Complexity of Abduction in the $\mathcal{EL}$ Family of Lightweight Description Logics

Meghyn Bienvenu
IRIT, Université Paul Sabatier
31062 Toulouse Cedex, France
bienvenu@irit.fr

Abstract

The complexity of logic-based abduction has been extensively studied for the case in which the background knowledge is represented by a propositional theory, but very little is known about abduction with respect to description logic knowledge bases. The purpose of the current paper is to examine the complexity of logic-based abduction for the $\mathcal{EL}$ family of lightweight description logics. We consider several minimality criteria for explanations (set inclusion, cardinality, prioritization, and weight) and three decision problems: deciding whether an explanation exists, deciding whether a given hypothesis appears in some acceptable explanation, and deciding whether a given hypothesis belongs to every acceptable explanation. We determine the complexity of these tasks for general TBoxes and also for $\mathcal{EL}$ and $\mathcal{EL}^+$ terminologies. We also provide results concerning the complexity of computing abductive explanations.

Introduction

Abduction is a form of reasoning that is used to generate possible explanations for a given observation. It has numerous applications in artificial intelligence, e.g. diagnosis, planning, natural language understanding, and computer vision (refer to (Eiter & Gottlob 1995) for references). There are a couple different approaches to abduction, but the one that interests us here is logic-based abduction. In this approach, an abduction problem consists of a background theory, a set of hypotheses, and an observation, all of which are represented by logical formulae. The objective is to find a set of hypotheses which are consistent with the background theory and which logically entail the observation when taken together with the background knowledge. Sets of hypotheses satisfying these conditions are called explanations. Very often a minimality criterion, like set inclusion or cardinality, is used to select preferred explanations.

Given a logic-based abduction problem, there are three main decision problems of interest: existence (does an explanation exist?), relevance (does a hypothesis appear in some preferred explanation?), and necessity (does a hypothesis appear in every preferred explanation?). There is also the search problems of generating some or all preferred explanations. The complexity of these problems has been extensively studied for different fragments of propositional logic (cf. (Bylander *et al.* 1991; Selman & Levesque 1996; Eiter & Gottlob 1995; Creignou & Zanuttini 2006)). There is very little known, however, about the complexity of abductive reasoning with respect to description logic knowledge bases.

The current paper aims to help fill this gap by providing an analysis of the complexity of abduction for the $\mathcal{EL}$ family of description logics (Baader, Brandt, & Lutz 2005). We chose to study this family of description logics for two reasons. First, we were attracted by their low complexity of reasoning. Given that abductive reasoning is generally of higher complexity than deductive reasoning, it seemed reasonable to start our investigation with description logics for which deductive reasoning is tractable. Our second motivation for studying the $\mathcal{EL}$ family stems from its usefulness in applications, particularly those in the biomedical domain.

Our paper is organized as follows. The first section recalls some basic notions from propositional logic, description logic, and computational complexity. In the following section, we introduce our abductive framework and the decision problems of interest. We then leverage results from propositional logic in order to determine the complexity of these decision problems for general $\mathcal{EL}$, $\mathcal{EL}^+$, and $\mathcal{EL}^{++}$ TBoxes. We next turn our attention to the special case in which the background knowledge is an $\mathcal{EL}^+$ or $\mathcal{EL}$ terminology, showing that the complexity improves for one of the decision problems but not the others. We then present some results concerning the problem of generating abductive explanations. At the end of the paper, we discuss possible extensions to our framework before concluding with related and future work.

Preliminaries

We review some basic notions from propositional logic, description logic, and computational complexity theory.

Propositional Logic

We assume a finite propositional language built from a set $\mathcal{V} = \{v_1, ..., v_n\}$ of atoms and the usual Boolean connectives. A *clause* is a disjunction $\lambda = \bigvee_{v_i \in Pos(\lambda)} v_i \vee \bigvee_{v_i \in Neg(\lambda)} \neg v_i$, where $Pos(\lambda)$ and $Neg(\lambda)$ are the sets of atoms which appear positively and negatively in λ and

$Pos(\lambda) \cap Neg(\lambda) = \emptyset$. We say that a clause λ is *negative* if $|Pos(\lambda)| = 0$, it is *definite Horn* if $|Pos(\lambda)| = 1$, and it is *Horn* if $|Pos(\lambda)| \leq 1$. A *Horn theory* is a set of Horn clauses, and a *definite Horn theory* is a set of definite Horn clauses.

The $\mathcal{EL}$ Family of Description Logics

We briefly review the syntax and semantics for the description logics $\mathcal{EL}$, $\mathcal{EL}^+$, and $\mathcal{EL}^{++}$. Concept descriptions are constructed inductively from a set N_C of atomic concepts and a set N_R of atomic roles using a set of concept constructors. For the description logics $\mathcal{EL}$ and $\mathcal{EL}^+$, the concept constructors are the top concept, conjunction, and existential restriction. For $\mathcal{EL}^{++}$, we have in addition the bottom concept, nominals, and a restricted form of concrete domains. In this paper, we will consider the version of $\mathcal{EL}^{++}$ without concrete domains. This is merely in order to simplify the exposition, as our complexity results for $\mathcal{EL}^{++}$ hold also in the presence of concrete domains. The syntax of the different constructors can be found in Table 1.

Name	Syntax	Semantics
top	$\top$	$\Delta^\mathcal{I}$
bottom	$\bot$	$\emptyset$
nominal	$\{a\}$	$\{a^\mathcal{I}\}$
conjunction	$C \sqcap D$	$C^\mathcal{I} \cap D^\mathcal{I}$
existential restriction	$\exists r.C$	$\{x \in \Delta^\mathcal{I} \mid \exists y \in \Delta^\mathcal{I} : (x,y) \in r^\mathcal{I}$ and $y \in C^\mathcal{I}\}$
GCI	$C \sqsubseteq D$	$C^\mathcal{I} \subseteq D^\mathcal{I}$
CD	$A \equiv C$	$A^\mathcal{I} = C^\mathcal{I}$
RI	$r_1 \circ ... \circ r_k \sqsubseteq r$	$r_1^\mathcal{I} \circ ... \circ r_k^\mathcal{I} \subseteq r^\mathcal{I}$

Table 1: Syntax and semantics of the $\mathcal{EL}$ family of DLs

The semantics of concept descriptions is defined in terms of interpretations. An *interpretation* $\mathcal{I}$ consists of a non-empty set $\Delta^\mathcal{I}$ and an interpretation function $\cdot^\mathcal{I}$ which assigns to each atomic concept name A a subset $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$, to each atomic role name r a relation $r^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$, and, in the case of $\mathcal{EL}^{++}$, to each nominal a an element $a^\mathcal{I} \in \Delta^\mathcal{I}$. The interpretation function is straightforwardly extended to complex concepts. Refer to Table 1 for details.

A *signature* is a set of atomic concept and role names. Two interpretations $\mathcal{I}$ and $\mathcal{J}$ *coincide* on a signature Σ if and only if $\Delta^\mathcal{I} = \Delta^\mathcal{J}$ and $S^\mathcal{I} = S^\mathcal{J}$ for every $S \in \Sigma$.

A *TBox* is a set of axioms describing the relationship between different concepts and roles. In $\mathcal{EL}$, a TBox can contain two types of axioms: *general concept inclusions* (GCIs) of the form $C \sqsubseteq D$ (where C and D are $\mathcal{EL}$ concepts) and *concept definitions* (CDs) of the form $A \equiv C$ (where A is an atomic concept and C an $\mathcal{EL}$ concept). For $\mathcal{EL}^+$, TBoxes may also contain *role inclusions* (RIs) of the form $r_1 \circ ... \circ r_m \sqsubseteq s$ (where r_i and s are role names). For $\mathcal{EL}^{++}$, we allow role inclusions and can use $\mathcal{EL}^{++}$ concepts in the concept axioms. The semantics of TBox axioms can be found in Table 1.

We call a TBox a *terminology* if it does not contain any GCIs and there is at most one CD for each atomic concept. A terminology is said to be *acyclic*[1] if for every CD $A \equiv C$, the concept C refers neither directly nor indirectly to A (cf. e.g. (Baader *et al.* 2003) for a formal definition of acyclicity). We say that an atomic concept is *defined* with respect to a terminology $\mathcal{T}$ if it appears on the left-hand side of a CD in $\mathcal{T}$. Atomic concepts which are not defined are called *primitive*.

The *signature* of a TBox $\mathcal{T}$, denoted $Sig(\mathcal{T})$, is the set of atomic concept and role names appearing in $\mathcal{T}$. Signatures of concepts are defined analogously.

We will use the term *propositional* to refer to TBoxes whose axioms are constructed using only atomic concept names and the $\top$, $\bot$, and $\sqcap$ concept constructors.

The main reasoning task for description logics is *subsumption* in which the problem is to decide for a TBox $\mathcal{T}$ and concepts C and D whether $\mathcal{T} \models C \sqsubseteq D$. Subsumption in $\mathcal{EL}^{++}$ is polynomial even with respect to general TBoxes (as shown in (Baader, Brandt, & Lutz 2005)).

