TECHNISCHE
UNIVERSITÄT
DRESDEN

Dresden University of Technology
Institute for Theoretical Computer Science
Chair for Automata Theory

LTCS–Report

On the Complexity of Axiom Pinpointing in Description Logics

Rafael Peñaloza
Theoretical Computer Science
TU Dresden, Germany
penaloza@tcs.inf.tu-dresden.de

Barış Sertkaya
SAP Research Center
Dresden, Germany
baris.sertkaya@sap.com

LTCS-Report 09-04

Postal Address:
Lehrstuhl für Automatentheorie
Institut für Theoretische Informatik
TU Dresden
01062 Dresden

http://lat.inf.tu-dresden.de

Visiting Address:
Nöthnitzer Str. 46
Dresden

On the Complexity of Axiom Pinpointing in Description Logics

Rafael Peñaloza
Theoretical Computer Science
TU Dresden, Germany
penaloza@tcs.inf.tu-dresden.de

Barış Sertkaya*
SAP Research Center
Dresden, Germany
baris.sertkaya@sap.com

January 8, 2010

Abstract

We investigate the computational complexity of axiom pinpointing in Description Logics, which is the task of finding minimal subsets of a knowledge base that have a given consequence. We consider the problems of enumerating such subsets with and without order, and show hardness results that already hold for the propositional Horn fragment, or for the Description Logic $\mathcal{EL}$. We show complexity results for several other related decision and enumeration problems for these fragments that extend to more expressive logics. In particular we show that hardness of these problems depends not only on expressivity of the fragment but also on the shape of the axioms used.

1 Introduction

Description Logics (DLs) [BCM$^+$03] are a well-established family of logic-based knowledge representation formalisms that are used to represent the conceptual knowledge of an application domain in a structured and formally well-understood way. DLs have proven successful in various application domains, but they have gained increased attention due to the fact that they provide the logical underpinning of OWL [HPSvH03], the standard ontology language for the semantic web. As a consequence of this standardization, several ontology editors [KFNM04, KPS$^+$06, HTR06], now support OWL and ontologies written in OWL are employed in more and more applications. As the size of these ontologies grows, tools that support knowledge engineers in maintaining their quality become more important. In real world applications often the knowledge engineer not only wants to know whether her ontology has a certain (unwanted) consequence or not, but also wants to know *why* it has this consequence. Even for KBs of moderate size, finding explanations for a given a consequence is not an easy task without getting support from an automated tool. The task of finding explanations for a given consequence, i.e., minimal subsets of the original KB that have the given consequence is called *axiom pinpointing* in the literature.

Existing work on axiom pinpointing in DLs can be classified under two main categories, namely the glass-box approach, and the black-box approach. The idea lying under the *glass-box approach* is to extend the existing reasoning algorithms such that while reasoning, at the same time they can keep track of the axioms used, and detect which of the axioms in the KB are responsible for a given consequence. In [SC03] a pinpointing extension of the well-known tableau-based satisfiability algorithm for the DL $\mathcal{ALC}$ [SSS91] has been introduced. Later in [PSK05], this approach has been further extended to DLs

*Part of this work has been done when the author was still employed at Institute of Theoretical Computer Science, TU Dresden in the DFG Project BA 1122/12-1.

that are more expressive than $\mathcal{ALC}$. In [MLBP06] a pinpointing algorithm for $\mathcal{ALC}$ with general concept inclusions (GCIs) has been presented by following the approach in [BH95]. In order to overcome the problem of developing a pinpointing extension for every particular tableau-based algorithm, a general pinpointing extension for tableau algorithms has been developed in [BP07, BP10]. Similarly, an automata-based general approach for obtaining glass-box pinpointing algorithms has been introduced in [BP08, BP09].

In contrast to the glass-box approach, the idea lying under the *black-box approach* is to make use of the existing highly optimized reasoning algorithms wihout having to modify them. The most naïve black-box approach would of course be to generate every subset of the originial KB, and ask a DL reasoner whether this subset has the given consequence or not, which obviously is very inefficient. In [KPHS07] more efficient approaches based on Reiter's hitting set tree algorithm [Rei87] have been presented. The experimental resuts in [KPHS07] demonstrate that this approach behaves quite well in practice on realistic KBs written in expressive DLs. A similar approach has successfully been used in [HPS09] for explaining inconsistencies in OWL ontologies. The main advantages of the black-box approach are that one can use existing DL reasoners, and that it is independent of the DL reasoner being used. In [HPS08] the black-box approach has been used for computing more fine grained explanations, i.e., not just the set of relevant axioms in the KB but parts of these axioms that actually lead to the given consequence.

Although various methods and aspects of axiom pinpointing have been considered in the literature, its computational complexity has not been investigated in detail yet. Obviously, axiom pinpointing is at least as hard as reasoning. Nevertheless, especially for tractable DLs it makes sense to investigate whether explanations for a consequence can efficiently be enumerated or not. In [BPS07] it has been shown that a given consequence can have exponentially-many explanations (there called *MinAs*, which stands for minimal axiom sets), and checking the existence of a MinA within a cardinality bound is NP-complete. There it has also been shown that in a setting where MinAs are required to contain certain (static) part of the KB, then the set of all MinAs cannot be computed in output polynomial time. In [PS09] among other results we have shown that without the static part this problem is at least as hard as computing minimal transversals of a hypergraph. We have also shown that if the MinAs are required to be output in a specified order, then the problem is not solvable with polynomial delay.

In the present paper we present several new interesting complexity results on axiom pinpointing. We give a polynomial delay algoritm for enumerating MinAs in the Horn setting, show that for dual-Horn KBs the problem is at least as hard as hypergraph transversal enumeration, and for $\mathcal{EL}$ KBs it is not output polynomial. We show that if MinAs are required to be output in a specified order, then for dual-Horn and $\mathcal{EL}$ KBs this cannot be done with polynomial delay. We also consider several other decision and enumeration problems on MinAs in different settings.

2 Preliminaries

We briefly recall basic notions from propositional logic, DLs, and complexity of enumeration. In propositional logic we build formulae using a set of *propositional variables* and the *Boolean connectives* $\neg$ (negation), $\vee$ (disjunction) and $\wedge$ (conjunction). A variable or its negation is called a *literal*, and a disjunction of literals is called a *clause*. A clause is called a *Horn* (*dual-Horn*) clause if it contains at most one positive (negative) literal, and a *definite Horn (dual-Horn)* clause if it contains exactly one positive (negative) literal. A Horn clause $p_1 \vee \neg p_2 \vee \neg p_3$ can also be written as an *implication* of the form $p_2 \wedge p_3 \rightarrow p_1$. Throughout the text we will call definite Horn (dual-Horn) clauses just Horn (dual-Horn) clauses for short. We will call clauses with exactly one positive and one negative literal like $p_1 \rightarrow p_2$ as *core* clauses.

In DLs one formalizes the relevant notions of an application domain with *concept descriptions*. Concept descriptions are inductively built with the help of a set of *constructors*, starting with a set $\mathsf{N_C}$ of *concept names* and a set $\mathsf{N_R}$ of *role names*. $\mathcal{EL}$ concept descriptions are formed using the three constructors $\sqcap, \exists$

Syntax	Semantics
$\top$	$\Delta^{\mathcal{I}}$
$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
$\exists r.C$	$\{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}} : (x, y) \in r^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\}$
$C \sqsubseteq D$	$C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$

Table 1: Syntax and semantics of $\mathcal{EL}$.

and $\top$ as shown in the upper part of Table 1. An *$\mathcal{EL}$ TBox* is a finite set of *general concept inclusion* axioms (*GCIs*), whose syntax is shown in the lower part of Table 1. The semantics of $\mathcal{EL}$ is defined in terms of *interpretations* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where the *domain* $\Delta^{\mathcal{I}}$ is a non-empty set of individuals, and the interpretation function $\cdot^{\mathcal{I}}$ maps each concept name $A \in \mathsf{N_C}$ to a subset $A^{\mathcal{I}}$ of $\Delta^{\mathcal{I}}$ and each role name $r \in \mathsf{N_R}$ to a binary relation $r^{\mathcal{I}}$ on $\Delta^{\mathcal{I}}$. The mapping $\cdot^{\mathcal{I}}$ can be extended to arbitrary concept descriptions as shown in the second colum of Table 1. An interpretation $\mathcal{I}$ is a *model* of a TBox $\mathcal{T}$ if, for every GCI in $\mathcal{T}$ the conditions on the semantics column of Table 1 are satisfied. The main inference problem for $\mathcal{EL}$ is the subsumption problem: Given two $\mathcal{EL}$ concept descriptions C, D and an $\mathcal{EL}$ TBox $\mathcal{T}$, check if C is *subsumed* by D w.r.t. $\mathcal{T}$ (written $\mathcal{T} \models C \sqsubseteq D$), i.e, check if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ holds in every model $\mathcal{I}$ of $\mathcal{T}$. We will call a concept description *simple* if it is of the form A or $\exists r.A$ for $A \in \mathsf{N_C}, r \in \mathsf{N_R}$, and a GCI a *Horn-$\mathcal{EL}$ GCI* if it is of the form $C_1 \sqcap \ldots \sqcap C_n \sqsubseteq D$, where C_i, D are simple concept descriptions, $1 \leq i \leq n$.

