doi:10.1017/xxxxx

Representative Sets in Propositional Abduction *

JOHANNES SCHMIDT
Jönköping University

MOHAMED MAIZIA
Jönköping University, Linköping University

VICTOR LAGERKVIST
Linköping University

JOHANNES K. FICHTE
Linköping University

Abstract

The propositional abduction problem is a well-known form of non-monotonic reasoning where we are asked to find an explanation of a given manifestation. Recently, there has been an influx of results asking more refined questions about the solution space rather than only individual solutions. For example, we might be interested in finding two solutions that are sufficiently far from each other (*diverse* solutions) in the solution space. In this paper we consider a related *representation* question where we ask if a given set of explanations S can represent any other explanation (that is, whether their symmetric difference is smaller than a given k). We first study this problem from a classical complexity perspective and obtain a complete classification. While only a handful of cases are tractable, the increase in complexity compared to classical abduction is often smaller than expected. We then study the parameterized complexity for several parameters and obtain new tractable and hard cases. Interestingly, a full parameterized complexity classification would require resolving the parameterized complexity of the *covering radius* problem from coding theory. To the best of our knowledge, no useful relationship between coding theory and non-monotonic reasoning has previously been established, but such connections seemingly become important when asking more complex questions about solution spaces.

KEYWORDS: Propositional Abduction, Computational Complexity, Post's Framework, Fine-grained Reasoning

1 Introduction

The *propositional abduction* problem is a well-known form of non-monotonic reasoning with many applications in e.g. AI and knowledge representation (Dellsén 2024; Dai et al. 2019; Ignatiev et al. 2019; Yu et al. 2023; Hu et al. 2025; Kakas et al. 1992; Minsky 1974). Here, we are asked to explain a given *manifestation*. The explanation thus needs to logically entail the manifestation, and, to avoid trivial explanations, be consistent with the given knowledge base. For example, consider a medical diagnosis setting where a patient may develop a certain symptom depending on underlying conditions. We could then have propositional variables a (the patient has a weakened immune system), b (the

* Author names are stated in reverse alphabetical order.

patient has a bacterial infection), c (the patient has a viral infection), d (the patient is exposed to severe environmental stress), and m (the patient develops a high fever). The knowledge base could then encode the following medical rules:

- If the patient has a weakened immune system and a bacterial infection, then they develop a high fever.
- If the patient has a weakened immune system and a viral infection, then they develop a high fever.
- If the patient has a weakened immune system and is exposed to severe environmental stress, then they develop a high fever.
- The patient cannot simultaneously have a bacterial and a viral infection.

Formally, we could represent this as $\mathrm{KB} = \{a \wedge b \rightarrow m,\ a \wedge c \rightarrow m,\ a \wedge d \rightarrow m,\ \neg(b \wedge c)\}$. The manifestation is the observation that the patient has developed a high fever, i.e., $M = \{m\}$, and the set of hypotheses is $H = \{a, b, c, d\}$. Then, for example, $\{a, b\}$ (weakened immune system and bacterial infection) and $\{a, b, d\}$ (additionally, stress) are both possible explanations, but, unless there is further evidence, one may argue that $\{a, b\}$ is preferable to $\{a, b, d\}$ since it makes fewer assumptions. In this scenario, it seems desirable to consider *all* minimal explanations, revealing that the fever can be explained by a weakened immune system combined with exactly one of the mutually exclusive infections, or with environmental stress.

As might be expected, computing/counting all (minimal) explanations is computationally expensive (Hermann and Pichler 2010; Creignou et al. 2019), and even deciding existence of just a single explanation is Σ_2^P-complete (Eiter and Gottlob 1995). Nevertheless, there has been many attempts to reason about the set of solutions, e.g. by identifying *facets* (Alrabbaa et al. 2018; Fichte et al. 2025; Rusovac et al. 2024), and finding *diverse* solutions (induced by a given distance metric between solutions). For example, diverse solutions have been considered for *answer set programming* (Eiter et al. 2013), abduction (Schmidt et al. 2025) *constraint satisfaction problems* (Hebrard et al. 2005), satisfiability problems (Misra et al. 2024), and a wealth of graph problems (Baste et al. 2020; Fomin et al. 2020; 2021).

Inspired by the success of this approach and recent work on answer set programming (Böhl et al. 2023) we in this paper consider a related question: does there exist a set of explanations that *represent* all (minimal) explanations? We formulate this as a decision problem and then qualify our problem with a parameter $k \geq 0$, a set of explanations S, and want to know if every explanation is within distance k from at least one explanation in S. Thus, unless k is large, we expect the set S to correlate with diverse solutions.

We denote this problem by $\textsc{RepABD}$, and the corresponding problem for representing subset minimal explanations, by $\textsc{RepABD}_{\subseteq}$. If no assumptions are imposed on the knowledge base KB it is easy to show that both these problems are Π_2^P-complete, and we therefore attempt a more fine-grained picture of the complexity with restricted knowledge bases (e.g., whether it is in Horn, or in 2-CNF). We write $\textsc{RepABD}(\Gamma)$ ($\textsc{RepABD}_{\subseteq}(\Gamma)$) for this problem where Γ is a set of relations, and then require that the knowledge base is given by a conjunctive Γ-formula, i.e., each atom is of the form $R(x_1, \ldots, x_k)$ for $R \in \Gamma$ and variables $x_1, \ldots, x_k$. We formally introduce this problem in Section 3 together with a few useful definability notions. Then, for our main technical contributions, we (in

Section 4) classify the classical complexity of RepABD(Γ) and RepABD$_{\subseteq}$(Γ) for all possible choices of Γ.

Our classification reveals that the two problems have few tractable cases. For example, if Γ can express the "trivial" unary relation $\mathsf{t} = \{(0), (1)\}$ then RepABD(Γ) is at least coNP-hard. Surprisingly, RepABD$_{\subseteq}$(Γ) fares marginally better in comparison and we prove that it is in P if each relation is *strictly essentially positive*, or the dual case of being *strictly essentially negative*. However, it should be noted that tractability in this case stems from rather trivial reasons, and to extend the tractable fragments we (in Section 5 turn to *parameterized complexity*. Here, we relax polynomial time to additionally allow a factor $f(p)$ where $p \in \mathbb{N}$ is a parameter (e.g., $|H|$, $|M|$, k, or a graph parameter of KB) and $f\colon \mathbb{N} \to \mathbb{N}$ a computable function. A problem admitting such a running time is said to be *fixed-parameter tractable* (FPT). The parameterized complexity of abduction is well understood for many natural parameters (Mahmood et al. 2021; Fellows et al. 2012) and admits non-trivial FPT cases, so there is reason for a certain optimism. We consider many different parameters (k, $|H|$, $|M|$, and $|S|$) and establish FPT for $|H|$ if Γ is *Schaefer* and for $|S|$ if Γ is strictly essentially positive. Importantly, we complement this with many lower bounds that rule out FPT under widely believed conjectures in parameterized complexity.

Interestingly, by considering such "fine-grained" questions about the set of explanations, we are able to make connections to problems previously unconnected to non-monotonic reasoning. For example, one of our main sources of hardness stems from the *covering radius* problem (Frances and Litman 1997), and one of our main FPT results are based on a reduction to the *closest string problem* (Gramm et al. 2003). As we show, a complete parameterized complexity classification of RepABD(Γ) would simultaneously need to resolve the parameterized complexity of the covering radius problem (with parameter r). We discuss this and other questions in Section 6.

2 Preliminaries

We follow standard notions in computational complexity theory (Downey et al. 2013), and propositional logic. Below, we briefly state the most important notions.

2.1 Computational Complexity

Let Σ and Σ' be some finite alphabets. We call $I \in \Sigma^*$ an *instance* and $||I||$ denotes the size of I. A *decision problem* is some subset $L \subseteq \Sigma^*$. Recall that P and NP are the complexity classes of all deterministically and non-deterministically polynomial-time solvable decision problems. A polynomial-time many-to-one reduction ($\leq^{\mathrm{P}}_m$) from L to L' is a function $r : \Sigma^* \to \Sigma'^*$ such that for all $I \in \Sigma^*$ we have $I \in L$ if and only if $r(I) \in L'$ and r is computable in time $\mathcal{O}(||I|| \cdot c)$ for some constant c. In other words, a polynomial-time many-to-one reduction transforms instances of the decision problem L into instances of decision problem L' in polynomial time. We also need the Polynomial Hierarchy (PH) (Stockmeyer and Meyer 1973; Stockmeyer 1976; Wrathall 1976). In particular, $\Delta^{\mathrm{P}}_0 \coloneqq \Pi^{\mathrm{P}}_0 \coloneqq \Sigma^{\mathrm{P}}_0 \coloneqq \mathrm{P}$ and $\Delta^{\mathrm{P}}_{i+1} \coloneqq P^{\Sigma^p_i}$, $\Sigma^{\mathrm{P}}_{i+1} \coloneqq \mathrm{NP}^{\Sigma^{\mathrm{P}}_i}$, and $\Pi^{\mathrm{P}}_{i+1} \coloneqq \mathrm{coNP}^{\Sigma^{\mathrm{P}}_i}$ for $i > 0$ where C^D is the class C of decision problems augmented by an oracle for some complete problem in class D. For a decision problem $X \subseteq \Sigma^*$ we write $\overline{X}$ for the complement of X (i.e., flipping yes- and no- answers).

We also introduce *parameterized complexity* (following standard notation (Downey et al. 2013)). For some alphabet Σ a *parameterized problem* L is then just a subset of $\Sigma^* \times \mathbb{N}$. The most desirable running time in parameterized complexity is a complete decoupling of the parameter, and we say that L is *fixed-parameter tractable* (FPT) if there is an algorithm deciding if $(I, k) \in \Sigma^* \times \mathbb{N}$ is in L in time $f(k) \cdot |I|^c$, where $f \colon \mathbb{N} \to \mathbb{N}$ is a computable function and c is a constant (independent of I and k).

Parameterized complexity also contains a hardness theory, and to state it we first introduce a suitable class of reductions. For two parameterized problems $L, L' \subseteq \Sigma^* \times \mathbb{N}$, a mapping $P \colon \Sigma^* \times \mathbb{N} \to \Sigma^* \times \mathbb{N}$ is a *fixed-parameter (*FPT*) reduction from L to L'* if there exist computable functions $f, p : \mathbb{N} \to \mathbb{N}$ and a constant c such that the following conditions hold (for every $(I, k) \in \Sigma^* \times \mathbb{N}$):

- $(I, k) \in L$ if and only if $P(I, k) = (I', k') \in L'$,
- $k' \leq p(k)$, and
- $P(I, k)$ can be computed in $f(k) \cdot |I|^c$ time.

There are then several hard classes, the most notable being the *weft hierarchy* $\mathrm{W}[1] \subseteq \mathrm{W}[2] \subseteq \ldots$. We only need a fragment of this hierarchy and say that a problem is W[1]-*hard* if it admits an FPT-reduction from Independent Set (parameterized by the number of vertices in the independent set). Another important W[1]-hard problem not believed to be FPT is the *weighted SAT* problem, i.e., we are given a formula φ over n variables and the question is whether there is a model of Hamming weight $\geq k$. We let $\mathrm{WSAT}(2{-}\mathrm{CNF}^-)$ be the (still W[1]-hard) restriction of this problem to formulas in negative 2-CNF.

Beyond the W-hierarchy, we consider parameterized versions of classical complexity classes. For a classical complexity class C, we let para-C be the class of all parameterized problems $P \subseteq \Sigma^* \times \mathbb{N}$ such that there is a computable function $f \colon \mathbb{N} \to \Delta^*$ and a language $L \in C$ with $L \subseteq \Sigma^* \times \Delta^*$ such that for all $(x, k) \in \Sigma^* \times \mathbb{N}$ we have that $(x, k) \in P \Leftrightarrow (x, f(k)) \in L$. This allows us to, for example, speak of para-NP para-coNP, para-DP, and para-coDP.

Propositional Logic. A *literal* is a variable x or its negation $\neg x$. A *clause* is a disjunction of literals, often represented as a set. A clause of arity 1, i.e., either (x) or $(\neg x)$, is a *unit clause*. We work in a general setting where atoms can be expressions of the form $R(x_1, \ldots, x_r)$ for variables $x_1, \ldots, x_r$ and an r-ary relation $R \subseteq \{0, 1\}^r$. A function $f \colon \{x_1, \ldots, x_r\} \to \{0, 1\}$ is then said to satisfy an atom $R(x_1, \ldots, x_r)$ if $(f(x_1), \ldots, f(x_r)) \in R$. A (conjunctive) *propositional formula* φ is a conjunction of atoms and we write $\mathrm{var}(\varphi)$ for its set of variables. A mapping $\sigma \colon \mathrm{var}(\varphi) \mapsto \{0, 1\}$ is called an *assignment* to the variables of φ and a *model* of a formula φ is an assignment to $\mathrm{var}(\varphi)$ that satisfies φ. For two formulas ψ and φ, we write $\psi \models \varphi$ if every model of ψ also satisfies φ.

