ECAI 2010
H. Coelho et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-606-5-895

The Complexity of Handling Minimal Solutions in Logic-Based Abduction[1]

Reinhard Pichler and **Stefan Woltran**[2]

Abstract. Logic-based abduction is an important reasoning method with many applications in Artificial Intelligence including diagnosis, planning, and configuration. The goal of an abduction problem is to find a "solution", i.e., an explanation for some observed symptoms. Usually, many solutions exist, and one is often interested in minimal ones only. Previous definitions of "solutions" to an abduction problem tacitly made an open-world assumption. However, as far as minimality is concerned, this assumption may not always lead to the desired behavior. To overcome this problem, we propose a new definition of solutions based on a closed-world approach. Moreover, we also introduce a new variant of minimality where only a part of the hypotheses is subject to minimization. A thorough complexity analysis reveals the close relationship between these two new notions as well as the differences compared with previous notions of solutions.

1 Introduction

Logic-based abduction is an important reasoning method with many applications in Artificial Intelligence. The goal of an abduction problem is to find a "solution", i.e., an explanation for some observed symptoms. Abduction is therefore very well suited for diagnosis problems – above all in the medical domain but also in system diagnosis, see e.g. [11, 2]. Further important applications of abduction include configuration, planning, and data mining, see e.g. [1, 8, 10].

Formally, logic-based abduction is defined 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 (built from H) such that $T \cup S$ is consistent and logically entails M. In this paper, we only consider *propositional abduction problems* (PAPs, for short), where the theory T is represented by a propositional formula and the sets H and M consist of propositional variables. There are two works[3] which provide a comprehensive complexity analysis of propositional abduction where the notion of "solution" of a PAP is defined slightly differently: In [5], solutions are considered as subsets of H, i.e., they contain only positive information; while in [4], solutions are subsets of $H \cup \{\neg h \mid h \in H\}$, i.e., they contain both positive and negative information.

Example 1 *Consider the following diagnosis problem in the football domain: If a match is lost, then the manager is either angry or sad but never both. There are several reasons for loosing a match, namely: the team is out of form, the team lacks motivation or the star is injured. The manager is sad if and only if the star is injured. The manager gets angry if there is a bad mood in the team and the team lacks motivation. Finally, a bad mood leads to lack of motivation. Now suppose that we observe that the manager is angry and that we want to find an explanation for this observation. This task can be modeled by a PAP with the following propositional variables V, theory T, hypotheses H, and manifestations M:*

$$V = \{\mathit{OF, LM, BM, SI, ML, MA, MS}\},$$

with the intended meaning OF = out_of_form, LM = lack_of_motivation, BM = bad_mood, SI = star_injured, ML = match_lost, MA = manager_angry, and MS = manager_sad.

$$\begin{aligned} T = \{&\neg \mathit{MA} \vee \neg \mathit{MS},\ \mathit{ML} \rightarrow \mathit{MA} \vee \mathit{MS}, \\ &(\mathit{OF} \vee \mathit{LM} \vee \mathit{SI}) \rightarrow \mathit{ML},\ \mathit{SI} \leftrightarrow \mathit{MS}, \\ &(\mathit{BM} \wedge \mathit{LM}) \rightarrow \mathit{MA},\ \mathit{BM} \rightarrow \mathit{LM}\} \end{aligned}$$

$$H = \{\mathit{OF, LM, BM, SI}\} \qquad M = \{\mathit{MA}\}$$

According to [5], this PAP has the following solutions:

$$S_1 = \{\mathit{BM}\}, \qquad S_2 = \{\mathit{BM, LM}\},$$
$$S_3 = \{\mathit{BM, OF}\}, \qquad S_4 = \{\mathit{BM, LM, OF}\}$$

The approach due to [4] yields many additional *solutions – each involving the literal ¬SI. We just mention a few examples below:*

$$S_5 = \{\neg \mathit{SI, OF}\}, \qquad S_6 = \{\neg \mathit{SI, LM}\},$$
$$S_7 = \{\neg \mathit{SI, OF, BM}\}, \qquad S_8 = \{\neg \mathit{SI, OF}, \neg \mathit{BM}\},$$
$$S_9 = \{\neg \mathit{SI, OF, LM}\}, \qquad \ldots \qquad \square$$

Obviously, not all solutions above 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, with $\subseteq$ (i.e., set-inclusion) being the most natural one. In the above example, only the solutions S_1, S_5, and S_6 are $\subseteq$-minimal.

Note that the minimization may not behave as intended if there are interdependencies between the hypotheses. In particular, suppose that for some solution S, additional hypotheses $S' \subseteq H \setminus S$ are entailed by $S \cup T$. In the above example, this is the case for $S = S_1$ and $S' = \{\mathrm{LM}\}$. More generally, in a system diagnosis problem, the failure of one component (e.g., a cooler) may cause the failure of other components (e.g., some parts which are particularly sensitive to heat). With the previous definitions of solutions, it clearly suffices to include the hypotheses of S in a solution and to ignore S', since the hypotheses in S' are entailed anyway. However, in system diagnosis, the solutions to an abduction problem are ultimately used to identify faulty components and to derive suitable repair actions (like

[1] This work was supported by the Austrian Science Fund (FWF) under grant P20704-N18 and the Vienna Science and Technology Fund (WWTF) under grant ICT08-028.
[2] Technische Universität Wien, email: {pichler, woltran}@dbai.tuwien.ac.at
[3] Actually, also in [9] a plethora of complexity results on propositional abduction is proved. But the focus there is on identifying tractable fragments.

the exchange of faulty components). For the identification of repair actions, the hypotheses in S' clearly must not be ignored. In Example 1, S_1 is a minimal solution according to the previous approaches to abduction, while S_2 is not. However, if S_1 is the right explanation of the manifestation MA, then the corresponding minimal set of "faulty components" is given by S_2 rather than S_1. To eliminate such undesired effects, we propose to define solutions with a *Closed World Assumption* (CWA) in mind, in order to make entailments between hypotheses explicit. More formally, for every solution $S \subseteq H$, we stipulate that all hypotheses in $H \setminus S$ do not hold. In Example 1, we would thus rule out the solution S_1, since the hypothesis LM $\in H \setminus S$ is assumed to be false by our CWA-style approach. However, this contradicts the fact that LM is entailed by $S_1 \cup T$. Now, since S_1 is no longer a solution, S_2 is indeed minimal. Note that previous approaches tacitly made an *Open World Assumption* (OWA), in that the truth value of hypotheses not contained in a solution S is left open.