We will refer to both propositional clauses and $\mathcal{EL}$ GCIs as *axioms*, and a set of axioms as a *knowledge base (KB)*. We will say that a KB is a Horn (core, dual-Horn, $\mathcal{EL}$) KB if it contains only Horn (core,dual-Horn,$\mathcal{EL}$) axioms. We are going to formulate our problems in a generic way without referring to a specific type of KB, and show our results for each KB type separately.

In complexity theory, we say that an algorithm runs with *polynomial delay* [JYP88] if the time until the first solution is generated, and thereafter the time between any two consecutive solutions is bounded by a polynomial in the size of the input. We say that it runs in *output polynomial time* if it outputs all solutions in time polynomial in the size of the input *and the output*.

3 Complexity of Enumerating All MinAs

The main problem we consider is, given a KB and a consequence of it, computing all MinAs for this consequence in the given KB. We start with the definition of a MinA.

Definition 1. Let $\mathcal{K}$ be a set of axioms and φ be a logical consequence of it, i.e., $\mathcal{K} \models \varphi$. We call a set $\mathcal{M} \subseteq \mathcal{K}$ a *minimal axiom set* or *MinA* for φ in $\mathcal{K}$ if $\mathcal{M} \models \varphi$ and it is minimal w.r.t. set inclusion.

Our problem is formally defined as follows:

Problem: MINA-ENUM
Input: A KB $\mathcal{K}$ and an axiom φ of the same type such that $\mathcal{K} \models \varphi$.
Output: The set of all MinAs for φ in $\mathcal{K}$.

Note that for core KBs, which are basically directed graphs, a MinA is a simple path between two given vertices, and enumerating all MinAs corresponds to enumerating all simple paths between two given vertices, which can easily be done with polynomial delay. However, the situation is not so clear for Horn KBs. To the best of our knowledge, only [NPA06] considers a problem related to ours on directed hypergraphs, but it is not exactly the one considered here.

3.1 Enumeration without a Specific Order

We start with the Horn setting and show that for this setting MinAs can efficiently be enumerated by giving a polynomial delay algorithm. The algorithm depends on the following particular notion.

Definition 2. Let $\mathcal{K}$ be a Horn KB, and $\phi = \bigwedge_{i=1}^{n} a_i \to b$ be an axiom in $\mathcal{K}$. We denote the left handside (lhs) of ϕ with $\mathsf{T}(\phi)$, and its right handside (rhs) with $\mathsf{h}(\phi)$, i.e., $\mathsf{T}(\phi) := \{a_1, \ldots, a_n\}$ and $\mathsf{h}(\phi) := b$. With $\mathsf{h}^{-1}(b)$ we denote the set of axioms in $\mathcal{K}$ whose rhs are b. Let $\mathcal{M} = \{t_1, \ldots, t_m\}$ be a MinA for $\bigwedge_{a \in \mathcal{A}} a \to c$. We call an ordering $t_1 < \ldots < t_m$ a *valid ordering* on $\mathcal{M}$ if for every $1 \leq i \leq m$, $\mathsf{T}(t_i) \subseteq \mathcal{A} \cup \{\mathsf{h}(t_1), \ldots, \mathsf{h}(t_{i-1})\}$ holds.[1]

It is easy to see that for every MinA $\mathcal{M}$ there is always at least one such valid ordering. In the following, we use this fact to construct from a given MinA a set of KBs that precisely contain the remaining MinAs.

Definition 3. Let $\mathcal{M}$ be a MinA in $\mathcal{K}$ with $|\mathcal{M}| = m$, and $<$ be a valid ordering on $\mathcal{M}$. For each $1 \leq i \leq m$ we obtain a KB $\mathcal{K}_i$ from $\mathcal{K}$ as follows: (i) for each j s.t. $i < j \leq m$ remove all axioms in $\mathsf{h}^{-1}(\mathsf{h}(t_j))$ except for t_j, i.e., remove all axioms with the same rhs as t_j except for t_j itself. (ii) remove t_i.

Lemma 4. *Let $\mathcal{M}$ be a MinA for ϕ in $\mathcal{K}$, and let $\mathcal{K}_1, \ldots, \mathcal{K}_m$ be constructed from $\mathcal{K}$ and $\mathcal{M}$ as in Definition 3. Then, for every MinA $\mathcal{N}$ for ϕ in $\mathcal{K}$ that is different from $\mathcal{M}$, there exists* **exactly one** i, *where $1 \leq i \leq m$, such that $\mathcal{N}$ is a MinA for ϕ in $\mathcal{K}_i$.*

Proof. Let $t_1 < \ldots < t_m$ be a valid ordering on $\mathcal{M}$, and $\mathcal{N}$ a MinA for ϕ in $\mathcal{K}$ such that $\mathcal{N} \neq \mathcal{M}$. Then, $\mathcal{M} \setminus \mathcal{N} \neq \emptyset$. Let t_k be the largest axiom in $\mathcal{M} \setminus \mathcal{N}$ w.r.t. the ordering $<$. We show that $\mathcal{N} \subseteq \mathcal{K}_k$ and $\mathcal{N} \not\subseteq \mathcal{K}_i$ for all $i \neq k$, $1 \leq i \leq m$.

Assume there is an axiom $t \in \mathcal{N}$ s.t. $t \notin \mathcal{K}_k$. t should be one of the axioms removed from $\mathcal{K}$ either in step (i), or in step (ii) of Definition 3. It cannot be step (ii) because $t_k \notin \mathcal{N}$ since $t_k \in \mathcal{M} \setminus \mathcal{N}$. Thus it should be step (i). This implies that there exists a j, $k < j \leq m$, such that t_j satisfies $\mathsf{h}(t) = \mathsf{h}(t_j)$. Recall that we chose j to be the largest axiom in $\mathcal{M} \setminus \mathcal{N}$ w.r.t. the valid ordering $<$ on $\mathcal{M}$. Then this t_j should be in $\mathcal{N}$. But then $\mathcal{N}$ contains two axioms with the rhs $\mathsf{h}(t)$, which contradicts with the fact that $\mathcal{N}$ is a MinA, and thus it is minimal. Hence, $\mathcal{N} \subseteq \mathcal{K}_k$.

Now take an i s.t. $i \neq k$. If $i > k$, then $t_i \in \mathcal{N}$ but $t_i \notin \mathcal{K}_i$, and hence $\mathcal{N} \not\subseteq \mathcal{K}_i$. If $i < k$, then there is an axiom $t \in \mathcal{N}$ such that $\mathsf{h}(t) = \mathsf{h}(t_k)$ since otherwise $\mathcal{M}$ and $\mathcal{N}$ would not be MinAs. By construction, $t \notin \mathcal{K}_i$, hence $\mathcal{N} \not\subseteq \mathcal{K}_i$. ☐

Lemma 4 gives an idea of how to compute the remaining MinAs from a given one. Algorithm 1 describes how we can use this lemma for enumerating all MinAs.

Theorem 5. *Algorithm 1 solves* MINA-ENUM *for Horn KBs with polynomial delay.*

Proof. The algorithm terminates since $\mathcal{K}$ is finite. It is sound since its outputs are MinAs for ϕ in $\mathcal{K}$. Completeness follows from Lemma 4.

