Forgetting-Based Abduction in $\mathcal{ALC}$-Ontologies (Extended Abstract)

Warren Del-Pinto, Renate A. Schmidt

University of Manchester, UK
{warren.del-pinto,renate.schmidt}@manchester.ac.uk

Abstract. This paper presents a method for abduction in description logic ontologies. The method is based on forgetting and contrapositive reasoning and can produce semantically minimal hypotheses for TBox and ABox abduction in the description logic $\mathcal{ALC}$. The method is not restricted to Horn clauses or atomic observations and hypotheses. Key considerations when using forgetting for abduction are addressed. A Java prototype has been implemented, making use of the resolution-based forgetting method implemented in the tool LETHE. Experimental results over a corpus of ontologies show the practicality of the method across a number of scenarios.

1 Introduction

This paper presents a resolution-based method for performing both TBox and ABox abduction in $\mathcal{ALC}$ ontologies, which utilises the forgetting method developed in [7–10]. The observations and hypotheses can contain atomic and complex concepts. Currently, the method is restricted to the expressibility of $\mathcal{ALC}$ and cannot compute hypotheses for problems involving role assertions. System properties, essential postprocessing steps and other considerations are discussed with an evaluation of the system on various ontologies.

The contributions of this paper are as follows: (1) Abduction in ontologies is framed in terms of forgetting and contrapositive reasoning. Important considerations, such as extracting hypotheses from uniform interpolants, are discussed with proposed solutions. (2) We show that the uniform interpolation method in [9] can not only be used for TBox abduction [10] but also ABox abduction. (3) A unified method for TBox and ABox abduction based on forgetting is presented. This method can compute complex hypotheses from complex observations, finding a semantically minimal hypothesis for each observation in terms of a set of abducible symbols defined by a forgetting signature. (4) The practicality of the system is evaluated on a corpus of real-world ontologies.

2 Abduction in DL Ontologies

The aim of abduction is to compute a hypothesis to explain a new observation. This task is usually split into the tasks of TBox and ABox abduction, for which

In: P. Koopmann, S. Rudolph, R. Schmidt, C. Wernhard (eds.): *SOQE 2017 – Proceedings of the Workshop on Second-Order Quantifier Elimination and Related Topics, Dresden, Germany, December 6–8, 2017*, published at http://ceur-ws.org.

the observation ψ and the hypothesis $\mathcal{H}$ take the form of sets of TBox axioms and ABox axioms respectively [2]. Two of the most common requirements are expressed in conditions (i) and (ii) of Definition 1 [13, 5], which is the form of the abduction problem considered in this paper.

Definition 1. ***Abduction in Ontologies.*** *Let $\mathcal{O}$ be an ontology, $\mathcal{S}_A$ be a set of abducible symbols, and ψ an ABox or TBox axiom such that $\mathcal{O}, \psi \not\models \bot$ and $\mathcal{O} \not\models \psi$. The* abduction problem *is to find a hypothesis $\mathcal{H}$ in the form of a set of axioms, consisting only of symbols in $\mathcal{S}_A$, such that:* (i) *$\mathcal{O}, \mathcal{H} \not\models \bot$,* (ii) *$\mathcal{O}, \mathcal{H} \models \psi$ and* (iii) *there is no other hypothesis $\mathcal{H}'$ such that $sig(\mathcal{H}') \subseteq \mathcal{S}_A$, $\mathcal{O}, \mathcal{H}' \models \psi$ and $\mathcal{H} \models \mathcal{H}'$, unless $\mathcal{O}, \mathcal{H}' \equiv \mathcal{O}, \mathcal{H}$.*

The condition $\mathcal{O}, \psi \not\models \bot$ ensures that neither the ontology itself nor the ontology with the observation entail false, otherwise everything follows. The condition $\mathcal{O} \not\models \psi$ imposes the constraint that the observation should not already follow from the original ontology, otherwise the problem is deductive and the abduction solution is simply $\mathcal{H} = \top$. For $\mathcal{ALC}$, it is also worth noting that for ABox abduction both ψ and $\mathcal{H}$ must take the form of ABox axioms, while in the case of TBox abduction both must take the form of TBox axioms.