Computational Complexity

We recall some basic definitions from computational complexity (cf. (Papadimitriou 1994)). The class NC consists of all problems which can be decided in polylogarithmic time on a parallel computer with a polynomial number of processors. The class P comprises all problems which can be decided in polynomial time by a deterministic Turing machine. The class NP contains all problems which can be decided in polynomial time by a non-deterministic Turing machine. The class co-NP is defined to be the set of all problems whose complement belongs to NP. The class $\Delta_2^P = P^{NP}$ is the set of all problems which can be decided in deterministic polynomial time given an NP oracle, and the class $\Delta_2^P[O(\log\, n)]$ refers to those problems which require only $O(\log n)$ queries to the NP oracle. The class $\Sigma_2^P = \text{NP}^{\text{NP}}$ consists of those problems which can be decided in polynomial time by a non-deterministic Turing machine which can query an NP oracle. The class Π_2^P comprises all problems whose complement is in Σ_2^P.

Our Abductive Framework

In this section, we introduce our abductive framework, which is a straightforward adaptation of the one studied in propositional logic (Eiter & Gottlob 1995).

Definition 1. An *abduction problem* is a tuple $\langle \mathcal{T}, \mathcal{H}, O \rangle$, where $\mathcal{T}$ is a TBox, $\mathcal{H}$ is a set of atomic concepts, and O is a single atomic concept.

Here the TBox $\mathcal{T}$ represents the background information, the set $\mathcal{H}$ represents the set of possible hypotheses, and the concept O represents the observation to be explained. We remark that requiring the observation to be an atomic concept is without loss of generality: if the observation is described by a complex concept D, we can simply add the concept

[1]The term *acyclic* is also used for Horn theories, but the usage is different. Indeed, acyclic Horn theories may not be expressible as acyclic TBoxes, and propositional acyclic TBoxes may not be expressible as acyclic Horn theories.

definition $O \equiv D$ to $\mathcal{T}$. Similarly, if one wishes to have arbitrary concepts as hypotheses, it suffices to add axioms of the form $A \equiv D$ to give names to these complex concepts.

A solution to an abduction problem is a consistent combination of hypotheses that explain the observation given the background knowledge. We will call sets of hypotheses satisfying these conditions *explanations*.

Definition 2. A set $\{A_1, ..., A_n\} \subseteq \mathcal{H}$ is an *explanation* for an abduction problem $\langle \mathcal{T}, \mathcal{H}, O \rangle$ if and only if

- $A_1 \sqcap ... \sqcap A_n$ is satisfiable with respect to $\mathcal{T}$
- $\mathcal{T} \models A_1 \sqcap ... \sqcap A_n \sqsubseteq O$

In general, not all explanations for an abduction problem may be equally satisfying. There are a number of different criteria that can be used to select preferred explanations. In this paper, we will concentrate on those most commonly studied in propositional logic, namely set inclusion, cardinality, prioritization, and weight.

Definition 3. Let $\mathcal{A}$ be an explanation for an abduction problem $\mathcal{P}$, let $P = \langle P_1, ..., P_n \rangle$ be a partition of $\mathcal{H}$ into priority classes, and let $w : \mathcal{H} \to \mathbb{N}$ be a function assigning numerical weights to the concepts in $\mathcal{H}$.

- $\mathcal{A}$ is $\subseteq$-*minimal* if there is no explanation $\mathcal{B}$ of $\mathcal{P}$ such that $\mathcal{B} \subsetneq \mathcal{A}$
- $\mathcal{A}$ is $\leq$-*minimal* if there is no explanation $\mathcal{B}$ of $\mathcal{P}$ such that $|\mathcal{B}| < |\mathcal{A}|$
- $\mathcal{A}$ is $\subseteq_P$-*minimal* if there is no explanation $\mathcal{B}$ of $\mathcal{P}$ and index i such that $\mathcal{B} \cap P_i \subsetneq \mathcal{A} \cap P_i$ and $\mathcal{B} \cap P_j = \mathcal{A} \cap P_j$ for every $1 \leq j < i$
- $\mathcal{A}$ is $\leq_P$-*minimal* if there is no explanation $\mathcal{B}$ of $\mathcal{P}$ and index i such that $|\mathcal{B} \cap P_i| < |\mathcal{A} \cap P_i|$ and $|\mathcal{B} \cap P_j| = |\mathcal{A} \cap P_j|$ for every $1 \leq j < i$
- $\mathcal{A}$ is $\sqsubseteq_w$-minimal if there is no explanation $\mathcal{B}$ of $\mathcal{P}$ such that $\sum_{B \in \mathcal{B}} w(B) < \sum_{A \in \mathcal{A}} w(A)$

Given an abduction problem $\mathcal{P}$ and a preference criterion, there are three main decision problems of interest:

Existence Does there exist an explanation for $\mathcal{P}$?

Relevance Is the hypothesis H *relevant*, i.e. does H appear in some preferred explanation for $\mathcal{P}$?

Necessity Is the hypothesis H *necessary*, i.e. does H appear in every preferred explanation for $\mathcal{P}$?

Also of interest are the search problems of generating one or all preferred explanations.

Complexity Results for General TBoxes

In this section, we establish the complexity of the decision problems just presented for general TBoxes from the $\mathcal{EL}$ family.

For our complexity bounds, we leverage previous work on abduction for propositional Horn and definite Horn theories, which can be straightforwardly embedded in $\mathcal{EL}$ and $\mathcal{EL}^{++}$ TBoxes:

Lemma 4. *Every definite Horn theory can be represented by a general $\mathcal{EL}$ TBox. Every Horn theory can be represented by a general $\mathcal{EL}^{++}$ TBox.*

Proof. We associate each atom v_i with an atomic concept A_i. We map a definite Horn clause $\neg v_{i_1} \vee ... \vee \neg v_{i_n} \vee v_{i_{n+1}}$ into the $\mathcal{EL}$ GCI $A_{i_1} \sqcap ... \sqcap A_{i_n} \sqsubseteq A_{i_{n+1}}$. Each negative clause $\neg v_{i_1} \vee ... \vee \neg v_{i_n}$ is mapped into the $\mathcal{EL}^{++}$ GCI $A_{i_1} \sqcap ... \sqcap A_{i_n} \sqsubseteq \bot$. □

As the complexity of the different abduction decision problems has already been established for Horn and definite Horn theories, we can use Lemma 4 to obtain lower bounds on the complexity of these problems for $\mathcal{EL}$, $\mathcal{EL}^{+}$, and $\mathcal{EL}^{++}$.

For upper bounds, we can also take advantage of work from propositional logic. It turns out that the proofs for the complexity upper bounds for Horn and definite Horn theories are very general and can be easily adapted to $\mathcal{EL}$, $\mathcal{EL}^{+}$, and $\mathcal{EL}^{++}$. Indeed, for Horn theories, essentially the only property used in the membership proofs is that satisfiability and deduction are tractable, which allows one to verify in polynomial time whether a given set of hypotheses is an explanation or not. This property is satisfied for $\mathcal{EL}^{++}$ (as well as for an extension of $\mathcal{EL}^{++}$ recently introduced in (Baader, Brandt, & Lutz 2008)). We thus find that the different abduction decision problems for $\mathcal{EL}^{++}$ TBoxes have exactly the same worst-case complexity as the corresponding decision problems for propositional Horn theories.

For some of the decision problems, like existence, definite Horn theories exhibit better complexity than Horn theories. For these cases, the upper bound proofs rely on the fact that any set of hypotheses is consistent with respect to a definite Horn theory. It is not hard to show that the same property holds for $\mathcal{EL}$ and $\mathcal{EL}^{+}$. We can thus conclude that the abduction decision problems for $\mathcal{EL}$ and $\mathcal{EL}^{+}$ TBoxes are of the same complexity as the corresponding problems for definite Horn theories.

Proposition 5. *The complexity results shown in Figure 1 are correct.*

Proof. This follows from complexity results for propositional Horn and definite Horn theories found in (Eiter & Gottlob 1995). Most results appeared there for the first time, but some were published earlier: the NP-completeness of the existence problem for Horn theories comes from (Selman & Levesque 1990), and the NP-completeness of relevance and tractability of necessity for definite Horn theories appeared in (Friedrich, Gottlob, & Nejdl 1990). □

Complexity Results for $\mathcal{EL}^{+}$ Terminologies

The embedding of definite Horn and Horn theories in $\mathcal{EL}$ and $\mathcal{EL}^{++}$ TBoxes which we used for our lower complexity bounds made use of general concept inclusions, which raises the question of whether we might achieve better complexity results if we considered terminologies instead of general TBoxes. In this section, we examine the case in which the background knowledge is represented by an (acyclic or cyclic) $\mathcal{EL}^{+}$ terminology. These results have a practical interest since some large real-world ontologies, like the biomedical ontology SNOMED (Spackman 2000) and the Gene Ontology (The Gene Ontology Consortium 2000), can be expressed as $\mathcal{EL}^{+}$ terminologies.