In each recursive call of the algorithm there is one consequence check (line 2), and one MinA computation (line 4). The consequence check can be done in polynomial time by the well-known linear-time algorithm in [DG84]. One MinA can be computed in polynomial time by iterating over the axioms in $\mathcal{K}$ and removing an axiom if remaining ones still have the consequence. Thus the algorithm spends at most polynomial time between each output, i.e., it is polynomial delay. ☐

[1] That is, each variable on the lhs of t_i is in $\mathcal{A}$, or it is the rhs of a previous axiom.

Algorithm 1 Enumerating all MinAs for Horn KBs

ALL-MINAS($\mathcal{K}$,ϕ)
  ▷ ($\mathcal{K}$ a Horn KB, ϕ an axiom s.t. $\mathcal{K} \models \phi$)
if $\mathcal{K} \not\models \phi$ **then return**
else
  $\mathcal{M}$:= a MinA in $\mathcal{K}$
  output $\mathcal{M}$
  for $1 \leq i \leq |\mathcal{M}|$ **do**
    compute $\mathcal{K}_i$ from $\mathcal{M}$ as in Definition 3
    ALL-MINAS($\mathcal{K}_i$,ϕ)
  end for
end if

Next we consider MINA-ENUM for dual-Horn KBs. For this, we first investigate the following decision problem which is is closely related to MINA-ENUM. As we will see, determining its complexity is important for determining the complexity of MINA-ENUM.

Problem: ALL-MINAS
Input: A KB $\mathcal{K}$ and an axiom φ of the same type such that $\mathcal{K} \models \varphi$, and a set of KBs $\mathscr{K} \subseteq \mathscr{P}(\mathcal{K})$.
Question: Is $\mathscr{K}$ precisely the set of all MinAs for φ in $\mathcal{K}$?

As Proposition 6 shows, if ALL-MINAS cannot be decided in polynomial time, then MINA-ENUM cannot be solved in output polynomial time.

Proposition 6. *If* ALL-MINAS *cannot be decided in polynomial time, then* MINA-ENUM *cannot be solved in output-polynomial time.*

Proof. Assume we have an algorithm A that solves MINA-ENUM in output-polynomial time. Let its runtime be bounded by a polynomial $p(IS, OS)$ where IS denotes the size of the input KB and OS denotes the size of the output, i.e., the set of all MinAs.

In order to decide ALL-MINAS for an instance given by $\mathcal{K}$, φ, and $\mathscr{K} \subseteq \mathscr{P}(\mathcal{K})$, we construct another algorithm A' that works as follows: it runs A on $\mathcal{K}$ and φ for at most $p(|\mathcal{K}|, |\mathscr{K}|)$-many steps. If A terminates within this many steps, then A' compares the output of A with $\mathscr{K}$ and returns *yes* if and only if they are equal. If they are not equal, A' returns *no*. If A has not yet terminated after $p(|\mathcal{K}|, |\mathscr{K}|)$-many steps, this implies that there is at least one MinA that is not contained in $\mathscr{K}$, so A' returns *no*. It is easy to see that the runtime of A' is bounded by a polynomial in $|\mathcal{K}|$ and $|\mathscr{K}|$, that is A' decides ALL-MINAS in polynomial time. □

This proposition shows that the complexity of ALL-MINAS is indeed closely related to the complexity of MINA-ENUM. It is not difficult to see that, for all types of axioms considered in this paper, ALL-MINAS is in coNP: given an instance of ALL-MINAS, a nondeterministic algorithm can guess a subset of $\mathcal{K}$ that is not in $\mathscr{K}$, and in polynomial time verify that this is a MinA, thus $\mathscr{K}$ is *not* the set of all MinAs. In the following we show that for dual-Horn KBs ALL-MINAS is at least as hard as recognizing the set of all minimal transversals of a given hypergraph. However, whether it is coNP-hard remains unfortunately open. We later show that ALL-MINAS is coNP-complete if Horn-$\mathcal{EL}$ axioms are considered.

First we briefly recall some basic notions on hypergraphs. A *hypergraph* [Ber89] $\mathcal{H} = (V, \mathcal{E})$ consists of a set of *vertices* $V = \{v_i \mid 1 \leq i \leq n\}$, and a set of *(hyper)edges* $\mathcal{E} = \{E_j \mid 1 \leq j \leq m\}$ where $E_j \subseteq V$. Following the convention in [Ber89] we assume that the set of edges as well as the set of vertices is nonempty, and the union of all edges yields the vertex set. A set $W \subseteq V$ is called a *transversal* of $\mathcal{H}$ if it intersects every edge of $\mathcal{H}$, i.e., $\forall E \in \mathcal{E}.\ E \cap W \neq \emptyset$. A transversal is called *minimal* if no proper subset of it is a transversal. The set of all minimal transversals of $\mathcal{H}$ constitutes another hypergraph on V

called the *transversal hypergraph* of $\mathcal{H}$, which is denoted by $Tr(\mathcal{H})$. Generating $Tr(\mathcal{H})$ is an important problem which has applications in many fields of computer science [GKMT97, EG02, Hag08]. It is defined as follows:

Problem: TRANSVERSAL ENUMERATION (TRANS-ENUM)
Input: A hypergraph $\mathcal{H} = (V, \mathcal{E})$ on a finite set V.
Output: The edges of the transversal hypergraph $Tr(\mathcal{H})$.

The well-known decision problem associated to this computation problem is defined as follows:

Problem: TRANSVERSAL HYPERGRAPH (TRANS-HYP)
Input: Two hypergraphs $\mathcal{H} = (V, \mathcal{E}_\mathcal{H})$ and $\mathcal{G} = (V, \mathcal{E}_\mathcal{G})$.
Question: Is $\mathcal{G}$ the transversal hypergraph of $\mathcal{H}$, i.e., does $Tr(\mathcal{H}) = \mathcal{G}$ hold?

Complexity of TRANS-HYP has been investigated in detail in the literature [EG91, EG95b, EGM03, EMG08, KS03]. It is known to be in coNP, but its lower bound is a prominent open problem. So far neither a polynomial time algorithm has been found, nor has it been proved to be coNP-hard. In a landmark paper [FK96] Fredman and Khachiyan proved that MONOTONE BOOLEAN DUALIZATION, which is another well known problem that is computatitionally equivalent to TRANS-HYP, can be solved in $n^{o(\log n)}$ time. This implies that TRANS-HYP is most likely *not* coNP-hard. It is conjectured that this problem, together with several computationally equivalent problems, forms a class properly contained between P and coNP [FK96].

In the following we say that a decision problem π is TRANS-HYP-hard if TRANS-HYP can be reduced to π by a standard polynomial transformation.We say that π is TRANS-HYP-complete if it is TRANS-HYP-hard and π can be reduced to TRANS-HYP by a polynomial transformation.

Theorem 7. ALL-MINAS *is* TRANS-HYP*-hard for dual-Horn KBs.*

Proof. Let an instance of TRANS-HYP be given by the hypergraphs $\mathcal{H} = (V, \mathcal{E}_\mathcal{H})$ and $\mathcal{G} = (V, \mathcal{E}_\mathcal{G})$. From $\mathcal{H}$ and $\mathcal{G}$ we construct an instance of ALL-MINAS as follows: for every vertex $v \in V$ we introduce a propositional variable p_v, for every edge $E \in \mathcal{E}_\mathcal{H}$ a propositional variable p_E, and finally one additional propositional variable a. For constructing a dual-Horn KB from $\mathcal{H}$ and a set of vertices $W \subseteq V$, we define the following operator, which is also going to be used in later proofs:

$$\mathcal{K}_{W,\mathcal{H}} := \{p_v \to \bigwedge_{v \in E, E \in \mathcal{E}_\mathcal{H}} p_E \mid v \in W\} \cup \{a \to \bigwedge_{v \in V} p_v\}.$$

Using these we construct the KB $\mathcal{K} := \mathcal{K}_{V,\mathcal{H}}$, a set of KBs $\mathscr{K} := \{\mathcal{K}_{E,\mathcal{H}} \mid E \in \mathcal{E}_\mathcal{G}\} \subseteq \mathscr{P}(\mathcal{K})$, and the axiom $\varphi := a \to \bigwedge_{E \in \mathcal{E}_\mathcal{H}} p_E$ that follows from $\mathcal{K}$. Obviously this construction creates an instance of ALL-MINAS for dual-Horn KBs and it can be done in time polynomial in the sizes of $\mathcal{H}$ and $\mathcal{G}$.

