}$ that is abstracted by $\mathfrak{I}$ entails Φ. If $\mathfrak{I}$ is $\mathcal{ALC}$-conform, Σ-complete, and explains Φ, we call it a *hypothesis abstraction*.

Lemma 1. *Every hypothesis* $\mathcal{H}$ *for* $\mathfrak{A}$ *can be translated into a hypothesis abstraction.*

Now that we can translate hypotheses into an alternative representation, the next step is to decrease their size while making sure they still explain the observation. For this, we can now use similar techniques as in Theorems 2 and 4, where we now identify nodes $v \in V$ based on the their label $\lambda(v)$ instead of on a single type.

Lemma 2. *Let* $\mathfrak{I}$ *be a hypothesis abstraction. Then,* $\mathfrak{I}$ *can be transformed into a hypothesis abstraction* $\mathfrak{I}' = \langle V', \lambda', \mathfrak{R}', F' \rangle$ *where, for* $\ell = |T_{\mathcal{K}\cup\Phi}| \cdot 2^{|T_{\mathcal{K}\cup\Phi}|}$*, i)* V' *contains at most* ℓ *internal nodes, ii) every* $v \in V'$ *has at most* ℓ *successors in* $\mathfrak{R}'$*, and iii) every path of outgoing nodes contains at most* ℓ *nodes.*

The final ingredient to establish Theorem 6 is the following lemma.

Lemma 3. *For every* $\mathcal{ALC}$*-conform,* Σ*-complete interpretation abstraction* $\mathfrak{I} = \langle V, \lambda, s, \mathfrak{R}, F \rangle$*, there exists an ABox* $\mathcal{H}$ *s.t. i) the models of* $\mathcal{K} \cup \mathcal{H}$ *are exactly the models of* $\mathcal{K}$ *accepted by* $\mathfrak{I}$*, ii)* $|\mathsf{ind}(\mathcal{H})| \leq |V|$*, and iii) for every* $a \in \mathsf{ind}(\mathcal{H})$*,* $\mathcal{H}$ *contains one assertion* $C(a)$*, with* $\mathsf{size}(C)$ *exponentially bounded by the path lengths between outgoing nodes in* $\mathfrak{I}$*.*

Unfortunately, interpretation abstractions cannot be as easily constructed by a deterministic procedure as we did for Theorems 3 and 4, as there can to be non-trivial interactions between connected internal nodes, and we only have a double exponential upper bound on their number. To decide $\mathcal{ALC}$ abduction, we can however guess an interpretation abstraction within the bounds of Lemma 2, and then guess assignments of types to nodes to obtain its models. We thus obtain the following theorem.

Theorem 7. $\mathcal{ALC}$ *ABox abduction is in* $\text{N2ExpTime}^{\text{NP}}$.

5 Size-Restricted Abduction

Because hypotheses can become very large, a natural requirement is to compute hypotheses of minimal or bounded size. We here obtain the following complexities.

Theorem 8. *Size restricted* $\mathcal{L}$ *ABox abduction is*

- NP*-complete for* $\mathcal{L} = \mathcal{EL}$*,*
- NExpTime*-complete for* $\mathcal{L} = \mathcal{EL}_\bot$*,*
- $\text{NExpTime}^{\text{NP}}$*-complete for the flat variant and* $\mathcal{L} \in \{\mathcal{ALC}, \mathcal{ALCI}\}$*, and*
- *in* 2ExpTime *for* $\mathcal{L} = \mathcal{ALCQI}$*.*

The upper bounds are based on guess-and-check algorithms. For $\mathcal{EL}$, we exploit the fact that, by Theorem 2, we can always find a solution of polynomial size. For $\mathcal{EL}_\bot$, we note that the size of the hypothesis is exponentially bounded by the number of bits used for the size bound k. The $\text{NExpTime}^{\text{NP}}$-upper bound can be obtained by a refinement of the procedure used in the proof for Theorem 3. For the double exponential upper bound, we iterate over the double exponentially many possible KBs within the size and signature bounds—independent on whether we are interested in flat or complex solutions—and then check for entailment in time exponential in the size of the current solution. The lower bounds are provided by the following lemmas.

