Counting Complexity of Propositional Abduction

Miki Hermann
LIX (CNRS, UMR 7161)
École Polytechnique
91128 Palaiseau, France

Reinhard Pichler
Institut für Informationssysteme
Technische Universität Wien
Favoritenstrasse 9-11, A-1040 Wien

Abstract

Abduction is an important method of non-monotonic reasoning with many applications in AI and related topics. In this paper, we concentrate on propositional abduction, where the background knowledge is given by a propositional formula. Decision problems of great interest are the existence and the relevance problems. The complexity of these decision problems has been systematically studied while the counting complexity of propositional abduction has remained obscure. The goal of this work is to provide a comprehensive analysis of the counting complexity of propositional abduction in various classes of theories.

1 Introduction

Abduction is a method of non-monotonic reasoning which has taken a fundamental importance in artificial intelligence and related topics. It is widely used to produce explanations for observed symptoms and manifestations, therefore it has an important application field in diagnosis – notably in the medical domain (see [12]). Other important applications of abduction can be found in planning, database updates, datamining and many more areas (see e.g. [9; 10; 11]).

Logic-based abduction can be formally described as follows. Given a logical theory T formalizing an application, a set M of manifestations, and a set H of hypotheses, find an explanation S for M, i.e., a suitable set $S \subseteq H$ such that $T \cup S$ is consistent and logically entails M. In this paper, we consider *propositional abduction problems* (PAPs, for short), where the theory T is represented by a propositional formula over a Boolean algebra $(\{0,1\}; \vee, \wedge, \neg, \rightarrow, \equiv)$ or a Boolean field $\mathbb{Z}_2 = (\{0,1\}; +, \wedge)$, and the sets of hypotheses H together with the manifestations M consist of variables.

Example 1 Consider the following football knowledge base.

$$\begin{aligned} T = \{\, & \textit{weak\_defense} \vee \textit{weak\_attack} \rightarrow \textit{match\_lost}, \\ & \textit{match\_lost} \rightarrow \textit{manager\_sad} \wedge \textit{press\_angry} \\ & \textit{star\_injured} \rightarrow \textit{manager\_sad} \wedge \textit{press\_sad} \,\} \end{aligned}$$

Moreover, let the set of observed manifestations be

$$M = \{\, \textit{manager\_sad, press\_angry} \,\}$$

Finally, let the set of hypotheses be given as

$$H = \{\, \textit{weak\_defense, weak\_attack, star\_injured} \,\}$$

This PAP has six abductive explanations (= "solutions").

$$\begin{aligned} \mathcal{S}_1 &= \{\, \textit{weak\_defense} \,\} \\ \mathcal{S}_2 &= \{\, \textit{weak\_attack} \,\} \\ \mathcal{S}_3 &= \{\, \textit{weak\_defense, weak\_attack} \,\} \\ \mathcal{S}_4 &= \{\, \textit{weak\_attack, star\_injured} \,\} \\ \mathcal{S}_5 &= \{\, \textit{weak\_defense, star\_injured} \,\} \\ \mathcal{S}_6 &= \{\, \textit{weak\_defense, weak\_attack, star\_injured} \,\} \end{aligned}$$

Obviously, in the above example, not all solutions are equally intuitive. Indeed, for many applications, one is not interested in *all* solutions of a given PAP $\mathcal{P}$ but only in *all acceptable* solutions of $\mathcal{P}$. *Acceptable* in this context means *minimal* w.r.t. some preorder $\preceq$ on the powerset 2^H. The most natural preorder is subset-minimality $\subseteq$. This criterion can be further refined by a hierarchical organization of our hypotheses according to some *priorities* (cf. [7]). In this context, priorities can be considered as a qualitative version of probability. The resulting preorder is denoted by $\subseteq_P$. On the other hand, if indeed all solutions are acceptable, then the corresponding preorder is the syntactic equality $=$.

In Example 1, only the solutions $\mathcal{S}_1$ and $\mathcal{S}_2$ are subset-minimal. Moreover, suppose that for some reason we know that (for a specific team) *weak_defense* is much less likely to occur than *weak_attack*. This judgment can be formalized by assigning lower priority to the former. Thus, only $\mathcal{S}_2$ is considered as $\subseteq_P$-minimal w.r.t. these priorities.

The usually observed algorithmic problem in logic-based abduction is the existence problem, i.e. deciding whether at least one solution $\mathcal{S}$ exists for a given abduction problem $\mathcal{P}$. Another well-studied decision problem is the so-called relevance problem, i.e. Given a PAP $\mathcal{P}$ and a hypothesis $h \in H$, is h part of at least one acceptable solution? However, this approach is not always satisfactory. Especially in database applications, in diagnosis, and in data-mining there exist situations where we need to know *all* acceptable solutions of the abduction problem or at least an important part of them. Consequently, the enumeration problem (i.e., the computation of all acceptable solutions) has received much interest (see e.g. [5; 6]). Another natural question is concerned with the total number of solutions to the considered problem. The latter problem refers to the *counting complexity* of abduction.