We claim that $\mathcal{G}$ is the transversal hypergraph of $\mathcal{H}$ if and only if $\mathscr{K}$ is precisely the set of all MinAs for φ in $\mathcal{K}$. Note that $a \to \bigwedge_{v \in V} p_v$ is the only axiom in $\mathcal{K}$ such that a appears on the lhs, which implies that every MinA must contain this axiom. Hence, every MinA is of the form $\mathcal{K}_{W,\mathcal{H}}$ for some $W \subseteq V$. To prove our claim, it suffices to show that a set of vertices $W \subseteq V$ is a minimal transversal of $\mathcal{H}$ if and only if the set of axioms $\mathcal{K}_{W,\mathcal{H}}$ is a MinA.

($\Rightarrow$) Assume that W is a minimal transversal of $\mathcal{H}$. By definition W satisfies $W \cap E \neq \emptyset$ for every $E \in \mathcal{E}_H$. This implies that $\mathcal{K}_{W,\mathcal{H}} \models \varphi$ holds. Moreover, $\mathcal{K}_{W,\mathcal{H}}$ is minimal since W is minimal, i.e., $\mathcal{K}_{W,\mathcal{H}}$ is a MinA.

($\Leftarrow$) Now assume that $\mathcal{K}_{W,\mathcal{H}}$ is a MinA. Then every p_E where $E \in \mathcal{E}_\mathcal{H}$ appears on the rhs of at least one of the axioms in $\mathcal{K}_{W,\mathcal{H}}$. This implies that W intersects every E, i.e., it is a transversal of $\mathcal{H}$. Moreover it is minimal since $\mathcal{K}_{W,\mathcal{H}}$ is minimal. $\square$

Corollary 8. MINA-ENUM *for dual-Horn KBs is at least as hard as enumerating hypergraph transversals.*

Next we show that ALL-MINAS becomes intractable for Horn-$\mathcal{EL}$ KBs.

Theorem 9. ALL-MINAS *is co*NP*-complete for Horn-$\mathcal{EL}$ TBoxes.*

Proof. We have already shown that it is in coNP. To show coNP-hardness, we present a reduction from the following coNP-hard problem [EG91, BPS07].

Problem: ALL-MV
Input: A monotone Boolean formula ϕ and a set $\mathscr{V}$ of minimal valuations satisfying ϕ.
Question: Is $\mathscr{V}$ precisely the set of all minimal valuations satisfying ϕ?

Let $\phi, \mathscr{V}$ be an instance of ALL-MV; we denote as $\mathsf{sub}(\phi)$ the set of all subformulas of ϕ, and define $\mathsf{csub}(\phi) := \mathsf{sub}(\phi) \setminus \{p \in \mathsf{sub}(\phi) \mid p$ is a propositional variable$\}$. We introduce three concept names B_ψ, C_ψ, D_ψ, and two role names r_ψ, s_ψ for every subformula ψ of ϕ and two additional concept names A and E. For each $\psi \in \mathsf{sub}(\phi)$ we define a TBox $\mathcal{T}_\psi$ as follows: if ψ is the propositional variable p, then $\mathcal{T}_\psi := \{A \sqsubseteq B_p\}$; if $\psi = \psi_1 \wedge \psi_2$, then $\mathcal{T}_\psi := \{A \sqsubseteq \exists r_\psi.C_\psi, C_\psi \sqsubseteq B_{\psi_1}, C_\psi \sqsubseteq B_{\psi_2}, \exists r_\psi.B_\psi \sqsubseteq D_\psi, B_{\psi_1} \sqcap B_{\psi_2} \sqsubseteq B_\psi\}$; if $\psi = \psi_1 \vee \psi_2$, then $\mathcal{T}_\psi := \{A \sqsubseteq \exists r_\psi.B_{\psi_1}, A \sqsubseteq \exists s_\psi.B_{\psi_2}, \exists r_\psi.B_\psi \sqcap \exists s_\psi.B_\psi \sqsubseteq D_\psi, B_{\psi_1} \sqsubseteq B_\psi, B_{\psi_2} \sqsubseteq B_\psi\}$. Finally, we set

$$\mathcal{T} := \bigcup_{\psi \in \mathsf{sub}(\phi)} \mathcal{T}_\psi \cup \{ \bigsqcap_{\psi \in \mathsf{csub}(\phi)} D_\psi \sqcap B_\phi \sqsubseteq E\}.$$

Notice that for every $\mathcal{T}' \subseteq \mathcal{T}$, if $\mathcal{T}' \models A \sqsubseteq E$, then also $A \sqsubseteq D_\psi$ for every $\psi \in \mathsf{csub}(\phi)$. But in order to have $A \sqsubseteq D_\psi$, all the axioms in $\mathcal{T}_\psi$ are necessary, and thus $\mathcal{T}_\psi \subseteq \mathcal{T}'$. In particular, if $\psi = \psi_1 \wedge \psi_2$, then $B_{\psi_1} \sqcap B_{\psi_2} \sqsubseteq B_\psi \in \mathcal{T}'$, and if $\psi = \psi_1 \vee \psi_2$, then $\{B_{\psi_1} \sqsubseteq B_\psi, B_{\psi_2} \sqsubseteq B_\psi\} \subseteq \mathcal{T}'$. Thus, a valuation $\mathcal{V}$ satisfies ϕ iff $\mathcal{T}_\mathcal{V} := \{A \sqsubseteq B_p \mid p \in \mathcal{V}\} \cup \bigcup_{\psi \in \mathsf{csub}(\phi)} \mathcal{T}_\psi \cup \{\bigsqcap_{\psi \in \mathsf{csub}(\phi)} D_\psi \sqcap B_\phi \sqsubseteq E\}$ entails $A \sqsubseteq E$. This in particular shows that $\mathscr{V}$ is the set of all minimal valuations satisfying ϕ iff $\{\mathcal{T}_\mathcal{V} \mid \mathcal{V} \in \mathscr{V}\}$ is the set of all MinAs for $A \sqsubseteq E$ in $\mathcal{T}$. ☐

The following is an immediate consequence of Theorem 9 and Proposition 6.

Corollary 10. *For Horn-$\mathcal{EL}$ TBoxes* MINA-ENUM *cannot be solved in output polynomial time, unless* P =NP.

3.2 Enumeration in a Specified Order

We now consider the case when MinAs are required to be output in a specified lexicographic order. The lexicographic order we use is defined as follows:

Definition 11. Let the elements of a set S be linearly ordered. This order induces a linear strict order on $\mathscr{P}(S)$, which is called the *lexicographic order*. We say that a set $R \subseteq S$ is *lexicographically smaller* than a set $T \subseteq S$ where $R \neq T$ if the first element at which they disagree is in R.

Problem: FIRST-MINA
Input: A KB $\mathcal{K}$ and an axiom φ of the same type such that $\mathcal{K} \models \varphi$, a MinA $\mathcal{M}$ for φ in $\mathcal{K}$, and a linear order on $\mathcal{K}$.
Question: Is $\mathcal{M}$ the first MinA w.r.t. the lexicographic order induced by the given linear order?

Theorem 12. FIRST-MINA *is co*NP*-complete for dual-Horn KBs.*

Proof. The problem is in coNP. If $\mathcal{M}$ is not the lexicographically first MinA, a proof of this can be given by guessing a subset of $\mathcal{K}$ and verifying in polynomial time that it is a MinA, and it is lexicographically smaller than $\mathcal{M}$.

In order to show coNP-hardness, we present a reduction from the problem of checking whether a given maximal independent set is the lexicographically last maximal independent set of a given graph. Recall that a *maximal independent set* of a graph $\mathcal{G} = (V, \mathcal{E})$ is a subset $V' \subseteq V$ of the vertices such that no two vertices in V' are joined by an edge in $\mathcal{E}$, and each vertex in $V \setminus V'$ is joined by an edge to some vertex in V'. This problem is known to be coNP-complete [JYP88].

Problem: LAST MAX. INDEPENDENT SET (LAST-MIS)
Input: A graph $\mathcal{G} = (V, \mathcal{E})$, a maximal independent set $S \subseteq V$, and a linear order on V.
Question: Is S the last maximal independent set w.r.t. the lexicographic order induced by the given linear order?