2.2 Restrictions of Constraint Languages

We work in a generalized setting where atoms can be formed by combining relations and variables. Then, a *constraint language* Γ is a set of Boolean relations, and a Γ-*formula* over some variables V is a propositional formula φ where $R \in \Gamma$ and $x_1, \ldots, x_r \in V$ for

co-clone	clauses/equation	name/indication
BR (II_2)	all clauses	all Boolean relations
II_0	at least one negative literal per clause	0-valid
IN_2	$\mathrm{NAE} = \{0,1\}^3 \setminus \{000, 111\}$	complementive
IN	$\mathrm{DUP} = \{0,1\}^3 \setminus \{101, 010\}$	complementive and 1- and 0-valid
IE_2	clauses with at most one positive literal	Horn
IV_2	clauses with at most one negative literal	dualHorn
IL_2	all affine clauses (all linear equations)	affine
IL	$(x_1 \oplus \cdots \oplus x_n = 0)$, n even	affine and 1- and 0-valid
ID_2	clauses of size 1 or 2	Krom, bijunctive, 2-CNF
ID	affine clauses of size 2	strict 2-affine
IM	$(x_1 \to x_2)$	implicative and 1- and 0-valid
IS_{12}	$(x_1), (\neg x_1 \vee \cdots \vee \neg x_n), n \geq 0, (x_1 = x_2)$	essentially negative (EN)
IS_{11}^k	$(x_1 \to x_2), (\neg x_1 \vee \cdots \vee \neg x_n), k \geq n \geq 0$	-
IS_1^k	$(\neg x_1 \vee \cdots \vee \neg x_n), k \geq n \geq 0, (x_1 = x_2)$	negative (N) of width k
IS_{02}	$(\neg x_1), (x_1 \vee \cdots \vee x_n), n \geq 0, (x_1 = x_2)$	essentially positive (EP)
IR_1	$(x_1), (x_1 = x_2)$	-
IR_0	$(\neg x_1), (x_1 = x_2)$	-
IR (IBF)	$(x_1 = x_2)$	-

Table 1: Overview of some co-clones and clause descriptions (Nordh and Zanuttini 2008).

each atom $R(x_1, \ldots, x_r)$. For a constraint language Γ, we write $\mathrm{SAT}(\Gamma)$ for the problem of deciding if a given Γ-formula admits at least one model. Usually, we do not distinguish between the relation or a clause defining the relation. For example, we simply write (x) for the unary relation $\{(1)\}$, $(\neg x)$ for $\{(0)\}$, $(x_1 \to x_2)$ or $(\neg x_1 \vee x_2)$ for $\{(0,0),(0,1),(1,1)\}$, and so on. The empty set $\emptyset$ is the (nullary) relation that is always false, we write $R_=$ for the equality relation $\{(0,0),(1,1)\}$ (but often written in infix form as $(x = y)$ instead of $R_=(x,y)$), and t for the trivial unary relation $\{(0),(1)\}$ that is always true.

For a constraint language Γ and $k \geq 1$, we often use the notation k-Γ for the set of relations/clauses of arity at most k (e.g., 2-CNF contains all clauses of arity 1 and 2). Additionally, for a language Γ we, let (1) $\Gamma^{-} = \Gamma \setminus \{(x), (\neg x)\}$ be Γ without the two unit clauses, and (2) $\Gamma^+ = \Gamma \cup \{(x), (\neg x)\}$ be Γ expanded with the two unit clauses. A language Γ is *b-valid* for $b \in \{0,1\}$, if $(b, \ldots, b) \in R$ for each $R \in \Gamma$. We introduce the most important constraint languages for the purpose of this paper in Table 1. We also write EP^- for the set of *strictly essentially positive clauses* $\mathrm{EP} \setminus \{(x = y)\}$ and $\mathrm{EN}^- = \mathrm{EN} \setminus \{(x = y)\}$ for the set of *strictly essentially negative clauses.* To avoid doing an exhaustive case analysis of *all* possible constraint languages we introduce a useful closure property on relations. Say that an r-ary relation R has a *primitive positive definition* (pp-definition) over Γ if

$$R(x_1, \ldots, x_r) := \exists y_1, \ldots, y_n \,.\, \varphi(x_1, \ldots, x_r, y_1, \ldots, y_n)$$

where φ is a $(\Gamma \cup \{\emptyset, \mathsf{t}, (x = y)\})$-formula. Thus, put otherwise, R can be defined as the set of models of $\exists y_1, \ldots, y_n \,.\, \varphi(x_1, \ldots, x_r, y_1, \ldots, y_n)$ with respect to the free variables $x_1, \ldots, x_r$.

Definition 2.1. *For a constraint language Γ we let $\langle \Gamma \rangle$ be the smallest set of relations containing Γ and where $R \in \Gamma$ for any pp-definable relation R over Γ.*

The set Γ is in this context said to be a *base*, and $\langle \Gamma \rangle$ is sometimes called a *relational clone*, or a *co-clone*. For details, we refer to the work by Böhler et al. (2005). We note that the complexity of SAT(Γ) is completely determined due to Schaefer's famous dichotomy result: it is polynomial if Γ is 0- or 1-valid or *Schaefer* (that is, Γ is Horn, or dualHorn, or affine, or 2-CNF) and NP-complete otherwise (Schaefer 1978).

2.3 Propositional Abduction

Let Γ be a constraint language, for example, a set of clauses. An instance I of the *positive propositional* abduction problem over Γ, ABD(Γ) for short, is a tuple $I = (\mathrm{KB}, H, M)$ with KB being a Γ-formula over a finite set of Boolean variables called the *knowledge base* (or *theory*), $H \subseteq \mathrm{var}(\mathrm{KB})$ called *hypotheses*, $M \subseteq \mathrm{var}(\mathrm{KB})$ called *manifestations*. Since we have defined a Γ-formula as a conjunctive formula with atoms from Γ we sometimes take the liberty of viewing the knowledge base as a set rather than as a formula. A *positive explanation* E, *explanation* for short, is a subset $E \subseteq H$ such that (i) $\mathrm{KB} \wedge E$ is satisfiable and (ii) $\mathrm{KB} \wedge E \models M$. An explanation E is *(subset-)minimal* if no other set $E' \subsetneq E$ is an explanation of I.

The problem ABD(Γ) asks whether there is an explanation, which in the decision context is the same as asking whether there is a minimal explanation. If Γ is arbitrary, we omit Γ from the problem and write ABD. Note that the complexity of ABD(Γ) is completely determined (Nordh and Zanuttini 2008). See Figure 1 for a visualization.

We write $\mathcal{E}(I)$ to refer to the set of all explanations and $\mathcal{E}_M(I)$ for the set of all subset-minimal explanations.

Example 2.2. *Consider the abduction example from Section 1 where $I = (KB, H, M)$ with $KB = \{a \wedge b \rightarrow m,\ a \wedge c \rightarrow m,\ a \wedge d \rightarrow m,\ \neg(b \wedge c)\}$, manifestation m, and the hypotheses $\{a, b, c, d\}$. Then $\mathcal{E}_M(I) = \{\{a, b\}, \{a, c\}, \{a, d\}\}$ and $\mathcal{E}(I) = \{\{a, b\}, \{a, c\}, \{a, d\}, \{a, b, d\}, \{a, c, d\}\}$.*

3 The Representative Explanation Problem

We begin the technical part of the paper by formally introducing our problem as well as the simplifying algebraic machinery. Let $I = (\mathrm{KB}, H, M)$ be an ABD instance. For a set $E \subseteq H$, define $S_k(E)$ as the set of all explanations within Hamming distance k, i.e.,

$$S_k(E) = \{E_1 \in \mathcal{E}(I) \mid d(E, E_1) \leq k\},$$

where $d(\cdot, \cdot)$ denotes the symmetric difference, i.e.,

$$\begin{aligned} d(E_1, E_2) &= |\, E_1 \triangle E_2 \,| = |\, (E_1 \cup E_2) \setminus (E_1 \cap E_2) \,| \\ &= |\, \{x \in H \mid x \in E_1 \text{ and } x \notin E_2 \text{ or } x \notin E_1 \text{ and } x \in E_2\} \,|. \end{aligned}$$

For a fixed k, an explanation $E \in \mathcal{E}(I)$ is called *k-representative*, if $|S_k(E)|$ is maximal. That is, such an E represents maximally many explanations. We extend this notion to sets of explanations as follows: for a fixed k, a set of explanations $S \subseteq \mathcal{E}(I)$ is called *k-representative*, if

$$\mathcal{E}(I) = \bigcup_{E \in S} S_k(E).$$

We then consider the following problem where the task is to verify if a set of explanations is k-representative or not.

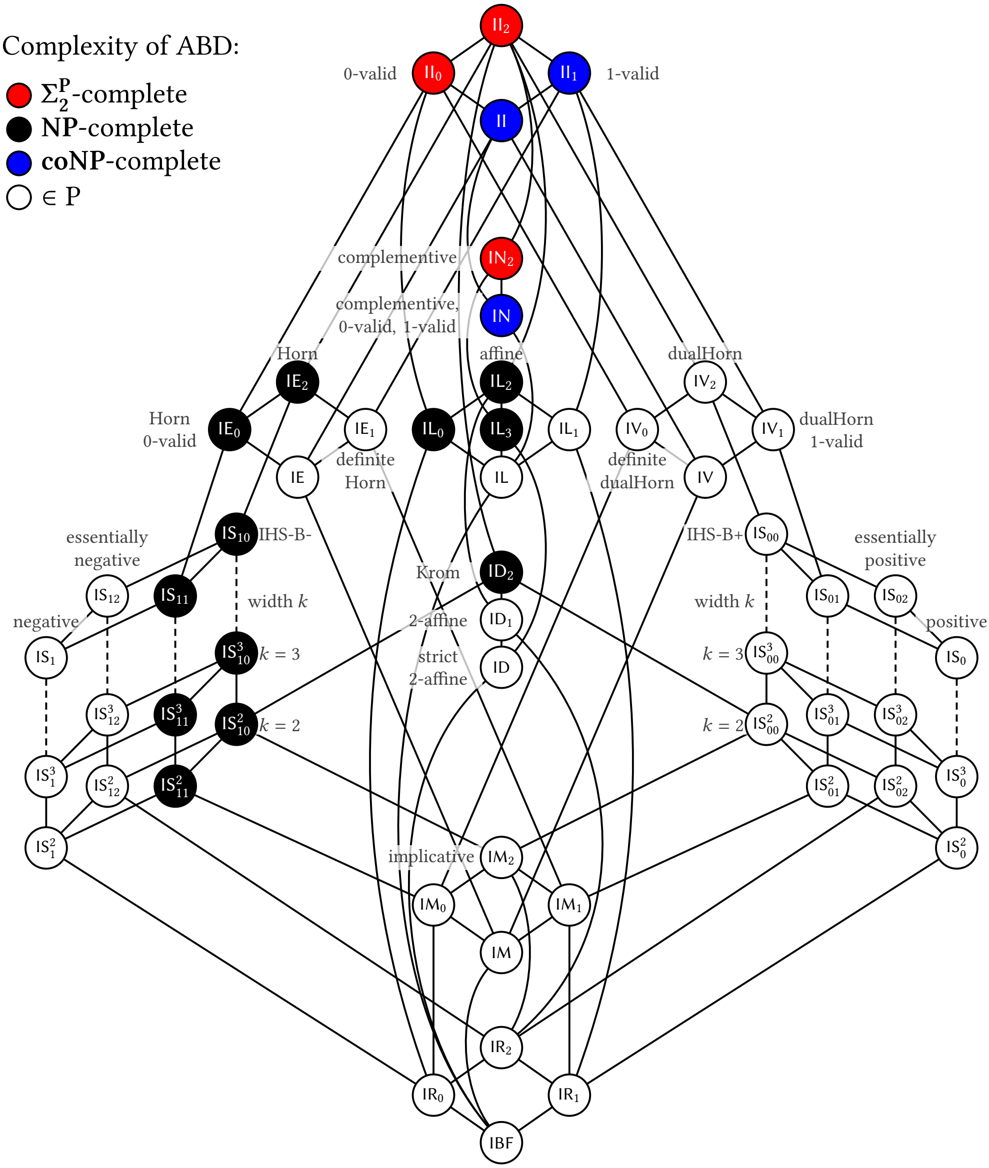

Figure 1: Complexity of propositional abduction from (Nordh and Zanuttini 2008), illustrated on Post's lattice.