Decision Problem		$\mathcal{EL}$	$\mathcal{EL}^+$	$\mathcal{EL}^{++}$
Existence		P	P	NP-complete
Relevance	$\subseteq$	NP-complete	NP-complete	NP-complete
	$\leq$	$\Delta_2^P[O(\log n)]$-complete	$\Delta_2^P[O(\log n)]$-complete	$\Delta_2^P[O(\log n)]$-complete
	$\subseteq_P$	NP-complete	NP-complete	Σ_2^P-complete
	$\leq_P$	Δ_2^P-complete	Δ_2^P-complete	Δ_2^P-complete
	$\sqsubseteq_w$	Δ_2^P-complete	Δ_2^P-complete	Δ_2^P-complete
Necessity	$\subseteq$	P	P	co-NP-complete
	$\leq$	$\Delta_2^P[O(\log n)]$-complete	$\Delta_2^P[O(\log n)]$-complete	$\Delta_2^P[O(\log n)]$-complete
	$\subseteq_P$	co-NP-complete	co-NP-complete	Π_2^P-complete
	$\leq_P$	Δ_2^P-complete	Δ_2^P-complete	Δ_2^P-complete
	$\sqsubseteq_w$	Δ_2^P-complete	Δ_2^P-complete	Δ_2^P-complete

Figure 1: Complexity of Abduction for General TBoxes

Reformulation in Terms of Hitting Sets

Our first step will be to reformulate the different abduction decision problems for $\mathcal{EL}^+$ terminologies as decision problems concerning hitting sets. We recall the definition of a hitting set:

Definition 6. Let $\{S_1, ..., S_n\}$ be a collection of non-empty sets. Then a set $T \subseteq \cup_{i=1}^n S_i$ is a *hitting set* for $\{S_1, ..., S_n\}$ if and only if $T \cap S_i \neq \emptyset$ for every $1 \leq i \leq n$.

We will abuse notation and use the terms $\subseteq$-, $\leq$-, $\subseteq_P$-, $\leq_P$-, and $\sqsubseteq_w$-minimal to describe hitting sets as well as explanations. For instance, we will say that a hitting set T is $\subseteq$-minimal if no proper subset of T is a hitting set.

In what follows, we will find it convenient to work with terminologies whose right-hand sides do not contain any defined concepts as conjuncts.

Definition 7. Let $\mathcal{T}$ be an $\mathcal{EL}^+$ terminology. We say that $\mathcal{T}$ is *reduced* if each of its axioms $A \equiv C$ is such that C is a conjunction of primitive concepts and existential restrictions.

Lemma 8 shows that every $\mathcal{EL}^+$ terminology can be transformed in polynomial time into a reduced terminology. The new terminology may contain some additional concept names, but it acts exactly the same as the initial terminology over the original signature.

Lemma 8. *Let $\mathcal{T}$ be an $\mathcal{EL}^+$ terminology. We can compute in polynomial time a reduced $\mathcal{EL}^+$ terminology $\mathcal{T}'$ such that (i) $\mathcal{T}' \models \mathcal{T}$ (ii) $Sig(\mathcal{T}) \subseteq Sig(\mathcal{T}')$ and (iii) for every model $\mathcal{I}$ of $\mathcal{T}$ there is a model $\mathcal{J}$ of $\mathcal{T}'$ which coincides with $\mathcal{I}$ on $Sig(\mathcal{T})$.*

Proof. Follows from the normalization procedure for $\mathcal{EL}$ terminologies introduced in (Baader 2003). ☐

Our next lemma shows that if a primitive concept or existential restriction is entailed by a conjunction of atomic concepts with respect to an $\mathcal{EL}^+$ terminology, it must be entailed by one of the conjuncts with respect to the terminology[2].

Lemma 9. *Let $\mathcal{T}$ be an $\mathcal{EL}^+$ terminology, A be a primitive concept with respect to $\mathcal{T}$, $\exists s.C$ an $\mathcal{EL}$ concept, and $B_1 \sqcap ... \sqcap B_n$ a conjunction of atomic concepts. Then*

1. *$\mathcal{T} \models B_1 \sqcap ... \sqcap B_n \sqsubseteq A$ if and only if $\mathcal{T} \models B_j \sqsubseteq A$ for some $1 \leq j \leq n$*
2. *$\mathcal{T} \models B_1 \sqcap ... \sqcap B_n \sqsubseteq \exists s.C$ if and only if $\mathcal{T} \models B_j \sqsubseteq \exists s.C$ for some $1 \leq j \leq n$*

Proof. Because of Lemma 8 we can assume without loss of generality that $\mathcal{T}$ is reduced. For (1), let us suppose that $\mathcal{T} \not\models B_j \sqsubseteq A$ for all $1 \leq j \leq n$. Consider the interpretation $\mathcal{I}$ defined as follows:

- $\Delta^{\mathcal{I}} = \{u_1, u_2\}$
- $r^{\mathcal{I}} = \{(u_1, u_2), (u_2, u_2)\}$ for every role r
- $E^{\mathcal{I}} = \{u_2\}$ if $E = A$ or $E \equiv C \in \mathcal{T}$ and $A \in C$
- $E^{\mathcal{I}} = \{u_1, u_2\}$ otherwise

Remark that $A^{\mathcal{I}} = \{u_2\}$ but $(B_1 \sqcap ... \sqcap B_n)^{\mathcal{I}} = \{u_1, u_2\}$ since all primitive concepts are interpreted as $\{u_1, u_2\}$ as are all defined concepts whose definition does not contain A as a conjunct (which must be the case since we assumed $\mathcal{T} \not\models B_j \sqsubseteq A$). This means that $\mathcal{I}$ is not a model of $B_1 \sqcap ... \sqcap B_n \sqsubseteq A$. We now show that $\mathcal{I}$ is a model of $\mathcal{T}$. First, consider any role inclusion axiom $r_1 \circ ... \circ r_m \sqsubseteq r$ of $\mathcal{T}$. Clearly this axiom must be satisfied by $\mathcal{I}$ since all roles are interpreted as the relation $\{\{u_1, u_2\}, \{u_2, u_2\}\}$ which is idempotent under composition. Next, consider a concept definition $F \equiv G_1 \sqcap ... \sqcap G_k \sqcap \exists r_1.H_1 \sqcap ... \sqcap \exists r_l.H_l$ where the G_i are primitive concepts and the H_i are arbitrary $\mathcal{EL}$ concepts. Notice that by construction $(\exists r_i.H_i)^{\mathcal{I}} = \{u_1, u_2\}$ for every $1 \leq i \leq l$. Now if there is no G_i such that $G_i = A$, then both F and $G_1 \sqcap ... \sqcap G_k \sqcap \exists r_1.H_1 \sqcap ... \sqcap \exists r_l.H_l$ are interpreted by

[2] A more general result was proven in (Konev, Walther, & Wolter 2008) but only for $\mathcal{EL}$.

$\{u_1, u_2\}$, so the definition is satisfied. If instead we have $G_i = A$, then both F and $G_1 \sqcap ... \sqcap G_k \sqcap \exists r_1.H_1 \sqcap ... \sqcap \exists r_l.H_l$ are interpreted by $\{u_2\}$. We have thus shown that $\mathcal{I}$ satisfies $\mathcal{T}$, and hence that $\mathcal{T} \not\models B_1 \sqcap ... \sqcap B_n \sqsubseteq A$. The other direction of (1) is trivial.

For (2), we suppose that $\mathcal{T} \not\models B_j \sqsubseteq \exists s.C$ for every $1 \le j \le n$. That means that we can find for each j an intepretation $\mathcal{I}_j = \langle \Delta^j, \cdot^j \rangle$ which is a model of $\mathcal{T}$ and which contains an element $w_j \in \Delta^j$ such that $w_j \in B_j^j$ but $w_j \notin (\exists s.C)^j$. We assume without loss of generality that each $\mathcal{I}_j$ is a tree-shaped model with root w_j. We now create a new interpretation $\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$ as follows:

- $\Delta^{\mathcal{I}} = \{w\} \cup \bigcup_{j=1}^n (\Delta^j \setminus \{w_j\})$
- $r^{\mathcal{I}} = \bigcup_{j=1}^n r^j[w_j : w]$ for every role r
- $A^{\mathcal{I}} = \{w\} \cup \bigcup_{j=1}^n (A^j \setminus \{w_j\})$ if A is primitive or $A \equiv D \in \mathcal{T}$ and for every conjunct $\exists r.E$ of D there is some j such that $w_j \in (\exists r.E)^j$
- $A^{\mathcal{I}} = \bigcup_{j=1}^n (A^j \setminus \{w_j\})$ otherwise