#-Abduction	$=$	$\subseteq$	$\subseteq_P$
General case	$\#\cdot$coNP	$\#\cdot$coNP	$\#\cdot\Pi_2$P
T is Horn	#P	#P	$\#\cdot$coNP
T is definite Horn	#P	#P	#P
T is dual Horn	#P	#P	#P
T is bijunctive	#P	#P	$\#\cdot$coNP
T is affine	FP	#P	#P

Table 1: Counting complexity of propositional abduction

Clearly, the counting complexity provides a lower bound for the complexity of the enumeration problem. Moreover, counting the number of abductive explanations can be useful for probabilistic abduction problems (see e.g. [13]). Indeed, in order to compute the probability of failure of a given component in a diagnosis problem (under the assumption that all preferred explanations are equiprobable), we need to count the number of preferred explanations as well as the number of preferred explanations that contain a given hypothesis.

The counting complexity has been started by Valiant [15; 16] and is now a well-established part of the complexity theory, where the most known class is #P. Many counting variants of decision problems have been proved #P-complete. Higher counting complexity classes do exist, but they are not commonly known. A counting equivalent of the polynomial hierarchy was defined by Hemaspaandra and Vollmer [8], whereas generic complete problems for these counting hierarchy classes were presented in [3].

Results. The goal of this work is to provide a comprehensive analysis of the counting complexity of propositional abduction in various settings. An overview of our results is given in Table 1. The columns of this table correspond to the three preorders on 2^H considered here for defining the notion of *acceptable* solutions, namely equality $=$, subset-minimality $\subseteq$, and subset-minimality with priorities $\subseteq_P$.

Apart from the general case where the theory T is an arbitrary propositional formula, we also consider the subclasses of Horn, definite Horn, dual Horn, bijunctive, and affine theories T. The aforementioned classes enjoy several favorable properties. For instance, they are closed under conjunction and existential quantification, i.e., a conjunction of two formulas from C belongs to the class C and a formula from C with an existentially quantified variable is logically equivalent to another formula from C. Moreover, they represent the most studied formulas in logic, complexity, constraint satisfaction problems, and artificial intelligence. This is mainly due to Schaefer's famous result that the satisfiability problem for them is polynomial as opposed to the NP-completeness of the general case (see [14]).

Structure of the Paper. The paper is organized as follows. After recalling some basic definitions and results in Section 2, we analyze the counting complexity of propositional abduction for general theories (Section 3), for Horn, definite Horn, dual Horn and bijunctive theories (Section 4) and finally for affine theories (Section 5). We conclude with Section 6.

2 Preliminaries

2.1 Propositional Abduction

A *propositional abduction problem* (PAP) $\mathcal{P}$ consists of a tuple $\langle V, H, M, T\rangle$, where V is a finite set of *variables*, $H \subseteq V$ is the set of *hypotheses*, $M \subseteq V$ is the set of *manifestations*, and T is a consistent *theory* in the form of a propositional formula. A set $S \subseteq H$ is a *solution* (also called *explanation*) to $\mathcal{P}$ if $T \cup S$ is consistent and $T \cup S \models M$ holds. *Priorities* $P = \langle H_1, \ldots, H_K\rangle$ are a stratification of the hypotheses $H = H_1 \cup \cdots \cup H_K$ into a fixed number of disjoint sets. The *subset-minimality with priorities* relation $A \subseteq_P B$ holds if $A = B$ or there exists an $i \in \{1, \ldots, K\}$ such that $A \cap H_j = B \cap H_j$ for all $j < i$ and $A \cap H_i \subsetneq B \cap H_i$.

In this paper, we follow the formalism of Eiter and Gottlob [4], allowing only positive literals in the solutions, except for the affine case. In contrast, Creignou and Zanuttini [1] also allow negative literals in the solutions S. We apply the latter to affine PAPs, where we need the possibility in the algebraic setting to assign a variable to 0.

Together with the general case where T can be an arbitrary propositional formula, we consider the special cases where T is Horn, definite Horn, dual Horn, bijunctive, and affine. Due to Schaefer's famous dichotomy result (see [14]), these are the most frequently studied sub-cases of propositional formulas. A propositional clause C is said to be *Horn*, *definite Horn*, *dual Horn*, or *bijunctive* if it has at most one positive literal, exactly one positive literal, at most one negative literal, or at most two literals, respectively. A clause C is *affine* if it can be written in the form of a linear equation $x_1 + \cdots + x_k = b$ over the Boolean field $\mathbb{Z}_2$. A theory T is Horn, definite Horn, dual Horn, bijunctive, or affine if it is a conjunction (or, equivalently, a set) of Horn, definite Horn, dual Horn, bijunctive, or affine clauses, respectively.

2.2 Counting Complexity

The *#-abduction* problem is the problem of counting the number of solutions of a PAP $\mathcal{P}$. The *#-$\subseteq$-abduction* problem counts the subset-minimal solutions of $\mathcal{P}$, whereas *#-$\subseteq_P$-abduction* counts the minimal solutions w.r.t. priorities P.