Bearing the intended correspondence between minimal solutions and minimal repairs in mind, another shortcoming of the previous approaches becomes apparent, namely: Some hypotheses may refer to external factors, which are important for explaining the manifestations but which are not subject to repair. For instance, in our football example, we may have an additional hypothesis BL (= "bad_luck" with the draw, yielding an overpowering opponent), which provides another explanation for the fact "match_lost" even though it is clearly not accessible to any repair action. It would therefore be desirable to consider only parts of the hypotheses for minimization. A more technical argument in favor of such a "partial minimization" comes from the observation that negated hypotheses in the approach of [4] (which can be easily represented by positive hypotheses in the approach of [5] as we shall see in Theorem 4) may in fact be used to state that certain components are *not* faulty. Hence, minimizing these negative hypotheses (or their representation by positive hypotheses) would clearly go directly against the goal of identifying the explanations with minimal sets of faulty components. We shall therefore investigate a refinement of minimality where only a subset of the hypotheses is subject to minimization. Note that, excluding certain hypotheses from the minimization is different from simply excluding these variables from the hypotheses. Indeed, in abduction, hypotheses are the only construct by which we can postulate facts which – together with the theory – may be used for the entailment (= the "explanation") of the manifestations. With the partial minimality approach, external factors may be part of the "explanation" without taking them into account for the minimization. This is not possible, if we exclude external factors from the hypotheses.

In this paper, we establish several important properties of these two new notions of *minimal solutions*, where we either consider a new type of solutions (namely, CWA-solutions) or a new type of minimization (namely, partial minimization). More specifically, we shall study the following decision problems:

- SOLVABILITY: Does a given PAP have a solution?
- MINIMAL SOLUTION: Given a PAP $\mathcal{P}$ and a set S, is S a minimal solution?
- $\subseteq$-RELEVANCE: Given a PAP $\mathcal{P}$ and a hypothesis h, is h contained in a minimal solution to $\mathcal{P}$?

Summary of results and structure of the paper. In Section 2, we recall some basic notions and present our CWA-style definition of solutions. A conclusion and an outlook to future work are given in Section 7. Our main results are detailed in Sections 3 – 6.

- In Section 3, we compare the solvability under our new CWA-style definition of solutions with the solvability according to the approaches of [4] and [5]. We shall show that the solvability under any of these formalisms either coincides with or can be efficiently reduced to the solvability under any of the other formalisms.

- In Sections 4 and 5, we show that the situation changes dramatically when minimality is taken into account. We provide a detailed complexity analysis of the MINIMAL SOLUTION problem and the $\subseteq$-RELEVANCE problem. Apart from the general case of propositional abduction, we consider the most important subclasses of propositional formulae representing the theory T, namely Horn, definite Horn, dual Horn, and Krom. An enormous difference between an OWA- and a CWA-definition of solutions will become apparent.

- In Section 6, we introduce the concept of *partial minimization* of the hypotheses. Due to lack of space, we only extend the approach of [5] (but not the approach of [4]) by partial minimization. Our complexity analysis of this variant of minimality will reveal many similarities with the minimality of CWA-solutions. However, we shall also point out a significant difference between the two approaches. Indeed, while the MINIMAL SOLUTION problem for definite Horn and dual Horn theories is coNP-complete for CWA-solutions, this problem becomes tractable if we consider partial minimization. In other words, for these kinds of theories, the additional expressive power of the partial minimization approach compared with the approach of [5] does not increase the computational complexity.

Note that closed-world assumption and partial minimization are classical methods in non-monotonic reasoning, see e.g. [3]. To the best of our knowledge, their application to abduction is new. Instead of modifying the definition of "solutions" we also could have introduced a corresponding non-standard consequence relation $\models_S$ and stipulate that solutions fulfill $T \cup S \models_S M$ (rather than $T \cup S \models M$). Such an approach was followed, e.g. for abduction in default logic [6].

2 Preliminaries

A *propositional abduction problem* (PAP) $\mathcal{P}$ is given by 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 propositional formula called the *theory* of $\mathcal{P}$.

For a set U of atoms, let $\overline{U}$ denote the set $\{\neg u \mid u \in U\}$. We recall two previous approaches of defining the "solutions" to an abduction problem: Following [5], a solution to a PAP $\mathcal{P}$ is a set $S \subseteq H$, s.t. $T \cup S$ is consistent and $T \cup S \models M$ holds. We call these solutions *EG-solutions* or *"positive solutions"*. We write $Sol_{EG}(\mathcal{P})$ to denote the set of all EG-solutions. Following [4], a solution to a PAP $\mathcal{P}$ is a set $S \subseteq H \cup \overline{H}$, s.t. $T \cup S$ is consistent and $T \cup S \models M$ holds. We call these solutions *CZ-solutions* or *"general solutions"*. We write $Sol_{CZ}(\mathcal{P})$ to denote the set of all CZ-solutions.

We introduce a new notion of solution as follows:

Definition 2 *Let $\mathcal{P} = \langle V, H, M, T\rangle$ be a PAP. A set $S \subseteq H$ is a* CWA-solution *to $\mathcal{P}$ if $T \cup S \cup \overline{(H \setminus S)}$ is consistent and $T \cup S \cup \overline{(H \setminus S)} \models M$ holds. We write $Sol_{CWA}(\mathcal{P})$ to denote the set of all CWA-solutions to $\mathcal{P}$. Moreover, we write $Ext(S)$ as a short-hand for $S \cup \overline{(H \setminus S)}$.*

As already mentioned in Section 1, the *minimal* solutions are usually the preferred ones. Formally, we define:

Definition 3 *Let $\tau \in \{EG, CZ, CWA\}$. Then a set S is a* minimal τ-solution *to $\mathcal{P}$, if $S \in Sol_\tau(\mathcal{P})$ and for each $S' \in Sol_\tau(\mathcal{P})$, $S' \not\subset S$ holds. Accordingly, we denote the set of* minimal τ-solutions *to a PAP $\mathcal{P}$ as $SolMin_\tau(\mathcal{P})$. A hypothesis $h \in H$ is called* τ-relevant *(resp.* $\subseteq_\tau$-relevant*), if* $h \in S$ *for at least one* $S \in Sol_\tau(\mathcal{P})$ *(resp.* $S \in SolMin_\tau(\mathcal{P})$*).*