where $r^j[w_j : w]$ is the relation obtained from r^j by replacing each instance of w_j by w. Notice that we have $w \in B_j^{\mathcal{I}}$ for every $1 \le j \le n$ since either B_j is a primitive concept, or it is a defined concept in which case by construction we have $w_j \in (\exists r.E)^j$ for every conjunct $\exists r.E$ on the right-hand side of the definition of B_j. We must also have $w \notin (\exists s.C)^{\mathcal{I}}$ since the successors of w are exactly the successors of the w_j and $w_j \notin (\exists s.C)^j$ for every j. It remains to be shown that $\mathcal{I}$ is a model of $\mathcal{T}$. Clearly every role inclusion is satisfied by $\mathcal{I}$ since the role axioms are satisfied by every $\mathcal{I}_j$. Consider some concept definition $F \equiv G_1 \sqcap ... \sqcap G_k \sqcap \exists r_1.H_1 \sqcap ... \sqcap \exists r_l.H_l$ where the G_i are primitive concepts and the H_i are arbitrary $\mathcal{EL}$ concepts. This definition must be verified for every element in $\mathcal{I}$ different from w because the interpretations $\mathcal{I}_j$ are all models of $\mathcal{T}$. Next suppose that $w \in F^{\mathcal{I}}$. Then by construction for each $\exists r_i.H_i$ there is some j such that $w_j \in (\exists r_i.H_i)^j$, which means that $w \in (\exists r_i.H_i)^{\mathcal{I}}$ since w is connected to every element which is connected to w_j in $\mathcal{I}_j$. As the primitive concepts are all true at w, it follows that both F and its right-hand side have the same interpretation in $\mathcal{I}$. Finally, suppose that $w \notin F^{\mathcal{I}}$. This means that there must be some $\exists r_i.H_i$ such that for all j we have $w_j \notin (\exists r_i.H_i)^j$. As w has no successors other than the successors of the w_j, it follows that $w \notin (\exists r_i.H_i)^{\mathcal{I}}$, so w does not belong to the interpretation of the right-hand side of definition for F. We have thus shown that $\mathcal{I}$ is a model of $\mathcal{T}$ and hence that $\mathcal{T} \not\models B_1 \sqcap ... \sqcap B_n \sqsubseteq \exists s.C$. The other direction of the equivalence is immediate. □

We are now ready to make explicit the relationship between hitting sets and explanations. Proposition 10 shows how abduction problems with respect to $\mathcal{EL}^+$ terminologies can be mapped into hitting set problems, and Proposition 11 gives a mapping in the other direction.

Proposition 10. *Consider an abduction problem $\langle \mathcal{T}, \mathcal{H}, O \rangle$ where $\mathcal{T}$ is a reduced $\mathcal{EL}^+$ terminology and $O \equiv D_1 \sqcap ... \sqcap D_n \in \mathcal{T}$. Let $Supp(D_i)$ be the set of atomic concepts E such that $\mathcal{T} \models E \sqsubseteq D_i$. Then a set $\mathcal{A}$ of atomic concepts is a standard (respectively $\star$-minimal, $\star \in \{\subseteq, \le, \subseteq_P, \le_P, \sqsubseteq_w\}$) explanation for $\langle \mathcal{T}, \mathcal{H}, O \rangle$ if and only if $\mathcal{A}$ is a standard (respectively $\star$-minimal) hitting set of $\{Supp(D_1) \cap \mathcal{H}, ..., Supp(D_n) \cap \mathcal{H}\}$.*

Proof. Let $\mathcal{A}$ be an explanation for $\mathcal{P}$. Then $\mathcal{T} \models \bigsqcap_{A \in \mathcal{A}} A \sqsubseteq D_1 \sqcap ... \sqcap D_n$. As the conjuncts D_i are all either primitive concepts or existential restrictions, it follows from Lemma 9 that for each D_i there is some $A \in \mathcal{A}$ such that $\mathcal{T} \models A \sqsubseteq D_i$. This means that $\mathcal{A} \cap (Supp(D_i) \cap \mathcal{H}) \neq \emptyset$ for every $1 \le i \le n$, i.e. $\mathcal{A}$ is a hitting set of $\{Supp(D_1) \cap \mathcal{H}, ..., Supp(D_n) \cap \mathcal{H}\}$.

For the other direction, suppose $\mathcal{A}$ is a hitting set of $\{Supp(D_1) \cap \mathcal{H}, ..., Supp(D_n) \cap \mathcal{H}\}$. Then $\mathcal{A}$ is a subset of $\mathcal{H}$ such that $\mathcal{A} \cap Supp(D_i) \neq \emptyset$ for every i. But that means that $\mathcal{T} \models \bigsqcap_{A \in \mathcal{A}} A \sqsubseteq D_1 \sqcap ... \sqcap D_n$, hence $\mathcal{T} \models \bigsqcap_{A \in \mathcal{A}} A \sqsubseteq O$, so $\mathcal{A}$ is an explanation for $\mathcal{P}$.

One can easily verify that $\mathcal{A}$ is a $\star$-minimal explanation of $\mathcal{P}$ just in the case that $\mathcal{A}$ is a $\star$-minimal hitting set (for every $\star \in \{\subseteq, \le, \subseteq_P, \le_P, \sqsubseteq_w\}$). □

The restriction to reduced terminologies in Proposition 10 is without loss of generality because of Lemma 8.

Proposition 11. *Let $S_1, ..., S_n$ be a collection of sets. A subset T of $\cup_{i=1}^n S_i$ is a standard (respectively $\star$-minimal for $\star \in \{\subseteq, \le, \subseteq_P, \le_P, \sqsubseteq_w\}$) hitting set for $\{S_1, ..., S_n\}$ if and only if T is a standard (respectively $\star$-minimal) explanation for the abduction problem $\langle \mathcal{T}, \mathcal{H}, O \rangle$ with $\mathcal{T} = \{E \equiv \bigsqcap_{E \in S_i} S_i \mid E \in \bigcup_{i=1}^n S_i\} \cup \{O \equiv S_1 \sqcap ... \sqcap S_n\}$ and $\mathcal{H} = \bigcup_{i=1}^n S_i$.*

Proof. For the first direction, suppose that T is a hitting set for $\{S_1, ..., S_n\}$. This means that $T \cap S_i \neq \emptyset$ for every $1 \le i \le n$, so $\mathcal{T} \models \bigsqcap_{E \in T} E \sqsubseteq O$, i.e. T is an explanation for $\mathcal{T}$. For the other direction, suppose that T is an explanation for $\langle \mathcal{T}, \mathcal{H}, O \rangle$ as defined above. Then by Lemma 9 for each S_i there must be some $E \in T$ such that $\mathcal{T} \models E \sqsubseteq S_i$. But that means that $T \cap Supp(S_i) \neq \emptyset$, and hence $T \cap S_i \neq \emptyset$, for all $1 \le i \le n$. It can be readily verified that if T is a preferred hitting set then so is the corresponding explanation, and vice-versa. □

Complexity for $\subseteq$-Minimality

We consider the complexity of abduction for $\mathcal{EL}^+$ terminologies using the $\subseteq$-minimality criterion. We will restrict our attention to the relevance problem since necessity is tractable even for general TBoxes.

It turns out that testing relevance for $\mathcal{EL}^+$ terminologies with respect to the $\subseteq$-minimality criterion is feasible in polynomial time, which contrasts with the NP-hardness of this task for general $\mathcal{EL}^+$ TBoxes. This result is a corollary of the corresponding result for hitting sets, which was proven in (Boros *et al.* 2000). As this result is not that well-known in the abduction literature, we include a sketch of the proof.

Proposition 12. *(Boros* et al. *2000) The problem of deciding whether an element belongs to some $\subseteq$-minimal hitting set of a collection of sets is in* P.

Proof. Let $\{S_1, ..., S_n\}$ be a collection of sets, and let e be an element appearing in some S_i. Suppose without loss of generality that there exists k such that $e \in S_i$ for all $1 \leq i \leq k$ and $e \notin S_j$ for $k+1 \leq j \leq n$. Consider the following procedure:

```
For each 1≤i≤k
   Return yes if Sj\Si ≠ ∅ for every k+1≤j≤n
Return no
```

Clearly this procedure terminates in time polynomial in $\Sigma_{i=1}^{n}|S_i|$. We now show that it outputs yes if and only if e belongs to some $\subseteq$-minimal hitting set.

First suppose the above procedure returns yes. Then there is some S_i, $1 \leq i \leq k$, such that $S_j \setminus S_i \neq \emptyset$ for every $k+1 \leq j \leq n$. Let $T = \{e\} \cup \bigcup_{j=k+1}^{n} S_j \setminus S_i$. Now clearly T is a hitting set for $\{S_1, ..., S_n\}$. As $T \cap S_i = \{e\}$, it follows that e must belong to every $\subseteq$-minimal hitting set which is a subset of T, and hence to at least one $\subseteq$-minimal hitting set.

Suppose next that e belongs to a $\subseteq$-minimal hitting set T of $\{S_1, ..., S_n\}$. Then there must be some S_i such that $T \cap S_i = \{e\}$, since otherwise $T \setminus \{e\}$ would be a hitting set. It follows that $T \cap S_j \subseteq S_j \setminus S_i$, hence $S_j \setminus S_i \neq \emptyset$, for every $k+1 \leq j \leq n$, so the procedure will return yes when examining S_i. □

Corollary 13. *The relevance problem for the $\subseteq$-minimality criterion is tractable for abduction problems with respect to $\mathcal{EL}^+$ terminologies.*

Corollary 13 is interesting since it shows that abduction w.r.t $\mathcal{EL}^+$ terminologies is better behaved than abduction with respect to either general $\mathcal{EL}$ TBoxes or definite Horn theories. This result has practical implications since it means that before starting to generate explanations, we can remove from consideration all hypotheses which are guaranteed not to appear in any $\subseteq$-minimal explanation, thereby reducing the search space.