Formally, a *counting problem* is presented using a *witness* function which for every input x returns a set of *witnesses* for x. A *witness* function is a function $w\colon \Sigma^* \to \mathcal{P}^{<\omega}(\Gamma^*)$, where Σ and Γ are two alphabets, and $\mathcal{P}^{<\omega}(\Gamma^*)$ is the collection of all finite subsets of Γ^*. Every such witness function gives rise to the following *counting problem*: given a string $x \in \Sigma^*$, find the cardinality $|w(x)|$ of the *witness* set $w(x)$. According to [8], if $\mathcal{C}$ is a complexity class of decision problems, we define $\#\cdot\mathcal{C}$ to be the class of all counting problems whose witness function w satisfies the following conditions.

1. There is a polynomial $p(n)$ such that for every $x \in \Sigma^*$ and every $y \in w(x)$ we have $|y| \leq p(|x|)$;
2. The problem "given x and y, is $y \in w(x)$?" is in $\mathcal{C}$.

It is easy to verify that $\#\mathrm{P} = \#\cdot\mathrm{P}$. The counting hierarchy is ordered by linear inclusion [8]. In particular, we have that

$\#\text{P} \subseteq \#\cdot\text{coNP} \subseteq \#\cdot\Pi_2\text{P} \subseteq \#\cdot\Pi_3\text{P}$, etc. Note that one can, of course, also consider the classes $\#\cdot\text{NP}$, $\#\cdot\Sigma_2\text{P}$, $\#\cdot\Sigma_3\text{P}$, etc. However, they play no role in this work.

The prototypical $\#\cdot\Pi_k\text{P}$-complete problem for $k \in \mathbb{N}$ is $\#\Pi_k\text{SAT}$ [3], defined as follows. Given a formula

$$\psi(X) = \forall Y_1 \exists Y_2 \cdots Q_k Y_k \; \varphi(X, Y_1, \ldots, Y_k),$$

where φ is a Boolean formula and $X, Y_1, \ldots, Y_k$ are sets of propositional variables, count the number of truth assignments to the variables in X that satisfy ψ.

Completeness of counting problems in #P is usually proved by means of Turing reductions. However, these reductions do not preserve the counting classes $\#\cdot\Pi_k\text{P}$. It is therefore better to use *subtractive reductions* [3] which preserve the aforementioned counting classes. We write $\#\cdot R$ to denote the following counting problem: given a string $x \in \Sigma^*$, find the cardinality $|R(x)|$ of the witness set $R(x)$ associated with x. The counting problem $\#\cdot A$ reduces to $\#\cdot B$ via a *strong subtractive reduction* if there exist two polynomial-time computable functions f and g such that for each $x \in \Sigma^*$:

$$B(f(x)) \subseteq B(g(x)) \text{ and } |A(x)| = |B(g(x))| - |B(f(x))|$$

A strong subtractive reduction with $B(f(x)) = \emptyset$ is called *parsimonious*. A *subtractive reduction* is a transitive closure of strong subtractive reductions.

3 General Propositional Theories

The decidability problem of propositional abduction was shown to be $\Sigma_2\text{P}$-complete in [4]. The hardness part was proved via a reduction from QSAT_2. A modification of this reduction yields the following counting complexity result.

Theorem 2 *The #-abduction problem and the #-$\subseteq$-abduction problem are $\#\cdot\text{coNP}$-complete.*

Proof: The $\#\cdot\text{coNP}$-membership is clear by the fact that it is in $\Delta_2\text{P}$ to test whether a subset $S \subseteq H$ is a solution (resp. a subset-minimal solution) of a given PAP (see [4], Proposition 2.1.5). The $\#\cdot\text{coNP}$-hardness is shown via the following parsimonious reduction from $\#\Pi_1\text{SAT}$. Let an instance of the $\#\Pi_1\text{SAT}$ problem be given by a formula

$$\psi(X) \quad = \quad \forall Y \; \varphi(X, Y)$$

with $X = \{x_1, \ldots, x_k\}$ and $Y = \{y_1, \ldots, y_l\}$. Moreover, let $x'_1, \ldots, x'_k, r_1, \ldots, r_k, t$ denote fresh, pairwise distinct variables and let $X' = \{x'_1, \ldots, x'_k\}$ and $R = \{r_1, \ldots, r_k\}$. We define the PAP $\mathcal{P} = \langle V, H, M, T \rangle$ as follows.

$$\begin{aligned}
V &= X \cup X' \cup Y \cup R \cup \{t\} \\
H &= X \cup X' \\
M &= R \cup \{t\} \\
T &= \{\neg x_i \vee \neg x'_i,\; x_i \rightarrow r_i,\; x'_i \rightarrow r_i \mid 1 \leq i \leq k\} \\
&\quad \cup \{\varphi(X, Y) \rightarrow t\}
\end{aligned}$$

Obviously, this reduction is feasible in polynomial time. We now show that the reduction is indeed parsimonious.

The manifestations R together with the formulas $x_i \rightarrow r_i$, $x'_i \rightarrow r_i$ in T enforce that in every solution S of the PAP, we have to select at least one of x_i and x'_i. The additional formula $\neg x_i \vee \neg x'_i$ enforces that we have to select at most one of x_i and x'_i. By these two conditions, the value of x'_i is fully determined by x_i, namely x'_i is the dual of x_i.