Together with the general case where T can be an arbitrary propositional formula, we consider in this paper the special cases where T is Horn, definite Horn, dual Horn, or Krom. These are the most frequently studied subcases of propositional formulas since testing their satisfiability is tractable. They also play a prominent role in the complexity analysis in [4]. Recall that a propositional clause is said to be *Horn*, *definite Horn*, *dual Horn*, or *Krom* if it has at most one positive literal, exactly one positive literal, at most one negative literal, or at most two literals, respectively. A theory is Horn, definite Horn, dual Horn, or Krom if it is a conjunction (or, equivalently, a set) of Horn, definite Horn, dual Horn, or Krom clauses, respectively.

3 Solvability

We now compare the sets of solutions to an abduction problem for the different notions of solutions. We are thus interested in inclusion properties between $Sol_{EG}(\mathcal{P})$, $Sol_{CZ}(\mathcal{P})$, and $Sol_{CWA}(\mathcal{P})$. Moreover, we compare these three approaches in terms of solvability, i.e., when are $Sol_{EG}(\mathcal{P})$, $Sol_{CZ}(\mathcal{P})$, and $Sol_{CWA}(\mathcal{P})$ non-empty?

Proposition 1 *$Sol_{CWA}(.)$ and $Sol_{EG}(.)$ are incomparable. More specifically, there exists a PAP $\mathcal{P}$, s.t. both $Sol_{CWA}(\mathcal{P}) \not\subseteq Sol_{EG}(\mathcal{P})$ and $Sol_{EG}(\mathcal{P}) \not\subseteq Sol_{CWA}(\mathcal{P})$ hold.*

Proof. Consider the PAP $\mathcal{P} = \langle V, H, M, T\rangle$ with $T = \{a \rightarrow b, a \rightarrow d, c \vee d\}$, $H = \{a, b, c\}$, $M = \{d\}$, and $V = \{a, b, c, d\}$. Then $\{a\}$ is an EG-solution to the PAP $\mathcal{P} = \langle V, H, M, T\rangle$, but it is not a CWA-solution to $\mathcal{P}$, since $T \cup \{a, \neg b, \neg c\}$ is inconsistent.

On the other hand, $\{b\}$ is a CWA-solution to $\mathcal{P}$, since $T \cup \{\neg a, b, \neg c\}$ is consistent and implies d (using $\{\neg c, c \vee d\}$), while $T \cup \{b\} \not\models d$, and thus $\{b\}$ is not an EG-solution. □

Proposition 2 *For every PAP $\mathcal{P}$, $Sol_{EG}(\mathcal{P}) \subseteq Sol_{CZ}(\mathcal{P})$, while the converse is, in general, not true.*

Proof. $Sol_{EG}(\mathcal{P}) \subseteq Sol_{CZ}(\mathcal{P})$ follows immediately from the definition of $Sol_{EG}(.)$ and $Sol_{CZ}(.)$. A counter-example for the converse is the PAP $\mathcal{P}$ in the proof of Proposition 1, where $\{\neg c\}$ is in $Sol_{CZ}(\mathcal{P})$ but not in $Sol_{EG}(\mathcal{P})$. □

Proposition 3 *$Sol_{CWA}(.)$ and $Sol_{CZ}(.)$ are incomparable. However, for every PAP $\mathcal{P}$, the following relationships hold: Every $S \in Sol_{CWA}(\mathcal{P})$ can be extended to a solution $S' \in Sol_{CZ}(\mathcal{P})$. Likewise, for every $S' \in Sol_{CZ}(\mathcal{P})$, the "positive part" $S' \cap H$ can be extended to a solution $S \in Sol_{CWA}(\mathcal{P})$.*

Proof. Let $\mathcal{P} = \langle V, H, M, T\rangle$ with $H = \{h_1, \ldots, h_n\}$ and let $S = \{h_1, \ldots, h_k\}$ be a CWA-solution. Then $S' = \{h_1, \ldots, h_k\} \cup \{\neg h_{k+1}, \ldots, \neg h_n\}$ is a CZ-solution.

Conversely, let S' be a CZ-solution. W.l.o.g., S' is of the form $S' = \{h_1, \ldots, h_i\} \cup \{\neg h_{i+1}, \ldots, \neg h_j\}$ with $0 \leq i \leq j \leq n$. By definition of CZ-solutions, $S' \cup T$ is consistent and $S' \cup T \models M$. Hence, there exists a truth assignment I to V, s.t. $S' \cup T$ is true in I. Then $S = \{h_1, \ldots, h_i\} \cup \{h_\alpha \mid j < \alpha \leq n \text{ and } I(h_\alpha) = \text{true}\}$ is a CWA-solution. □

In summary, we get the following relationships between the solvability of PAPs under the three notions of solutions:

Theorem 4 *For every PAP $\mathcal{P}$, the following properties hold:*

(1) *$Sol_{CWA}(\mathcal{P}) \neq \emptyset$ iff $Sol_{CZ}(\mathcal{P}) \neq \emptyset$.*
(2) *If $Sol_{EG}(\mathcal{P}) \neq \emptyset$ then $Sol_{CZ}(\mathcal{P}) \neq \emptyset$ and $Sol_{CWA}(\mathcal{P}) \neq \emptyset$, while the converse is, in general, not true.*
(3) *The solvability w.r.t. any of the three formalisms can be reduced to any of the others in polynomial time.*