Let an instance of LAST-MIS be given with the graph $\mathcal{G} = (V, \mathcal{E})$ and the maximal independent set S. From $\mathcal{G}$ and S we construct an instance of FIRST-MINA as follows: We construct the sets $\mathcal{K}_{W,\mathcal{G}}$ as in the proof of Theorem 7, and consider the axiom $\varphi := a \rightarrow \bigwedge_{E \in \mathcal{E}} p_E$ that follows from $\mathcal{K}_{V,\mathcal{G}}$. Additionally by using S we construct the set of axioms $\mathcal{M} := \mathcal{K}_{V \setminus S, \mathcal{G}}$. Note that $\mathcal{K}_{V,\mathcal{G}}$ contains exactly $|V|+1$ axioms. We order these axioms such that an axiom with premise p_v comes before the axiom with premise $p_{v'}$ if and only if the vertex v comes before the vertex v' in the originally given linear order on V. Finally we place φ as the last one. It is easy to see that this construction indeed creates an instance of FIRST-MINA for dual-Horn KBs, and it can be done in time polynomial in the sizes of $\mathcal{G}$ and S. We claim that S is lexicographically the last maximal independent set if and only if $\mathcal{M}$ is lexicographically the first MinA.

($\Rightarrow$) Assume S is the lexicographically last maximal independent set. Then $V \setminus S$ contains at least one vertex from every edge (i.e., it is a vertex cover), since otherwise S would not be an independent set. Thus every p_E, for $E \in \mathcal{E}$, appears on the rhs of at least one axiom in $\mathcal{M}$. That is $\mathcal{M} \models \varphi$ holds. Since S is maximal, $V \setminus S$ and thus $\mathcal{M}$ is minimal, i.e., $\mathcal{M}$ is a MinA. Moreover it is lexicographically the first one since S is lexicographically the last maximal independent set.

($\Leftarrow$) Assume $\mathcal{M}$ is lexicographically the first MinA. Then every p_E, for $E \in \mathcal{E}$, appears on the rhs of at least one axiom in $\mathcal{M}$ since otherwise $\mathcal{M} \models \varphi$ would not hold. That is, $V \setminus S$ contains at least one vertex from every edge. Then S contains at most one vertex from every edge, i.e., it is an independent set. Since $\mathcal{M}$ is minimal, $V \setminus S$ is also minimal, and thus S is maximal. That is, S is a maximal independent set. Moreover it is lexicographically the last one since $\mathcal{M}$ is the lexicographically first MinA. $\square$

Since generating the lexicographically first MinA is already intractable, Theorem 12 has the following consequence:

Corollary 13. *Unless* P = NP, *MinAs cannot be enumerated for dual-Horn KBs in lexicographic order with polynomial delay.*

Next we consider the problem for Horn-$\mathcal{EL}$ KBs.

Theorem 14. FIRST-MINA *is coNP-complete for Horn-$\mathcal{EL}$ KBs.*

Proof. The problem is clearly in coNP. To show hardness, we give a reduction from LAST-MIS. Let $\mathcal{G} = (V, \mathcal{E})$ and S be an instance of LAST-MIS. From $\mathcal{G}$ we construct a Horn-$\mathcal{EL}$ TBox $\mathcal{T}$ as follows: first we introduce a concept P_E for every $E \in \mathcal{E}$, and concepts P_v, Q_v and role name r_v for each $v \in V$, and additionally two concept names A, B. For every $v \in V$ we construct the TBox $\mathcal{T}_v := \{P_v \sqsubseteq P_E \mid v \in E, E \in \mathcal{E}_\mathcal{G}\} \cup \{A \sqsubseteq \exists r_v.P_v, \bigsqcap_{v \in E, E \in \mathcal{E}_\mathcal{G}} \exists r_v.P_E \sqsubseteq Q_v\}$. We then define the

Algorithm 2 Enumerating all MinAs in reverse lex. order

1: ALL-MINAS-REV-ORDER($\mathcal{K}$,ϕ)
2: ▷ ($\mathcal{K}$ a Horn KB, ϕ an axiom s.t. $\mathcal{K} \models \phi$)
3: $\mathcal{Q} := \{\mathcal{K}\}$
4: **while** $\mathcal{Q} \neq \emptyset$ **do**
5: $\quad\mathcal{J}$:= maximum element of $\mathcal{Q}$
6: $\quad$remove $\mathcal{J}$ from $\mathcal{Q}$
7: $\quad\mathcal{M}$:= the lex. largest MinA in $\mathcal{J}$
8: $\quad$output $\mathcal{M}$
9: $\quad$**for** $1 \leq i \leq |\mathcal{M}|$ **do**
10: $\quad\quad$compute $\mathcal{K}_i$ from $\mathcal{M}$ as in Definition 3
11: $\quad\quad$insert $\mathcal{K}_i$ into $\mathcal{Q}$ if $\mathcal{K}_i \models \phi$
12: $\quad$**end for**
13: **end while**

set $\mathcal{T}_f := \bigcup_{v \in V} \mathcal{T}_v \cup \{\bigsqcap_{E \in \mathcal{E}_\mathcal{H}} P_E \sqcap \bigsqcap_{v \in V} Q_v \sqsubseteq B\}$, and finally, for a set of $W \subseteq V$, we define $\mathcal{T}_W := \mathcal{T}_f \cup \{A \sqsubseteq P_v \mid v \in W\}$.

Notice that for every $\mathcal{T}' \subseteq \mathcal{T}$, if $\mathcal{T}' \models A \sqsubseteq Q_v$, then $\mathcal{T}_v \subseteq \mathcal{T}'$. Hence, if $\mathcal{T}' \models A \sqsubseteq B$, then $\mathcal{T}_f \subseteq \mathcal{T}'$. Furthermore, $S \subseteq V$ is an independent set iff $\mathcal{T}_{V \setminus S} \models A \sqsubseteq B$.

We now order the axioms in $\mathcal{T}_V$ as follows: first appear all the axioms $A \sqsubseteq P_v$ using the same order of V, and afterwards are all the axioms in $\mathcal{T}_f$ in any order. Then S is the last maximal independent set iff $\mathcal{T}_{V \setminus S}$ is the first MinA for $A \sqsubseteq B$ in $\mathcal{T}_V$. $\square$

Although computing the first MinA is coNP-hard for both dual-Horn and Horn-$\mathcal{EL}$ KBs, interestingly computing the last MinA is polynomial for all types of KBs we consider here. We start iterating over the axioms of the KB with the axiom that is the smallest one w.r.t. the linear order on KB, and remove an axiom if the remaining ones still have the given conseqence. The resulting set of axioms is lexicographically the last MinA. Even more interestingly, we now give an algorithm for Horn KBs that enumerates MinAs in reverse lexicographic order with polynomial delay.

Our algorithm keeps a set of KBs in a priority queue $\mathcal{Q}$. These KBs are the "candidates" from which the MinAs are going to be computed. Each KB can contain zero or more MinAs. They are inserted into $\mathcal{Q}$ by the algorithm at a cost of $O(n \cdot log(M))$ per insertion, where n is the size of the original KB and M is the total number of such KBs inserted. Note that M can be exponentially bigger than n since there can be exponentially many MinAs. That is the algorithm uses potentially exponential space. The other operation that the algorithm performs on $\mathcal{Q}$ is to find and delete the maximum element of $\mathcal{Q}$. The maximum element of $\mathcal{Q}$ is the KB in $\mathcal{Q}$ that contains the lexicographically largest MinA among the MinAs contained in all other KBs in $\mathcal{Q}$. This operation can also be performed within $O(n \cdot log(M))$ time bound. The time bounds for insertion and deletion depend also on n since they require a last MinA computation.

Theorem 15. *Algorithm 2 enumerates MinAs in the Horn setting in reverse lexicographic order with polynomial delay.*

Proof. The algorithm terminates since $\mathcal{K}$ is finite. Soundness is shown as follows: $\mathcal{Q}$ contains initially only the original KB $\mathcal{K}$. Thus the first output is lexicographically the last MinA in $\mathcal{K}$. By Lemma 4 the MinA that comes just before the last one is contained in exactly one of the $\mathcal{K}_i$s that are computed and inserted into $\mathcal{Q}$ in lines 10 and 11. In line 5 $\mathcal{J}$ is assigned the KB that contains this MinA. Thus the next output will be the MinA that comes just before the lexicographically last one. It is not difficult to see that in this way the MinAs will be enumerated in reverse lexicographic order. By Lemma 4 it is guaranteed that the algorithm enumerates all MinAs.