Moreover, it is easy to check that there is a one-to-one relationship between the solutions $S \subseteq X$ of $\mathcal{P}$ and the models of $\forall Y \; \varphi(X, Y)$. Hence, this reduction is indeed parsimonious. The complementarity of X and X' enforces each solution to be incomparable with the others and, therefore, to be subset-minimal. $\square$

According to the above theorem, #-abduction and #-$\subseteq$-abduction have the same counting complexity. Intuitively, this is due to the following equivalence (cf. [4]): S is a $\subseteq$-minimal solution of the PAP $\mathcal{P}$, if and only if S is a solution of $\mathcal{P}$ and for every $h \in S$, $S \setminus \{h\}$ is not a solution. Hence, taking the $\subseteq$-minimality into account makes things only polynomially harder. In contrast, as soon as there are at least 2 priority levels, the following effect may occur. Suppose that S is a solution of the PAP and that $S \setminus \{h\}$ is not a solution for every $h \in S$. Then it might well happen that, for some $h \in S$, some set of the form $S' = (S \setminus \{h\}) \cup X$ is a solution, where all hypotheses in X have higher priority than h. Checking if such a set S' (and, in particular, if such a set X) exists comes down to yet another non-deterministic guess. Formally, we thus get the following complexity result.

Theorem 3 *The #-$\subseteq_P$-abduction problem is $\#\cdot\Pi_2\text{P}$-complete via subtractive reductions.*

Proof: The $\subseteq_P$-minimal solutions of a PAP can obviously be computed by a non-deterministic polynomial-time Turing machine that generates all subsets $S \subseteq H$ and (i) checks whether S is a solution of the PAP and (ii) if so, checks whether S is $\subseteq_P$-minimal. The latter test – which is the most expensive part – can be done by a $\Pi_2\text{P}$-oracle. Indeed, the problem of testing that S is *not* $\subseteq_P$-minimal can be done by the following $\Sigma_2\text{P}$-algorithm: guess a subset $S' \subseteq H$ s.t. S' is $\subseteq_P$-smaller than S and check that S' is a solution of the PAP. Hence, the #-$\subseteq_P$-abduction problem is in $\#\cdot\Pi_2\text{P}$.

The $\#\cdot\Pi_2\text{P}$-hardness is shown by the following (strong) subtractive reduction from $\#\Pi_2\text{SAT}$. Let an instance of the $\#\Pi_2\text{SAT}$ problem be given by a formula

$$\psi(X) \quad = \quad \forall Y \; \exists Z \; \varphi(X, Y, Z)$$

with the variables $X = \{x_1, \ldots, x_k\}$, $Y = \{y_1, \ldots, y_l\}$, and $Z = \{z_1, \ldots, z_m\}$. Moreover, let $x'_1, \ldots, x'_k, p_1, \ldots, p_k$, $y'_1, \ldots, y'_l$, $q_1, \ldots, q_l, r, t$ be fresh, pairwise distinct variables and $X' = \{x'_1, \ldots, x'_k\}$, $P = \{p_1, \ldots, p_k\}$, $Y' = \{y'_1, \ldots, y'_l\}$, and $Q = \{q_1, \ldots, q_l\}$. Then we define two PAPs $\mathcal{P}_1$ and $\mathcal{P}_2$ as follows.

$$\begin{aligned}
V &= X \cup X' \cup Y \cup Y' \cup Z \cup P \cup Q \cup \{r, t\} \\
H &= X \cup X' \cup Y \cup Y' \cup \{r\} \\
&\quad \text{with priorities } H_1 = H \setminus Y' \quad \text{and} \quad H_2 = Y' \\
M &= P \cup Q \cup \{t\} \\
T_1 &= \{\neg x_i \vee \neg x'_i,\; x_i \rightarrow p_i,\; x'_i \rightarrow p_i \mid 1 \leq i \leq k\} \\
&\quad \cup \{\neg y_i \vee \neg y'_i,\; y_i \rightarrow q_i,\; y'_i \rightarrow q_i \mid 1 \leq i \leq l\} \\
&\quad \cup \{\neg\varphi(X, Y, Z) \rightarrow t\} \\
T_2 &= T_1 \cup \{r \wedge y_1 \wedge \cdots \wedge y_l \rightarrow t\}
\end{aligned}$$

Finally we set $\mathcal{P}_1 = \langle V, H, M, T_1 \rangle$ and $\mathcal{P}_2 = \langle V, H, M, T_2 \rangle$.

Obviously, this reduction is feasible in polynomial time. Now let $A(\psi)$ denote the set of all satisfying assignments of a $\#\Pi_2$SAT-formula ψ and let $B(\mathcal{P})$ denote the set of $\subseteq_P$-minimal solutions of a PAP $\mathcal{P}$. We claim that the above definition of the PAPs $\mathcal{P}_1$ and $\mathcal{P}_2$ is indeed a (strong) subtractive reduction, i.e. that

$$B(\mathcal{P}_1) \subseteq B(\mathcal{P}_2) \quad \text{and} \quad |A(\psi)| = |B(\mathcal{P}_2)| - |B(\mathcal{P}_1)|$$

Due to lack of space, the proof of this claim is omitted. □

4 Horn, Dual Horn, and Bijunctive Theories

In this section, we consider the special case where the theory T is a set of (arbitrary or definite) Horn, dual Horn, or bijunctive clauses. If no minimality criterion is applied to the solutions then we get the following result.