REPABD(Γ)
Given: An ABD(Γ) instance $I = (\mathrm{KB}, H, M)$, a set $S \subseteq \mathcal{E}(I)$, and $k \geq 1$.
Task: Is S k-representative for I?

We also consider REPABD(Γ) restricted to $\subseteq$-minimal explanations (given now $S \subseteq \mathcal{E}_M(I)$ check if $\mathcal{E}_M(I) = \bigcup_{E \in S} S_k^M(E)$, where $S_k^M(E) = \{E_1 \in \mathcal{E}_M(I) \mid d(E, E_1) \leq k\}$) and write REPABD$_\subseteq$($\Gamma$) for this variant.

Example 3.1. *We continue with Example 2.2 where we had* $KB = \{a \wedge b \to m,\ a \wedge c \to m,\ a \wedge d \to m,\ \neg(b \wedge c)\}$*,* $M = \{m\}$*,* $H = \{a, b, c, d\}$*, and* $\mathcal{E}(I) =$

$\{\{a,b\},\{a,c\},\{a,d\},\{a,b,d\},\{a,c,d\}\}$. *Then* $S_1 = \{\{a,b\}\}$ *is not 2-representative, since* $d(\{a,b\},\{a,c,d\}) = 3$. *But* $S_2 = \{\{a,d\}\}$ *is 2-representative, since* $d(\{a,d\},E) \leq 2$ *for all* $E \in \mathcal{E}(I)$*, as is easily verified.*

Example 3.2. *We also consider an example that illustrates the difference between minimal explanations. Here,* $KB = \{a \wedge b \rightarrow m,\ a \wedge c \rightarrow m,\ b \wedge d \rightarrow m\}$*, with manifestation* m *and hypothesis* $\{a,b,c,d\}$.
Then $\mathcal{E}_M(I) = \{\{a,b\},\{a,c\},\{b,d\}\}$, $S_1 = \{\{a,c\}\}$ *is not 2-representative, since* $d(\{a,c\},\{b,d\}) = 4$*, but* $S_2 = \{\{a,b\}\}$ *is 2-representative, since* $d(\{a,b\},\{a,c\}) = 2$ *and* $d(\{a,b\},\{b,d\}) = 2$.

Before turning to the complexity of $\textsc{RepABD}(\Gamma)$ and $\textsc{RepABD}_{\subseteq}(\Gamma)$ we show how to apply the algebraic approach — with a minor modification. First, we say that a pp-definition $R(x_1,\ldots,x_r) := \exists y_1,\ldots,y_n \,.\, \varphi(x_1,\ldots,x_r,y_1,\ldots,y_n)$ is *equality-free* if each atom in φ is from $\Gamma \cup \{\emptyset\}$, i.e., we do not allow (1) the equality relation, or (2) the full relation $\mathtt{t}$. We let $\langle\Gamma\rangle_{\neq}$ be the smallest set of relations containing Γ closed under such definitions. We have the following basic characterization (where a relation R is said to be constant if $|R| = 1$).

Before stating and proving the lemma we need a few additional preliminaries. For a tuple $t = (x_1,\ldots,x_i,\ldots,x_n) \in \{0,1\}^n$ and $i \in [n] = \{1,\ldots,n\}$ we write $t[i] = x_i$ for the ith component. For $R \subseteq \{0,1\}^n$ of arity n we say that $i \in [n]$ is *fictitious* if $(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_n) \in R$ if and only if $(x_1,\ldots,x_{i-1},\bar{x}_i,x_{i+1},\ldots,x_n) \in R$, and it is *redundant* if there exists $j \in [n]$, $j \neq i$, such that $t[i] = t[j]$ for any $t \in R$. Furthermore, say that R is *irredundant* if it has no redundant coordinates, and that it is *prime* if it is irredundant and has no fictive coordinates. We can now relate these notions to efpp-definability as follows.

Lemma 3.3. *($\star$) Let* Γ *be a set of Boolean relations and let* $R \in \langle\Gamma\rangle$*. If*

1. R *is prime then* $R \in \langle\Gamma\rangle_{\neq}$,
2. R *is irredundant then* $R \in \langle\Gamma \cup \{\mathtt{t}\}\rangle_{\neq}$*, and*
3. *otherwise* $R \in \langle\Gamma \cup \{R_{=}\}\rangle_{\neq}$.

Proof. We consider each case in turn. First, if R is prime, then any pp-definition $R(x_1,\ldots,x_n) := \exists y_1,\ldots,y_m.\varphi(x_1,\ldots,x_n,y_1,\ldots,y_m)$ can be rewritten into an efpp-definition by considering the following cases for an atom in φ over $\mathtt{t}$ or $R_{=}$, where $i \neq j$ are distinct indices over $[n]$ or $[m]$.

1. $\mathtt{t}(x_i)$ can be removed (and x_i must then occur elsewhere, since otherwise it would witness that R has a fictitious argument),
2. $\mathtt{t}(y_i)$ can be removed, and if y_i does not occur anywhere else the existential quantifier $\exists y_i$ can be dropped,
3. $R_{=}(x_i,x_j)$ cannot occur (this would witness that R is not prime),
4. $R_{=}(y_i,y_j)$ can be removed by taking all such existentially quantified variables and order them into equivalence classes, picking one representative for each set and replace all other occurrences with this representative, and, last, by removing all such equality constraints. For a more formal treatment, cf. (Lagerkvist 2020).

The second case can be proved with a similar case analysis, and the third case follows trivially. □

This in turn leads to the following classification of efpp-closed sets.

Lemma 3.4. *($\star$) Let Γ be a set of Boolean relations. If*

1. *if $R_= \in \langle\Gamma\rangle_{\neq}$ then $\mathtt{t} \in \langle\Gamma\rangle_{\neq}$,*
2. *if $R_= \in \langle\Gamma\rangle_{\neq}$ then $\langle\Gamma\rangle_{\neq} = \langle\Gamma\rangle$,*
3. *if Γ contains a non-constant relation then $\langle\Gamma\rangle_{\neq} = \langle\Gamma \cup \{\mathtt{t}\}\rangle_{\neq}$, and*
4. *if $\mathtt{t} \notin \langle\Gamma\rangle_{\neq}$ then $\langle\Gamma\rangle_{\neq} \subsetneq \langle\Gamma \cup \{\mathtt{t}\}\rangle_{\neq} \subsetneq \langle\Gamma \cup \{R_=\}\rangle_{\neq} = \langle\Gamma\rangle$, and $\langle\Delta\rangle_{\neq} \in \{\langle\Gamma\rangle_{\neq}, \langle\Gamma \cup \mathtt{t}\rangle_{\neq}, \langle\Gamma \cup \{R_=\}\rangle_{\neq}\}$ for any Δ such that $\langle\Gamma\rangle = \langle\Delta\rangle$.*

Proof. The first claim follows immediately since we can define $\mathtt{t}(x)$ via the definition $(x = x)$. For the second claim: if we can efpp-define $R_=$ then we (by the first claim) can also efpp-define $\mathtt{t}$, and any pp-definition (possibly using $R_=$- or $\mathtt{t}$-constraints) can be converted into a suitable efpp-definition. For the third claim, let $R \in \Gamma$ be a non-constant relation, of arity, say r, and let $1 \leq i \leq r$ be an argument such that $|\{x_i \mid (x_1, \ldots, x_i, \ldots, x_r) \in R\}| = 2$. It follows that $\mathtt{t}(x) := \exists x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_r . R(x_1, \ldots, x_{i_1}, x, x_{i+1}, \ldots, x_r)$.

For the fourth and last claim, $\langle\Gamma\rangle_{\neq} \subsetneq \langle\Gamma \cup \{\mathtt{t}\}\rangle_{\neq}$ follows from the assumption that $\mathtt{t} \notin \langle\Gamma\rangle_{\neq}$. For the second inclusion $\langle\Gamma \cup \{\mathtt{t}\}\rangle_{\neq} \subsetneq \langle\Gamma \cup \{R_=\}\rangle_{\neq}\}$, observe that $R_=$ has no fictitious argument, and, hence, if $R_= \in \langle\Gamma \cup \{\mathtt{t}\}\rangle_{\neq}$ then $R_= \in \langle\Gamma\rangle_{\neq}$ (by Lemma 3.3). But then (as established in the first item of this lemma) $\langle\Gamma\rangle_{\neq} = \langle\Gamma\rangle$ which contradicts the assumption that $\mathtt{t} \notin \langle\Gamma\rangle$. Hence, $R_= \notin \langle\Gamma \cup \{\mathtt{t}\}\rangle_{\neq}$ and the inclusion $\langle\Gamma \cup \{\mathtt{t}\}\rangle_{\neq} \subsetneq \langle\Gamma \cup \{R_=\}\rangle_{\neq}$ must be proper. Now, consider a Δ such that $\langle\Delta\rangle = \langle\Gamma\rangle$. The claim that $\langle\Delta\rangle_{\neq} \in \{\langle\Gamma\rangle_{\neq}, \langle\Gamma \cup \mathtt{t}\rangle_{\neq}, \langle\Gamma \cup \{R_=\}\rangle_{\neq}\}$ then follows through a similar case analysis: if every relation is prime, then $\langle\Delta\rangle_{\neq} = \langle\Gamma\rangle_{\neq}$, if every relation is irredundant then $\langle\Delta\rangle_{\neq} \in \{\langle\Gamma\rangle_{\neq}, \langle\Gamma \cup \{\mathtt{t}\}\rangle_{\neq}\}$, and if these two cases do not apply then $\langle\Delta\rangle_{\neq} = \langle\Gamma\rangle$. □

We remark that disallowing equality is a fairly standard assumption for certain problems (Creignou et al. 2014; Mahmood et al. 2021; 2023; Creignou et al. 2023) but $\mathtt{t}$ is normally such a harmless relation that it is not explicitly acknowledged in definitions. However, we will later see that their presence *do* make a difference for the RepABD(Γ) problem in the sense that there are languages Γ such that RepABD(Γ) is in P but RepABD($\Gamma \cup \{\mathtt{t}\}$) is intractable. With this in mind we obtain the following basic reducibility result.

Lemma 3.5. *($\star$) Let Γ and Γ' be two constraint languages. If $\Gamma' \subseteq \langle\Gamma\rangle_{\neq}$, then* RepABD($\Gamma'$) $\leq^{\mathrm{P}}_m$ RepABD(Γ) *and* RepABD$_{\subseteq}$(Γ') $\leq^{\mathrm{P}}_m$ RepABD$_{\subseteq}$(Γ).

Proof (Idea). We omit details, since the construction is exactly the same as in previous work (Nordh and Zanuttini 2008, Lemma 22), but the basic idea is simply to replace each relation by the set of constraints prescribed by the efpp-definition, and introducing fresh variables (kept outside the hypothesis) for any existentially quantified variables. This exactly preserves the set of (minimal) explanations since we, importantly, do not need to identify variables occurring in equality constraints. Therefore, the answer to the question whether a given set S is k-representative does not change. □

We further need the following expressiveness result.

Lemma 3.6. *(⋆) Let Γ be a constraint language. If $\Gamma \not\subseteq \langle \mathrm{EP}^- \rangle_{\neq}$ and $\Gamma \not\subseteq \langle \mathrm{EN}^- \rangle_{\neq}$ then $(x = y) \in \langle \Gamma \rangle_{\neq}$ and $\langle \Gamma \rangle = \langle \Gamma \rangle_{\neq}$.*

Proof. This follows immediately from the literature, namely from Proposition 3.1, Proposition 3.2, and Lemma 3.3 in (Mahmood et al. 2023). □

4 Classical Complexity

We begin by analyzing the "classical" complexity of $\textsc{RepABD}(\Gamma)$, i.e., whether it is in P, or in an intractable class, and thus study the complexity of the problem up to polynomial-time reductions. We first observe a straightforward upper bound.