Lemma 4. *Size-restricted $\mathcal{EL}$ abduction is* NP*-hard.*

Proof sketch. We reduce the NP-complete problem CNF-SAT to deciding whether a given signature-based problem has a hypothesis of size at most k. Let $\phi = c_1 \wedge \ldots \wedge c_n$ be a CNF formula over propositional variables $p_1, \ldots, p_m$. $\mathcal{K}$ contains the following axioms:

$$\mathsf{True} \sqsubseteq P \quad \mathsf{False} \sqsubseteq P \quad \exists r.\mathsf{True} \sqsubseteq C \quad \exists s.\mathsf{False} \sqsubseteq C$$
$$r(c_i, p_j) \text{ for every } i \in [\![1,n]\!], j \in [\![1,m]\!], \text{ if } p_j \in c_j$$
$$s(c_i, p_j) \text{ for every } i \in [\![1,n]\!], j \in [\![1,m]\!], \text{ if } \neg p_j \in c_j$$

Φ contains $P(p_i)$ for every $i \in [\![1,m]\!]$ and $C(c_i)$ for every $i \in [\![1,n]\!]$. Finally, $\Sigma = \{\mathsf{True}, \mathsf{False}\}$. $\langle \mathcal{K}, \Phi, \Sigma \rangle$ has a hypothesis of size at most $2m$ iff ϕ is satisfiable. □

Lemma 5. *Size restricted $\mathcal{EL}_\bot$ abduction is* NEXPTIME*-hard.*

Proof sketch. The hardness follows from a reduction of the NEXPTIME-complete *exponential tiling problem*, which is given by a tuple $\langle T, T_I, t_e, V, H, n \rangle$ of a set T of tile types, a sequence $T_I \in T^*$ of initial tiles, a final tile t_e, vertical and horizontal tiling conditions $V, H \subseteq T \times T$, and a number n in unary encoding. A solution to this problem is then a *tiling*, as a function $f : [\![1, 2^n]\!] \times [\![1, 2^n]\!] \to T$ assigning tiles to coordinates, s.t. the first tiles are as in T_I, $f(2^n, 2^n) = t_e$, and that obeys the vertical and horizontal tiling conditions [van Emde Boas, 1997].

In the reduction, concept names Start and End respectively mark the initial and the final tile. We implement two binary counters X and Y as for Theorem 1 which are decremented over the roles x and y, and encode the coordinates of the tiles. Each tile type $t \in T$ is represented by a concept name A_t. We enforce the horizontal tiling conditions using CIs

$$\exists x.A_t \sqcap A_{t'} \sqsubseteq \bot \qquad \text{for each } \langle t, t' \rangle \in (T \times T) \setminus H$$

and correspondingly for the vertical conditions. The (hidden) concept name $B \notin \Sigma$ is used to ensure that the hypothesis contains at least one individual per coordinate. This name is initialised by the individual satisfying Start, and then propagated in x and y direction, provided that a tiling type is associated. The observation to be explained is $\mathsf{End}(a)$, where End occurs in the following CI:

$$\prod_{i=1}^{n} X_i \sqcap \prod_{i=1}^{n} Y_i \sqcap B \sqcap A_{t_e} \sqsubseteq \mathsf{End}$$

and the abducibles are

$$\Sigma = \{\mathsf{Start}, x, y\} \cup \{A_t \mid t \in T\}.$$

Without the size restriction, a valid hypothesis corresponds to a binary tree with tile types associated to each node, and tiling conditions ensured along the x- and y-successors. To make sure it forms a $2^n \times 2^n$ grid, we choose the size k appropriately in a way that every coordinate can be used at most once. Valid hypotheses of size k then correspond to solutions to the tiling problem. □

To present the proof idea for the NEXPTIME$^{\text{NP}}$-hardness result more concisely, we introduce a new tiling problem.

Definition 5. *A* NEXPTIME$^{\text{NP}}$-tiling problem *is given by a tuple $\langle T, T_I, t_e, H_1, V_1, H_2, V_2, n \rangle$, where $\langle T, T_I, t_e, H_1, V_1, n \rangle$ is an exponential tiling problem, $H_2, V_2 \subseteq T \times T$ are additional tiling conditions, and for which we want to decide the existence of a valid tiling $f : [\![1, 2^n]\!] \times [\![1, 2^n]\!] \to T$ for the tiling problem $\langle T, T_I, t_e, V_1, H_1, n \rangle$, s.t. for no $i \in [\![1, 2^n]\!]$, there exists a valid tiling for the tiling problem $\langle T, f(i), t_e, V_2, H_2, n \rangle$, where $f(i)$ denotes the ith row of the tiling f.*

In other words, we have to find a tiling using conditions H_1 and V_1, while avoiding any rows that can be first row of any tiling for conditions H_2 and V_2.