Theorem 4 *The #-abduction problem of Horn, definite Horn, dual Horn, or bijunctive clauses is* $\#\mathrm{P}$*-complete.*

Proof: The $\#\mathrm{P}$-membership is easily seen by the fact that it can be checked in polynomial time whether some subset $S \subseteq H$ is a solution, since the satisfiability and also the unsatisfiability of a set of (dual) Horn or bijunctive clauses can be checked in polynomial time.

For the $\#\mathrm{P}$-hardness, we reduce the #POSITIVE-2SAT problem (which is known to be $\#\mathrm{P}$-complete by [16]) to it and show that this reduction is parsimonious. Let an arbitrary instance of #POSITIVE-2SAT be given as a 2CNF-formula

$$\psi = (p_1 \vee q_1) \wedge \cdots \wedge (p_n \vee q_n),$$

where the p_i's and q_i's are propositional variables from the set $X = \{x_1, \ldots, x_k\}$. Moreover, let $g_1, \ldots, g_n$ denote fresh, pairwise distinct variables and let $G = \{g_1, \ldots, g_n\}$. Then we define the PAP $\mathcal{P} = \langle V, H, M, T \rangle$ as follows.

$$\begin{aligned} V &= X \cup G \\ H &= X \\ M &= G \\ T &= \{p_i \rightarrow g_i \mid 1 \le i \le n\} \cup \{q_i \rightarrow g_i \mid 1 \le i \le n\} \end{aligned}$$

Obviously, this reduction is feasible in polynomial time. Moreover, it is easy to check that there is a one-to-one relationship between the solutions $S \subseteq X$ of $\mathcal{P}$ and the models of ψ. Note that the clauses in T are at the same time definite Horn, bijunctive, and dual Horn. □

Analogously to the case of general theories, the counting complexity remains unchanged when we restrict our attention to subset-minimal solutions.

Theorem 5 *The #-$\subseteq$-abduction problem of Horn, definite Horn, dual Horn, or bijunctive clauses is* $\#\mathrm{P}$*-complete.*

Proof: The $\#\mathrm{P}$-membership holds analogously to the case of abduction without subset-minimality. This is due to following property (see [4], Proposition 2.1.5). $\mathcal{S} \subseteq H$ is a subset-minimal solution of $\mathcal{P}$, if and only if $\mathcal{S}$ is a solution and for all $h \in \mathcal{S}$, the set $\mathcal{S} \smallsetminus \{h\}$ is not a solution of $\mathcal{P}$.

For the $\#\mathrm{P}$-hardness, we modify the reduction from the #POSITIVE-2SAT problem in Theorem 4. Let ψ, X, and G be defined as before. Moreover, let $X' = \{x'_1, \ldots, x'_k\}$ and $R = \{r_1, \ldots, r_k\}$ be fresh, pairwise distinct variables. Then we define $\mathcal{P} = \langle V, H, M, T \rangle$ as follows.

$$\begin{aligned} V &= X \cup X' \cup G \cup R \\ H &= X \cup X' \\ M &= R \cup G \\ T &= \{p_i \rightarrow g_i,\ q_i \rightarrow g_i \mid 1 \le i \le n\} \\ &\quad \cup \{\neg x_i \vee \neg x'_i, x_i \rightarrow r_i, x'_i \rightarrow r_i \mid 1 \le i \le k\} \end{aligned}$$

The idea of the variables X' and the additional manifestations G is exactly the same as in the proof of Theorem 2.

Actually, the formula $\neg x_i \vee \neg x'_i$ can even be omitted. This is due to the fact that, whenever a subset $S \subseteq H$ with $x_i, x'_i \in S$ is a solution of $\mathcal{P}$, then $S \smallsetminus \{x'_i\}$ is also a solution since x'_i is useless as soon as x_i is present (note that the only use of x'_i is to derive r_i in the absence of x_i). Therefore in a subset-minimal solution of the PAP $\mathcal{P}$, we will never select both x_i and x'_i even without the formula $\neg x_i \vee \neg x'_i$. The remaining formulas are indeed definite Horn and dual Horn. □

Below we consider PAPs with $\subseteq_P$-minimality. It turns out that for definite Horn and dual Horn clauses, the priorities leave the counting complexity unchanged. In all other cases, the counting complexity increases.

Theorem 6 *The #-$\subseteq_P$-abduction problem of definite Horn and of dual Horn clauses is* $\#\mathrm{P}$*-complete.*

Proof: The $\#\mathrm{P}$-hardness is clear, since it holds even without priorities. The $\#\mathrm{P}$-membership for definite Horn clauses is proved as follows. Let $\mathcal{P} = \langle V, H, M, T \rangle$ where T consists only of Horn clauses. According to [4], Theorem 5.3.3, for any $S \subseteq H$, we can check in polynomial time whether S is a $\subseteq_P$-minimal solution. The $\#\mathrm{P}$-membership for definite Horn clauses is thus proved.