In one iteration, the algorithm performs one find operation and one delete operation on $\mathcal{Q}$, which both take time $O(n \cdot log(M))$, and a MinA computation that takes $O(n)$ time. In addition it performs at most n $\mathcal{K}_i$ computations, and at most n insertions into $\mathcal{Q}$. Each $\mathcal{K}_i$ computation takes $O(n^2)$ time, and each insertion takes $O(n{\cdot}log(M))$ time. The total delay is thus $O(2{\cdot}(n{\cdot}log(M))+n+n{\cdot}(n^2+n{\cdot}log(M))) = O(n^3)$. □

4 Preferred and Unwanted Axioms

Next we investigate the problem of existence of a MinA that does not contain any of the given sets of axioms. This problem can be useful in applications where one wants to avoid certain combinations of axioms in the MinAs.

Problem: MINA-IRRELEVANCE
Input: A KB $\mathcal{K}$ and an axiom φ of the same type such that $\mathcal{K} \models \varphi$, and a set $\mathscr{K} \subseteq \mathscr{P}(\mathcal{K})$.
Question: Is there a MinA $\mathcal{M}$ for φ in $\mathcal{K}$ such that $\mathcal{S} \not\subseteq \mathcal{M}$ for every $\mathcal{S} \in \mathscr{K}$?

Theorem 16. MINA-IRRELEVANCE *is* NP*-complete for dual-Horn KBs.*

Proof. The problem is clearly in NP. A nondeterministic algorithm for solving it first guesses a set $\mathcal{M} \subseteq \mathcal{K}$, then tests in polynomial time whether it is a MinA that does not contain any of the $\mathcal{S}$ in $\mathscr{K}$. For showing hardness we give a reduction from the NP-hard hypergraph 2-coloring problem [GJ90].

Problem: HYPERGRAPH 2-COLORING
Input: A hypergraph $\mathcal{H} = (V, \mathcal{E})$.
Question: Is $\mathcal{H}$ 2-colorable, i.e., is there a $W \subseteq V$ such that for all $E \in \mathcal{E}$, $W \cap E \neq \emptyset$ and $(V \setminus W) \cap E \neq \emptyset$?

Let an instance of HYPERGRAPH 2-COLORING be given with the hypergraph $\mathcal{H} = (V, \mathcal{E})$. We construct an instance of MINA-IRRELEVANCE as follows: as in the proof of Theorem 7, we construct the KB $\mathcal{K} := \mathcal{K}_{V,\mathcal{H}}$ and the axiom φ contructed there, as well as a set of KBs $\mathscr{K} = \{\mathcal{K}_{E,\mathcal{H}} \mid E \in \mathcal{E}\}$. It is easy to see that this construction indeed creates an instance of MINA-IRRELEVANCE for dual-Horn KBs and it can be done in time polynomial in the size of $\mathcal{H}$. We claim that $\mathcal{H}$ is 2-colorable if and only if there is a MinA $\mathcal{M}$ for φ in $\mathcal{K}$ such that $\mathcal{M}$ satisfies $\mathcal{S} \not\subseteq \mathcal{M}$ for every $\mathcal{S} \in \mathscr{K}$.

($\Rightarrow$) Assume $\mathcal{H}$ is 2-colorable. Then there is a $W \subseteq V$ such that $W \cap E \neq \emptyset$ and $(V \setminus W) \cap E \neq \emptyset$ for every $E \in \mathcal{E}$, i.e., both W and its complement are transversals of $\mathcal{H}$. Assume w.l.o.g. that W is minimal. We claim that $\mathcal{K}_{W,\mathcal{H}}$ is the MinA we are looking for. Since W is a transversal, every p_E for $E \in \mathcal{E}$, appears on the rhs of at least one axiom in $\mathcal{K}_{W,\mathcal{H}}$. That is $\mathcal{K}_{W,\mathcal{H}} \models \varphi$ holds. $\mathcal{K}_{W,\mathcal{H}}$ is minimal since W is minimal. Moreover, since $V \setminus W$ is also a transversal, every edge $E \in \mathcal{E}$ contains at least one vertex that is *not* in W. Thus every $\mathcal{S} \in \mathscr{K}$ contains at least one axiom that is *not* in $\mathcal{K}_{W,\mathcal{H}}$. In other words, $\mathcal{K}_{W,\mathcal{H}}$ is a MinA that is not a superset of any $\mathcal{S} \in \mathscr{K}$.

($\Leftarrow$) Assume $\mathcal{M}$ is a MinA that is not a superset of any $\mathcal{S} \in \mathscr{K}$. Define the set $W_{\mathcal{M}} = \{v \mid p_v \rightarrow \bigwedge_{v \in E, E \in \mathcal{E}} p_E \in \mathcal{M}\}$. Since $\mathcal{M}$ is a MinA for φ, for every $E \in \mathcal{E}$ it contains at least one axiom on whose rhs p_E occurs. That is, $W_{\mathcal{M}}$ intersects every $E \in \mathcal{E}$. Since $\mathcal{M}$ is not a superset of any $\mathcal{S} \in \mathscr{K}$, every $\mathcal{S}$ contains at least one axiom that is *not* in $\mathcal{M}$. This that every $E \in \mathcal{E}$ contains at least one vertex that is *not* in $W_{\mathcal{M}}$. That is, $V \setminus W_{\mathcal{M}}$ intersects every $E \in \mathcal{E}$. Thus we have shown that $W_{\mathcal{M}}$ is a 2-coloring of $\mathcal{H}$. □

Theorem 17. MINA-IRRELEVANCE *is* NP*-complete for Horn-$\mathcal{EL}$ TBoxes*

Proof. The problem is clearly in NP. We show NP-hardness by a reduction from the HYPERGRAPH 2-COLORING problem. Let $\mathcal{H} = (V, \mathcal{E})$ be a hypergraph; we construct the TBoxes $\mathcal{T}_v, \mathcal{T}_f$ and $\mathcal{T}_V$ as in the proof of Theorem 14. It is easy to see that $\mathcal{T} := \mathcal{T}_V, \phi := A \sqsubseteq B$ and the set of TBoxes $\mathscr{K} := \{\mathcal{T}_E \mid E \in \mathcal{E}\}$ form an instance of MINA-IRRELEVANCE for Horn-$\mathcal{EL}$ TBoxes. Furthermore, we know that for every $W \subseteq V$, W is a transversal of $\mathcal{H}$ iff $\mathcal{T}_W$ is a MinA for ϕ in $\mathcal{T}$. The hypergraph $\mathcal{H}$ is 2-colorable iff there is a transversal W of $\mathcal{H}$ such that for all $E \in \mathcal{E}, E \not\subseteq W$. Hence, $\mathcal{H}$ is 2-colorable iff there is a MinA $\mathcal{T}'$ for ϕ in $\mathcal{T}$ such that $\mathcal{T}_E \not\subseteq \mathcal{T}'$ for all $E \in \mathcal{E}$. □

Next we consider the dual problem, which is checking the existence of a MinA that contains a certain axiom.

Problem: MINA-RELEVANCE
Input: A KB $\mathcal{K}$ and an axiom φ of the same type such that $\mathcal{K} \models \varphi$, and an axiom $\psi \in \mathcal{K}$.
Question: Is there a MinA $\mathcal{M}$ for φ in $\mathcal{K}$ such that $\psi \in \mathcal{M}$?

Theorem 18. MINA-RELEVANCE *is* NP*-complete for Horn KBs.*

Proof. The problem is clearly in NP. A nondeterministic algorithm for solving it first guesses a subset of $\mathcal{K}$, then tests in polynomial time whether it is a MinA containing ψ. For showing hardness we are going to give a reduction from the following NP-complete problem [EG95a]:

Problem: HORN-RELEVANCE
Input: Two sets of propositional variables H and M, a set $\mathcal{C}$ of definite Horn clauses over $H \cup M$, and a propositional variable $p \in H$.
Question: Is there a minimal $G \subseteq H$ such that $G \cup \mathcal{C} \models M$ and $p \in G$?