Lemma 4.1. *(⋆) $\textsc{RepABD}(\Gamma)$ is in Π_2^P.*

Proof. The following non-deterministic algorithm shows that $\overline{\textsc{RepABD}(\Gamma)}$ (recall that this is the complement of $\textsc{RepABD}(\Gamma)$) is in Σ_2^P. We are given an instance (KB, H, M), a set $S \subseteq \mathcal{E}(I)$, and $k \geq 1$. We then guess an $E' \subseteq H$, and verify that
(1) $E' \in \mathcal{E}(I)$ — feasible with an NP- and a coNP-oracle (first check if $\mathrm{KB} \land E'$ is satisfiable, and then if $\mathrm{KB} \land E' \models M$)
(2) $E' \notin \bigcup_{E \in S} S_k(E)$ — feasible with a coNP-oracle (since the complement question is in NP: guess an $E \in S$ and verify that $d(E, E') \leq k$). □

Recall that for a language Γ we let Γ^+ be Γ expanded with the two constant Boolean relations. We have the following useful conditional upper bound.

Lemma 4.2. *(⋆) If $\mathrm{SAT}(\Gamma^+) \in \mathrm{P}$ then $\textsc{RepABD}(\Gamma) \in \mathrm{coNP}$, for any constraint language Γ.*

Proof. Let $I = (\mathrm{KB}, H, M, S, k)$ be an instance of $\textsc{RepABD}(\Gamma)$. We guess a candidate explanation $E \subseteq H$ and can verify that it is an explanation in P. We can then, still in polynomial time (with respect to I) verify that $d(E, E') > k$ for each $E' \in S$. □

The following lemma lets us inherit numerous hardness results from ABD.

Lemma 4.3. *(⋆) Let Γ be a constraint language. Then $\overline{\mathrm{ABD}(\Gamma)} \leq_m^{\mathrm{P}} \textsc{RepABD}(\Gamma)$.*

Proof. Given an instance (KB, H, M) of $\mathrm{ABD}(\Gamma)$, we map it to $(\mathrm{KB}, H, M, \emptyset, 1)$ of $\textsc{RepABD}(\Gamma)$.

One easily verifies that with $S = \emptyset$ it holds that

$$\bigcup_{E \in S} S_k(E) = \bigcup_{E \in \emptyset} S_k(E) = \emptyset.$$

Now, let (KB, H, M) be a positive instance of $\mathrm{ABD}(\Gamma)$, and $E \subseteq H$ an explanation. This E is not represented by $S = \emptyset$, as shown above. Thus, $(\mathrm{KB}, H, M, \emptyset, 1)$ is a negative instance of $\textsc{RepABD}$. Conversely, let (KB, H, M) be a negative instance of $\mathrm{ABD}(\Gamma)$. That is, $\mathcal{E}(I) = \emptyset$. Since $\bigcup_{E \in S} S_k(E) = \emptyset$, it holds that $\bigcup_{E \in S} S_k(E) = \mathcal{E}(I)$. Thus, $(\mathrm{KB}, H, M, \emptyset, 1)$ is a positive instance of $\textsc{RepABD}(\Gamma)$. □

However, the following lemma proves that $\textsc{RepABD}$ is generally harder than ABD since it establishes coNP-hardness for many fragments where ABD is tractable. Recall that $\mathsf{t} = \{(0), (1)\}$ and that $T = \{(1)\}$. In the reduction we use the NP-complete *covering* *radius problem* (Frances and Litman 1997), where an instance is given by a binary code $C \subseteq \{0,1\}^n$ and an integer r, and the question is whether there is a vector $v \in \{0,1\}^n$ that has Hamming distance greater than r from every element in C (thus the *covering radius* of C is greater than r).

Lemma 4.4. RepABD(Γ) *is coNP-hard for any* Γ *such that* $\mathsf{t} \in \langle\Gamma\rangle_{\neq}$ *or* $T \in \langle\Gamma\rangle_{\neq}$.

Proof. We give a reduction from the covering radius problem to the complement of RepABD(Γ) (i.e., proving coNP-hardness). Let an instance of the covering radius problem be given by (C, r). Let $x_1, \ldots, x_n$ denote fresh variables, and for a $v \in \{0,1\}^n$ let $v[i]$ denote the i^{th} coordinate of v. For an $E \subseteq \{x_1, \ldots, x_n\}$ let E_{vec} denote the characteristic vector of E, that is, $E_{vec}[i] = 1$ if $x_i \in E$, and $E_{vec}[i] = 0$ if $x_i \notin E$. For a $v \in \{0,1\}^n$ let v_{set} denote the set representation of v, that is, $v_{set} = \{x_i \mid v[i] = 1\}$. In the following, the constraint D denotes either t or T (the construction is valid for either relation). We map (C, r) to $(\mathrm{KB}, H, M, S, k)$ as follows.

$$\begin{aligned}
\mathrm{KB} &= \bigwedge_{i=1}^{n} D(x_i) \\
H &= \{x_1, \ldots, x_n\} \\
M &= \emptyset \\
S &= \{E \subseteq H \mid E_{vec} \in C\} \\
k &= r
\end{aligned}$$

We first observe that the constraints $D(x_i)$ in KB together with $M = \emptyset$ imply that $\mathcal{E}((\mathrm{KB}, H, M)) = 2^H$. Now, let (C, r) be a positive instance. That is, there is a vector $v \in \{0,1\}^n$ that has hamming distance greater than r from every element in C. One easily verifies that $\mathrm{KB} \wedge v_{set}$ is consistent and that $\mathrm{KB} \wedge v_{set} \models M$. Therefore, v_{set} is an explanation for (KB, H, M). However, v_{set} is of greater distance than $k = r$ from every element in C, that is, v_{set} is not k-represented by S. Thus $(\mathrm{KB}, H, M, S, k)$ is a negative instance of RepABD.

Conversely, let (C, r) be a negative instance. That is, every vector $v \in \{0,1\}^n$ has hamming distance equal or less than r from some element in C. Consequently, any $E \subseteq H$ (which by construction is a valid explanation), is k-represented by S. Thus $(\mathrm{KB}, H, M, S, k)$ is a positive instance of RepABD.

□

We obtain the following complexity classification of RepABD(Γ).

Theorem 4.5. *Let* Γ *be a constraint language. Then* RepABD(Γ) *is*

1. Π_2^P*-complete if* $\mathsf{IN}_2 \subseteq \langle\Gamma\rangle \subseteq \mathsf{II}_2$ *or* $\mathsf{II}_0 \subseteq \langle\Gamma\rangle \subseteq \mathsf{II}_2$,
2. coNP*-hard and* NP*-hard if* $\mathsf{IN} \subseteq \langle\Gamma\rangle$,
3. coNP*-complete if* $C \subseteq \langle\Gamma\rangle \subseteq D$ *for* $C \in \{\mathsf{IS}_1^2, \mathsf{IM}, \mathsf{IR}_1, \mathsf{ID}, \mathsf{IL}\}$ *and* $D \in \{\mathsf{IE}_2, \mathsf{IL}_2, \mathsf{IV}_2, \mathsf{ID}_2\}$,
4. coNP*-complete if* $\mathsf{IBF} \subseteq \langle\Gamma\rangle \subseteq \mathsf{IR}_0$ *and* $\mathsf{t} \in \langle\Gamma\rangle_{\neq}$, *and*
5. $\in$ P *otherwise*

Proof. First, assume that $\mathsf{IN}_2 \subseteq \langle\Gamma\rangle \subseteq \mathsf{II}_2$ or that $\mathsf{II}_0 \subseteq \langle\Gamma\rangle \subseteq \mathsf{II}_2$. From Lemma 4.1 we know that $\text{RepABD}(\Gamma) \in \Pi_2^P$, and from Figure c we know that $\text{ABD}(\Gamma)$ is Σ_2^P-complete, which in combination with Lemma 4.3 gives the desired Π_2^P-hardness. Second, assume that $\mathsf{IN} \subseteq \langle\Gamma\rangle$. Then $\text{ABD}(\Gamma)$ is coNP-hard, and Lemma 4.3 therefore gives NP-hardness for $\text{RepABD}(\Gamma)$. For the coNP-hardness claim we first observe that Γ must contain a non-constant relation, which from Lemma 3.4 implies that $\mathsf{t} \in \langle\Gamma\rangle_{\neq}$, and we finally get coNP-hardness from Lemma 4.4. Third, assume that $C \subseteq \langle\Gamma\rangle \subseteq D$ for $C \in \{\mathsf{IS}_1^2, \mathsf{IM}, \mathsf{IR}_1, \mathsf{ID}, \mathsf{IL}\}$ and $D \in \{\mathsf{IE}_2, \mathsf{IL}_2, \mathsf{IV}_2, \mathsf{ID}_2\}$. Then $\text{SAT}(\Gamma^+) \in \text{P}$ (Schaefer 1978) and we therefore (via Lemma 4.2) conclude that $\text{RepABD}(\Gamma)$ is in coNP. For hardness, first assume that $\langle\Gamma\rangle \neq \mathsf{IR}_1$. Then we similarly to the above case observe that Γ must contain a non-constant relation, and we apply Lemma 4.4 for the desired result. For $\langle\Gamma\rangle = \mathsf{IR}_1$ we observe that any pp-definition of T (possibly using $R_=$ and t) can be simplified into an equivalent pp-definition still defining T, and we then apply Lemma 4.4. Fourth, the case when $\mathsf{IBF} \subseteq \langle\Gamma\rangle \subseteq \mathsf{IR}_0$ and $\mathsf{t} \in \langle\Gamma\rangle_{\neq}$ also follows from Lemma 4.4.

It can be verified that the only remaining case is when $\mathsf{t} \notin \langle\Gamma\rangle_{\neq}$ and if $\langle\Gamma\rangle = \mathsf{IBF}$ or $\langle\Gamma\rangle = \mathsf{IR}_0$. We assume that $\langle\Gamma\rangle = \mathsf{IR}_0$ since it subsumes the other case. We now apply Lemma 3.4 and conclude that $\langle\Gamma\rangle_{\neq} = \langle\{\emptyset, F\}\rangle_{\neq}$. In this case $\text{RepABD}(\Gamma)$ can be solved in polynomial time. For an instance (KB, H, M, S, k) we first check if $\text{KB} = \emptyset$. If so, $H = M = \emptyset$, and the only possible explanation is $E = \emptyset$, and we answer yes if $S = \{\emptyset\}$ and no if $S = \emptyset$. If, on the other hand, $\text{KB} \neq \emptyset$ then we first check if there exists a constraint $\emptyset(x)$ in KB. If so, then $\mathcal{E}((\text{KB}, H, M)) = \emptyset$ since KB is not satisfiable, and we can answer yes or no via a simple case analysis. Otherwise, we must have $F(x)$ for each atom in KB, but then we either have no explanation, or $E = \emptyset = M$ as the only possible explanation, and we can easily answer yes or no. □

See Figure 2 for a visualization of this classification.

We continue by considering the subsetminimal variant $\text{RepABD}_{\subseteq}$ of RepABD, which, surprisingly, turns out to be easier for certain fragments. We obtain tractability for essentially negative and essentially positive fragments, as long as the equality constraint can not be expressed. Recall that EP^- and EN^- denote the set of strictly essentially positive, respectively negative, clauses.

Lemma 4.6. $\text{RepABD}_{\subseteq}(\Gamma) \in \text{P}$ *if* $\Gamma \subseteq \left\langle \text{EP}^- \right\rangle_{\neq}$ *or* $\Gamma \subseteq \left\langle \text{EN}^- \right\rangle_{\neq}$.

Proof. Let (KB, H, M, S, k) be an instance of $\text{RepABD}_{\subseteq}(\Gamma)$. We assume that each $R \in \Gamma$ is represented by a conjunctive formula over EP^- (or EN^-). We can then without loss of generality assume that each atom in KB is from EP^- (respectively, from EN^-).

In EN^-, variables in M can either appear as positive singletons, or as negative literals in a negative clause. If the abduction problem has an explanation, then there exists a unique subset-minimal explanation $E = \{M \cap H\} \setminus \{m \mid (m) \in \text{KB}\}$. This is because every $m \in M$ is either in H and entails itself, or m is already true by being a positive singleton, or there is no solution.

Similarly, in EP^-, any $m \in M$ is either a singleton that is already always true, or $m \in H$ entails itself, or the problem has no solution. Again, there can only be a single subset-minimal explanation $E = \{M \cap H\} \setminus \{m \mid (m) \in \text{KB}\}$.