Even with these conditions, the number of possible hypotheses is often intractably large. There is also the problem of finding the preferred hypotheses among these possible solutions. Thus, a variety of additional constraints are often considered to further restrict the space of abductive hypotheses [2].

The proposed method computes consistent, explanatory and semantically minimal hypotheses in accordance with Definition 1. The semantic minimality constraint in condition (iii) restricts the hypotheses to those that make the fewest assumptions, and is the "strong" semantic minimality constraint in [4]. This is a desirable characteristic for comparing hypotheses in many applications [14].

A set of *abducibles* is usually defined to further restrict the abductive hypotheses. The set of abducibles defines a subset of symbols in $\mathcal{O}$ that may appear in the hypothesis $\mathcal{H}$. Here, the abducibles are defined by a forgetting signature, as the proposed method utilises forgetting to compute hypotheses that satisfy the conditions outlined for the abduction task in Definition 1.

3 Forgetting and Uniform Interpolation

Forgetting, also known as uniform interpolation, is a process of finding a compact representation of an ontology by hiding or removing subsets of symbols within it. Here, the term *symbols* refers to concept and role names in the signature of the ontology. The symbols to be hidden are defined by a forgetting signature $\mathcal{F}$, which consists of a subset of symbols in the ontology $\mathcal{O}$. The symbols in $\mathcal{F}$ should be removed from $\mathcal{O}$, while preserving all entailments of $\mathcal{O}$ that can be represented using the restricted signature $sig(\mathcal{O}) \backslash \mathcal{F}$, resulting in a new ontology.

Definition 2. ***Uniform Interpolation in $\mathcal{ALC}$.*** *Let $\mathcal{O}$ be an $\mathcal{ALC}$-ontology and $\mathcal{F}$ a signature of symbols to be forgotten from $\mathcal{O}$. Let $\mathcal{S}_A = sig(\mathcal{O}) \backslash \mathcal{F}$ be the* *complement of $\mathcal{F}$. The* uniform interpolation problem [12] *is the task of finding an ontology $\mathcal{V}$ such that the following conditions hold:* (i) $sig(\mathcal{V}) \subseteq \mathcal{S}_A$, (ii) *for any set of axioms β: $\mathcal{O} \models \beta$ iff $\mathcal{V} \models \beta$ provided that $sig(\beta) \subseteq \mathcal{S}_A$. The ontology $\mathcal{V}$ is a* uniform interpolant *of $\mathcal{O}$ for the signature $\mathcal{S}_A$.*

Theorem 1. *$\mathcal{V}$ is the uniform interpolant of ontology $\mathcal{O}$ for signature $\mathcal{S}_A$ iff $\mathcal{V}$ is the strongest necessary entailment of $\mathcal{O}$ in the signature $\mathcal{S}_A$.*

It is not necessarily the case that the source and target languages are the same. In [9] the uniform interpolant (or forgetting solution) of an ontology in $\mathcal{ALC}$ may need to be expressed in an extension of $\mathcal{ALC}$, which includes the following additions: (i) fixpoint operators or definer symbols for representing cycles and (ii) disjunctions of ABox axioms. Here, "definer symbols" refer to new symbols that are not present in the original ontology. "Cycles" refer to uniform interpolants that are not finitely representable in $\mathcal{ALC}$. However, evaluations on real-world ontologies have shown that the majority of uniform interpolants can be represented in pure $\mathcal{ALC}$. In the case where the result does contain cycles, we represent this result using fixpoints. Thus, definer symbols do not appear in any of the uniform interpolants computed.

The proposed abduction method utilises the resolution-based forgetting method developed in [7–10], which can compute uniform interpolants of $\mathcal{ALC}$ ontologies by forgetting both concept and role symbols in the original ontology. Here, this method is referred to as $Int_{\mathcal{ALC}}$. Two key characteristics for computing uniform interpolants that are essential to the proposed abduction method are as follows.