What would be even nicer is to be able to decide whether a *set of hypotheses* is contained in some $\subseteq$-minimal explanation, as this would enable us to selectively enumerate explanations. Unfortunately, this problem is not tractable: the corresponding problem for hitting sets was proven NP-complete in (Boros *et al.* 2000). We include a brief sketch of the proof.

Proposition 14. *(Boros* et al. *2000) The problem of deciding whether a given set is contained in some $\subseteq$-minimal hitting set of a collection of sets is* NP*-hard.*

Proof Sketch. The proof is via a reduction from SAT. Let $\varphi = \lambda_1 \wedge ... \wedge \lambda_m$ be a propositional CNF over the set of atoms $\{v_1, ..., v_n\}$. Consider the collection composed of the following sets:

- the sets $\{T_i, P_i\}$ and $\{F_i, P_i\}$ for each $1 \leq i \leq n$
- the set $\{T_j \mid v_j \in Pos(\lambda_k)\} \cup \{F_j \mid v_j \in Neg(\lambda_k)\}$ for each $1 \leq k \leq m$

One can then verify that φ is satisfiable just in the case that $\{P_1, ..., P_n\}$ is contained in some $\subseteq$-minimal hitting set of the above collection of sets. □

Corollary 15. *For abduction problems with respect to $\mathcal{EL}^+$ terminologies, deciding whether a given subset of $\mathcal{H}$ is included in some $\subseteq$-minimal explanation is* NP*-complete.*

In the proof of Proposition 14, the reduction hinges on our ability to test whether sets of arbitrary size can be extended to $\subseteq$-minimal hitting sets. If we place a bound on the size of sets considered, then the complexity drops back down to P. In fact, this problem has been shown to belong to the class NC which means that it can be efficiently solved on a parallel computer.

Proposition 16. *(Boros* et al. *2000) Fix $k \in \mathbb{N}$. The problem of deciding whether a given set with cardinality less than k is contained in some $\subseteq$-minimal hitting set of a collection of sets is in* NC.

Corollary 17. *Fix $k \in \mathbb{N}$. For abduction problems with respect to $\mathcal{EL}^+$ terminologies, the problem of deciding whether a given subset S of $\mathcal{H}$ with $|S| \leq k$ is included in some $\subseteq$-minimal explanation belongs to* NC.

Complexity for Other Minimality Criteria

We now consider the complexity of the relevance and necessity tasks for the other minimality criteria. We demonstrate that for these criteria even a restriction to acyclic propositional terminologies is not sufficient to yield a drop in complexity.

We begin by showing the $\Delta_2^P[O(\log n)]$-completeness of relevance and necessity for hitting sets. Our proof is a modified version of the proof in (Eiter & Gottlob 1995) of the corresponding result for definite Horn theories.

Proposition 18. *The problem of deciding whether an element appears in some or every cardinality-minimal hitting set is $\Delta_2^P[O(\log n)]$-hard.*

Proof. We reduce the following problem which was shown $\Delta_2^P[O(\log n)]$-complete in (Eiter & Gottlob 1995): given a set $C = \{\lambda_1, ..., \lambda_m\}$ of propositional clauses over variables $\{v_1, ..., v_n\}$ and a distinguished clause λ_l, decide whether λ_l is verified by every model which satisfies a maximum number of clauses from C. We create a collection of sets consisting of:

- $\{X_{i,j}, X'_{i,j}\}$ for every $1 \leq i \leq n$ and $1 \leq j \leq m$
- $\{X_{i,j} \mid v_i \in Pos(\lambda_k)\} \cup \{X'_{i,j} \mid v_i \in Neg(\lambda_k)\} \cup \{U_k\}$ for every $1 \leq k \leq m$ and $1 \leq j \leq m$
- $\{U_l, U_{m+1}\}$

We will show that U_l appears in every $\leq$-minimal hitting set just in the case that there is some model which satisfies a maximum number of clauses in C but does not satisfy λ_l. This yields the $\Delta_2^P[O(\log n)]$-hardness of the necessity problem, and it also gives us $\Delta_2^P[O(\log n)]$-hardness for the relevance problem since U_{m+1} appears in some $\leq$-minimal hitting set if and only if U_l does not belong to every $\leq$-minimal hitting set.

We first remark that the cardinality of the $\leq$-minimal hitting sets is at most $nm+z+1$, where z is the minimum number of clauses from C unsatisfied by a model. To see why, let $\mathcal{I}$ be a model satisfying a maximum number of clauses from

C, and set $S = \{X_{i,j} \,|\, \mathcal{I} \models v_i, 1 \leq j \leq m\} \cup \{X'_{i,j} \,|\, \mathcal{I} \models \neg v_i, 1 \leq j \leq m\} \cup \{U_k \,|\, \mathcal{I} \not\models \lambda_k\} \cup \{U_{m+1}\}$. It is easy to see that S is a hitting set and $|S| = nm + z + 1$. We also obtain a hitting set if we replace U_{m+1} by U_l. In this case the hitting set has either cardinality $nm + z + 1$ or $nm + z$ depending on whether or not $\mathcal{I} \models \lambda_l$.

We next remark that if S is a hitting set with cardinality $nm + z$, then (i) there is no pair of indices i and j such that $\{X_{i,j}, X'_{i,j}\} \in S$ and (ii) $U_l \in S$. For (i), suppose to the contrary that S contains both $X_{i,j}$ and $X'_{i,j}$ for some i and j. Then the set $\{\lambda_k \,|\, U_k \notin S\}$ is unsatisfiable since S will contain less than z elements of type U_k, as for every $1 \leq i \leq n$ and $1 \leq j \leq m$ either $X_{i,j}$ or $X_{i,j'}$ must belong to S. That means that in order to satisfy the sets in the second bullet, we must have at least $n + 1$ elements from $\{X_{i,j} \,|\, 1 \leq i \leq n\} \cup \{X'_{i,j} \,|\, 1 \leq i \leq n\}$ for each $1 \leq j \leq m$. It follows that S has cardinality at least $(n + 1)m$ which is a contradiction since $z < m$. For (ii), suppose for a contradiction that $U_l \notin S$. Then we must have $U_{m+1} \in S$, so the set $\{\lambda_k \,|\, U_k \notin S\}$ must be unsatisfiable. As S does not contain any pair of the form $\{X_{i,j}, X'_{i,j}\}$, there must be some sets in the second bullet which have an empty intersection with S, contradicting the fact that S is a hitting set.

Similarly we can show that any hitting set S must have cardinality at least $nm + z$ since otherwise S will contain less than z elements of type U_i, so $\{\lambda_j \,|\, U_j \notin S\}$ will be unsatisfiable. But then there must be some set $\{X_{i,j} \,|\, v_i \in Pos(\lambda_k)\} \cup \{X'_{i,j} \,|\, v_i \in Neg(\lambda_k)\} \cup \{U_k\}$ with an empty intersection with S, contradicting our assumption that S is a hitting set.

We are now ready to prove the result. For the first direction, we suppose that U_l appears in every $\leq$-minimal hitting set. This means that the minimal cardinality must be $nm+z$. Pick some $\leq$-minimal hitting set S, and let the interpretation $\mathcal{I}$ be defined as follows: $\mathcal{I} \models v_i$ if and only if $X_{1,i} \in S$. We know that $\mathcal{I}$ is well-defined since there is no index i such that both $X_{1,i}$ and $X'_{1,i}$ are in S. It is easy to see that $\mathcal{I}$ satisfies λ_i just in the case that $U_i \notin S$, and hence satisfies a maximum number of clauses from C. Moreover, since $U_l \in S$, we have $\mathcal{I} \not\models \lambda_l$, as desired.

For the other direction, suppose that there is some model $\mathcal{I}$ which satisfies a maximum number of clauses from C but falsifies λ_l. Consider the set $S = \{X_{i,j} \,|\, \mathcal{I} \models v_i, 1 \leq j \leq m\} \cup \{X'_{i,j} \,|\, \mathcal{I} \models \neg v_i, 1 \leq j \leq m\} \cup \{U_i \,|\, \mathcal{I} \not\models \lambda_i\}$. It can be easily verified that S is a hitting set of cardinality $nm + z$. But that means that all $\leq$-minimal hitting sets have cardinality $nm + z$, so U_l must belong to every $\leq$-minimal hitting set. □

Corollary 19. *Deciding whether a given hypothesis is $\leq$-necessary or $\leq$-relevant to an abduction problem is $\Delta_2^P[O(\log n)]$-hard even when the background theory is an acyclic propositional $\mathcal{EL}$ terminology.*

We next show that relevance and necessity are respectively NP-complete and co-NP-complete when the minimality criterion is prioritized set inclusion.