It suffices to check if $\forall c_i \in S : d(E, c_i) > k$, which can be done in linear time. This concludes the proof. □

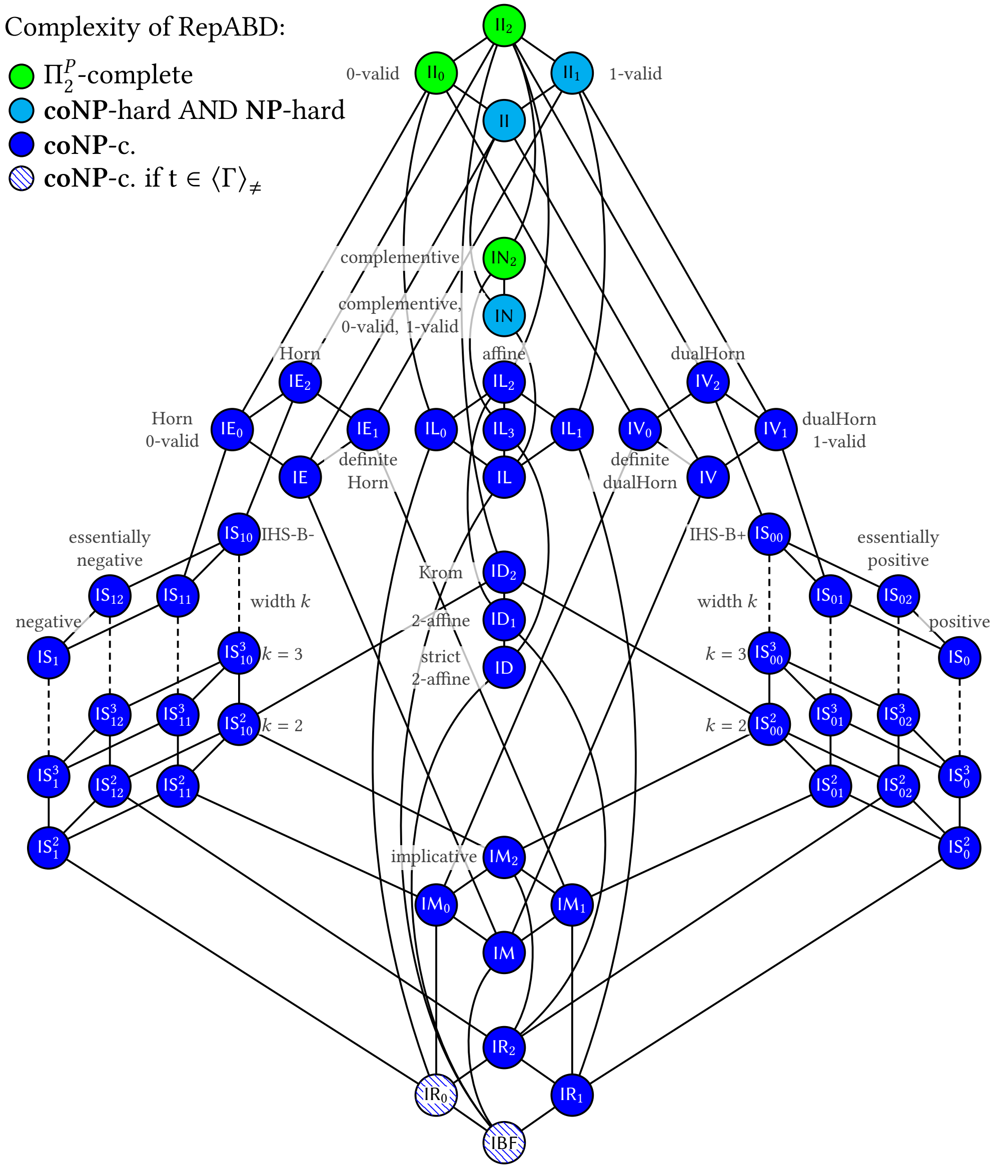

Figure 2: Illustration of complexity via Post's lattice.

If equality can be expressed, however, we obtain coNP-hardness.

Lemma 4.7. *($\star$)* $\textsc{RepABD}_{\subseteq}(\Gamma)$ *is coNP-hard for any language* Γ *such that* $R_= \in \langle\Gamma\rangle_{\neq}$.

Proof. We reduce from the covering radius problem. Let (C, r) be an arbitrary instance of the covering radius problem, where $C \subseteq \{0,1\}^n$ is a set of codewords and $r \leq n$ the radius.

We construct a $\textsc{RepABD}_{\subseteq}(\Gamma)$ instance $(\mathrm{KB}, H, M, S, k)$ as follows. We have a set of code words C that can be interpreted as truth assignments on n variables. For every $x \in \{1, \ldots, n\}$ we introduce two fresh variables x_1, x_2 and add the following rules to KB : $(x_1 = x_2)$ and $(x_2 = M_x)$, and we add x_1 and x_2 to H and M_x to M. We reduce the set C to a new set C' as such: for every code word $c \in C$ we make a code word $c' \in C' \subseteq \{0,1\}^{2n}$: for every coordinate $1 \leq i \leq n$ where $c[i] = 1$ we set two coordinates $c'[i] = 1$ and $c'[i+n] = 0$ in c', and the reverse if $c[i] = 0$ (note that these are coordinates of c' being set to values not to be confused with the rules in KB, we do not need constants in Γ). We put $S = \{(c_i)_{set} \mid c_i \in C'\}$ ($(c_i)_{\text{set}}$ is the set representation of the vector c_i as in Lemma 4.4). Finally, we put $k = 2r$. This concludes the reduction. For correctness: we have simply duplicated all coordinates in all the codewords, in a manner that forces one and only one of the two copies to be part of a subset minimal explanation. If neither is part of the explanation, the variable $M_x \in M$ cannot be entailed, if both are chosen, then it is not a subset-minimal explanation. These subset-minimal explanations are in a one-to-one correspondence with the codewords from $\{0,1\}^n$. A difference in a variable in between codewords in $\{0,1\}^n$ will result in a difference of weight 2 in our reduction, thus $k = 2r$. □

Furthermore, we also obtain all hardness results from ABD again, analogously to Lemma 4.3.

Lemma 4.8. *($\star$) Let Γ be a constraint language. Then* $\overline{\text{ABD}(\Gamma)} \leq_m^{\text{P}} \text{RepABD}_{\subseteq}(\Gamma)$.

Proof. Given an instance (KB, H, M) of $\text{ABD}(\Gamma)$, we map it to $(\text{KB}, H, M, \emptyset, 1)$ of $\text{RepABD}_{\subseteq}(\Gamma)$.

One easily verifies that with $S = \emptyset$ it holds that

$$\bigcup_{E \in S} S_k^M(E) = \bigcup_{E \in \emptyset} S_k^M(E) = \emptyset.$$

Now, let (KB, H, M) be a positive instance of $\text{ABD}(\Gamma)$, and $E \subseteq H$ a (subset-minimal) explanation. This E is not represented by $S = \emptyset$, as shown above. Thus, $(\text{KB}, H, M, \emptyset, 1)$ is a negative instance of $\text{RepABD}_{\subseteq}$. Conversely, let (KB, H, M) be a negative instance of $\text{ABD}(\Gamma)$. That is, $\mathcal{E}(I) = \mathcal{E}_M(I) = \emptyset$. Since $\bigcup_{E \in S} S_k^M(E) = \emptyset$, it holds that $\bigcup_{E \in S} S_k^M(E) = \mathcal{E}_M(I)$. Thus, $(\text{KB}, H, M, \emptyset, 1)$ is a positive instance of $\text{RepABD}_{\subseteq}(\Gamma)$. □

The following Theorem summarizes the results on $\text{RepABD}_{\subseteq}$.

Theorem 4.9. *($\star$) Let Γ be a constraint language. Then* $\text{RepABD}_{\subseteq}(\Gamma)$ *is*

1. Π_2^P*-complete if* $\mathsf{IN}_2 \subseteq \langle\Gamma\rangle \subseteq \mathsf{II}_2$ *or* $\mathsf{II}_0 \subseteq \langle\Gamma\rangle \subseteq \mathsf{II}_2$
2. coNP*-hard and* NP*-hard if* $\mathsf{IN} \subseteq \langle\Gamma\rangle$
3. coNP*-complete if* $C \subseteq \langle\Gamma\rangle \subseteq D$ *for* $C \in \{\mathsf{IM}, \mathsf{ID}, \mathsf{IL}\}$ *and* $D \in \{\mathsf{IE}_2, \mathsf{IL}_2, \mathsf{IV}_2, \mathsf{ID}_2\}$
4. coNP*-complete if* $\mathsf{IBF} \subseteq \langle\Gamma\rangle \subseteq D$ *for* $D \in \{\mathsf{IS}_{12}, \mathsf{IS}_{02}\}$*, and* $\Gamma \not\subseteq \langle \text{EP}^- \rangle_{\neq}$ *and* $\Gamma \not\subseteq \langle \text{EN}^- \rangle_{\neq}$
5. $\in$ P *otherwise (*$\Gamma \subseteq \langle \text{EP}^- \rangle_{\neq}$ *or* $\Gamma \subseteq \langle \text{EN}^- \rangle_{\neq}$*)*

Proof. We address each case separately.

1. Π_2^P-membership: analogously to Lemma 4.1, replacing $\mathcal{E}(I)$ by $\mathcal{E}_M(I)$ and adding to step (1) the check of subset minimality of E': for all $x \in E'$ verify that $E' \setminus \{x\}$ is not an explanation (using another NP-oracle). For Π_2^P-hardness we use Lemma 4.8.
2. coNP-hardness: via Lemma 4.7 and Lemma 3.6. We get NP-hardness via Lemma 4.8.

3. coNP-membership: here $\mathrm{SAT}(\Gamma^{+}) \in \mathrm{P}$ (Schaefer 1978) and the analogue of Lemma 4.2 holds: we only need to add a check for subset minimality which is in P this time. For coNP-hardness we use Lemma 4.7 and Lemma 3.6.
4. Membership and hardness follow exactly as in the previous case.
5. Confer Lemma 4.6.

□

See Figure 3 for a visualization of this classification.

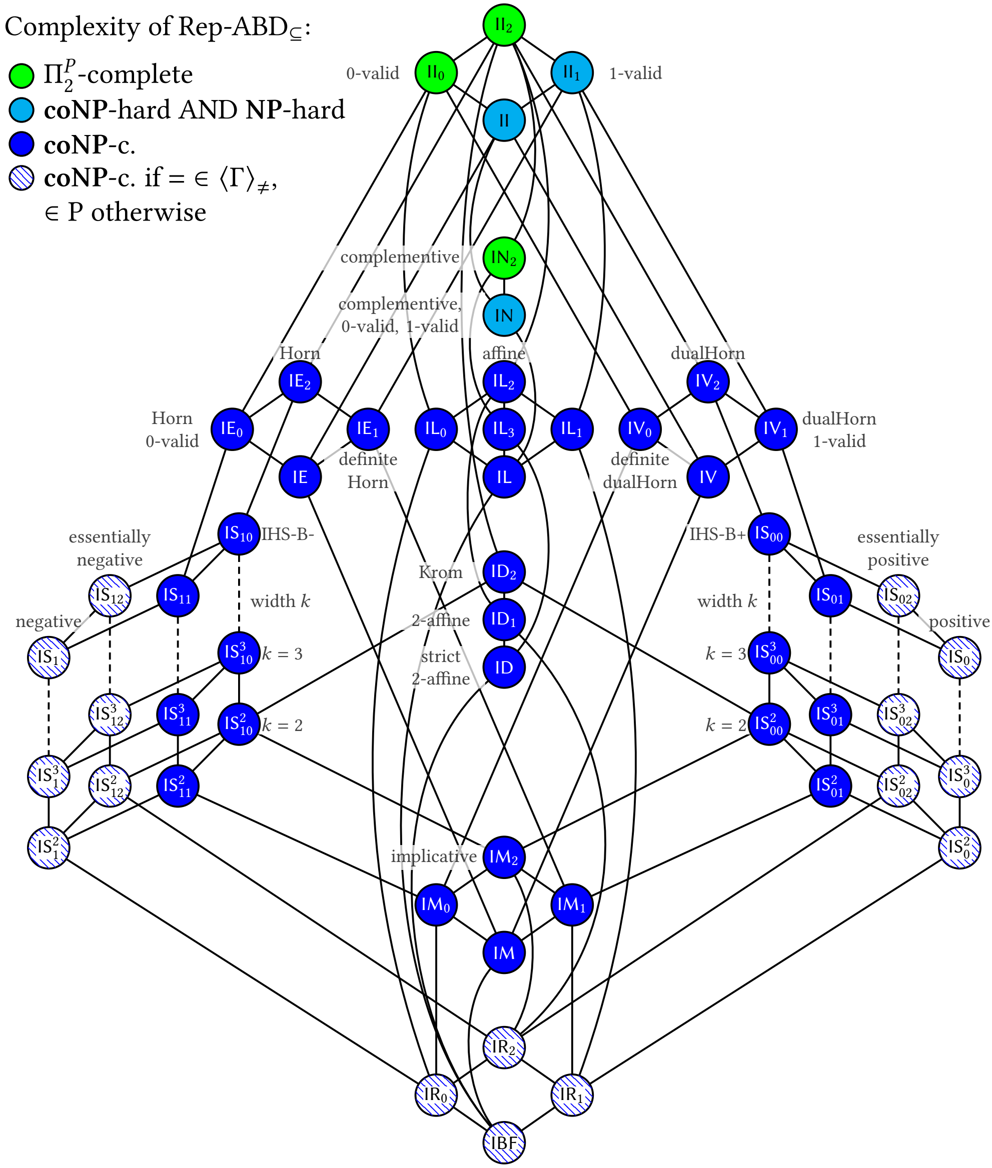

Figure 3: Illustration of complexity via Post's lattice.