Theorem 2. *The uniform interpolation method has the following properties:*
(1) ***Soundness:*** *any ontology $\mathcal{O}'$ returned by applying $Int_{\mathcal{ALC}}$ to an ontology $\mathcal{O}$ is a uniform interpolant and hence satisfies criteria* (i) *and* (ii) *in Definition 2.*
(2) ***Interpolation Completeness:*** *if there exists a uniform interpolant $\mathcal{O}'$ of ontology $\mathcal{O}$, then $Int_{\mathcal{ALC}}$ returns an ontology $\mathcal{V}$ such that $\mathcal{V} \equiv \mathcal{O}'$.*

For any combination of an $\mathcal{ALC}$ ontology $\mathcal{O}$ and forgetting signature $\mathcal{F}$, $Int_{\mathcal{ALC}}$ returns a finite uniform interpolant [9, 6].

The method $Int_{\mathcal{ALC}}$ relies on the transformation of the ontology to a normal form given by a set of clauses of concept literals. The rules of the forgetting calculus utilised in $Int_{\mathcal{ALC}}$ can be found in [9]. Definer symbols are introduced to represent concepts that occur under the scope of a quantifier. Resolution inferences are then made on literals including the symbols present in $\mathcal{F}$ and the definer symbols. Once all possible inferences have been made, any clauses containing symbols in $\mathcal{F}$ are removed and the definer symbols are eliminated resulting in an ontology $\mathcal{O}'$ that is free of all symbols in $\mathcal{F}$.

4 A Forgetting-Based Abduction Method

The resolution-based nature of the $Int_{\mathcal{ALC}}$ makes it well suited to abduction via contrapositive reasoning. The calculus is applicable not only to TBox abduction [10], but also to ABox abduction in a single unified framework. Algorithm 1 below outlines our forgetting-based method which utilises $Int_{\mathcal{ALC}}$ to compute semantically minimal abductive hypotheses by exploiting contrapositive reasoning as shown in the following theorem.

Theorem 3. *Let $\mathcal{O}$ be an ontology, ψ an observation and $\mathcal{H}$ a hypothesis. Then the following holds: $\mathcal{O}, \mathcal{H} \models \psi$ iff $\mathcal{O}, \neg\psi \models \neg\mathcal{H}$.*

Algorithm 1 Forgetting-Based Abduction. *The algorithm computes hypothesis $\mathcal{H}$ for an observation ψ relative to ontology $\mathcal{O}$. It is assumed that ψ is a single axiom, which can also be a conjunction of assertions over a single individual "a" which does not occur in the ontology $\mathcal{O}$. Two cases are considered:* (i) *ψ is an ABox axiom $C(a)$ or* (ii) *ψ is a TBox axiom $C \sqsubseteq D$ where C and D can be any atomic or complex $\mathcal{ALC}$ concepts. The steps for both cases* (i) *and* (ii) *are as follows:*