Proof. (1) follows from Proposition 3. The implication in (2) is clear by Propositions 2 and 3. As a counter-example for the converse of (2), we consider the PAP $\mathcal{P} = \langle V, H, M, T\rangle$ with $V = \{a, b\}$, $H = \{a\}$, $M = \{b\}$, and $T = \{a \vee b\}$. Clearly, $\emptyset$ is a CWA-solution (since $\{\neg a, a \vee b\} \models b\}$), and $\{\neg a\}$ is a CZ-solution. However, $\mathcal{P}$ has no EG-solution. For (3), the mutual PTIME-reducibility between EG-solvability and CZ-solvability is implicit in the $\Sigma_2 P$-completeness of both problems according to [4] and [5], respectively. The mutual PTIME-reducibility between any of the three formalisms is then a consequence of property (1) of this theorem. We only make the reduction from CZ-solvability to EG-solvability explicit here: Let $\mathcal{P} = \langle V, H, M, T\rangle$ be an arbitrary PAP. Let $H' = \{h' \mid h \in H\}$ and $T' = \{h' \rightarrow \neg h \mid h \in H\}$, i.e., the additional hypotheses $h' \in H'$ can be used to enforce a negative truth value for any $h \in H$. Then we define the PAP $\mathcal{P}' = \langle V \cup H', H \cup H', M, T \cup T'\rangle$. It is easy to check that $Sol_{CZ}(\mathcal{P}) \neq \emptyset$ iff $Sol_{EG}(\mathcal{P}') \neq \emptyset$. □

Remarks. Property (1) in Theorem 4 together with the $\Sigma_2 P$-completeness of the CZ-solvability problem immediately yields the $\Sigma_2 P$-completeness of the solvability problem in case of the CWA-approach. Likewise, all complexity results proved in [4] for restricted theories (like Horn, definite Horn, dual Horn, Krom) carry over from the CZ-approach to the CWA-approach. As far as the interreducibility between the EG- and CZ-approach is concerned, we have only considered the case of general theories in the proof of Property (3) in Theorem 4. Of course, adding the set of formulae $T' = \{h' \rightarrow \neg h\}$ to T preserves the restriction to Krom and Horn theories. Moreover, for definite Horn theories, one can easily verify that negative literals in solutions have no effect, i.e.: S is a CZ-solution iff $S \setminus \overline{H}$ is an EG-solution, provided that T is definite Horn. The only case where *no polynomial-time reduction from CZ-solvability to EG-solvability exists* (unless PTIME $=$ NP) are dual Horn theories. This is due to the fact that the CZ-solvability problem in the dual Horn case is NP-complete (see [4]) while the EG-solvability problem for dual Horn theories can be shown to be in PTIME (by using ideas similar to the PTIME-membership proof in Theorem 13 below).

4 Minimal Solution Problem

We start our complexity analysis with the problem of recognizing minimal solutions.

Problem: MINIMAL SOLUTION
Input: PAP $\mathcal{P} = \langle V, H, M, T\rangle$ and a set S.
Output: Is S a *minimal* solution?

For the EG- and CZ-approach, this problem is tractable for Horn, dual Horn or Krom theories, and in Δ_2P for unrestricted theories. We show that, for CWA-solutions, the complexity increases to coNP-completeness for the special cases and to Π_2P-completeness for the general case. This jump in the complexity is due to the non-monotonicity introduced by the CWA-approach: For the EG- and CZ-approach, the minimality of some solution S is checked by verifying that none of the "direct subsets" S' of S (i.e., the sets $S' = S \setminus \{h\}$ for some $h \in S \cap H$) is a solution. Since the CWA-approach destroys this non-monotonicity, another non-deterministic guess of an arbitrary subset S' of S is required.

Theorem 5 *The* MINIMAL SOLUTION *problem of CWA-solutions is* Π_2P*-complete.*

Proof. For the membership, consider $\mathcal{P} = \langle V, H, M, T\rangle$ and $S \subseteq H$. S is a minimal CWA-solution to $\mathcal{P}$ if S is a CWA-solution and there exists no strictly smaller CWA-solution. We can check in Δ_2P whether S is a CWA-solution, i.e.: the consistency of $Ext(S) \cup T$ and the entailment $Ext(S) \cup T \models M$ can be checked by two calls to an NP-oracle. It remains to show that the minimality of S can be checked in Π_2P. In fact, the co-problem of checking that there does not exist a smaller solution can be decided in Σ_2P: Guess a subset $S' \subset S$ and check by two calls to an NP-oracle that $Ext(S') \cup T$ is consistent and $Ext(S') \cup T \models M$.

The Π_2P-hardness is shown by reducing QSAT$_2$ to the co-problem: Let $\psi = \exists X \forall Y \varphi(X, Y)$ with $X = \{x_1, \dots, x_k\}$ and $Y = \{y_1, \dots, y_\ell\}$ denote an arbitrary instance of QSAT$_2$. Let h, t denote fresh variables. We define an instance of the co-problem of MINIMAL SOLUTION by $\mathcal{P} = \langle V, H, M, T\rangle$ and $S = \{x_1, \dots, x_k, h\}$ with

$$\begin{aligned} V &= X \cup Y \cup \{h, t\} \\ H &= X \cup \{h\} \qquad\qquad M = \{t\} \\ T &= \{\varphi(X, Y) \to t\} \cup \{(x_1 \wedge \cdots \wedge x_k \wedge h) \to t\} \end{aligned}$$

Obviously, this reduction is feasible in polynomial time. It remains to show its correctness: $\psi = \exists X \forall Y \varphi(X, Y)$ is valid $\Leftrightarrow$ S is not a minimal solution of $\mathcal{P}$. We only show the "$\Leftarrow$"-direction. The "$\Rightarrow$"-direction is analogous. Suppose that S is *not* a minimal solution. Clearly, S is a solution. Hence, there exists a strictly smaller solution $S' \subset S$. Then $S' \cup T$ implies t via the subformula $\varphi(X, Y) \to t$. We define the assignment I on X with $I(x_i) =$ true if $x_i \in S'$ and $I(x_i) =$ false otherwise. Then I is the desired assignment on X, s.t. $\varphi(X, Y)$ is true in every extension J of I to Y. □

Theorem 6 *The* MINIMAL SOLUTION *problem of CWA-solutions is* coNP*-complete for dual Horn and Krom theories.*