Let an instance of HORN RELEVANCE be given with $H, M, \mathcal{C}$ and p. We construct an instance of MINA-RELEVANCE as follows: In addition to the propositional variables in $H \cup M$, we introduce two more fresh ones a, and b. Using these we construct the Horn KB $\mathcal{K} := \{a \to h \mid h \in H\} \cup \mathcal{C} \cup \{\bigwedge_{m \in M} m \to b\}$, the axiom $\varphi := a \to b$, and the axiom $\psi := a \to p$. It is easy to see that this construction indeed creates an instance of MINA-RELEVANCE and it can be done in polynomial time. We claim that there is a minimal $G \subseteq H$ such that $G \cup \mathcal{C} \models M$ and $p \in G$ if and only if there is a MinA $\mathcal{M}$ for φ in $\mathcal{K}$ such that $\psi \in \mathcal{M}$.

($\Rightarrow$) Assume that there is such a minimal G. From G we construct $\mathcal{K}_G := \{a \to g \mid g \in G\} \cup \mathcal{C} \cup \{\bigwedge_{m \in M} m \to b\}$. $\mathcal{K}_G \models a \to b$ since $G \cup \mathcal{C} \models M$. Thus, there is a MinA $\mathcal{M}$ for ϕ in $\mathcal{K}_G$. Furthermore, since G is minimal, for every $g \in G$ the axiom $a \to g$ is in $\mathcal{M}$. In particular, $\phi \in \mathcal{M}$.

($\Leftarrow$) Assume that there is such a MinA $\mathcal{M}$. It contains the axiom $\bigwedge_{m \in M} m \to b$, and also contains axioms from $\mathcal{C}$ such that every $m \in M$ occurs on the rhs of at least one axiom. Additionally $\mathcal{M}$ contains axioms of the form $a \to h$ such that $\mathcal{M} \models a \to \bigwedge_{m \in M} m$. Then the set $G := \{h \mid a \to h \in \mathcal{M}\}$ satisfies $G \cup \mathcal{C} \models M$. Moreover $p \in G$ since $a \to p \in \mathcal{M}$, and G is minimal since $\mathcal{M}$ is minimal. □

5 Counting MinAs

In applications where one is interested in computing all MinAs, it might also be useful to know in advance how many of them exist. Next we consider this counting problem.

Problem: #MINA
Input: A KB $\mathcal{K}$ and an axiom ϕ of the same type such that $\mathcal{K} \models \phi$.
Output: The number of all MinAs for ϕ in $\mathcal{K}$.

If $\mathcal{K}$ is a core KB, the problem boils down to the problem of counting simple paths between two vertices of a given directed graph. This problem called S-T CONNECTEDNESS has already been considered in [Val79b].

Problem: S-T CONNECTEDNESS
Input: A directed graph $G = (V, E)$, and two vertices $s, t \in V$.
Output: The number of subgraphs of G in which there is a path from s to t.

In [Val79b] it has been shown that this problem is #P-complete. #P is defined [Val79a] as the class of functions counting the accepting paths of nondeterministic Turing machines. Typical members of this class are the problems of counting the number of solutions of NP-complete problems. Among them, the most well-known one is #SAT, which is the problem of counting the distinct truth assignments that satisfy a given Boolean formula in CNF.

Since core KBs are the simplest type of KB, the hardness result applies to the other KB types we consider here. Moreover for the most expressive fragment we consider, namely $\mathcal{EL}$, the problem of checking whether a given set of axioms is a MinA is polynomial. This implies that for this fragment, and all others considered here, #MINA is is in #P, thus it is #P-complete.

Corollary 19. #MINA *is #P-complete for core,Horn,dual-Horn,Bool and $\mathcal{EL}$ KBs.*

Next we consider another counting problem. Instead of the number of all MinAs, one can also be interested in the number of MinAs that contain a specific axiom. If we are trying to explain an unwanted consequence, the solution of this counting problem will allow us to detect axioms that are most likely to be faulty, i. e. those that appear in the most MinAs. This idea has been proposed in [SHCH07] as a heuristic for correcting an error while minimizing the changes in the set of axioms.

Problem: #MINA-RELEVANCE
Input: A KB $\mathcal{K}$ and an axiom ϕ of the same type such that $\mathcal{K} \models \phi$, and an axiom $\psi \in \mathcal{K}$.
Output: The number of all MinAs for ϕ in $\mathcal{K}$ that contain ψ.

Theorem 20. #MINA-RELEVANCE *is #P-complete for Horn KBs.*

Proof. The problem is in #P since given a Horn KB $\mathcal{K}$, an axiom ϕ that follows from $\mathcal{K}$, an axiom $\psi \in \mathcal{K}$, and a candidate solution $\mathcal{K}' \subseteq \mathcal{K}$, we can in polynomial time verify that $\mathcal{K}$' is a MinA and it contains ψ.

For showing #P-hardness we give a parsimonious reduction from #MINA for core KBs, which has been shown to be #P-hard above. Given an instance of #MINA with the core KB $\mathcal{K}$ and the axiom $a \to b$ we construct the Horn KB $\mathcal{K}' := \mathcal{K} \cup \mathcal{S}$, where $\mathcal{S} = \{a \to c, b \wedge c \to d\}$, and c and d are two fresh propositional variable names not occurring in $\mathcal{K}$. It is not difficult to see that a set $\mathcal{M} \subseteq \mathcal{K}$ is a MinA for $a \to b$ in $\mathcal{K}$ if and only if $\mathcal{M} \cup \mathcal{S}$ is a MinA for $a \to d$ in $\mathcal{K}'$. Moreover, every MinA for $a \to d$ in $\mathcal{K}'$ contains the axioms in $\mathcal{S}$. Thus, there are exactly as many MinAs for $a \to b$ in $\mathcal{K}$ as there are for $a \to d$ in $\mathcal{K}$' containing the axiom $a \to c$. □

Obviously, Theorem 20 implies that #MINA is #P-complete for Bool and Horn-$\mathcal{EL}$ KBs.

6 Concluding Remarks and Future Work

We have analyzed the complexity of axiom pinpointing and many related problems in the propositional Horn fragment and in the DL $\mathcal{EL}$. Our hardness results extend to more expressive DLs. Tables 2 and 3 summarize our results where TH stands for TRANS-HYP, TE stands for transversal enumeration, '-h'

	FIRST-MINA	LAST-MINA	ALL-MINAS	MINA-REL	MINA-IRREL	#MINA	#MINA-REL
core		poly	poly			#P-c	
Horn		poly	poly	NP-c		#P-c	#P-c
dual-Horn	coNP-c	poly	TRANS-HYP-h		NP-c	#P-c	
Bool	coNP-c	poly	TRANS-HYP-h	NP-c	NP-c	#P-c	#P-c
Horn-$\mathcal{EL}$	coNP-c	poly	coNP-c	NP-c	NP-c	#P-c	#P-c

Table 2: Complexity of related decision and counting problems

	MINA-ENUM		
	in lexicographic order		unordered
	forward	backward	
core	output polynomial	polynomial delay	polynomial delay
Horn	output polynomial	polynomial delay	polynomial delay
dual-Horn	not polynomial delay	TRANS-ENUM-h	TRANS-ENUM-h
Bool	not polynomial delay	TRANS-ENUM-h	TRANS-ENUM-h
Horn-$\mathcal{EL}$	not output polynomial		

Table 3: Complexity of MINA-ENUM in different settings

stands for hard, and '-c' stands for complete. As future work we are going to work on determining the exact complexity of ALL-MINAS problem for dual-Horn KBs. We are going to check whether it is equivalent to the TRANS-HYP problem. We are also going to investigate the complexity of ALL-MINAS for more expressive DLs to see whether it remains in the same complexity class as reasoning.

References

[BCM+03] Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, 2003.

[Ber89] Claude Berge. *Hypergraphs*. Elsevier Science Publishers B.V. (North Holland), 1989.

[BH95] Franz Baader and Bernhard Hollunder. Embedding defaults into terminological representation systems. *Journal of Automated Reasoning*, 14:149–180, 1995.

[BP07] Franz Baader and Rafael Peñaloza. Axiom pinpointing in general tableaux. In Nicola Olivetti, editor, *Proceedings of the 16th International Conference on Automated Reasoning with Analytic Tableaux and Related Methods, (TABLEAUX 2007)*, volume 4548 of *Lecture Notes in Computer Science*, pages 11–27. Springer-Verlag, 2007.