1. *Negate the observation to obtain $\neg\psi$. In case* (i) *$\neg\psi = \neg C(a)$, while in case* (ii) *$\neg\psi = (C \sqcap \neg D)(a)$.*
2. *Choose a forgetting signature set $\mathcal{F}$ such that $\mathcal{F} \cap sig(\neg\psi) \neq \emptyset$, where at least one of the symbols in both $\mathcal{F}$ and $\neg\psi$ occurs with opposite polarity in the ontology $\mathcal{O}$. Let $\mathcal{S}_A = sig(\mathcal{O}) \setminus \mathcal{F}$.*
3. *Use $Int_{\mathcal{ALC}}$ to compute a uniform interpolant of $(O, \neg\psi)$ for $\mathcal{S}_A$ by forgetting the symbols in $\mathcal{F}$.*
4. *Let $\mathcal{V}$ be the uniform interpolant computed. Apply filtering to $\mathcal{V}$ to obtain the set of axioms $\mathcal{V}^* \subseteq \mathcal{V}$ that are dependent on $\neg\psi$. This means that the axioms in $\mathcal{V}^*$ are conclusions of inferences in $Int_{\mathcal{ALC}}$ with clauses from $\neg\psi$.*
5. *Assuming $\mathcal{V}^* = \{\overline{\alpha}_1(a), ..., \overline{\alpha}_k(a)\}$, let* (i) *$\mathcal{H}_I = (\alpha_1 \sqcup ... \sqcup \alpha_k)(a)$ when ψ is an ABox axiom, and* (ii) *$\mathcal{H}_I = \top \sqsubseteq (\alpha_1 \sqcup ... \sqcup \alpha_k)$ when ψ is a TBox axiom, where $\alpha_i \equiv \neg\overline{\alpha}_i$ for any $1 \leq i \leq k$.*
6. *In the case* (i) *of ABox abduction, let $\mathcal{H}_f = \mathcal{H}_I$. In the case* (ii) *of TBox abduction, an additional check is needed to ensure consistency of the hypothesis: $\mathcal{O}, \mathcal{H}_I \not\models \bot$. If this succeeds, then $\mathcal{H}_f = \mathcal{H}_I$.*

This procedure can be performed iteratively over a set of observations. In the event that cycles involving definer symbols occur in $\mathcal{V}$, these will be represented using fixpoint operators. It is important to note that $\mathcal{F}$ must contain at least one symbol in the observation ψ, as described in step 2. This enables the computation of inferences between the ontology $\mathcal{O}$ and the negated observation $\neg\psi$ using $Int_{\mathcal{ALC}}$, ensuring that the set of axioms $\mathcal{V}^*$ is computed. Otherwise, the trivial hypothesis $\mathcal{H}_f = \psi$ will be obtained. It is also worth noting the choice of representation for the negated observation $\neg\psi$. For ABox abduction, the negation of an ABox axiom $\psi = C(a)$ is simply $\neg\psi = \neg C(a)$. For TBox abduction, the negation of a TBox axiom $\psi = C \sqsubseteq D$ can be represented in several ways [10]: we choose to include a fresh individual name a and represent as an ABox axiom $\neg\psi = (C \sqcap \neg D)(a)$. Thus, for both TBox and ABox abduction $\neg\psi$ takes the form of an ABox axiom. This choice is exploited in the extraction of $\mathcal{V}^*$ from the uniform interpolant $\mathcal{V}$, as described in the next section.

The exclusion of role assertions is due to the fact that the method $Int_{\mathcal{ALC}}$ does not cater directly for negated role assertions. As a result, observations containing role assertions cannot currently be handled by our abduction method.

5 Extraction of $\mathcal{V}^*$ from $\mathcal{V}$

Many of the entailments in $\mathcal{V}$ do not involve inferences with the negated observation $\neg\psi$. These entailments do not contribute towards an explanation for ψ, and must be removed to reduce redundancy and guarantee consistency of the hypothesis returned. This leaves the set of axioms $\mathcal{V}^*$, which consists of axioms obtained by inferences in $Int_{\mathcal{ALC}}$ with either clauses in $\neg\psi$ or with clauses previously derived by inferences with $\neg\psi$.

There are several ways to remove the unnecessary axioms in $\overline{\mathcal{V}^*}$. These include checking the consistency of each disjunct in $\mathcal{H}_I$ with the ontology $\mathcal{O}$ with an external reasoner or performing subsumption deletion between axioms in $\mathcal{V}$ and those in $\mathcal{O}$. Both of these methods are computationally expensive, particularly as there are often a large number of axioms in $\mathcal{V}$. A third possibility is to trace the dependency on $\neg\psi$ as the inferences are performed in $Int_{\mathcal{ALC}}$. However, an alternate method was devised to eliminate these guaranteed redundancies without relying on an external reasoner, subsumption deletion or dependency tracing. This method utilises a property of the forgetting calculus $Int_{\mathcal{ALC}}$.