5 Parameterized complexity

Having exhausted all possible sources of polynomial-time solvability we now turn our attention to parameterized complexity. We write p-RepABD(Γ, ℓ) where ℓ is the type of parameter in question. We consider several different parameters beginning with k itself (i.e., the k in the given RepABD(Γ) instance), and then continuing with $|H|$, $|M|$, and $|S|$.

5.1 Parameter k

Given a RepABD(Γ) instance $(\mathrm{KB}, H, M, S, k)$ we first consider the k itself as parameter. While this may feel like an obvious parameter choice we will soon prove that the parameter is not likely to help much (coW[1]-hard) for most choices of Γ.

Lemma 5.1. *($\star$) p-RepABD(Γ, k) is coW[1]-hard for any Γ such that $\mathsf{IS}_1^2 \subseteq \langle\Gamma\rangle$.*

Proof. We begin by giving a reduction from the W[1]-complete problem WSAT(2−CNF$^-$) to the complement of p-RepABD(Δ, k), where $\Delta = \{(\neg x_1 \vee \neg x_2)\}$ contains a single negative 2-clause. In the end, we show why this also gives coW[1]-hardness for p-RepABD(Γ, k).

Hence, let (φ, k) be an instance of WSAT(2−CNF$^-$), that is, φ is a negative 2-CNF formula over n variables, $x_1, \ldots, x_n$, and the question is whether there is a model of weight $\geq k$. We map (φ, k) to the instance $(\mathrm{KB}, H, M, S, k')$ of p-RepABD(Δ, k), where

$$\begin{aligned} \mathrm{KB} &= \varphi \\ H &= \{x_1, \ldots, x_n\} \\ M &= \emptyset \\ S &= \{\emptyset\} \\ k' &= k + 1 \end{aligned}$$

Note that KB only uses the constraint $(\neg x \vee \neg y) \in \mathsf{IS}_1^2$.

To prove correctness we use the following observation.

$$\bigcup_{E \in S} S_{k'}(E) = S_{k'}(\emptyset) = \text{all explanations of weight} \leq k' < k \tag{1}$$

In other words, S represents precisely all explanations of weight less than k.

Assume (φ, k) is a positive instance of WSAT(2−CNF$^-$). That is, there is a model σ of φ of weight at least k. We define $E = \{x \in H \mid \sigma(x) = 1\}$ and note that $|E| \geq k$. By construction, $\mathrm{KB} \wedge E$ is consistent and entails $M = \emptyset$. Therefore, E is an explanation. Since $|E| \geq k$, E is not represented by S (confer observation 1). Therefore, $(\mathrm{KB}, H, M, S, k')$ is a negative instance.

Conversely, assume $(\mathrm{KB}, H, M, S, k')$ is a negative instance. That is, there is an explanation E that is not represented. By observation 1 we conclude that $|E| \geq k$. Since E is an explanation, $\mathrm{KB} \wedge E$ must be consistent. We conclude that φ must admit a model of weight $\geq k$, thus (φ, k) is a positive instance of WSAT(2−CNF$^-$).

Last, we observe that Δ contains a single prime relation (no fictitious or redundant arguments). Hence, if $\Delta \subseteq \langle\Gamma\rangle$ then $\Delta \subseteq \langle\Gamma\rangle_{\neq}$ (Lemma 3.3), and we then apply the reduction in Lemma 3.5 together with the observation that this reduction does not affect the parameter k at all. □

We prove an analogous bound for any language Γ that can express implication.

Lemma 5.2. *($\star$) p-RepABD(Γ, k) is coW[1]-hard for any Γ such that $\mathsf{IM} \subseteq \langle \Gamma \rangle$.*

Proof. We give a reduction from the W[1]-complete problem WSAT(2−CNF$^-$) to the complement of p-RepABD. Let (φ, k) be an instance of WSAT(2−CNF$^-$), that is, φ is a negative 2-CNF formula over n variables $x_1, \ldots, x_n$. Let $\varphi = \bigwedge_{i=1}^{m} C_i$ where each clause is of the form $C_i = (\neg x_i \vee \neg y_i)$. We introduce a fresh variable c_i for each clause C_i and map (φ, k) to the instance $(\mathrm{KB}, H, M, S, k')$ as follows.

$$\begin{aligned} \mathrm{KB} &= \bigwedge_{i=1}^{m} (x_i \to c_i \wedge y_i \to c_i) \\ H &= \{x_1, \ldots, x_n\} \\ M &= \{c_1, \ldots, c_m\} \\ S &= \{H\} \\ k' &= k - 1 \end{aligned}$$

Note that KB only uses the constraint $(x \to y) \in \mathsf{IM}$ and we therefore prove hardness with respect to the language $\{(x \to y)\}$. However, the results carry over to p-RepABD(Γ, k) by combining Lemma 3.5 and Lemma 3.6 (again, note that the reduction does not affect k).

To prove correctness we use the following observation.

$$\bigcup_{E \in S} S_{k'}(E) = S_{k'}(H) = \text{all explanations of weight} \geq |H| - k' > |H| - k \qquad (2)$$

In other words, S represents precisely all explanations of weight greater than $|H| - k$.

Assume (φ, k) is a positive instance of WSAT(2−CNF$^-$). That is, there is a model σ of φ of weight at least k. We define $E = \{x \in H \mid \sigma(x) = 0\}$ and note that $|E| \leq H - k$. Since σ satisfies each clause $C_i = (\neg x_i \vee \neg y_i)$, it assigns 0 to either x_i or y_i or both. We conclude that E contains for each C_i either x_i or y_i or both. By construction it holds therefore that $\mathrm{KB} \wedge E$ entails each c_i, and therefore M. Furthermore, $\mathrm{KB} \wedge E$ is consistent, therefore E is an explanation. Since $|E| \leq H - k$, we conclude that E is not represented by S (confer observation 2). Therefore, $(\mathrm{KB}, H, M, S, k')$ is a negative instance.

Conversely, assume $(\mathrm{KB}, H, M, S, k')$ is a negative instance. That is, there is an explanation E that is not represented. By observation 2 we conclude that $|E| \leq |H| - k$. Since E is an explanation, $\mathrm{KB} \wedge E$ entails M and thus each c_i. By construction of KB this implies that E contains for each i either x_i or y_i or both. Define $\sigma(x) = 0$ if $x \in E$ and $\sigma(x) = 1$ otherwise. We observe that σ is an assignment of weight $\geq k$ and, for each i, assigns 0 to either x_i or y_i or both. Thus, σ is a model of φ, of weight $\geq k$. In conclusion, (φ, k) is a positive instance of WSAT(2−CNF$^-$). □

The three main classes that we are missing for a complete classification are linear equations ($\Gamma \subseteq \mathsf{IL}_2$), complementive languages ($\Gamma \subseteq \mathsf{IN}_2$), and any language below IS_{02} (essentially positive). We do not fully manage to describe these cases but can for the latter at least prove that its parameterized complexity essentially coincides with the parameterized complexity of the covering radius problem (with parameter r). We prove a slightly stronger result and prove that we only need to consider strictly essentially positive clauses, i.e., we do not need the equality relation in the reduction.

Lemma 5.3. *($\star$) The covering radius problem (with parameter r) is FPT-equivalent with the complement of* p-RepABD(EP^-, k).

Proof. For the first direction we use Lemma 4.4 (observe that the parameter is unchanged) and the observation that $\mathsf{t} \in \langle \mathrm{EN}^- \rangle_{\neq}$ (this follows from Lemma 3.4).

For the other direction let $I = (\mathrm{KB}, H, M, S, k)$ an arbitrary instance of RepABD. Note that KB contains unit clauses (positive and negative) and positive clauses of size ≥ 2. With this one easily observes that given an $E \subseteq H$ and an $m \in M$ it holds $\mathrm{KB} \wedge E \models m$ if and only if $\mathrm{KB} \models m$ or $E \models m$. We conclude that, provided there are explanations, there is a unique subset-minimal explanation, namely

$$E_{min} = \{m \in M \mid \mathrm{KB} \not\models m\},$$

and that $\mathcal{E}(I) = \{E \subseteq H \mid E_{min} \subseteq E\}$ in this case. It can happen that there are no explanations because either 1) $\mathrm{KB} \wedge E_{min}$ is unsatisfiable, or because 2) $E_{min} \not\subseteq H$. In this case $\mathcal{E}(I) = \emptyset$ and also the given $S \subseteq \mathcal{E}(I)$ is empty, and therefore k-represents all explanations by definition. We map in this case to a fixed negative instance of the covering radius problem.

Suppose now that $\mathcal{E}(I) \neq \emptyset$. We define $n = |H \setminus E_{min}|$, rename the variables in $H \setminus E_{min}$ as $x_1, \ldots, x_n$ and define $S' = \{E \setminus E_{min} \mid E \in S\}$. We map the instance I to (C, r) as follows. Recall that E_{vec} denotes the characteristic vector of the set $E \subseteq \{x_1, \ldots, x_n\}$ (as in Lemma 4.4).

$$r = k$$
$$C = \{E_{vec} \mid E \in S'\}$$

For correctness, suppose first that $I = (\mathrm{KB}, H, M, S, k)$ is a positive instance. That means any $E \in \mathcal{E}(I) = \{E \subseteq H \mid E_{min} \subseteq E\}$ is k-represented by S, or in other words, for any $E \in \mathcal{E}(I)$ there is an $E' \in S$ such that $d(E, E') \leq k$. Observe that E_{min} has no influence on the distance $d(E, E')$ for any $E, E' \in \mathcal{E}(I)$, since all explanations contain E_{min}. Therefore, also the following holds true: for any $E \subseteq H \setminus E_{min} = \{x_1, \ldots, x_n\}$ there is an $E' \in S' = \{E \setminus E_{min} \mid E \in S\}$ such that $d(E, E') \leq k$. By definition of r and C this implies immediately that for any vector $v \in \{0,1\}^n$ there is a vector $u \in C$ within hamming distance $r = k$. In other words, the covering radius of C is less than r, hence (C, r) is a negative instance.

Suppose now (C, r) is a negative instance of the covering radius problem. That is, for any vector $v \in \{0,1\}^n$ there is a vector $u \in C$ within hamming distance r. By definition of r and C this implies that for any $E \subseteq H \setminus E_{min} = \{x_1, \ldots, x_n\}$ there is an $E' \in S' = \{E \setminus E_{min} \mid E \in S\}$ such that $d(E, E') \leq k$. If we add the set E_{min} to all considered sets, and reverse the renaming of variables, we obtain that for any $E \in \mathcal{E}(I)$ there is an $E' \in S$ such that $d(E, E') \leq k$. That is, $I = (\mathrm{KB}, H, M, S, k)$ is a positive instance. $\square$

5.2 Parameter $|H|$

Next, we consider the size of the hypothesis, $|H|$, as parameter, where we obtain a straightforward parameterized dichotomy. The FPT case can be proven as follows.