Now suppose that T is dual Horn, i.e. the clauses in T are either of the form $\neg p$ or $\neg p \vee q_1 \vee \cdots \vee q_m$, or $q_1 \vee \cdots \vee q_m$ (for $p, q_1, \ldots, q_m \in V$ and $m \ge 1$). Moreover, let N denote the propositional variables occurring in negative unit clauses in T, i.e. $N = \{p \mid \neg p \in T\}$. Then for every solution S of $\mathcal{P}$ we have $S \subseteq H \smallsetminus N$, since otherwise $T \cup S$ would be inconsistent. Moreover, for any S' with $S \subseteq S' \subseteq H \smallsetminus N$, the set S' is also a solution of $\mathcal{P}$, since (by the special form of dual Horn) $S' \cup T$ is also consistent and (by the monotonicity of $\models$) $S' \cup T$ also implies M.

Let $H_1, \ldots, H_K$ denote the priorities of H. Now S is a $\subseteq_P$-minimal solution of $\mathcal{P}$ if and only if S is a solution of $\mathcal{P}$ and for all $i \in \{1, \ldots, k\}$ and for all $x \in (S \cap H_i)$ the set

$$S' \quad = \quad (S \smallsetminus \{x\}) \cup (H_{i+1} \cap N) \cup \cdots \cup (H_K \cap N)$$

is *not* a solution of $\mathcal{P}$. The latter test is clearly feasible in polynomial time in the dual Horn case. □

Recall from our remark preceding Theorem 3 that the effect of at least 2 priority levels is as follows. In order to check that some solution S is *not* $\subseteq_P$-minimal, we have to test that there exists some solution of the form $S' = (S \smallsetminus \{h\}) \cup X$, where all hypotheses in X have higher priority than h. In general, the difficulty of determining if such a set X exists is the following one. If we choose X too small, then S' might not entail the manifestations M. If we choose X too big, then $S' \cup T$ might be inconsistent. The intuition underlying Theorem 6 is that the problem of choosing X too big disappears for definite Horn and dual Horn clauses. For definite Horn, the only candidate X that has to be checked is $X = H_{i+1} \cup \cdots \cup H_K$. For dual Horn, the only candidate X is $X = (H_{i+1} \cup \cdots \cup H_K) \cap N$, where N takes care that if the theory T contains a negative unit clause $\neg p$, then p must not be included in any solution.

Theorem 7 *The #-$\subseteq_P$-abduction problem of Horn or bijunctive clauses is* $\#\cdot\mathrm{coNP}$*-complete.*

Proof: The $\#\cdot\mathrm{coNP}$-membership is established as follows: Given a set of variables S, we have to (i) check whether S is a solution of the PAP and (ii) if so, check whether S is $\subseteq_P$-minimal. The latter test, which dominates the overall complexity, can be done by a coNP-oracle. Indeed, the problem of testing that S is *not* $\subseteq_P$-minimal can be done by the following NP-algorithm: guess a subset $S' \subseteq H$ s.t. S' is $\subseteq_P$-smaller than S and check (in polynomial time for a Horn or bijunctive theory), that S' is a solution of the PAP. Hence, in this case, #-$\subseteq_P$-abduction in $\#\cdot\mathrm{coNP}$.

The $\#\cdot\mathrm{coNP}$-hardness is shown by a (strong) subtractive reduction from $\#\Pi_1\mathrm{SAT}$. Let an instance of the $\#\Pi_1\mathrm{SAT}$ problem be given by a formula $\psi(X) = \forall Y\ \varphi(X,Y)$ with $X = \{x_1, \ldots, x_k\}$ and $Y = \{y_1, \ldots, y_l\}$. W.l.o.g. (see [17]), we may assume that $\varphi(X,Y)$ is in 3DNF, i.e., it is of the form $C_1 \vee \cdots \vee C_n$ where each C_i is of the form $C_i = l_{i1} \wedge l_{i2} \wedge l_{i3}$ and the l_{ij}'s are propositional literals over $X \cup Y$.

Let $x'_1, \ldots, x'_k$, $p_1, \ldots, p_k$, $y'_1, \ldots, y'_l$, $q_1, \ldots, q_l$, $g_1, \ldots, g_n$, r, t denote fresh, pairwise distinct variables and let $X' = \{x'_1, \ldots, x'_k\}$, $Y' = \{y'_1, \ldots, y'_l\}$, $P = \{p_1, \ldots, p_k\}$, $Q = \{q_1, \ldots, q_l\}$ and $G = \{g_1, \ldots, g_n\}$. Then we define two PAPs $\mathcal{P}_1$ and $\mathcal{P}_2$ as follows.

$$\begin{aligned}
V &= X \cup X' \cup Y \cup Y' \cup P \cup Q \cup G \cup \{r\} \\
H &= X \cup X' \cup Y \cup Y' \cup \{r\} \\
&\quad \text{with priorities } H_1 = H \setminus Y' \text{ and } H_2 = Y' \\
M &= P \cup Q \cup G \\
T_1 &= \{\neg x_i \vee \neg x'_i,\ x_i \to p_i,\ x'_i \to p_i \mid 1 \le i \le k\} \\
&\quad \cup \{\neg y_i \vee \neg y'_i,\ y_i \to q_i,\ y'_i \to q_i \mid 1 \le i \le l\} \\
&\quad \cup \{z_{ij} \to g_i \mid 1 \le i \le n \text{ and } 1 \le j \le 3\}
\end{aligned}$$

where z_{ij} is either of the form x_k, x'_k, y_l, or y'_l depending on whether the literal l_{ij} in C_i is of the form $\neg x_k$, x_k, $\neg y_l$, or y_l, respectively. In other words, z_{ij} encodes the dual of l_{ij}.

$$T_2 = T_1 \cup \{r \to g_i \mid 1 \le i \le n\} \cup \{r \to y_j \mid 1 \le j \le l\}$$

Finally, we set $\mathcal{P}_1 = \langle V, H, M, T_1\rangle$ and $\mathcal{P}_2 = \langle V, H, M, T_2\rangle$.