Proposition 20. *The problem of testing if an element belongs to some (respectively every) prioritized inclusion-minimal hitting set is* NP*-hard (respectively co-*NP*-hard).*

Proof. For necessity, we give a reduction from UNSAT. Let $\varphi = \lambda_1 \wedge ... \wedge \lambda_m$ be a propositional CNF over the set of atoms $\{v_1, ..., v_n\}$. Consider the collection composed of the following sets:

- the sets $\{T_i, F_i\}$ for each $1 \leq i \leq n$
- the set $\{T_j \,|\, v_j \in Pos(\lambda_k)\} \cup \{F_j \,|\, v_j \in Neg(\lambda_k)\} \cup \{U\}$ for each $1 \leq k \leq m$

and the priority ordering $P = \langle \{T_i, F_i\}_{i=1}^n, \{U\} \rangle$ We claim that φ unsatisfiable just in the case that U is contained in every $\subseteq_P$-minimal hitting set of the above collection of sets. Clearly any $\subseteq_P$-minimal hitting set must contain exactly one of T_i and F_i for each i, i.e. it corresponds to a propositional interpretation. It follows that if φ is unsatisfiable, then every hitting set must contain U, and conversely, if φ is satisfiable, there are $\subseteq_P$-minimal hitting sets which do not contain U.

To show NP-hardness of relevance, we show how to reduce the complement of necessity to relevance. Let $\Sigma = \{S_1, ..., S_n\}$ be a collection of sets, and let $P = \langle P_1, ..., P_k \rangle$ be a partition of the elements in $\cup_{i=1}^n S_i$. Then an element E is not a member of every $\subseteq_P$-minimal hitting set of Σ just in the case that E' belongs to some $\subseteq_{P'}$-minimal hitting set of $\Sigma' = \Sigma \cup \{\{E, E'\}\}$ where $P' = \langle P_1, ..., P_k, \{E'\} \rangle$. □

Corollary 21. *Deciding whether a hypothesis is $\subseteq_P$-relevant (respectively $\subseteq_P$-necessary) is* NP*-hard (respectively co-*NP*-hard) even when the background theory is an acyclic propositional $\mathcal{EL}$ terminology.*

Finally, for prioritized cardinality minimality, we prove Δ_2^P-hardness by translating the reduction used for definite Horn theories (Eiter & Gottlob 1995) into hitting sets. The Δ_2^P-hardness for the $\sqsubseteq_w$-minimality criterion follows as a corollary.

Proposition 22. *The problem of deciding whether an element belongs to some or every $\leq_P$-minimal hitting set of a collection of sets is Δ_2^P-hard.*

Proof. Following (Eiter & Gottlob 1995), we give a reduction of the following Δ_2^P-complete problem: given a satisfiable set of clauses $C = \{C_1, ..., C_m\}$ on variables $\{v_1, ..., v_n\}$, decide whether the lexicographically maximum (with respect to to $v_1, ..., v_n$) interpretation satisfying C verifies v_n. Consider the following collection of sets:

- the sets $\{X_i, X'_i\}$ and $\{X_i, X''_i\}$ for each $1 \leq i \leq n$
- the set $\{X_i \,|\, v_i \in Pos(C_j)\} \cup \{X'_i \,|\, v_i \in Neg(C_j)\}$ for each $1 \leq j \leq m$

and the following prioritization:

$$P = \{\{X_i, X'_i \,|\, 1 \leq i \leq n\}, \{X''_1\}, ..., \{X''_n\}\}$$

We intend to show that X_n belongs to some (or every) $\leq_P$-minimal hitting set just in the case that the lexicographically maximum model of C verifies v_n. We remark that every $\leq_P$-minimal hitting set must contain either X_i or both X'_i

and X_i'' for every $1 \leq i \leq n$. This means that every $\leq_P$-minimal hitting set S corresponds to a model $\mathcal{I}_S$ of C defined by $\mathcal{I}_S \models v_i$ if and only if $X_i \in S$. Moreover, hitting sets which do not contain X_i'' (and hence contain X_i) are preferred over those containing X_i'' (and hence not containing X_i), with preference given to those not containing X_1'', then those not containing X_2'', etc. It follows that if $\mathcal{I}_S$ is the interpretation associated with a $\leq_P$-minimal hitting set S then $\mathcal{I}_S$ is the unique lexicographically maximal model of C. It follows that there is a single $\leq_P$-minimal hitting set. We thus find that the lexicographically maximum model of C verifies v_n just in the case that some $\leq_P$-minimal hitting set contains X_n just in the case that all $\leq_P$-minimal hitting sets contains X_n. □

Corollary 23. *The problem of deciding whether an element belongs to some (respectively every) $\sqsubseteq_w$-minimal hitting set of a collection of sets is Δ_2^P-hard.*

Proof. One can verify that a hitting set is $\leq_P$-minimal, $P = \langle P_1, ..., P_k \rangle$, just in the case that it is $\sqsubseteq_w$-minimal where w is defined as follows:

$$w(H) = d^{k-i} \text{ if } H \in P_i \text{ and } d = 1 + max\{|P_i|, 1 \leq i \leq k\}$$

Refer to (Eiter & Gottlob 1995) for more details. □

Corollary 24. *The relevance and necessity problems with respect to either the $\leq_P$ or $\sqsubseteq_w$ minimality criteria are Δ_2^P-hard even when the background theory is an acyclic propositional $\mathcal{EL}$ terminology.*

Generating Abductive Explanations

In this section we consider the complexity of producing some or all abductive explanations.

We begin with the complexity of generating a single preferred explanation. This problem is tractable for $\mathcal{EL}^+$ abduction problems with respect to the $\subseteq$- and $\subseteq_P$- criteria, but is NP-hard in all other cases.

Proposition 25. *It is possible to generate a single $\subseteq$- or $\subseteq_P$-minimal explanation in polynomial time for $\mathcal{EL}^+$ abduction problems. Computing a $\subseteq$- or $\subseteq_P$-minimal explanation for $\mathcal{EL}^{++}$ abduction problems is* NP*-hard, as is the problem of generating a single $\leq$-, $\leq_P$-, or $\sqsubseteq_w$-minimal explanation, even when the background theory is an acyclic propositional $\mathcal{EL}$ terminology.*

Proof. To generate a $\subseteq_P$-minimal explanation for $\mathcal{EL}^+$ abduction problems, we remove one by one elements from P_1 until we find a subset S_1 of P_1 such that $S_1 \cup P_2 \cup ... \cup P_k$ is an explanation but no proper subset of S_1 satisfies this property. We then move on to P_2, choosing a subset S_2 of P_2 such that $S_1 \cup S_2 \cup P_2 \cup ... \cup P_k$ is an explanation but no proper subset of S_2 has this property. It can be easily verified that the set $S_1 \cup ... \cup S_k$ obtained by iterating this process is a $\subseteq_P$-minimal explanation. For $\mathcal{EL}^{++}$ problems, NP-hardness follows directly from the NP-hardness of the existence problem. For $\leq$-minimal explanations with respect to acyclic propositional $\mathcal{EL}$ terminologies, NP-hardness follows from the NP-completeness of deciding the existence of a hitting set of less than a given cardinality (Garey & Johnson 1979). This also yields the NP-hardness of computing a single $\leq_P$- or $\sqsubseteq_w$-minimal explanation, since $\leq$-minimal explanations are just $\leq_P$-minimal explanations where $P = \mathcal{H}$ or $\sqsubseteq_w$-minimal explanations where $w = 1$. □

Given that it is difficult to produce even one $\subseteq$- or $\subseteq_P$-minimal explanation for $\mathcal{EL}^{++}$ abduction problems, it follows that the problem of generating all such explanations is not feasible in output-polynomial time (assuming of course P$\neq$ NP). The same holds for computing the set of all $\leq$-, $\leq_P$, or $\sqsubseteq_w$-minimal explanations even for acyclic propositional terminologies.

We can also show there is no output-polynomial time algorithm for generating all $\subseteq$- and $\subseteq_P$-minimal explanations for $\mathcal{EL}$ abduction problems. This is a corollary of the corresponding result for definite Horn theories, which was shown in (Friedrich, Gottlob, & Nejdl 1990) using a similar result from (Bylander *et al.* 1989).

Proposition 26. *If* P$\neq$NP*, then it is not possible to generate all $\subseteq$-minimal explanations of a definite Horn abduction problem in output-polynomial time.*

Proof. We show that the problem of deciding whether an additional explanation exists for a definite Horn abduction problem is NP-hard via a reduction from SAT. Let $\varphi = \lambda_1 \wedge ... \wedge \lambda_m$ be a CNF formula over the variables $\{v_1, ..., v_n\}$. We assume that there is no literal which appears in every clause in φ. This is without loss of generality since otherwise φ is trivially satisfiable. Consider the definite Horn theory T composed of the following clauses:

- $\neg v_j \vee c_i$ for $i = 1, ..., m$ and $v_j \in Pos(\lambda_i)$
- $\neg v_j' \vee c_i$ for $i = 1, ..., m$ and $v_j \in Neg(\lambda_i)$
- $\neg v_j \vee \neg v_j' \vee z$ for every $j = 1, ..., n$
- $\neg c_1 \vee ... \vee \neg c_m \vee z$

We will be interested in the explanations of z with respect to T using the hypotheses $\cup_{i=1}^n \{v_i, v_i'\}$. Because there is no literal common to all of the clauses in φ, there can be no explanation for z of the form $\{v_i\}$ or $\{v_i'\}$. This means that for every i the set $\{v_i, v_i'\}$ is a $\subseteq$-minimal explanation for z with respect to T. Any additional $\subseteq$-minimal explanation for z must correspond to a consistent set of literals which implies φ. It follows that such an explanation exists just in the case that φ is satisfiable. □