[BP08] Franz Baader and Rafael Peñaloza. Automata-based axiom pinpointing. In Alessandro Armando, Peter Baumgartner, and Gilles Dowek, editors, *Proceedings of the 4th International Joint Conference on Automated Reasoning, (IJCAR 2008)*, volume 5195 of *Lecture Notes in Computer Science*, pages 226–241. Springer-Verlag, 2008.

[BP09] Franz Baader and Rafael Peñaloza. Automata-based axiom pinpointing. *Journal of Automated Reasoning*, 2009. To appear.

[BP10] Franz Baader and Rafael Peñaloza. Axiom pinpointing in general tableaux. *Journal of Logic and Computation*, 2010. To appear.

[BPS07] Franz Baader, Rafael Peñaloza, and Boontawee Suntisrivaraporn. Pinpointing in the description logic $\mathcal{EL}^{+}$. In Joachim Hertzberg, Michael Beetz, and Roman Englert, editors, *Proceedings of the 30th German Conference on Artificial Intelligence (KI2007)*, volume 4667 of *Lecture Notes in Artificial Intelligence*, pages 52–67. Springer-Verlag, 2007.

[DG84] William F. Dowling and Jean H. Gallier. Linear-time algorithms for testing the satisfiability of propositional Horn formulae. *Journal of Logic Programming*, 3:267–284, 1984.

[EG91] Thomas Eiter and Georg Gottlob. Identifying the minimal transversals of a hypergraph and related problems. Technical Report CD-TR 91/16, Christian Doppler Laboratory for Expert Systems, TU Vienna, 1991.

[EG95a] Thomas Eiter and Georg Gottlob. The complexity of logic-based abduction. *Journal of the ACM*, 42(1):3–42, 1995.

[EG95b] Thomas Eiter and Georg Gottlob. Identifying the minimal transversals of a hypergraph and related problems. *SIAM Journal on Computing*, 24(6):1278–1304, 1995.

[EG02] Thomas Eiter and Georg Gottlob. Hypergraph transversal computation and related problems in logic and AI. In Sergio Flesca, Sergio Greco, Nicola Leone, and Giovambattista Ianni, editors, *Proceedings of the European Conference on Logics in Artificial Intelligence (JELIA 2002)*, volume 2424 of *Lecture Notes in Computer Science*, pages 549–564. Springer-Verlag, 2002.

[EGM03] Thomas Eiter, Georg Gottlob, and Kazuhisa Makino. New results on monotone dualization and generating hypergraph transversals. *SIAM Journal on Computing*, 32(2):514–537, 2003.

[EMG08] Thomas Eiter, Kazuhisa Makino, and Georg Gottlob. Computational aspects of monotone dualization: A brief survey. *Discrete Applied Mathematics*, 156(11):2035–2049, 2008.

[FK96] Michael L. Fredman and Leonid Khachiyan. On the complexity of dualization of monotone disjunctive normal forms. *Journal of Algorithms*, 21(3):618–628, 1996.

[GJ90] Michael R. Garey and David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. W. H. Freeman & Company, New York, NY, USA, 1990.

[GKMT97] Dimitrios Gunopulos, Roni Khardon, Heikki Mannila, and Hannu Toivonen. Data mining, hypergraph transversals, and machine learning. In *Proceedings of the Sixteenth Symposium on Principles of Database Systems, (PODS 97)*, pages 209–216, 1997.

[Hag08] Matthias Hagen. *Algorithmic and Computational Complexity Issues of MONET*. Ph.D. dissertation, Institut für Informatik, Friedrich-Schiller-Universität Jena, 2008.

[HPS08] Matthew Horridge, Bijan Parsia, and Ulrike Sattler. Laconic and precise justifications in owl. In Amit P. Sheth, Steffen Staab, Mike Dean, Massimo Paolucci, Diana Maynard, Timothy W. Finin, and Krishnaprasad Thirunarayan, editors, *Proceedings of the 7th International Semantic Web Conference, (ISWC 2008)*, volume 5318 of *Lecture Notes in Computer Science*, pages 323–338. Springer-Verlag, 2008.

[HPS09] Matthew Horridge, Bijan Parsia, and Ulrike Sattler. Explaining inconsistencies in owl ontologies. In Lluis Godo and Andrea Pugliese, editors, *Proceedings of the Third International Conference on Scalable Uncertainty Management, (SUM 2009)*, volume 5785 of *Lecture Notes in Computer Science*, pages 124–137. Springer-Verlag, 2009.

[HPSvH03] Ian Horrocks, Peter F. Patel-Schneider, and Frank van Harmelen. From SHIQ and RDF to OWL: the making of a web ontology language. *Journal of Web Semantics*, 1(1):7–26, 2003.

[HTR06] Matthew Horridge, Dmitry Tsarkov, and Timothy Redmond. Supporting early adoption of OWL 1.1 with Protege-OWL and FaCT++. In *Proceedings of the Second International Workshop OWL: Experiences and Directions (OWLED 2006)*. CEUR-WS, 2006.

[JYP88] David S. Johnson, Mihalis Yannakakis, and Christos H. Papadimitriou. On generating all maximal independent sets. *Information Processing Letters*, 27(3):119–123, 1988.

[KFNM04] Holger Knublauch, Ray W. Fergerson, Natalya Fridman Noy, and Mark A. Musen. The protégé OWL plugin: An open development environment for semantic web applications. In Sheila A. McIlraith, Dimitris Plexousakis, and Frank van Harmelen, editors, *Proceedings of the 3rd International Semantic Web Conference, (ISWC 2004)*, volume 3298 of *Lecture Notes in Computer Science*, pages 229–243. Springer-Verlag, 2004.

[KPHS07] Aditya Kalyanpur, Bijan Parsia, Matthew Horridge, and Evren Sirin. Finding all justifications of owl dl entailments. In *Proceedings of the 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, (ISWC 2007 + ASWC 2007)*, volume 4825 of *Lecture Notes in Computer Science*, pages 267–280. Springer-Verlag, 2007.

[KPS+06] Aditya Kalyanpur, Bijan Parsia, Evren Sirin, Bernardo Cuenca Grau, and James A. Hendler. Swoop: A web ontology editing browser. *Journal of Web Semantics*, 4(2):144–153, 2006.

[KS03] Dimitris J. Kavvadias and Elias C. Stavropoulos. Checking monotone boolean duality with limited nondeterminism. Technical Report TR2003/07/02, Computer Technology Institute, Patras, Greece, 2003.

[MLBP06] Thomas Meyer, Kevin Lee, Richard Booth, and Jeff Z. Pan. Finding maximally satisfiable terminologies for the description logic $\mathcal{ALC}$. In *Proceedings of the 21st Natonal Conference on Artificial Intelligence (AAAI 2006)*, pages 269–274. AAAI Press/The MIT Press, 2006.

[NPA06] Lars Relund Nielsen, Daniele Pretolani, and Kim Allan Andersen. Finding the shortest hyperpaths using reoptimization. *Oper. Res. Lett.*, 34(2):155–164, 2006.

[PS09] Rafael Peñaloza and Barış Sertkaya. Axiom pinpointing is hard. In Bernardo Cuenca Grau, Ian Horrocks, Boris Motik, and Ulrike Sattler, editors, *Proceedings of the 2009 International Workshop on Description Logics (DL2009)*, volume 477 of *CEUR-WS*, 2009.

[PSK05] Bijan Parsia, Evren Sirin, and Aditya Kalyanpur. Debugging OWL ontologies. In Allan Ellis and Tatsuya Hagino, editors, *Proceedings of the 14th international conference on World Wide Web (WWW 2005)*, pages 633–640. ACM, 2005.

[Rei87] R. Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32(1):57–95, 1987.

[SC03] Stefan Schlobach and Ronald Cornet. Non-standard reasoning services for the debugging of description logic terminologies. In Georg Gottlob and Toby Walsh, editors, *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI'03)*, pages 355–362. Morgan Kaufmann, 2003.

[SHCH07] S. Schlobach, Z. Huang, R. Cornet, and F. Harmelen. Debugging incoherent terminologies. *Journal of Automated Reasoning*, 39(3):317–349, 2007.

[SSS91] Manfred Schmidt-Schauß and Gert Smolka. Attributive concept descriptions with complements. *Artificial Intelligence*, 48(1):1–26, 1991.

[Val79a] Leslie G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8(2):189–201, 1979.

[Val79b] Leslie G. Valiant. The complexity of enumeration and reliability problems. *SIAM Journal on Computing*, 8(3):410–421, 1979.