Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning
Main Track

Abductive Reasoning in a Paraconsistent Framework

Meghyn Bienvenu[1,2] , **Katsumi Inoue**[3] , **Daniil Kozhemiachenko**[1]
[1]Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, Talence, France
[2] Japanese-French Laboratory for Informatics, CNRS, NII, IRL 2537, Tokyo, Japan
[3]National Institute of Informatics, Tokyo, Japan
{meghyn.bienvenu,daniil.kozhemiachenko}@u-bordeaux.fr, inoue@nii.ac.jp

Abstract

We explore the problem of explaining observations starting from a classically inconsistent theory by adopting a paraconsistent framework. We consider two expansions of the well-known Belnap–Dunn paraconsistent four-valued logic BD: $\mathsf{BD}_{\circ}$ introduces formulas of the form $\circ\phi$ ('the information about ϕ is reliable'), while $\mathsf{BD}_{\triangle}$ augments the language with formulas $\triangle\phi$ ('there is information that ϕ is true'). We define and motivate the notions of abduction problems and explanations in $\mathsf{BD}_{\circ}$ and $\mathsf{BD}_{\triangle}$ and show that they are not reducible to one another. We analyse the complexity of standard abductive reasoning tasks (solution recognition, solution existence, and relevance / necessity of hypotheses) in both logics. Finally, we show how to reduce abduction in $\mathsf{BD}_{\circ}$ and $\mathsf{BD}_{\triangle}$ to abduction in classical propositional logic, thereby enabling the reuse of existing abductive reasoning procedures.

1 Introduction

Logic-based abduction is an important form of reasoning with multiple applications in artificial intelligence, including diagnosis and commonsense reasoning (Eiter and Gottlob 1995). An *abduction problem* can be generally formulated as a pair $\langle\Gamma,\chi\rangle$ consisting of a set of formulas Γ (*theory*) and a formula χ (*observation*) s.t. $\Gamma\not\models\chi$, and the task is to find an *explanation*, i.e., a formula ϕ s.t. $\Gamma,\phi\models\chi$. Of course, not *every* formula is intuitively acceptable as an explanation which is why there are usually some restrictions on ϕ. In particular, $\Gamma\cup\{\phi\}$ should be consistent, ϕ should not entail χ by itself nor contain atoms not occurring in $\Gamma\cup\{\phi\}$, ϕ should be syntactically restricted so as to be easily understandable, and ϕ should constitute a weakest possible (or minimal) explanation, cf. discussion in (Marquis 2000, §4.2) or (Aliseda 2006, §3.3). Most commonly, the third desiderata is enforced by requiring abductive solutions to take the form of *terms* (conjunctions of literals), in which case the logically weakest solutions are simply the subset-minimal ones.

Note, however, that in classical propositional logic (CPL) any *contradictory* theory is inconsistent. Thus, there is no explanation of χ from a contradictory Γ. This can be circumvented in two ways. First, by *repairing* Γ, i.e., making it consistent and then proceeding as usual (cf., e.g., (Du, Wang, and Shen 2015)). Second, by moving to a *paraconsistent* logic. The characteristic feature of such logics is the failure of the explosion principle — $p,\neg p\not\models q$.

Abduction in Paraconsistent Logics The question of how to employ paraconsistent logics to perform abductive reasoning on classically inconsistent theories has already generated interest in the philosophical logic community. For example, (Carnielli 2006) considers abduction in a three-valued logic called *Logic of Formal Inconsistency* ($\mathbf{LFI1}$), obtaining by expanding the language of CPL (classical propositional logic) with new connectives $\bullet\phi$ and $\circ\phi$ (read 'ϕ has a non-classical value' or 'the information about ϕ is unreliable' and 'ϕ has a classical value' or 'the information about ϕ is reliable', respectively). More recently, (Bueno-Soler et al. 2017) and (Chlebowski, Gajda, and Urbański 2022) considered abductive explanations in the *minimal Logic of Formal Inconsistency* ($\mathbf{mbC}$), and (Rodrigues et al. 2023) consider abduction in a four-valued *Logic of Evidence and Truth* ($\mathsf{LET}_{\mathsf{K}}$).