Proof. The coNP-membership is clear, since the satisfiability of a dual Horn or Krom theory can be decided in PTIME. We show the hardness by reduction from co-3-SAT: Let $\varphi = C_1 \wedge \cdots \wedge C_n$ be a propositional formula in 3-CNF with variables $X = \{x_1, \dots, x_k\}$, i.e., every C_i is a clause of the form $C_i = l_{i1} \vee l_{i2} \vee l_{i3}$, s.t. the l_{ij}'s are literals over X. Let $G = \{g_1, \dots, g_n\}$ be a set of new variables. Then we construct the PAP $\mathcal{P}$ as follows:

$$\begin{aligned} V &= X \cup G \cup \{h\} \\ H &= X \cup \{h\} \qquad\qquad M = G \\ T &= \{h \to g_i \mid 1 \leq i \leq n\} \cup \{h \to x_j \mid 1 \leq j \leq k\} \cup \\ &\quad\ \{\bar{l}_{ij} \vee g_i \mid 1 \leq i \leq n, 1 \leq j \leq 3\}, \end{aligned}$$

where we define $\bar{l}_{ij} = \neg x_\alpha$ if l_{ij} is of the form x_α and $\bar{l}_{ij} = x_\alpha$ if l_{ij} is of the form $\neg x_\alpha$. Finally, we set $S = X \cup \{h\}$. We claim that S is a minimal solution of $\mathcal{P} \Leftrightarrow \varphi$ is unsatisfiable.

We only show the "$\Leftarrow$"-direction. The "$\Leftarrow$"-direction is shown analogously. Suppose that S is *not* a minimal CWA-solution. As easily verified, S is a CWA-solution. Hence, there exists a strictly smaller CWA-solution $S' \subset S$. Clearly, $h \notin S'$ since otherwise (by the rules $h \to x_j$ in T), $S' = S$. We define the truth assignment I on X as $I(x_i) =$ true if $x_i \in S'$ and $I(x_i) =$ false otherwise.

It remains to show that I is a model of φ, i.e., for every $i \in \{1, \dots, n\}$, the clause C_i is true in I. Since S' is a solution and $g_i \in M$, we have $Ext(S') \cup T \models g_i$. There are 4 clauses containing g_i in T. Clearly, $h \to g_i$ cannot be used to imply g_i, since $h \notin S'$. Hence, g_i is implied via a clause $\bar{l}_{ij} \vee g_i$.

We distinguish the 2 cases of the definition of $\bar{l}_{ij}$: Suppose that $\bar{l}_{ij} = \neg x_\alpha$ and $x_\alpha \in Ext(S')$. Then l_{ij} is of the form x_α and $I(x_\alpha) =$ true by definition of I. Likewise, if $\bar{l}_{ij} = x_\alpha$ and $\neg x_\alpha \in Ext(S')$, then l_{ij} is of the form $\neg x_\alpha$ and $I(x_\alpha) =$ false. In either case, the clause C_i is true in I. □

Theorem 7 *The* MINIMAL SOLUTION *problem of CWA-solutions is* coNP*-complete for Horn theories. The* coNP*-completeness even holds for definite Horn theories.*

Proof. Again, only the hardness part is non-trivial. The proof proceeds by a reduction from the co-3-SAT-problem: Let $\varphi = C_1 \wedge \cdots \wedge C_n$ be an arbitrary propositional formula in 3-CNF over the variables $X = \{x_1, \dots, x_k\}$. Let $X' = \{x'_i \mid x_i \in X\}$ and let $G = \{g_1, \dots, g_n\}$. Then we construct the PAP $\mathcal{P}$ as follows:

$$\begin{aligned} V &= X \cup X' \cup G \cup \{h\} \\ H &= X \cup X' \cup \{h\} \qquad\qquad M = G \\ T &= \{x_j \wedge x'_j \to h,\ h \to x_j,\ h \to x'_j \mid 1 \leq j \leq k\} \cup \\ &\quad\ \{h \to g_i \mid 1 \leq i \leq n\} \cup \\ &\quad\ \{l^*_{ij} \to g_i \mid 1 \leq i \leq n, 1 \leq j \leq 3\}, \end{aligned}$$

where we define $l^*_{ij} = x_\alpha$ if l_{ij} is of the form x_α and $l^*_{ij} = x'_\alpha$ if l_{ij} is of the form $\neg x_\alpha$. Finally, we set $S = X \cup X' \cup \{h\}$, which is easily verified to be a CWA-solution to $\mathcal{P}$. It remains to prove that S is a *minimal* CWA-solution to $\mathcal{P} \Leftrightarrow \varphi$ is unsatisfiable.

We only show the "$\Leftarrow$"-direction. The "$\Rightarrow$"-direction is analogous. Suppose that S is *not* a minimal CWA-solution to $\mathcal{P}$, i.e. there exists a strictly smaller CWA-solution $S' \subset S$. Clearly, for every $i \in \{1, \dots, k\}$, S' does not contain both x_i and x'_i since, otherwise, $S' = S$ would hold. For the same reason, $h \notin S'$. We define the truth assignment I on X as $I(x_i) =$ true if $x_i \in S'$ and $I(x_i) =$ false otherwise. We claim that I is a model of φ. This is due to the fact that every $g_i \in G$ is implied by $Ext(S') \cup T$ via the last line of the definition of T. By the correspondence that $I(x_i) =$ true (resp. false) iff $x_i \in S'$ (resp. $x'_i \notin S'$), it follows immediately that, in every clause C_i of φ, at least one literal l_{ij} is true in I. □

5 $\subseteq$-Relevance Problem

We now extend our complexity analysis to the problem of recognizing if some hypothesis is $\subseteq$-relevant.

Problem: $\subseteq$-RELEVANCE
Input: PAP $\mathcal{P} = \langle V, H, M, T\rangle$ and hypothesis $h \in H$.
Output: Is h contained in some *minimal* solution to $\mathcal{P}$?

In [5], the relevance problem of $\subseteq$-abduction was shown to be Σ_2P-complete – exactly as if we did not require the subset-minimality. Likewise, for Horn theories, it was shown that the complexity remains unchanged if we consider $\subseteq$-RELEVANCE rather than RELEVANCE. In this section we show that, with the CWA-notion of solutions, the complexity goes one level up in the polynomial hierarchy.

As a short-hand, we write $\subseteq_{CWA}$-RELEVANCE to denote the $\subseteq$-RELEVANCE problem for CWA-solutions.

Theorem 8 *The* $\subseteq_{CWA}$*-Relevance problem is* Σ_3P*-complete.*

Proof. For the Σ_3P-membership, we give a non-deterministic algorithm: Given a PAP $\mathcal{P} = \langle V, H, M, T\rangle$, guess a subset $S \subseteq H$ with $h \in S$ and check by means of an oracle that S is a *minimal* CWA-solution to $\mathcal{P}$. This oracle works in Π_2P by Theorem 5.