Theorem 4. Efficient Filtering of $\mathcal{V}$. *Let $\mathcal{O}$, ψ and $\mathcal{V}$ be defined as in Algorithm 1, where $\neg\psi$ is an ABox axiom $\neg C(a)$. For any α in the uniform interpolant $\mathcal{V}$, $\alpha \in \mathcal{V}^*$ iff the signature $sig_I(\alpha)$ of individuals contains a.*

After computing the set of axioms $\mathcal{V}^*$, this set is negated to obtain a hypothesis $\mathcal{H}_I$, exploiting contrapositive reasoning as in Theorem 3. This is outlined in step 5 of Algorithm 1. We have that $\mathcal{O}, \neg\psi \models \mathcal{V}^*$ iff $\mathcal{O}, \mathcal{H}_I \models \psi$ where $\neg\mathcal{V}^* \equiv \mathcal{H}_I$.

In the ABox abduction case, the unnecessary axioms in $\mathcal{V} \setminus \mathcal{V}^*$ account for all possible inconsistencies in $\mathcal{H}_I$, and no further processing is required. This was confirmed empirically by the lack of any difference between $\mathcal{H}_I$ and $\mathcal{H}_f$ in the experimental evaluations in Tables 2 and 3. $\mathcal{H}_I$ represents the hypothesis prior to the following additional check of $\mathcal{O}, \mathcal{H}_I \not\models \bot$. For TBox abduction, this test is needed to ensure that the hypothesis returned by the system is not inconsistent with the original ontology. This is due to the transformation from an ABox assertion to a TBox axiom $\top \sqsubseteq (\alpha_1 \sqcup ... \sqcup \alpha_k)$ described in step 5 of Algorithm 1. This transformation is necessary as in $\mathcal{ALC}$, if the observation ψ is a TBox axiom then the hypothesis must also be a TBox axiom to ensure the condition in Definition 1(ii) is satisfied.

6 Properties of Method

Key properties of the abduction method are presented here. These properties hold for consistent $\mathcal{ALC}$ ontologies and the characteristics of the computed hypotheses are relative to the signature defined by eliminating the symbols contained within the chosen forgetting signature $\mathcal{F}$. If cycles occur in uniform interpolants, these are represented using fixpoint operators.

The proposed abduction method computes semantically minimal hypotheses. This can be seen in terms of strongest necessary and weakest sufficient conditions [11, 1]. The uniform interpolant computed by $Int_{\mathcal{ALC}}$ is the strongest necessary set of entailments of the original ontology, as in Theorem 1. Thus, the axioms $\mathcal{V}^*$ can be seen as a set of strongest necessary entailments of $(\mathcal{O}, \neg\psi)$ that depend upon the observation ψ. As discussed in [11], strongest necessary and weakest sufficient conditions are dual conditions. Thus, by negating the set of axioms $\mathcal{V}^*$ under contrapositive reasoning, a weakest sufficient hypothesis $\mathcal{H}_f$ is obtained.

Theorem 5. *Let $\mathcal{O}$, ψ, $\mathcal{H}_I$ and $\mathcal{H}_f$ be defined as in Algorithm 1. The following conditions hold for the hypothesis $\mathcal{H}_f$:* (i) $\mathcal{O}, \mathcal{H}_f \not\models \bot$, (ii) $\mathcal{O}, \mathcal{H}_f \models \psi$ *and* (iii) $\mathcal{H}_f$ *is a weakest sufficient explanation, i.e, if there is a $\mathcal{H}$ such that* (i) *and* (ii) *hold and $\mathcal{H}_f \models \mathcal{H}$, then $\mathcal{O}, \mathcal{H}_f \equiv \mathcal{O}, \mathcal{H}$.*