These studies showcase the interest of paraconsistent abduction, but as they issue from a different research community, the formulation of abductive solutions and the questions that are explored depart from those typically considered in knowledge representation and reasoning (KR). In particular, these works allow *arbitrary* formulas as solutions (rather than terms). Moreover, to the best of our knowledge, there are no results on the complexity of paraconsistent reasoning tasks (e.g., solution existence). Also, in $\mathbf{mbC}$ and $\mathsf{LET}_{\mathsf{K}}$, $\circ$ and $\bullet$ do not have truth-functional semantics, which complicates the comparison with classical logic and the reuse of established techniques.

Abduction in Belnap–Dunn Logic The preceding considerations motivate us to revisit paraconsistent abduction by taking a KR perspective and adopting the well-known paraconsistent propositional logic BD by (Dunn 1976; Belnap 1977a; Belnap 1977b). The main idea of BD is to treat the values of formulas as the information an agent (or a computer as in (Belnap 1977b)) might have w.r.t. a given statement ϕ. This results in four 'Belnapian' values:

- $\mathbf{T}$ — 'the agent is only told that ϕ is true';
- $\mathbf{F}$ — 'the agent is only told that ϕ is false';
- $\mathbf{B}$ — 'the agent is told that ϕ is false and that it is true';
- $\mathbf{N}$ — 'the agent is not told that ϕ is false nor that it is true'.

The truth and falsity conditions of $\neg$, $\wedge$, and $\vee$ are defined in a classical manner but assumed to be *independent*.

	is true when	**is false when**
$\neg\phi$	ϕ is false	ϕ is true
$\phi_1 \wedge \phi_2$	ϕ_1 and ϕ_2 are true	ϕ_1 or ϕ_2 is false
$\phi_1 \vee \phi_2$	ϕ_1 or ϕ_2 is true	ϕ_1 and ϕ_2 are false

One can see from the table above that there are no *valid formulas* (that always have value in $\{\mathbf{T},\mathbf{B}\}$, i.e., *at least true*) over $\{\neg,\wedge,\vee\}$. Likewise, there are no formulas that are always *at least non-true* (have value in $\{\mathbf{N},\mathbf{F}\}$). Thus, defining $\phi \models_{\mathsf{BD}} \chi$ as 'in every valuation v s.t. $v(\phi) \in \{\mathbf{T},\mathbf{B}\}$, $v(\chi) \in \{\mathbf{T},\mathbf{B}\}$ as well', we obtain that $p \wedge \neg p \not\models_{\mathsf{BD}} q$.

Note also that, as already observed in (Rodrigues et al. 2023), the $\{\neg,\wedge,\vee\}$-language is too weak for abduction: there is no solution for $\langle\{p \vee q\}, q\rangle$ except for q itself because $p \vee q, \neg p \not\models_{\mathsf{BD}} q$. However, if one assumes that p has value $\mathbf{F}$, this will explain q provided $p \vee q$. To do this, one needs to expand the language of BD with new connectives.

One option is to formulate abductive solutions in $\mathsf{BD}_\circ$, the expansion of BD with the *truth-functional* version of $\circ$ by (Omori and Waragai 2011; Omori and Sano 2014). The connective $\circ$ is interpreted as follows: $\circ\phi$ has value $\mathbf{T}$ if ϕ has value $\mathbf{T}$ or $\mathbf{F}$, and has value $\mathbf{F}$, otherwise. In this expanded language, the formula $\neg p \wedge \circ p$, which expresses that *there is reliable information that p is false*, yields a solution to $\langle\{p \vee q\}, q\rangle$. Another possibility is to adopt $\mathsf{BD}_\triangle$, the expansion of BD with $\triangle$ from (Sano and Omori 2014). Here, $\triangle\phi$ can be interpreted as 'there is information that ϕ is true' and has the following semantics: $v(\triangle\phi) = \mathbf{T}$ if $v(\phi) \in \{\mathbf{T},\mathbf{B}\}$ and $v(\triangle\phi) = \mathbf{F}$, otherwise. In this case, the formula $\neg p \wedge \neg\triangle p$ (which reads: *p is false and there is no information that it is true*) is a solution.

Importantly, $\mathsf{BD}_\triangle$ and $\mathsf{BD}_\circ$ will allow us to solve some abduction problems that do not admit any solutions in classical logic. The following example adapted from (Rodrigues et al. 2023, Example 5) shows how one can deal with explanations from *classically inconsistent theories*.