The Σ_3P-hardness is shown by a reduction from QSAT$_3$: Let an arbitrary instance of QSAT$_3$ be given by the formula $\psi = \exists X \forall Y \exists Z\ \varphi(X, Y, Z)$ with $X = \{x_1, \ldots, x_k\}$, $Y = \{y_1, \ldots, y_\ell\}$, and $Z = \{z_1, \ldots, z_m\}$. Let $X' = \{x'_1, \ldots, x'_k\}$ and h, t denote fresh, pairwise distinct variables. We define an instance of $\subseteq_{CWA}$-RELEVANCE by the distinguished hypothesis h and PAP $\mathcal{P} = \langle V, H, M, T\rangle$ with

$$
\begin{aligned}
V &= X \cup X' \cup Y \cup Z \cup \{h, t\} \\
H &= X \cup X' \cup Y \cup \{h\} \qquad\qquad M = \{t\} \\
T &= \{\neg x_i \leftrightarrow x'_i \mid 1 \le i \le k\} \cup \\
&\quad\ \{\neg\varphi(X, Y, Z) \to t\} \cup \\
&\quad\ \{(y_1 \wedge \cdots \wedge y_\ell \wedge h) \to t\}.
\end{aligned}
$$

Obviously, this reduction is feasible in polynomial time. It remains to show its correctness: ψ is valid $\Leftrightarrow$ h is contained in a $\subseteq$-minimal CWA-solution S of $\mathcal{P}$. We only work out the "$\Rightarrow$"-direction here. The "$\Leftarrow$"-direction is shown analogously. Suppose that $\psi = \exists X \forall Y \exists Z\ \varphi(X, Y, Z)$ is valid, i.e., there exists an assignment I on X, s.t. for any assignment J on Y, there exists an assignment K on Z, s.t. the formula $\varphi(X, Y, Z)$ is true in the overall assignment.

Starting from the assignment I on X, we define a subset $S \subseteq H$ as follows: $S = \{x_i \mid I(x_i) = \text{true}\} \cup \{x'_i \mid I(x_i) = \text{false}\} \cup Y \cup \{h\}$. It is easy to check that S is a CWA-solution to $\mathcal{P}$. We claim that S is subset-minimal. Suppose to the contrary that there exists a smaller CWA-solution S'. Actually, S' restricted to $X \cup X'$ cannot be smaller than S restricted to $X \cup X'$. This is due to the clauses $\neg x_i \leftrightarrow x'_i$ in T which ensure that any solution to $\mathcal{P}$ sets precisely k out of the $2k$ variables in $X \cup X'$ to true. Hence, any two CWA-solutions of $\mathcal{P}$ coincide or are incomparable on $X \cup X'$.

So suppose that S' is smaller than S because of one of the hypotheses in $Y \cup \{h\}$, i.e., S' does not contain either h or some y_j. Then the subformula $(y_1 \wedge \cdots \wedge y_\ell \wedge h)$ in T is clearly false in the assignment defined by S'. Hence, the only way to force t to true is via the implication $\neg\varphi(X, Y, Z) \to t$ in T. Note that the variables in Z are not contained in the hypotheses. This means, that $\neg\varphi(X, Y, Z)$ must be true in the assignment defined by S' for any values of Z, i.e.: for this particular assignment I on X, there exists an assignment J on Y (with $J^{-1}(\text{true}) = S' \cap Y$)), s.t. for all assignments K on Z, $\neg\varphi(X, Y, Z)$ is true or, equivalently, $\varphi(X, Y, Z)$ is false. In other words, for the assignments I and J on X and Y, there is no assignment on Z to make $\varphi(X, Y, Z)$ true. This contradicts the assumption that for the assignment I on X we have that for all assignments on Y, there exists an assignment on Z, s.t. $\varphi(X, Y, Z)$ is true. $\square$

Theorem 9 *The $\subseteq_{CWA}$-RELEVANCE problem is Σ_2P-complete for dual Horn and Krom theories.*

Proof. For the Σ_2P-membership, we give a non-deterministic algorithm: Guess a subset $S \subseteq H$ with $h \in S$ and check by a coNP-oracle (cf. Theorem 6) that S is a *minimal* CWA-solution to $\mathcal{P}$.

The Σ_2P-hardness is shown by the following problem reduction from QSAT$_2$: Let $\psi = \exists X \forall Y\ \varphi(X, Y)$ with $X = \{x_1, \ldots, x_k\}$ and $Y = \{y_1, \ldots, y_\ell\}$. W.l.o.g., $\varphi(X, Y)$ is in 3-DNF, i.e., it is of the form $D_1 \vee \cdots \vee D_n$ with $D_i = l_{i1} \wedge l_{i2} \wedge l_{i3}$, s.t. the l_{ij}'s are literals over X and Y. Let $X' = \{x'_1, \ldots, x'_k\}$ and $G = \{g_1, \ldots, g_n\}$. We construct $\mathcal{P} = \langle V, H, M, T\rangle$ as follows:

$$
\begin{aligned}
V &= X \cup X' \cup Y \cup G \cup \{h\} \\
H &= X \cup X' \cup Y \cup \{h\} \qquad\qquad M = G \\
T &= \{x_j \vee x'_j \mid 1 \le j \le k\} \cup \\
&\quad\ \{h \to g_i, h \to y_j \mid 1 \le i \le n, 1 \le j \le \ell\} \cup \\
&\quad\ \{l_{ij} \vee g_i \mid 1 \le i \le n, 1 \le j \le 3\}.
\end{aligned}
$$

It remains to show that ψ is valid $\Leftrightarrow$ h is contained in a minimal CWA-solution of $\mathcal{P}$. The proof combines ideas from the proofs of Theorems 6 and 8. $\square$

Theorem 10 *The $\subseteq_{CWA}$-RELEVANCE problem is Σ_2P-complete for Horn theories. The Σ_2P-completeness still holds even if we further restrict the theories to definite Horn.*