Theorem 6. Completeness with respect to $\mathcal{S}_A$. *For an ontology $\mathcal{O}$ and observation ψ, if there exists a consistent, semantically minimal hypothesis $\mathcal{H}'$ within the signature $\mathcal{S}_A = sig(\mathcal{O}) \setminus \mathcal{F}$ such that $(O, \mathcal{H}') \models \psi$, then the proposed method returns a hypothesis $\mathcal{H}_f$ such that $\mathcal{H}_f \equiv \mathcal{H}'$.*

Several other properties of the method are worth noting. Firstly, in the case that ψ is an ABox axiom, each disjunct $\alpha_i(a)$ in the final hypothesis $\mathcal{H}_f$ is also a hypothesis since $\mathcal{O}, \alpha_i(a) \models \mathcal{O}, (\alpha_1 \sqcup ... \sqcup \alpha_k)(a)$ for $1 \leq i \leq k$. Secondly, it is possible to iteratively compute semantically stronger hypotheses due to the fact that symbols are iteratively eliminated in the forgetting loop of $Int_{\mathcal{ALC}}$.

Below is an example of an ABox abduction problem by [14, 15], used to illustrate the abduction procedure given by Algorithm 1 and the semantic minimality of the hypothesis returned.

Example 1. *Consider the following ontology $\mathcal{O}$, consisting of two TBox axioms β_1 and β_2, and the observation ψ:*

β_1: Professor $\sqcup$ Scientist $\sqsubseteq$ Academic
β_2: AssocProfessor $\sqsubseteq$ Professor
ψ : Academic(Jack)

The steps in applying the proposed method are as follows: (1) *Negate ψ to obtain $\neg$Academic(Jack).* (2) *Choose a forgetting signature $\mathcal{F}$ such that $\mathcal{F} \cap sig(\neg\psi) \neq \emptyset$, in this case: $\mathcal{F} = \{Academic\}$.* (3) *Obtain the uniform interpolant $\mathcal{V}$ by applying $Int_{\mathcal{ALC}}$ to $(\mathcal{O}, \neg\psi)$ with the forgetting signature $\mathcal{F}$. Using $\mathcal{F} = \{Academic\}$, the following uniform interpolant $\mathcal{V}$ is obtained with $Int_{\mathcal{ALC}}$:*

$\overline{\alpha}_1$: AssocProfessor $\sqsubseteq$ Professor
$\overline{\alpha}_2$: $\neg$ Professor(Jack)
$\overline{\alpha}_3$: $\neg$ Scientist(Jack)

(4) *Obtain the set of axioms $\mathcal{V}^*$ that are dependent on $\neg\psi$ by applying the filtering described in Theorem 4 to the uniform interpolant $\mathcal{V}$. The first entailment $\overline{\alpha}_1$ is filtered out as it follows directly from the ontology $\mathcal{O}$ and does not contain the*

individual observed: Jack. The entailments $\overline{\alpha}_2$ *and* $\overline{\alpha}_3$ *are dependent on* $\neg\psi$ *and are retained.* (5) *Negate* $\mathcal{V}^*$ *to obtain a hypothesis* $\mathcal{H}_I$*:*

$\mathcal{H}_I = (\text{Professor} \sqcup \text{Scientist})(\text{Jack})$.

(6) *In this case,* ψ *is an ABox axiom as in case* (i) *of Algorithm 1. Thus, no further checks are required and* $\mathcal{H}_I = \mathcal{H}_f$*.*

The disjuncts of the hypothesis in this example are also (stronger) hypotheses: Professor(Jack) and Scientist(Jack), as is the conjunction of these two.

7 Experimental Evaluation

A Java prototype was implemented using the OWL-API[1] and the library of the utilised forgetting method: LETHE[2]. A set of experiments was then carried out over ontologies from the NCBO BioPortal[3] and OBO [4] repositories, plus the LUBM benchmark [3] and Semintec [5] financial ontologies. The ontologies were converted to their $\mathcal{ALC}$ fragments: axioms not representable in $\mathcal{ALC}$ were deleted while others, such as domain and range restrictions, were replaced by equivalent $\mathcal{ALC}$ axioms. The characteristics of the resulting corpus are shown in Table 1. The experiments were performed on a machine using a 4.00GHz Intel Core i7-6700K CPU with 16GB of RAM.