Obviously, this reduction is feasible in polynomial time. Now let $A(\psi)$ denote the set of all satisfying assignments of a $\#\Pi_1\mathrm{SAT}$-formula ψ and let $B(P)$ denote the set of $\subseteq_P$-minimal solutions of a PAP P. We claim that $\mathcal{P}_1$ and $\mathcal{P}_2$ have the following property.

$$B(\mathcal{P}_1) \subseteq B(\mathcal{P}_2) \quad \text{and} \quad |A(\psi)| = |B(\mathcal{P}_2)| - |B(\mathcal{P}_1)|$$

Due to lack of space, the proof of this claim is omitted. □

5 Affine Theories

In this section, we consider the special case where the theory T is a set of affine clauses. Recall that such a set can be written as a linear system $AX = b$ over the Boolean field $\mathbb{Z}_2$. We change to the Creignou and Zanuttini approach [1] in this case, since we need the possibility to set a variable to 0. If no minimality criterion is applied to the solutions then we get the following result.

Theorem 8 *The #-abduction of affine clauses is in* FP.

Proof: Given an affine PAP $\mathcal{P} = \langle V, H, M, T\rangle$ where T is an affine system $AX = b$, we reduce it to the problem of counting the satisfying assignments of linear systems over $\mathbb{Z}_2$. Recall that $T \cup S \models M$ means that $x = 1$ for each $x \in S$ and $x = 0$ for each $x \in H \setminus (S \cup M)$ entails $y = 1$ for each $y \in M$ in the system T. First we check whether $T \cup S \models M$ can hold. If $V \setminus (H \cup M)$ is nonempty then $\mathcal{P}$ has no solution, since we cannot force all $y \in M$ to $y = 1$. Otherwise transform the system $AX = b$ to $EY = b + FZ$, where $Y \subseteq M$ and $Z \cap M = \emptyset$, such that $EY + FZ = AX$. Transform EY by Gaussian elimination to Smith normal form giving $E'Y'$. If $E'Y'$ has a row with more than one variable, say $y_{i1} + \cdots + y_{il}$ for $l \geq 2$, then either $\mathcal{P}$ has no solution, or each solution S compatible with $y_{i1} = \cdots = y_{il} = 1$ is also compatible with $y_{ij} = 0$ for some $j \in \{1, \ldots, l\}$. If $E'Y'$ has only rows with one variable, then add to $AX = b$ the equations $y = 1$ for each $y \in M$, resulting in the new system $A'X = b'$. Check whether the last system is satisfiable and transform it by Gaussian elimination into the Smith normal form $(I\ B)X = b''$. Each truth assignment I of the variables $H \setminus M$ satisfying the linear system determines a solution S of $\mathcal{P}$, i.e., $S = \{y \in H \mid I(y) = 1\} \cup \{\neg y \mid I(y) = 0\}$. Let the linear system $(I\ B)X = b''$ have k rows. Then there are $2^{|H \setminus M| - k}$ different solutions of $\mathcal{P}$. □

Theorem 9 *The #-$\subseteq$-abduction problem of affine clauses is* $\#\mathrm{P}$*-complete.*

Proof: The $\#\mathrm{P}$-membership is clear from the fact that it can be checked in polynomial time whether a set $S \subseteq H$ is a subset-minimal solution of an affine system according to Proposition 1 in [2]. The problem of minimal affine extension, namely that given an affine system $AX = b$ and a partial assignment s to the variables X, count the number of extensions s which are minimal solutions of $AX = b$, is proved to be $\#\mathrm{P}$-complete in [2] even if the partial assignment contains no 0. There is a parsimonious reduction to an affine PAP $\mathcal{P} = \langle V, H, M, T\rangle$ as follows. Let $V = H = X$, $M = \{x_i \in X \mid s(x_i) = 1\}$, and $T = \{AX = b\}$. Let $Y \subseteq X$ be the variables *not* assigned by s and let $\bar{s}$ be an extension of s satisfying the affine system. Then the set of variables $S = \{x_i \in Y \mid \bar{s}(x_i) = 1\}$ is a subset-minimal solution of $\mathcal{P}$ if and only if the extension $\bar{s}$ is minimal. □

Theorem 10 *The #-$\subseteq_P$-abduction problem of affine clauses is* $\#\mathrm{P}$*-complete.*