Lemma 5.4. p-RepABD$(\Gamma, |H|) \in$ FPT *if* Γ *is Schaefer.*

Proof. We can loop over all explanation candidates $E \subseteq H$ in brute force time $2^{|H|}$. Since the parameter is $|H|$, this is FPT-time. Now we only need to check for each such candidate E 1) whether it is an explanation, and 2) if so, whether it is k-represented by S or not. Step 1) amounts to checking whether $\mathrm{KB} \wedge E$ is satisfiable and whether $\mathrm{KB} \wedge E \models M$. Both checks can be achieved in polynomial time since for Schaefer languages $\mathrm{SAT}(\Gamma^{+}) \in \mathrm{P}$. Step 2) amounts to looping over the elements of S and computing for each $E' \in S$ whether $d(E, E') \leq k$ or not. This can be done in polynomial time, since S is part of the input. ☐

Lemma 5.5. p-RepABD$(\Gamma, |H|)$ *is* para-coDP-*hard for any* Γ *such that* $\mathsf{IN}_2 \subseteq \langle\Gamma\rangle$ *or* $\mathsf{II}_0 \subseteq \langle\Gamma\rangle$.

Proof. According to (Mahmood et al. 2021) the problem p-ABD$(\Gamma, |H|)$ is para-DP-hard for any Γ such that $\mathsf{IN}_2 \subseteq \langle\Gamma\rangle$ or $\mathsf{II}_0 \subseteq \langle\Gamma\rangle$. The result follows now with the reduction of Lemma 4.3. ☐

Lemma 5.6. p-RepABD$(\Gamma, |H|)$ *is* para-NP-*hard for any* Γ *such that* $\mathsf{IN} \subseteq \langle\Gamma\rangle$.

Proof. According to (Mahmood et al. 2021) the problem p-ABD$(\Gamma, |H|)$ is para-coNP-hard for any Γ such that $\mathsf{IN} \subseteq \langle\Gamma\rangle$. The result follows now with the reduction of Lemma 4.3. ☐

This leads to the following classification (note, however, that the two hardness cases are not fully exhaustive since we do not have inclusion, only hardness).

Theorem 5.7. *Let* Γ *be a constraint language. Then* p-RepABD$(\Gamma, |H|)$ *is*

1. para-coDP-*hard if* $\mathsf{IN}_2 \subseteq \langle\Gamma\rangle$ *or* $\mathsf{II}_0 \subseteq \langle\Gamma\rangle$
2. para-NP-*hard if* $\mathsf{IN} \subseteq \langle\Gamma\rangle$
3. $\in$ FPT *otherwise (that is,* $\Gamma \subseteq D$ *for* $D \in \{\mathsf{IE}_2, \mathsf{IL}_2, \mathsf{IV}_2, \mathsf{ID}_2\}$

5.3 Parameter $|M|$

For the size of the manifestation ($|M|$) we currently lack FPT cases, but, on the other hand, obtain a general and non-trivial coW[1]-hard case.

Lemma 5.8. p-RepABD$(\Gamma, |M|)$ *is coW[1]-hard for any* Γ *such that* $\mathsf{IS}_{11}^{2} \subseteq \langle\Gamma\rangle$.

Proof. Due to the reduction of Lemma 4.3 it suffices to establish W[1]-hardness of $\mathrm{ABD}(\Gamma)$ for parameter $|M|$. We give a reduction from the W[1]-complete problem $\mathrm{WSAT}(2{-}\mathrm{CNF}^{-})$ to $\mathrm{ABD}(\Gamma)$. Be (φ, k) an instance of $\mathrm{WSAT}(2{-}\mathrm{CNF}^{-})$, that is, φ is a negative 2-CNF formula over n variables, and the question is whether there is a model of weight k. We construct the target instance (KB, H, M) starting with KB. Denote $\varphi = \bigwedge_{r=1}^{m} C_r$, where $C_r = (\neg x_r \vee \neg y_r)$, and let $V = \mathrm{var}(\varphi)$. Introduce fresh variables

$m_1, \dots, m_k$ and for each $x \in V$ introduce k copies, denoted $x^1, \dots, x^k$. The knowledge base KB consists of the following parts.

1. $\varphi' = \bigwedge_{r=1}^{m} C'_r$, where C'_r is obtained from $C_r = (\neg x \vee \neg y)$ via 'multiplying through all copies of x and y'. That is, $C'_r = \bigwedge_{i,j=1}^{k}(\neg x^i \vee \neg y^j)$. Hence, we have satisfiability equivalence with the original φ.
2. $\bigwedge_{x\in V} \bigwedge_{i,j=1,i\neq j}^{k}(\neg x^i \vee \neg x^j)$. Hence, for each x at most one copy can be selected.
3. $\bigwedge_{x\in V} \bigwedge_{i=1}^{k}(x^i \to m_i)$. Hence, each m_i can be explained. Indeed, any $x \in V$ can do so, via the copy x^i.
4. $\bigwedge_{x,y\in V,x\neq y} \bigwedge_{i=1}^{k}(\neg x^i \vee \neg y^i)$. Hence, each m_i can be explained at most once.

Observe that all formulas in KB are built form the constraints $(x \to y)$ and $(\neg x \vee \neg y)$ and are thus from IS_{11}^2. In other words we prove hardness with respect to the language $\{(x \to y), (\neg x \vee \neg y)\}$, but the results carry over to p-RepABD$(\Gamma, |M|)$ by combining Lemma 3.5 and Lemma 3.6 (the reduction in Lemma 3.5 does not affect $|M|$). Furthermore, all formulas can be constructed in FPT-time (even polynomial time). Define

$$H = \{x^i \mid i \in \{1, \dots, k\}, x \in V\},$$
$$M = \{m_1, \dots, m_k\}.$$

Assume now (φ, k) is a positive instance of WSAT(2−CNF$^-$). Let W.L.O.G. $x_1, \dots, x_k$ be a model of φ of weight k. Define $E = \{x_1^1, \dots, x_k^k\} \subseteq H$. One verifies that E is consistent with KB, in particular parts 1, 2 and 4. Furthermore, because of part 3, we obtain that KB $\wedge E$ entails M.

For the converse direction, let $E \subseteq H$ be an explanation for M. That is, KB $\wedge E$ is consistent and KB $\wedge E$ entails M. By construction (part 3) each m_i can only be explained by the ith copy of a variable. Because of part 2 (for each x at most one copy is selected) there must be at least k different variables from φ responsible (through their copies) for explaining the k different m_i. Because of part 4, this must be at most k different variables. In summary, $|E| = k$. Hence, there are exactly k variables responsible for explaining the m_i, and each through a different copy. Denote them W.L.O.G. $x_1, \dots, x_k$. Since E is consistent with KB, and in particular with part 1, we conclude that $x_1, \dots, x_k$ is satisfying φ, thus a model of weight k. □

5.4 Parameter $|S|$

The last parameter that we consider is the number of sets $|S|$. We first observe a hard case by using Lemma 4.3.

Lemma 5.9. $(\star)$ p-RepABD$(\Gamma, |S|)$ *is para-coNP-hard for any* Γ *such that* $\mathsf{IS}_{11}^2 \subseteq \langle\Gamma\rangle$.

Proof. We can use once more the reduction from Lemma 4.3: it shows para-NP-hardness for the slice $|S| = 0$, for any language such that $\mathsf{IS}_{11}^2 \subseteq \langle\Gamma\rangle$ (since here ABD(Γ) is NP-hard). □

Lemma 5.10. p-RepABD(EP$^-$, $|S|$) *is in FPT.*

Proof. We first observe that p-RepABD(EP$^-$, $|S|$) is FPT-equivalent to the covering radius problem by the same reduction used in Lemma 5.3.

Next, we show FPT for this problem by giving an FPT-reduction to the *closest string problem* with alphabet $\Sigma = \{0, 1\}$. Here, we are given a set of strings C, a radius r and want to know if there exists $x \in \{0, 1\}^n$ such that $d(x, c_i) \leq r$ for every $c_i \in C$. This problem is known to be FPT with respect to $|C|$ (Gramm et al. 2003). Given an instance $C \subseteq \{0, 1\}^n$ and $r \geq 1$ of the covering radius problem we map it to an instance (C', r') of the closest string problem as follows.

- $C' = \{\overline{c_i} \mid c_i \in C\}$,
- $r' = n - r$.

The correctness is as follows: assume $\exists x : \forall c_i \in C : d(x, c_i) > r \leftrightarrow |x \cap c_i| > r \leftrightarrow |x \cap \overline{c_i}| \leq n - r$. Moreover, the reduction can clearly be carried out in FPT time (indeed, even polynomial time) and preserves the parameter since $|C'| = |C|$. This concludes the proof. $\square$

This result extends to p-RepABD$(\Gamma, |S|)$ for any $\Gamma \subseteq \langle \mathrm{EP}^- \rangle_{\neq}$ since the reduction in Lemma 3.5 does not affect the parameter $|S|$.

6 Conclusion

We introduced the problem of representative sets in propositional abduction. We illustrated that this problem gives a better understanding of the space of possible explanations, for example, if we return to the example in Section 1, in a medical situation, representative abduction can tell doctors if their set of hypothesis has covered all major possibilities to explain the symptoms or if they might have missed something crucial. Complexity wise, we have established an almost complete classification of the problem for constraint language restrictions (Post's lattice). We show that the representative abduction problem is still in the second level of the polynomial hierarchy, making it harder than abduction only in the lower parts of the lattice. Although the problem admits almost no tractable cases in its normal form, the subset minimal variant and the parametrized complexity approach give us several additional tractable cases. Let us now discuss some potential future research directions.

Subset minimal representative abduction. We have made an almost complete classification of the subset-minimal version of the representative abduction problem. Although it has a very similar classification in the lattice with the non-subset minimal variant, it is surprisingly easier for EP^- and EN^-. A possible future research direction is to consider different parameters for a parametrized complexity approach for this variant of the problem.

Completing the classification. We have three open cases shared by representative abduction and its subset minimal variant. These are the 1-valid languages. We have already established that they are both NP- and coNP-hard. It would be interesting to investigate their exact complexity class, and a good candidate for these languages is the class DP. The techniques that would be used to prove it would most likely be novel for abduction related problems.

Parametrized complexity. We have investigated the parametrized complexity of representative abduction with different parameters, namely k (the Hamming distance), $|H|$, $|M|$, and $|S|$ (size of the representative set). For parameter $|H|$ we have established a complete classification, obtained FPT-results for a large portion of languages (Schaefer languages), and proven hardness for the rest. We have established a non-trivial coW[1]-hardness result for the majority of languages for parameter $|M|$. In our opinion, the parameters k and $|S|$ are perhaps the most interesting ones. For parameter $|S|$ we have established both hardness for a large portion of the lattice, but also non-trivial FPT results for EP^--languages, but which still leaves some interesting open cases for complementive and affine languages. Finally, for parameter k we have established non-trivial coW[1]-hardness results for languages IM and EN^-, leaving what are perhaps the most interesting open cases for future research. The parametrized complexity approach seems to us to be extremely interesting. It establishes possibly useful FPT results, as well as interesting reductions for proving hardness, and it links our problem to know problems from coding theory.

Relation to coding theory. We have established a strong connection between our representative abduction problem and the covering radius problem from coding theory, that as far as we know has not been established before. A lot of classical and parametrized complexity results from coding theory were instrumental to our lattice classifications, such as NP-completeness of the covering radius problem (Frances and Litman 1997) and the FPT result for the closest string problem with parameter $|C|$ (Gramm et al. 2003). Future research on parametrized complexity with parameter k for representative abduction and parameter r (radius) for the covering radius problem, will greatly benefit both fields.

References

Alrabbaa, C., Rudolph, S., and Schweizer, L. Faceted answer-set navigation. In *Proc. RuleML+RR'18* 2018, pp. 211–225.

Baste, J., Fellows, M. R., Jaffke, L., Masařík, T., Oliveira Oliveira, M. d., Philip, G., and Rosamond, F. A. Diversity of solutions: An exploration through the lens of fixed-parameter tractability theory. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)* 2020, pp. 1119–1125.

Böhl, E., Gaggl, S. A., and Rusovac, D. Representative answer sets: Collecting something of everything. In *ECAI* 2023, pp. 271–278. IOS Press.

Böhler, E., Reith, S., Schnoor, H., and Vollmer, H. 2005. Bases for Boolean co-clones. *Inf. Process. Lett.*, *96*, 2, 59–66.

Creignou, N., Egly, U., and Schmidt, J. 2014. Complexity classifications for logic-based argumentation. *ACM Trans. Comput. Log.*, *15*, 3, 19:1–19:20.

Creignou, N., Kröll, M., Pichler, R., Skritek, S., and Vollmer, H. 2019. A complexity theory for hard enumeration problems. *Discrete Applied Mathematics*, *268*, 191–209.