Proof. The Σ_2P-membership is shown as in Theorem 9. The Σ_2P-hardness is shown by reduction from QSAT$_2$: Let an arbitrary instance of QSAT$_2$ be given by the formula $\psi = \exists X \forall Y \varphi(X, Y)$ with $X = \{x_1, \ldots, x_k\}$ and $Y = \{y_1, \ldots, y_\ell\}$. Moreover, let $\varphi(X, Y)$ be in 3-DNF as in the proof of Theorem 9 and let $X' = \{x'_1, \ldots, x'_k\}$, $Y' = \{y'_1, \ldots, y'_\ell\}$, $R = \{r_1, \ldots, r_k\}$ and $G = \{g_1, \ldots, g_n\}$. We construct the PAP $\mathcal{P} = \langle V, H, M, T\rangle$ as follows:

$$
\begin{aligned}
V &= X \cup X' \cup Y \cup Y' \cup G \cup R \cup \{h\} \\
H &= X \cup X' \cup Y \cup Y' \cup \{h\} \qquad\qquad M = G \cup R \\
T &= \{x_j \to r_j, x'_j \to r_j \mid 1 \le j \le k\} \cup \\
&\quad\ \{x_j \wedge x'_j \to g_i \mid 1 \le i \le n, 1 \le j \le k\} \cup \\
&\quad\ \{y_j \wedge y'_j \to h,\ h \to y_j,\ h \to y'_j \mid 1 \le j \le \ell\} \cup \\
&\quad\ \{h \to g_i \mid 1 \le i \le n\} \cup \\
&\quad\ \{l^*_{ij} \to g_i \mid 1 \le i \le n, 1 \le j \le 3\},
\end{aligned}
$$

where $l^*_{ij} = x'_\alpha$ (resp. x_α, y'_β, or y_β) if l_{ij} is of the form x_α (resp. $\neg x_\alpha$, y_β, or $\neg y_\beta$). Note that, in contrast to the proof of Theorem 7, l^*_{ij} encodes the dual of l_{ij}. It remains to show that ψ is valid $\Leftrightarrow$ h is contained in a minimal CWA-solution of $\mathcal{P}$. The proof combines ideas from the proofs of Theorems 7 and 8. $\square$

6 Partial Minimization

We now consider another variant of *minimal* solutions, namely: We assume that a PAP $\mathcal{P} = \langle V, H, M, T\rangle$ is given together with a subset $H_0 \subseteq H$ and that the preferred solutions S are those where $S \cap H_0$ is minimal. We shall call such solutions *partially minimal*. Partially minimal solutions can of course be studied in the context of all notions of solutions considered here. Due to lack of space, we restrict ourselves to the EG-approach. For the CWA-approach, it is straightforward to verify that the complexity results (in particular, the membership results) for the MINIMAL SOLUTION and $\subseteq$-RELEVANCE problem are not affected by partial minimization. The CZ-approach would be treated similarly to the EG-approach detailed here.

Definition 11 *Let a PAP $\mathcal{P} = \langle V, H, M, T\rangle$ be given together with a subset $H_0 \subseteq H$. A set $S \subseteq H$ is a partial minimal solution if (1) $S \cup T$ is consistent, (2) $S \cup T \models M$, and (3) for every S' with $S' \cap H_0 \subset S \cap H_0$, either $S' \cup T$ is inconsistent or $S' \cup T \not\models M$.*

The MINIMAL SOLUTION problem and the $\subseteq$-RELEVANCE problem are modified in the obvious way so as to cover *partially* minimal solutions. Interestingly, the complexity increases for general theories, Horn, and Krom (as for the CWA-solutions) but not for definite Horn and dual Horn.

Theorem 12 *The* MINIMAL SOLUTION *problem of partially minimal solutions is* $\Pi_2\mathrm{P}$*-complete for general theories and* coNP*-complete for Horn or Krom theories.*

The $\subseteq$-RELEVANCE *problem of partially minimal solutions is* $\Sigma_3\mathrm{P}$*-complete for general theories and* $\Sigma_2\mathrm{P}$*-complete for Horn or Krom theories.*

Proof. The non-trivial part is the hardness. We present a reduction from CWA-solutions to partially minimal solutions. Let $\mathcal{P} = \langle V, H, M, T\rangle$ be an arbitrary PAP with $H = \{h_1, \dots, h_k\}$. Let $H' = \{h'_1, \dots, h'_k\}$ and $G = \{g_1, \dots, g_k\}$ be fresh, pairwise distinct variables and let $T' = \{\neg h_j \vee \neg h'_j, h_j \rightarrow g_j, h'_j \rightarrow g_j \mid 1 \leq j \leq k\}$. Clearly, all clauses in T' are Horn and Krom. Then we define the PAP $\mathcal{P}' = \langle V \cup H' \cup G, H \cup H', M \cup G, T \cup T'\rangle$ and set $H_0 = H$. There exists a one-to-one-correspondence between the minimal CWA-solutions of $\mathcal{P}$ and the partially minimal solutions of $\mathcal{P}'$, namely: for every $S \subseteq H$, if S is a minimal CWA-solution to $\mathcal{P}$ then $S' = S \cup \{h'_j \mid h_j \notin S\}$ is a partially minimal solution to $\mathcal{P}'$. Likewise, if $S' \subseteq H \cup H'$ is a partially minimal solution to $\mathcal{P}'$ then $S \cap H$ is a minimal CWA-solution to $\mathcal{P}$. A problem reduction from the MINIMAL SOLUTION resp. $\subseteq$-RELEVANCE problem of CWA-solutions to the corresponding problem of partially minimal solutions is now immediate. Thus, the hardness results follow from Theorems 5 – 9. □

Theorem 13 *The* MINIMAL SOLUTION *problem of partially minimal solutions is in PTIME for definite Horn and dual Horn theories. The* $\subseteq$-RELEVANCE *problem of partially minimal solutions is* NP*-complete for definite Horn and dual Horn theories.*

Proof. In [5], the NP-hardness for definite Horn theories was shown even for general subset minimality (rather than partial minimality). The NP-hardness for dual Horn can be shown similarly. We therefore restrict ourselves here to the membership proof. It suffices to show that the MINIMAL SOLUTION problem of partially minimal solutions is in PTIME. The NP-membership of $\subseteq$-RELEVANCE is then shown by the obvious guess and check algorithm. Now let $\mathcal{P} = \langle V, H, M, T\rangle$, $H_0 \subseteq H$, and $S \subseteq H$. It can be checked by the following PTIME-algorithm whether S is a partially minimal solution to $\mathcal{P}$:

First, let T be a definite Horn theory: Clearly, checking if S is an EG-solution (i.e., $S \cup T$ is consistent and $S \cup T \models M$) can be done in polynomial time. It remains to check that there exists no EG-solution S' with $S' \cap H_0 \subset S \cap H_0$. We claim that it suffices to show this property for the (polynomially many!) sets $S' = (S \setminus \{h\}) \cup (H \setminus H_0)$ for every $h \in S$. Suppose that there exists an arbitrary EG-solution S'' with $S'' \cap H_0 \subset S \cap H_0$. By $S'' \cap H_0 \subset S \cap H_0$, there exists some $h \in (S \cap H_0) \setminus S''$. Now consider $S' = (S \setminus \{h\}) \cup (H \setminus H_0)$. Then $S'' \subseteq S'$ and, therefore, $S' \cup T \models M$ by the monotonicity of $\models$. Moreover, $S' \cup T$ is consistent since $S'' \cup T$ is consistent and – for definite Horn T – additional positive atoms cannot lead to inconsistency.

Now suppose that T is dual Horn. Let $N = \{x \mid x \in V$ and $T \models \neg x\}$. Clearly, since T is dual Horn, N can be computed in polynomial time. Every solution S of $\mathcal{P}$ fulfills $S \subseteq H \setminus N$ since $T \cup S$ is consistent. Moreover, for every S' with $S \subseteq S' \subseteq H \setminus N$, the set S' is also a solution to $\mathcal{P}$, 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. Now let S be an arbitrary EG-solution. To check that there exists no EG-solution S' with $S' \cap H_0 \subset S \cap H_0$, it suffices to show this property for the (polynomially many!) sets $S' = (S \setminus \{h\}) \cup (H \setminus (H_0 \cup N))$ for every $h \in (S \cap H_0)$. □

7 Conclusion

In this paper, we have introduced a new, CWA-style definition of solutions to propositional abduction problems and proved several important properties of these solutions. Moreover, we have considered the effect of minimizing only a part of the hypotheses in the solutions. The main contribution of this paper is a comprehensive complexity analysis. The results are summarized in the following table (all entries refer to completeness results, except the "in $\mathcal{C}$" entries).

	EG/CZ		**CWA**		**partial min.**	
	min.	$\subseteq$-rel.	**min.**	**$\subseteq$-rel.**	**min.**	**$\subseteq$-rel.**
general	in $\Delta_2\mathrm{P}$	$\Sigma_2\mathrm{P}$	$\mathbf{\Pi_2P}$	$\mathbf{\Sigma_3P}$	$\mathbf{\Pi_2P}$	$\mathbf{\Sigma_3P}$
Horn	in P	NP	**coNP**	$\mathbf{\Sigma_2P}$	**coNP**	$\mathbf{\Sigma_2P}$
def.	in P	NP	**coNP**	$\mathbf{\Sigma_2P}$	**in P**	**NP**
dual	in P	NP	**coNP**	$\mathbf{\Sigma_2P}$	**in P**	**NP**
Krom	in P	NP	**coNP**	$\mathbf{\Sigma_2P}$	**coNP**	$\mathbf{\Sigma_2P}$

Our new results (in boldface) are juxtaposed with already known results from [4] and [5]. The first block of two columns displays the complexity of the MINIMAL SOLUTION problem and the $\subseteq$-RELEVANCE problem for the previous approaches. The second and third block show our new results for the CWA-approach and for partial minimization, respectively. It is interesting to note that, in the last two blocks of the table, the complexity mostly but not always jumps one level up in the polynomial hierarchy. In particular, we have identified two interesting cases (namely dual Horn and definite Horn abduction), where the complexity does not increase.

An important task for future work concerns the search for tractable subclasses of MINIMAL SOLUTION and RELEVANCE. Note that, in case of the OWA-approach, several tractable cases of abduction have been recently identified in [7] and [9]. This study clearly should be extended to the CWA-approach and the partial minimization.

REFERENCES

[1] J. Amilhastre, H. Fargier, and P. Marquis, 'Consistency restoration and explanations in dynamic CSPs – application to configuration', *Artif. Intell.*, **135**(1-2), 199–234, (2002).

[2] T. Bylander, D. Allemang, M.C. Tanner, and J.R. Josephson, 'Some results concerning the computational complexity of abduction', in *Proc. KR'89*, pp. 44–54. Morgan Kaufmann, (1989).

[3] Marco Cadoli and Maurizio Lenzerini, 'The complexity of propositional closed world reasoning and circumscription', *J. Comput. Syst. Sci.*, **48**(2), 255–310, (1994).

[4] N. Creignou and B. Zanuttini, 'A complete classification of the complexity of propositional abduction', *SIAM J. Comput.*, **36**(1), 207–229, (2006).

[5] T. Eiter and G. Gottlob, 'The complexity of logic-based abduction', *J. ACM*, **42**(1), 3–42, (1995).

[6] T. Eiter, G. Gottlob, and N. Leone, 'Semantics and Complexity of Abduction from Default Theories', *Artificial Intelligence*, **90**(1–2), 177–223, (1997).

[7] G. Gottlob, R. Pichler, and F. Wei, 'Bounded treewidth as a key to tractability of knowledge representation and reasoning', in *Proc. AAAI'06*. AAAI Press, (2006).

[8] A. Herzig, J. Lang, P. Marquis, and T. Polacsek, 'Updates, actions, and planning.', in *Proc. IJCAI'01*, pp. 119–124. Morgan Kaufmann, (2001).

[9] G. Nordh and B. Zanuttini, 'What makes propositional abduction tractable', *Artif. Intell.*, **172**(10), 1245–1284, (2008).

[10] I. Papatheodorou, A.C. Kakas, and M.J. Sergot, 'Inference of gene relations from microarray data by abduction.', in *Proc. LPNMR'05*, volume 3662 of *LNCS*, pp. 389–393. Springer, (2005).

[11] Y. Peng and J.A. Reggia, *Abductive Inference Models for Diagnostic Problem Solving*, Springer, 1990.