Ontology Name	TBox Size	ABox Size	Number of Concepts	Number of Roles
BFO	52	0	35	0
HOM	83	0	66	0
LUBM	87	0	44	24
Semintec	199	65189	61	16
DOID	7892	0	11663	15
ICF	1910	6597	1597	41
OBI	28888	196	3691	67
NATPRO	68565	42763	9464	12

Table 1. Characteristics of the experimental corpus.

For each ontology, 30 observations were generated. For TBox abduction, each set of observations contained random TBox axioms from the associated ontology, consisting of atomic or complex concepts. For each individual test, the TBox axiom was first removed from the ontology then used as an observation. For ABox abduction, observations were randomly generated using assertions or arbitrary concepts in the ontology. This was done to emulate information that may be observed in practice. $\mathcal{F}$ was limited to the smallest possible signature: a randomly

[1] http://owlapi.sourceforge.net/
[2] http://www.cs.man.ac.uk/~koopmanp/lethe/index.html
[3] https://bioportal.bioontology.org/
[4] http://www.obofoundry.org/
[5] http://www.cs.put.poznan.pl/alawrynowicz/semintec.htm

selected concept symbol from ψ. The assumption was that a user may begin by finding the most general hypotheses, which has the benefit that stronger hypotheses can be found using the iterative abduction process described earlier.

The results are shown in Table 2. In almost all cases a semantically minimal hypothesis was computed within the time limit on LETHE. For the OBI ontology, for two TBox and five ABox cases, LETHE timed out before a uniform interpolant was computed. Given more time a uniform interpolant would be found. The time taken to compute $\mathcal{H}_f$ varied considerably with the ontology size, as did the size of $\mathcal{H}_f$. The difference between the number of axioms filtered and the size of $\mathcal{H}_f$, particularly for the larger ontologies, supports the need for efficient filtering such as the proposal in Theorem 4. The sizes of $\mathcal{H}_I$ and $\mathcal{H}_f$ were equal for the ABox abduction results, indicating that all unnecessary entailments in the uniform interpolants were removed by the proposed filtering. For TBox abduction the two values were different in all cases, leaving room for improvement in the filtering used to avoid the need for additional checks on $\mathcal{H}_I$.

ABox abduction						**TBox abduction**					
Ont. Name	Mean Time /s	Max. Time /s	Axioms Filtered from $\mathcal{V}$	Mean size /disjuncts		Ont. Name	Mean Time /s	Max. Time /s	Axioms Filtered from $\mathcal{V}$	Mean size /disjuncts	
				$\mathcal{H}_\mathcal{I}$	$\mathcal{H}_f$					$\mathcal{H}_\mathcal{I}$	$\mathcal{H}_f$
BFO	0.03	0.25	51	1.0	1.0	BFO	0.03	0.35	51	2.1	1.5
HOM	0.03	0.25	82	1.8	1.8	HOM	0.03	0.32	81	3.3	3.0
LUBM	0.04	0.24	92	2.0	2.0	LUBM	0.04	0.31	89	2.3	1.8
Semin.	2.13	5.73	66757	1.3	1.3	Semin.	4.06	9.28	69900	5.2	0.5
DOID	0.57	1.68	8508	4.8	4.8	DOID	0.51	1.46	7890	3.2	2.6
ICF	1.16	4.11	8490	3.3	3.3	ICF	0.50	1.38	8504	4.3	3.9
OBI*	5.34	22.44	29360	5.6	5.6	OBI*	53.16	94.74	29059	92.4	92.3
NATP	46.44	399.31	111329	9.4	9.4	NATP	412.34	685.60	111196	130.0	125.1

Table 2. Results obtained for (i) ABox and (ii) TBox abduction over 30 observations with a forgetting signature of size 1. * indicates ontologies for which LETHE did not terminate within 120 seconds in at least one case.