We should point out that in the proof of Proposition 26 the reduction depends crucially on the fact that only some of the propositional variables are included as hypotheses. Indeed, it was shown in (Eiter & Makino 2002) that if all propositional variables are considered hypotheses then the entire set of $\subseteq$-minimal explanations for a Horn abduction problem can be computed in output-polynomial time. This positive result cannot be extended to $\mathcal{EL}$ abduction problems:

Proposition 27. *If* P$\neq$NP*, then it is not possible to generate all $\subseteq$-minimal explanations to an $\mathcal{EL}$ abduction problem $\langle \mathcal{T}, \mathcal{H}, O \rangle$ in output-polynomial time, even when $\mathcal{H} = Sig(\mathcal{T}) \cap N_C$.*

Proof. The proof is almost identical to the proof of Proposition 26, except that we replace T by the following TBox:

- $V_i \sqsubseteq \exists r.C_i$ for $i = 1, ..., m$ and $v_i \in Pos(\lambda_i)$
- $V_i' \sqsubseteq \exists r.C_i$ for $i = 1, ..., m$ and $v_i \in Neg(\lambda_i)$
- $V_i \sqcap V_i' \sqsubseteq Z$ for every $i = 1, ..., n$
- $\exists r.C_1 \sqcap ... \sqcap \exists r.C_m \sqsubseteq Z$ □

For $\mathcal{EL}$ and $\mathcal{EL}^+$ terminologies, the results in this paper show that generating all $\subseteq$-minimal abductive explanations is equivalent to generating all inclusion-minimal hitting sets. It is a longstanding open question in complexity theory whether the latter problem is solvable in output-polynomial time (cf. (Eiter & Gottlob 2002; Eiter, Gottlob, & Makino 2006) for discussion), but there is evidence that this may be the case.

In any case, one advantage to rephrasing abduction problems for $\mathcal{EL}^+$ terminologies in terms of hitting sets is that we can exploit existing techniques for enumerating hitting sets to generate explanations. There are a variety of different hitting set algorithms available, including some which run in quasi-polynomial total time (cf. (Eiter, Gottlob, & Makino 2006) for references).

Extending our Abductive Framework

In (Elsenbroich, Kutz, & Sattler 2006), a number of different types of abduction problems for description logic are discussed. Our abductive framework corresponds in their terminology to a form of *simple concept abduction*, in which the target language for explanations is conjunctions of atomic concepts from $\mathcal{H}$. A more general type of concept abduction, called *conditionalized concept abduction*, is also discussed, in which explanations only need to entail the observation when taken together with a given concept C. In our framework, this would mean replacing the second bullet in Definition 2 by $\mathcal{T} \models A_1 \sqcap ... \sqcap A_n \sqcap C \sqsubseteq O$.

This suggests two possible ways of extending our abductive framework: (1) move to a richer target language (2) incorporate conditions. We will only discuss the first option, as our framework and results can be straightforwardly modified to handle conditionalized concept abduction.

Probably the most obvious and the most interesting way of extending our notion of explanation is to replace the set of hypotheses $\mathcal{H}$ by a signature S and let explanations be arbitrary concepts over S rather than conjunctions of concept names from $\mathcal{H}$.

Our minimality criteria must be generalized accordingly. Three criteria for general concept explanations were proposed in (Di Noia, Di Sciascio, & Donini 2007): irreducibility, length, and subsumption. Irreducibility generalizes inclusion-minimality: an explanation C is said to be irreducible, or $\sqcap$-minimal, if no concept obtained from C by removing a conjunct from C (at any level) is an explanation. Length-minimality generalizes our cardinality criterion: an explanation is C is length-minimal, written $\leq$-minimal (abusing notation), if there is no explanation D with $|D| < |C|$. Finally, another possible criterion is subsumption with respect to the background theory $\mathcal{T}$: an explanation C is $\sqsubseteq_{\mathcal{T}}$-minimal if there is no explanation D such that $\mathcal{T} \models C \sqsubseteq D$ and $\mathcal{T} \not\models D \sqsubseteq C$.

The principal difficulty in allowing general concepts as explanations is that explanations may be much larger than in the hypothesis-based setting. For instance, the next example (borrowed from (Konev, Walther, & Wolter 2008)) demonstrates that even for acyclic $\mathcal{EL}$ terminologies, the smallest explanations may have exponential size.

Example 28. Consider the acyclic TBox $\mathcal{T} = \{A \equiv \overline{A} \sqcap B_0\} \cup \{B_{i+1} \equiv \exists r.B_i \sqcap \exists s.B_i \mid 0 \leq i \leq n-1\} \cup \{O \equiv B_n\}$, and let $\mathcal{P} = \langle \mathcal{T}, \{A, r, s\}, O\rangle$. Construct the concept C_n as follows: $C_0 = A$, and $C_{i+1} = \exists r.C_i \sqcap \exists s.C_i$ for every $0 \leq i \leq n$. It is not hard to show that the concept C_n is a $\sqcap$-, $\sqsubseteq_{\mathcal{T}}$-, and $\leq$-minimal explanation for $\mathcal{P}$ and is exponential size in $|\mathcal{P}|$.

If we allow general TBoxes, the shortest explanations may have doubly-exponential size.

Example 29. (Borrowed from (Lutz & Wolter 2007)) Let $\mathcal{P} = \langle \mathcal{T}, \{A, r, s\}, O\rangle$ where $\mathcal{T}$ is composed of the following axioms:

- $A \sqsubseteq \overline{X}_0 \sqcap ... \sqcap \overline{X}_{n-1}$
- For $i < n$: $\bigsqcap_{\rho \in \{r,s\}} \exists \rho.(\overline{X}_i \sqcap X_0 \sqcap ... \sqcap X_{i-1}) \sqsubseteq X_i$
- For $i < n$: $\bigsqcap_{\rho \in \{r,s\}} \exists \rho.(X_i \sqcap X_0 \sqcap ... \sqcap X_{i-1}) \sqsubseteq \overline{X}_i$
- For $j < i < n$: $\bigsqcap_{\rho \in \{r,s\}} \exists \rho.(\overline{X}_i \sqcap \overline{X}_j) \sqsubseteq \overline{X}_i$
- For $j < i < n$: $\bigsqcap_{\rho \in \{r,s\}} \exists \rho.(X_i \sqcap \overline{X}_j) \sqsubseteq X_i$
- $X_0 \sqcap ... \sqcap X_{n-1} \sqsubseteq O$

Define C_i as follows: $C_0 = A$, $C_{i+1} = \exists r.C_i \sqcap \exists s.C_i$. It can be verified that C_{2^n-1} is a $\sqcap$-, $\sqsubseteq_{\mathcal{T}}$-, and $\leq$-minimal explanation and has doubly-exponential size in $|\mathcal{P}|$.

Clearly the fact that the smallest explanations for an abduction problem may have superpolynomial size means that we will not be able to employ the same methods as in the hypothesis-based setting, since all of the proofs require us to produce or guess explanations and check that they satisfy certain properties. One might however take a more pragmatic approach and simply require that explanations have only polynomial size. In this case, we find that the worst-case complexity of abduction with general explanations is not much higher than with hypothesis-based explanations.

Definition 30. A *signature-bounded abduction problem* is a tuple $\langle \mathcal{T}, S, O\rangle$, where $\mathcal{T}$ is a TBox, S a signature, and O a concept name. Given such a problem and a polynomial function f, an *f-explanation* is a concept C satisfying: $Sig(C) \subseteq S$, $|C| \leq f(|\mathcal{T}|)$, $\mathcal{T} \models C \sqsubseteq O$, and C is satisfiable with respect to $\mathcal{T}$. An element of S is f-relevant (respectively f-necessary) if it belongs to the signature of some (respectively all) minimal f-explanations.

Proposition 31. *Fix some polynomial function f. Let $\mathcal{P} = \langle \mathcal{T}, S, O\rangle$ be a signature-bounded abduction problem, where $\mathcal{T}$ is an $\mathcal{EL}^{++}$ TBox. Deciding whether $\mathcal{P}$ admits an f-explanation is* NP*-complete. Deciding if an element from S is f-relevant to $\mathcal{P}$ is*

- NP*-complete for $\sqcap$-minimality*
- $\Delta_2^P[O(\log n)]$*-complete for $\leq$-minimality*

Deciding whether an element in S is f-necessary to $\mathcal{P}$ is

- *co*-NP-*complete for* $\sqcap$-*minimality*
- $\Delta_2^P[O(\log n)]$-*complete for* $\leq$-*minimality*

Hardness holds even when $\mathcal{T}$ is an acyclic propositional $\mathcal{EL}$ terminology.

Proof. The existence problem is in NP since we can just guess some concept on S with size at most $f(\mathcal{T})$ and check whether it is an explanation. NP-hardness follows from the NP-hardness of deciding whether a hitting set of size at most k exists (just let f be the constant function k).

Membership of the f-relevance (respectively f-necessity) problem for $\sqcap$-minimality in NP (respectively co-NP) is straightforward: we guess a concept C with $|C| \leq f(|\mathcal{T}|)$ and $Sig(C) \subseteq S$ which contains or excludes the symbol in question and verify in polynomial time that it is an explanation and that no concept obtained by removing a conjunct from C is an explanation.