Proof: The $\#\mathrm{P}$-hardness is clear, since it holds without priorities. The $\#\mathrm{P}$-membership for affine clauses is proved as follows. Let $\mathcal{P} = \langle V, H, M, T\rangle$ where T is a linear system $AX = b$ over $\mathbb{Z}_2$. Let $H_1, \ldots, H_K$ denote the priorities of H. Now S is a $\subseteq_P$-minimal solution of $\mathcal{P}$ if and only if S is a solution of $\mathcal{P}$, there exists an $i \in \{1, \ldots, k\}$ such that $S \cap H_i$ is subset-minimal , and for each $j < i$ and all other solutions S' of $\mathcal{P}$ we have $S \cap H_j = S' \cap H_j$. We can decide in polynomial time if $S \cap H_i$ is a subset minimal solution of an affine system. The second condition is tested in polynomial time as follows. For each $j < i$ we set the variables $H \setminus H_j$ in the system $AX = b$ equal to 0, resulting in a system $A'_j Y_j = b'_j$. Then the identity $S \cap H_j = S' \cap H_j$ holds if and only if the resulting system $A'_j Y_j = b'_j$ has at most one solution, what can be tested in polynomial time. Hence the overall test can be performed in polynomial time. □

6 Conclusion

Eiter and Gottlob proved in [4] a plethora of complexity results for propositional abduction. Their results were extended to a trichotomy of PAPs without minimality-criterion by Creignou and Zanuttini [1]. The use of complexity results is usually twofold. Theoretically, they give us a better understanding of the nature of the considered problem class. Practically, they give us a hint as to which subclass of the problem we should aim at, provided that the application in mind admits such a restriction. In this sense, the counting complexity results shown here are important in complementing the already known decision complexity results. Note that our results reveal significant differences between the counting complexity behavior of PAPs and the decision complexity. For instance, definite Horn abduction and bijunctive abduction were shown to be tractable in [1]. In contrast, by our Theorem 4, the corresponding counting problems are $\#\mathrm{P}$-complete. This is one more example of the often observed "easy to decide, hard to count" phenomenon.

From a complexity theoretic point of view, there is another interesting aspect to the counting complexity results shown here. The class $\#\mathrm{P}$ has been studied intensively and many completeness results for this class can be found in the literature. In contrast, for the higher counting complexity classes $\#\cdot\Pi_k\mathrm{P}$ (with $k \geq 1$), very few complete problems are known apart from the generic problems $\#\Pi_k\mathrm{SAT}$. In fact, to the best of our knowledge, our $\#\cdot\Pi_2\mathrm{P}$-completeness result in Theorem 3 is the first one apart from $\#\Pi_2\mathrm{SAT}$.

In this work, we have considered the complexity of determining the number of all $\preceq$-minimal explanations of a propositional abduction problem, where $\preceq \in \{=, \subseteq, \subseteq_P\}$. Note that in [4], complexity problems related to abduction with further notions of minimality were analyzed, namely "minimal cardinality" with or without priorities and "minimal weight". Abduction with "minimal weight" can also be considered as cost-based abduction. We are planning to extend our counting complexity analysis to these notions of minimality.

References

[1] N. Creignou and B. Zanuttini. A complete classification of the complexity of propositional abduction. Submitted, 2004.

[2] A. Durand and M. Hermann. The inference problem for propositional circumscription of affine formulas is coNP-complete. In *20th STACS*, LNCS 2607, pages 451–462, 2003.

[3] A. Durand, M. Hermann, and P. G. Kolaitis. Subtractive reductions and complete problems for counting complexity classes. *Theoretical Computer Science*, 340(3):496–513, 2005.

[4] T. Eiter and G. Gottlob. The complexity of logic-based abduction. *Journal of the Association for Computing Machinery*, 42(1):3–42, 1995.

[5] T. Eiter and K. Makino. On computing all abductive explanations. In *Proc. AAAI/IAAI 2002*, pages 62–67, 2002.

[6] T. Eiter and K. Makino. Generating all abductive explanations for queries on propositional Horn theories. In *Proc. CSL 2003*, LNCS 2803, pages 197–211, 2003.

[7] R. Fagin, J. D. Ullman, and M. Y. Vardi. On the semantics of updates in databases. In *2nd PODS*, pages 352–365, 1983.

[8] L. A. Hemaspaandra and H. Vollmer. The satanic notations: Counting classes beyond #P and other definitional adventures. *SIGACT News, Complexity Theory Column 8*, 26(1):2–13, 1995.

[9] A. Herzig, J. Lang, P. Marquis, and T. Polacsek. Updates, actions, and planning. In *17th IJCAI*, pages 119–124, 2001.

[10] A. C. Kakas and P. Mancarella. Database updates through abduction. In *16th VLDB*, pages 650–661, 1990.

[11] I. Papatheodorou, A. C. Kakas, and M. J. Sergot. Inference of gene relations from microarray data by abduction. In *8th LPNMR*, LNCS 3662, pages 389–393, 2005.

[12] Y. Peng and J. A. Reggia. *Abductive inference models for diagnostic problem solving*. Springer-Verlag, 1990.

[13] D. Poole. Probabilistic Horn abduction and Bayesian networks. *Artificial Intelligence*, 64(1):81–129, 1993.

[14] T. J. Schaefer. The complexity of satisfiability problems. In *10th STOC*, pages 216–226, 1978.

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

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

[17] C. Wrathall. Complete sets and the polynomial-time hierarchy. *Theoretical Computer Science*, 3(1):23–33, 1976.