Lemma 6. *The* NEXPTIME$^{\text{NP}}$*-tiling problem is* NEXPTIME$^{\text{NP}}$*-hard.*

Lemma 7. *Size-restricted $\mathcal{ALC}$ abduction is* NEXPTIME$^{\text{NP}}$*-hard.*

Proof sketch. We modify the construction for Lemma 5 to encode the NEXPTIME$^{\text{NP}}$-tiling problem. We now use 3 roles x, y, z and corresponding binary counters so that, together with the size restriction, each hypothesis will have the shape of a cube. The bottom side of this cube has to correspond to a tiling for $\langle T, T_I, t_e, V_1, H_1, n \rangle$, which can be achieved using similar axioms as for Lemma 5. For nodes outside of the bottom side of the cube, we require the use a different set of concept names for the tile types, which are of the form A_t^*, and for which we have the axiom $T^* \sqsubseteq \bigsqcup_{t \in T} A_t^*$. We use

$$\Sigma = \{\mathsf{Start}, x, y, z, T^*\} \cup \{A_t \mid t \in T\},$$

and again require every coordinate to be assigned some tile type. For the coordinates outside the bottom side, we have to use the concept name T^* to assign tile types, which leaves the precise selection of the tile type to the different models of the hypothesis. We detect tiling errors in the different $x \times z$-squares with the following axioms

$$\exists x.A_t^* \sqcap A_{t'}^* \sqsubseteq B_3 \qquad \text{for } \langle t, t' \rangle \in (T \times T) \setminus H_2$$

and correspondingly for V_2. This information is propagated along the succeeding coordinates so that the observation $\mathsf{End}(a)$ is only entailed if every model of the hypothesis encodes a tiling error on each of the $x \times z$ squares. □

6 Outlook

We believe that our results for complex abduction in $\mathcal{ALC}$ can be extended to $\mathcal{ALCI}$, and that the bound for size restricted $\mathcal{ALCQI}$ abduction in is tight. A question is whether we can improve the N2EXPTIME$^{\text{NP}}$-bound for the most general variant of our abduction problem. Apart from that, we want to investigate our setting for observations formulated as conjunctive queries, which would allow us to explain negative query answers [Calvanese *et al.*, 2013]. Another interesting question is what happens if we allow fresh individual names for abduction with ontologies formulated using existential rules. For the $\mathcal{EL}_\bot$-variant, we are currently working on a practical method for computing size-minimal flat hypotheses.

Acknowledgements

Patrick Koopmann is supported by DFG grant 389793660 of TRR 248 (see https://www.perspicuous-computing.science/).

References

[Baader *et al.*, 2016] Franz Baader, Meghyn Bienvenu, Carsten Lutz, and Frank Wolter. Query and predicate emptiness in ontology-based data access. *J. Artif. Intell. Res.*, 56:1–59, 2016.

[Baader *et al.*, 2017] Franz Baader, Ian Horrocks, Carsten Lutz, and Ulrike Sattler. *An Introduction to Description Logic*. Cambridge University Press, 2017.

[Bienvenu, 2008] Meghyn Bienvenu. Complexity of abduction in the $\mathcal{EL}$ family of lightweight description logics. In *Proceedings of KR 2008*, pages 220–230. AAAI Press, 2008.

[Calvanese *et al.*, 2013] Diego Calvanese, Magdalena Ortiz, Mantas Simkus, and Giorgio Stefanoni. Reasoning about explanations for negative query answers in DL-Lite. *J. Artif. Intell. Res.*, 48:635–669, 2013.

[Ceylan *et al.*, 2020] İsmail İlkan Ceylan, Thomas Lukasiewicz, Enrico Malizia, Cristian Molinaro, and Andrius Vaicenavicius. Explanations for negative query answers under existential rules. In Diego Calvanese, Esra Erdem, and Michael Thielscher, editors, *Proceedings of KR 2020*, pages 223–232. AAAI Press, 2020.

[Del-Pinto and Schmidt, 2019] Warren Del-Pinto and Renate A. Schmidt. ABox abduction via forgetting in $\mathcal{ALC}$. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019*, pages 2768–2775. AAAI Press, 2019.

[Du *et al.*, 2012] Jianfeng Du, Guilin Qi, Yi-Dong Shen, and Jeff Z. Pan. Towards practical ABox abduction in large description logic ontologies. *Int. J. Semantic Web Inf. Syst.*, 8(2):1–33, 2012.