8 Conclusion and Future Work

In this paper, a method for performing both TBox and ABox abduction in $\mathcal{ALC}$ ontologies was presented. The method uses forgetting and can compute complex hypotheses to explain both atomic and complex observations. The computed hypotheses were shown to be semantically minimal within a specified set of symbols. The practicality of the method was illustrated empirically across a corpus of real-world ontologies.

The method will be extended to perform abduction in more expressive description logics and to handle statements involving role assertions. These aims will likely be achieved by extending the $Int_{\mathcal{ALC}}$ calculus [9] to handle negated role assertions. Another option would be to investigate other methods for forgetting [16]. Another aim is to investigate the use of the iterative abduction procedure described earlier to compute increasingly stronger hypotheses.

References

1. Doherty, P., Łukaszewicz, W., Szałas, A.: Computing strongest necessary and weakest sufficient conditions of first-order formulas. In: Proc. IJCAI'01. pp. 145–151. AAAI Press (2001)
2. Elsenbroich, C., Kutz, O., Sattler, U.: A case for abductive reasoning over ontologies. In: Proc. OWL: Experiences and Directions (2006)
3. Guo, Y., Pan, Z., Heflin, J.: LUBM: A benchmark for OWL knowledge base systems. Journal of Web Semantics: Science, Services and Agents on the World Wide Web 3, 158–182 (2005)
4. Halland, K., Britz, A., Klarman, S.: Tbox abduction in $\mathcal{ALC}$ using a DL tableau. In: Proc. DL'14. pp. 556–566. CEUR-WS.org (2014)
5. Klarman, S., Endriss, U., Schlobach, S.: Abox abduction in the description logic $\mathcal{ALC}$. Journal of Automated Reasoning 46, 43–80 (2011)
6. Koopmann, P.: Practical Uniform Interpolation for Expressive Description Logics. Ph.D. thesis, The University of Manchester, UK (2015)
7. Koopmann, P., Schmidt, R.A.: Forgetting concept and role symbols in $\mathcal{ALCH}$-ontologies. In: LPAR'13. LNCS, vol. 8312, pp. 552–567. Springer (2013)
8. Koopmann, P., Schmidt, R.A.: Uniform interpolation of $\mathcal{ALC}$ ontologies using fixpoints. In: Proc. FroCoS13. LNCS, vol. 8152, pp. 87–102. Springer (2013)
9. Koopmann, P., Schmidt, R.A.: Uniform interpolation and forgetting for $\mathcal{ALC}$ ontologies with ABoxes. In: Proc. AAAI'15. pp. 175–181. AAAI Press (2015)
10. Koopmann, P., Schmidt, R.A.: LETHE: Saturation based reasoning for non-standard reasoning tasks. In: Proc. ORE'15. pp. 23–30. CEUR-WS.org (2015)
11. Lin, F.: On strongest necessary and weakest sufficient conditions. Artificial Intelligence 128 pp. 143–159 (2001)
12. Lutz, C., Wolter, F.: Foundations for uniform interpolation and forgetting in expressive description logics. In: Proc. IJCAI'11. pp. 989–995. AAAI Press (2011)
13. Mooney, R.J.: Integrating abduction and induction in machine learning. In: Abduction and Induction. pp. 181–191. P. A. Flach and A. C. Kakas, Eds. Applied Logic Series. Kluwer (2000)
14. Pukancová, J., Homola, M.: Tableau-based ABox abduction for description logics: preliminary report. In: Proc. DL'16. CEUR-WS.org (2016)
15. Pukancová, J., Homola, M.: Tableau-based ABox abduction for the $\mathcal{ALCHO}$ description logic. In: Proc. DL'17,. CEUR-WS.org (2017)
16. Zhao, Y., Schmidt, R.: Forgetting concept and role symbols in $\mathcal{ALCOIH}\mu^+(\nabla,\sqcap)$-ontologies. In: Proc. IJCAI'16. AAAI Press (2016)