Creignou, N., Olive, F., and Schmidt, J. Complexity of reasoning with cardinality minimality conditions. In Williams, B., Chen, Y., and Neville, J., editors, *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence, IAAI 2023, Thirteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2023, Washington, DC, USA, February 7-14, 2023* 2023, pp. 3932–3940. AAAI Press.

Dai, W.-Z., Xu, Q., Yu, Y., and Zhou, Z.-H. Bridging machine learning and logical reasoning by abductive learning. In Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., and Garnett, R., editors, *NeurIPS'19* 2019, volume 32. Curran Associates, Inc.

Dellsén, F. 2024. *Abductive Reasoning in Science*. Elements in the Philosophy of Science. Cambridge University Press.

Downey, R. G., Fellows, M. R., and others 2013. *Fundamentals of parameterized complexity*, volume 4. Springer.

Eiter, T., Erdem, E., Erdogan, H., and Fink, M. 2013. Finding similar/diverse solutions in answer set programming. *Theory and Practice of Logic Programming*, *13*, 3, 303–359.

Eiter, T. and Gottlob, G. 1995. The complexity of logic-based abduction. *J. ACM*, *42*, 1, 3–42.

Fellows, M. R., Pfandler, A., Rosamond, F. A., and Rümmele, S. The parameterized complexity of abduction. In *Proc. AAAI* 2012, pp. 743–749. AAAI Press.

Fichte, J., Fröhlich, N., Hecher, M., Lagerkvist, V., Mahmood, Y., Meier, A., and Persson, J. Facets in argumentation: A formal approach to argument significance. In *Proceedings of the 34th International Joint Conference on Artificial Intelligence (IJCAI-2025)* 2025. ijcai.org.

Fomin, F. V., Golovach, P. A., Jaffke, L., Philip, G., and Sagunov, D. Diverse Pairs of Matchings. In Cao, Y., Cheng, S.-W., and Li, M., editors, *31st International Symposium on Algorithms and Computation (ISAAC 2020)* 2020, volume 181 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 26:1–26:12. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

Fomin, F. V., Golovach, P. A., Panolan, F., Philip, G., and Saurabh, S. Diverse Collections in Matroids and Graphs. In Bläser, M. and Monmege, B., editors, *38th International Symposium on Theoretical Aspects of Computer Science (STACS 2021)* 2021, volume 187 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 31:1–31:14. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

Frances, M. and Litman, A. 1997. On covering problems of codes. *Theory of Computing Systems*, *30*, 2, 113–119.

Gramm, J., Niedermeier, R., and Rossmanith, P. 2003. Fixed-parameter algorithms for CLOSEST STRING and related problems. *Algorithmica*, *37*, 1, 25–42.

Hebrard, E., Hnich, B., O'Sullivan, B., and Walsh, T. Finding diverse and similar solutions in constraint programming. In *Proceedings of the 20th National Conference on Artificial Intelligence (AAAI)* 2005, pp. 372–377.

Hermann, M. and Pichler, R. 2010. Counting complexity of propositional abduction. *Journal of Computer and System Sciences*, *76*, 7, 634–649.

Hu, W.-C., Dai, W.-Z., Jiang, Y., and Zhou, Z.-H. Efficient rectification of neuro-symbolic reasoning inconsistencies by abductive reflection. In *AAAI* 2025, pp. 17333–17341.

Ignatiev, A., Narodytska, N., and Marques-Silva, J. Abduction-based explanations for machine learning models. In *Proc. AAAI* 2019. AAAI Press.

Kakas, A. C., Kowalski, R. A., and Toni, F. 1992. Abductive logic programming. *Journal of Logic and Computation*, *2*, 6, 719–770.

Lagerkvist, V. A new characterization of restriction-closed hyperclones. In *Proceedings of the 50th International Symposium on Multiple-Valued Logic (ISMVL-2020)* 2020, pp. 303–308. IEEE Computer Society.

Mahmood, Y., Meier, A., and Schmidt, J. 2021. Parameterized complexity of abduction in schaefer's framework. *J. Log. Comput.*, *31*, 1, 266–296.

Mahmood, Y., Meier, A., and Schmidt, J. 2023. Parameterized complexity of logic-based argumentation in schaefer's framework. *ACM Trans. Comput. Log.*, *24*, 3, 26:1–26:25.

Minsky, M. 1974. A framework for representing knowledge. Technical report, Massachusetts Institute of Technology, USA.

Misra, N., Mittal, H., and Rai, A. On the parameterized complexity of diverse SAT. In *ISAAC* 2024, pp. 50:1–50:18. Dagstuhl.

Nordh, G. and Zanuttini, B. 2008. What makes propositional abduction tractable. *Artif. Intell.*, *172*, 10, 1245–1284.

Rusovac, D., Hecher, M., Gebser, M., Gaggl, S. A., and Fichte, J. K. Navigating and Querying Answer Sets: How Hard Is It Really and Why? In *Proc. of KR* 2024, pp. 642–653.

Schaefer, T. J. The complexity of satisfiability problems. In *Proc. STOC* 1978, pp. 216–226. ACM.

Schmidt, J., Maizia, M., Lagerkvist, V., and Fichte, J. 2025. Complexity of faceted explanations in propositional abduction. *Theory and Practice of Logic Programming*, *25*, 4, 775–793.

Stockmeyer, L. J. 1976. The polynomial-time hierarchy. *Theoretical Computer Science*, *3*, 1, 1–22.

Stockmeyer, L. J. and Meyer, A. R. Word problems requiring exponential time: Preliminary report. In *STOC* 1973, pp. 1–9. ACM.

Wrathall, C. 1976. Complete sets and the polynomial-time hierarchy. *Theoretical Computer Science*, *3*, 1, 23–33.

Yu, J., Ignatiev, A., Stuckey, P. J., Narodytska, N., and Marques-Silva, J. Eliminating the impossible, whatever remains must be true: On extracting and applying background knowledge in the context of formal explanations. In *Proc. AAAI* 2023, pp. 4123–4131.

co-clone	base	clauses/equation	name/indication
BR (II_2)	1-IN-3 $= \{001, 010, 100\}$	all clauses	all Boolean relations
II_1	$x \vee (y \oplus z)$	at least one positive literal per clause	1-valid
II_0	DUP, $x \to y$	at least one negative literal per clause	0-valid
II	EVEN^4, $x \to y$	at least one negative and one positive literal per clause	1- and 0-valid
IN_2	NAE $= \{0,1\}^3 \setminus \{000, 111\}$	cf. previous column	complementive
IN	DUP $= \{0,1\}^3 \setminus \{101, 010\}$	cf. previous column	complementive and 1- and 0-valid
IE_2	$x \wedge y \to z, x, \neg x$	clauses with at most one positive literal	Horn
IE_1	$x \wedge y \to z, x$	clauses with exactly one positive literal	definite Horn
IE_0	$x \wedge y \to z, \neg x$	$(x_1 \vee \neg x_2 \vee \cdots \vee \neg x_n), n \geq 2, (\neg x_1 \vee \cdots \vee \neg x_n), n \geq 1$	Horn and 0-valid
IE	$x \wedge y \to z$	$(x_1 \vee \neg x_2 \vee \cdots \vee \neg x_n), n \geq 2$	Horn and 1- and 0-valid
IV_2	$x \vee y \vee \neg z, x, \neg x$	clauses with at most one negative literal	dualHorn
IV_1	$x \vee y \vee \neg z, x$	$(\neg x_1 \vee x_2 \vee \cdots \vee x_n), n \geq 2, (x_1 \vee \cdots \vee x_n), n \geq 1$	dualHorn and 1-valid
IV_0	$x \vee y \vee \neg z, \neg x$	clauses with exactly one negative literal	definite dualHorn
IV	$x \vee y \vee \neg z$	$(\neg x_1 \vee x_2 \vee \cdots \vee x_n), n \geq 2$	dualHorn and 1- and 0-valid
IL_2	EVEN^4, x, $\neg x$	all affine clauses (all linear equations)	affine
IL_1	EVEN^4, x	$(x_1 \oplus \cdots \oplus x_n = a)$, $n \geq 0, a = n \pmod 2$	affine and 1-valid
IL_0	EVEN^4, $\neg x$	$(x_1 \oplus \cdots \oplus x_n = 0)$, $n \geq 0$	affine and 0-valid
IL_3	EVEN^4, $x \oplus y$	$(x_1 \oplus \cdots \oplus x_n = a)$, n even, $a \in \{0, 1\}$	-
IL	EVEN^4	$(x_1 \oplus \cdots \oplus x_n = 0)$, n even	affine and 1- and 0-valid
ID_2	$x \oplus y, x \to y$	clauses of size 1 or 2	Krom, bijunctive, 2-CNF
ID_1	$x \oplus y, x, \neg x$	affine clauses of size 1 or 2	2-affine
ID	$x \oplus y$	affine clauses of size 2	strict 2-affine
IM_2	$x \to y, x, \neg x$	$(x_1 \to x_2), (x_1), (\neg x_1)$	implicative
IM_1	$x \to y, x$	$(x_1 \to x_2), (x_1)$	implicative and 1-valid
IM_0	$x \to y, \neg x$	$(x_1 \to x_2), (\neg x_1)$	implicative and 0-valid
IM	$x \to y$	$(x_1 \to x_2)$	implicative and 1- and 0-valid
IS_{10}	cf. next column	$(x_1), (x_1 \to x_2), (\neg x_1 \vee \cdots \vee \neg x_n), n \geq 0$	IHS-B-
IS_{10}^k	cf. next column	$(x_1), (x_1 \to x_2), (\neg x_1 \vee \cdots \vee \neg x_n), k \geq n \geq 0$	IHS-B- of width k
IS_{12}	cf. next column	$(x_1), (\neg x_1 \vee \cdots \vee \neg x_n), n \geq 0, (x_1 = x_2)$	essentially negative (EN)
IS_{12}^k	cf. next column	$(x_1), (\neg x_1 \vee \cdots \vee \neg x_n), k \geq n \geq 0, (x_1 = x_2)$	essentially negative (EN) of width k
IS_{11}	cf. next column	$(x_1 \to x_2), (\neg x_1 \vee \cdots \vee \neg x_n), n \geq 0$	-
IS_{11}^k	cf. next column	$(x_1 \to x_2), (\neg x_1 \vee \cdots \vee \neg x_n), k \geq n \geq 0$	-
IS_1	cf. next column	$(\neg x_1 \vee \cdots \vee \neg x_n), n \geq 0, (x_1 = x_2)$	negative (N)
IS_1^k	cf. next column	$(\neg x_1 \vee \cdots \vee \neg x_n), k \geq n \geq 0, (x_1 = x_2)$	negative (N) of width k
IS_{00}	cf. next column	$(\neg x_1), (x_1 \to x_2), (x_1 \vee \cdots \vee x_n), n \geq 0$	IHS-B+
IS_{00}^k	cf. next column	$(\neg x_1), (x_1 \to x_2), (x_1 \vee \cdots \vee x_n), k \geq n \geq 0$	IHS-B+ of width k
IS_{02}	cf. next column	$(\neg x_1), (x_1 \vee \cdots \vee x_n), n \geq 0, (x_1 = x_2)$	essentially positive (EP)
IS_{02}^k	cf. next column	$(\neg x_1), (x_1 \vee \cdots \vee x_n), k \geq n \geq 0, (x_1 = x_2)$	essentially positive (EP) of width k
IS_{01}	cf. next column	$(x_1 \to x_2), (x_1 \vee \cdots \vee x_n), n \geq 0$	-
IS_{01}^k	cf. next column	$(x_1 \to x_2), (x_1 \vee \cdots \vee x_n), k \geq n \geq 0$	-
IS_0	cf. next column	$(x_1 \vee \cdots \vee x_n), n \geq 0, (x_1 = x_2)$	positive (P)
IS_0^k	cf. next column	$(x_1 \vee \cdots \vee x_n), k \geq n \geq 0, (x_1 = x_2)$	positive (P) of width k
IR_2	$x_1, \neg x_2$	$(x_1), (\neg x_1), (x_1 = x_2)$	-
IR_1	x_1	$(x_1), (x_1 = x_2)$	-
IR_0	$\neg x_1$	$(\neg x_1), (x_1 = x_2)$	-
IR (IBF)	$\emptyset$	$(x_1 = x_2)$	-

Table A 1: Overview of bases (Böhler et al. 2005) and clause descriptions (Nordh and Zanuttini 2008) for co-clones, where $\text{EVEN}^4 = x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus 1$.

Appendix A Omitted Definitions

See Table A 1 for a comprehensive list of all Boolean co-clones.