[Du *et al.*, 2014] Jianfeng Du, Kewen Wang, and Yi-Dong Shen. A tractable approach to ABox abduction over description logic ontologies. In Carla E. Brodley and Peter Stone, editors, *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 1034–1040. AAAI Press, 2014.

[Du *et al.*, 2017] Jianfeng Du, Hai Wan, and Huaguan Ma. Practical TBox abduction based on justification patterns. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pages 1100–1106, 2017.

[Elsenbroich *et al.*, 2006] Corinna Elsenbroich, Oliver Kutz, and Ulrike Sattler. A case for abductive reasoning over ontologies. In *Proceedings of the OWLED'06 Workshop on OWL: Experiences and Directions*, 2006.

[Fidelibus *et al.*, 2011] M Dolores Fidelibus, Francisco Gutiérrez, and Giuseppe Spilotro. 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.

[Ghilardi *et al.*, 2006] Silvio Ghilardi, Carsten Lutz, and Frank Wolter. Did I damage my ontology? A case for conservative extensions in description logics. In Patrick Doherty, John Mylopoulos, and Christopher A. Welty, editors, *Tenth International Conference on Principles of Knowledge Representation and Reasoning (KR 2006)*, pages 187–197. AAAI Press, 2006.

[Halland and Britz, 2012] Ken Halland and Katarina Britz. ABox abduction in $\mathcal{ALC}$ using a DL tableau. In *2012 South African Institute of Computer Scientists and Information Technologists Conference, SAICSIT '12*, pages 51–58, 2012.

[Klarman *et al.*, 2011] Szymon Klarman, Ulle Endriss, and Stefan Schlobach. ABox abduction in the description logic $\mathcal{ALC}$. *Journal of Automated Reasoning*, 46(1):43–80, 2011.

[Koopmann and Schmidt, 2015] Patrick Koopmann and Renate A. Schmidt. Uniform interpolation and forgetting for $\mathcal{ALC}$ ontologies with ABoxes. In *Proceedings of AAAI 2015*, pages 175–181. AAAI Press, 2015.

[Koopmann *et al.*, 2020] Patrick Koopmann, Warren Del-Pinto, Sophie Tourret, and Renate A. Schmidt. Signature-based abduction for expressive description logics. In Diego Calvanese, Esra Erdem, and Michael Thielscher, editors, *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020*, pages 592–602. AAAI Press, 2020.

[Koopmann, 2021] Patrick Koopmann. Signature-based abduction with fresh individuals and complex concepts for description logics (extended version), 2021. arXiv:2105.00274 [cs.AI].

[Lutz and Wolter, 2011] Carsten Lutz and Frank Wolter. Foundations for uniform interpolation and forgetting in expressive description logics. In *Proceedings of IJCAI 2011*, pages 989–995. IJCAI/AAAI, 2011.

[Obeid *et al.*, 2019] Mariam Obeid, Zeinab Obeid, Asma Moubaiddin, and Nadim Obeid. Using description logic and Abox abduction to capture medical diagnosis. In Franz Wotawa, Gerhard Friedrich, Ingo Pill, Roxane Koitz-Hristov, and Moonis Ali, editors, *Advances and Trends in Artificial Intelligence. From Theory to Practice*, pages 376–388, Cham, 2019. Springer International Publishing.

[Pukancová and Homola, 2017] Júlia Pukancová and Martin Homola. Tableau-based ABox abduction for the $\mathcal{ALCHO}$ description logic. In *Proceedings of the 30th International Workshop on Description Logics*, 2017.

[van Emde Boas, 1997] Peter van Emde Boas. The convenience of tilings. *Lecture Notes in Pure and Applied Mathematics*, pages 331–363, 1997.

[Wei-Kleiner *et al.*, 2014] Fang Wei-Kleiner, Zlatan Dragisic, and Patrick Lambrix. Abduction framework for repairing incomplete $\mathcal{EL}$ ontologies: Complexity results and algorithms. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 1120–1127. AAAI Press, 2014.

[Zhao and Schmidt, 2017] Yizheng Zhao and Renate A. Schmidt. Role forgetting for $\mathcal{ALCOQH}(\nabla)$-ontologies using an Ackermann-based approach. In Carles Sierra, editor, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017*, pages 1354–1361. ijcai.org, 2017.