Hardness is shown via a reduction from the hitting set problem. Let $\Sigma = \{S_1, ..., S_n\}$ be a collection of sets and k an integer. Let $\Sigma' = \Sigma \cup \{\{X, Y\}\}$ where X and Y do not appear in Σ. Then Σ possesses a hitting set of size at most k if and only if X appears in some $\subseteq$-minimal hitting set of Σ' of size at most $k+1$ if and only if X does not appear in every $\subseteq$-minimal hitting set of Σ' of size at most $k+1$.

For $\leq$-minimality, membership is similar to the proof for Horn or $\mathcal{EL}^{++}$ theories: we use binary search to determine the minimal size of an explanation using a logarithmic number of calls to an NP-oracle. We then make a final call to the NP-oracle to determine whether there exists an explanation of this size containing or excluding the hypothesis in question. Hardness follows from the $\Delta_2^P[O(\log n)]$ hardness of $\leq$-minimality in the hypothesis-based setting (cf. Proposition 18). □

Related Work

There has been surprisingly little research addressing the problem of abduction for description logics, even though it was argued in (Elsenbroich, Kutz, & Sattler 2006) that this topic is quite relevant from an applications point of view. One notable exception is the work of Di Noia and colleagues (Di Noia *et al.* 2003; Di Noia, Di Sciascio, & Donini 2007) who investigate conditionalized concept abduction for the description logic $\mathcal{ALN}$. Their abductive framework is more expressive than ours it allows for general explanations of arbitrary size. However, it is also less expressive than ours in that there is no way to restrict the signature of explanations. This is actually a rather significant difference since it means that in their framework an observation is always its own explanation, a fact which is exploited in their complexity results as well as in their algorithm for computing a single irreducibility-minimal explanation. Indeed the explanation returned by their algorithm is the observation with some subconcepts removed. While such an explanation is appropriate in their chosen application (matchmaking), it would be considered uninformative in more traditional applications of abduction.

We can also cite the work of Cialdea Mayer and Pirri (1995) on abduction for modal logic. They provide an algorithm for computing the set of abductive explanations for several modal logics. However, their work is less relevant to our own since their framework does not take into account TBoxes.

Our work also bears some similarity with axiom pinpointing (Baader, Peñaloza, & Suntisrivaraporn 2007), in which the goal is to find inclusion-minimal subsets of a TBox which imply a given axiom when taken together with a (possibly empty) fixed TBox.

Conclusion and Future Work

In this paper, we studied the computational complexity of the main abductive decision problems for the case in which the background knowledge is expressed by a description logic in the $\mathcal{EL}$ family. By leveraging results from propositional logic, we showed that for general TBoxes, the complexity of abduction for $\mathcal{EL}$ and $\mathcal{EL}^+$ is exactly the same as for propositional definite Horn theories; for $\mathcal{EL}^{++}$, the complexity matches that of propositional Horn theories. Our results are quite positive since they show that the additional expressivity afforded by the $\mathcal{EL}$ description logics over propositional Horn and definite Horn theories does not incur any additional computational cost, either for deductive or for abductive reasoning. This serves as further evidence to the interest of the $\mathcal{EL}$ family of description logics.

We also investigated the special case of abduction with respect to $\mathcal{EL}$ and $\mathcal{EL}^+$ terminologies, showing how the different abduction decision problems could be reformulated in terms of hitting sets. We then used this correspondence to prove that for $\subseteq$-minimality, the relevance problem becomes tractable, thereby showing that abduction with respect to $\mathcal{EL}^+$ terminologies is actually better-behaved than for definite Horn theories. Again, we take this as a promising sign especially since some large real world ontologies are expressible as $\mathcal{EL}^+$ terminologies.

Although the focus of this paper was abductive reasoning for description logics, some of our results are also of interest for propositional logic. In particular, our hardness results for deciding whether an element appears in some or every $\leq$-, $\subseteq_P$-, $\leq_P$-, or $\sqsubseteq_w$-minimal hitting set strengthen the corresponding results for definite Horn theories, and make it clear that hitting sets are the real source of complexity. As hitting sets provide about the simplest abductive framework one could imagine, our results suggest that for these minimality criteria the complexity for definite Horn theories, and for $\mathcal{EL}^+$ TBoxes, can be considered optimal.

There are several problems to be addressed in future work. First, we are interested in experimenting with different algorithms for hitting sets to see what kind of performance they give on large-scale $\mathcal{EL}^+$ terminologies. Second, we want to develop algorithms for computing explanations with respect to general $\mathcal{EL}$, $\mathcal{EL}^+$, and $\mathcal{EL}^{++}$ TBoxes. Finally, we would like to investigate the case in which general concepts of arbitrary size are admitted as explanations.

References

Baader, F.; McGuiness, D. L.; Nardi, D.; and Patel-Schneider, P., eds. 2003. *The Description Logic Handbook*. Cambridge University Press.

Baader, F.; Brandt, S.; and Lutz, C. 2005. Pushing the $\mathcal{EL}$ envelope. In *Proceedings of IJCAI'05*, 364–369.

Baader, F.; Brandt, S.; and Lutz, C. 2008. Pushing the $\mathcal{EL}$ envelope further. In *Proceedings of the OWLED Workshop*.

Baader, F.; Peñaloza, R.; and Suntisrivaraporn, B. 2007. Pinpointing in the description logic $\mathcal{EL}^+$. In *Proceedings of KI 2007*, number 4667 in LNAI, 52–67.

Baader, F. 2003. Terminological cycles in a description logic with existential restrictions. In *Proceedings of IJ-CAI'03*, 325–330.

Boros, E.; Elbassioni, K. M.; Gurvich, V.; and Khachiyan, L. 2000. An efficient incremental algorithm for generating all maximal independent sets in hypergraphs of bounded dimension. *Parallel Processing Letters* 10(4):253–266.

Bylander, T.; Allemang, D.; Tanner, M. C.; and Josephson, J. R. 1989. Some results concerning the computational complexity of abduction. In *Proceedings of KR'89*, 44–54.

Bylander, T.; Allemang, D.; Tanner, M.; and Josephson, J. 1991. The computational complexity of abduction. *Artificial Intelligence* 49(25-60).

Cialdea Mayer, M., and Pirri, F. 1995. Propositional abduction in modal logic. *Logic Journal of the IGPL* 3(6):907–919.

Creignou, N., and Zanuttini, B. 2006. A complete classification of the complexity of propositional abduction. *SIAM Journal on Computing* 36(1):207–229.

Di Noia, T.; Di Sciascio, E.; Donini, F. M.; and Mongiello, M. 2003. Abductive matchmaking using description logics. In *Proceedings of IJCAI'03*, 337–342.

Di Noia, T.; Di Sciascio, E.; and Donini, F. M. 2007. Semantic matchmaking as non-monotonic reasoning: A description logic approach. *Journal of Artifical Intelligence Research* 29:269–307.

Eiter, T., and Gottlob, G. 1995. The complexity of logic-based abduction. *Journal of the ACM* 42(1):3–42.

Eiter, T., and Gottlob, G. 2002. Hypergraph transversal computation and related problems in logic and AI. In *Proceedings of JELIA'02*, 549–564.

Eiter, T., and Makino, K. 2002. On computing all abductive explanations. In *Proceedings of AAAI'02*, 62–67.

Eiter, T.; Gottlob, G.; and Makino, K. 2006. Computational aspects of monotone dualization: A brief survey. Technical report, Vienna Technical University.

Elsenbroich, C.; Kutz, O.; and Sattler, U. 2006. A case for abductive reasoning over ontologies. In *Proceedings of OWLED'06*.

Friedrich, G.; Gottlob, G.; and Nejdl, W. 1990. Hypothesis classification, abductive diagnosis, and therapy. In *Proceedings of the Expert Systems in Engineering Workshop*, number 462 in LNAI, 69–78.

Garey, M., and Johnson, D. S. 1979. *Computers and intractability. A guide to the theory of NP-completeness*. W. H. Freeman.

Konev, B.; Walther, D.; and Wolter, F. 2008. The logical difference problem for description logic terminologies. In *Proceedings of IJCAR'08*.

Lutz, C., and Wolter, F. 2007. Conservative extensions in the lightweight description logic $\mathcal{EL}$. In *Proceedings of CADE-21*, volume 4603 of *LNAI*, 84–99.

Papadimitriou, C. 1994. *Computational Complexity*. Addison Welsey.

Selman, B., and Levesque, H. J. 1990. Abductive and default reasoning: A computational core. In *Proceedings of AAAI'90*, 343–348.

Selman, B., and Levesque, H. J. 1996. Support set selection for abductive and default reasoning. *Artificial Intelligence* 82:259–272.

Spackman, K. 2000. Managing clinical terminology hierarchies using algorithmic calculation of subsumption: Experience with SNOMED-RT. *Journal of the American Medical Informatics Association* Fall Symposium Special Issue.

The Gene Ontology Consortium. 2000. Gene ontology: Tool for the unification of biology. *Nature Genetics* 25:25–29.