Example 1. *Assume that a valuable item was stolen and we have contradictory information about the theft: it was stolen by either Paula or Quinn, but both of them claim to have an alibi. This can be represented with $\Gamma = \{p \vee q, \neg p, \neg q\}$. Classically, there is no explanation for p nor q w.r.t. the theory Γ. However, in $\mathsf{BD}_\circ$, we can explain q by assuming that* Paula's alibi was confirmed by a reliable source, *which can be represented by the formula $\circ p$. This abduction problem can also be solved in $\mathsf{BD}_\triangle$ as follows: q can be explained by assuming that* Paula's alibi was not disputed — *$\neg\triangle p$.*

Contributions In this paper, we formalize abductive reasoning in BD and explore its computational properties. We consider abductive solutions defined as terms in $\mathsf{BD}_\triangle$ and $\mathsf{BD}_\circ$ and compare the classes of abduction problems that can be solved in $\mathsf{BD}_\triangle$ and $\mathsf{BD}_\circ$. We establish an (almost complete) picture of the complexity of standard abductive reasoning tasks (solution recognition, solution existence, and relevance and necessity of hypotheses) in $\mathsf{BD}_\triangle$ and $\mathsf{BD}_\circ$, for two entailment-based notions of minimality. We show how BD abduction problems can be reduced to abduction in classical propositional logic (and vice-versa), which makes it possible to apply classical consequence-finding methods to generate abductive solutions in $\mathsf{BD}_\triangle$ and $\mathsf{BD}_\circ$.

Plan of the Paper The remainder of the text is organised as follows. In Section 2, we formally introduce BD, $\mathsf{BD}_\circ$, and $\mathsf{BD}_\triangle$ and discuss their semantical and computational properties. In Section 3, we present the notions of abduction problem and explanation in expansions of BD. Sections 4 and 5 are dedicated to the complexity of solving BD abduction problems. Section 6 discusses embeddings of BD abduction problems into the classical framework. Finally, in Section 7, we summarise the paper's results and outline a plan for future work. Due to limited space, some proofs have been put in the appendix of the extended version (Bienvenu, Inoue, and Kozhemiachenko 2024).

2 BD and its Expansions

Since BD, $\mathsf{BD}_\circ$, and $\mathsf{BD}_\triangle$ use the same set of truth values and the same $\neg$, $\wedge$, and $\vee$, we present their syntax and semantics together in the following definition.

Definition 1. *The language $\mathcal{L}_{\circ,\triangle}$ is constructed from a fixed countable set of propositional variables* Prop *via the following grammar:*

$$\mathcal{L}_{\circ,\triangle} \ni \phi \coloneqq p \in \mathsf{Prop} \mid \neg\phi \mid (\phi \wedge \phi) \mid (\phi \vee \phi) \mid \circ\phi \mid \triangle\phi$$

In what follows, we use $\mathcal{L}_{\mathsf{BD}}$, $\mathcal{L}_\circ$ and $\mathcal{L}_\triangle$ to denote the fragments of $\mathcal{L}_{\circ,\triangle}$ over $\{\neg,\wedge,\vee\}$, $\{\neg,\wedge,\vee,\circ\}$, and $\{\neg,\wedge,\vee,\triangle\}$, respectively. We will also use $\bullet\phi$ as a shorthand for $\neg\circ\phi$ and use $\mathsf{Prop}(\phi)$ to denote the set of all variables occurring in a formula ϕ.

We set $\mathbf{4} = \{\mathbf{T},\mathbf{B},\mathbf{N},\mathbf{F}\}$ and define a BD valuation *as a mapping $v : \mathsf{Prop} \to \mathbf{4}$ that is extended to complex formulas as follows.*

$\wedge$	**T**	**B**	**N**	**F**
T	**T**	**B**	**N**	**F**
B	**B**	**B**	**F**	**F**
N	**N**	**F**	**N**	**F**
F	**F**	**F**	**F**	**F**

	$\neg$
T	**F**
B	**B**
N	**N**
F	**T**

	$\circ$
T	**T**
B	**F**
N	**F**
F	**T**

$\vee$	**T**	**B**	**N**	**F**
T	**T**	**T**	**T**	**T**
B	**T**	**B**	**T**	**B**
N	**T**	**T**	**N**	**N**
F	**T**	**B**	**N**	**F**

	$\triangle$
T	**T**
B	**T**
N	**F**
F	**F**

	$\bullet$
T	**F**
B	**T**
N	**T**
F	**F**

A formula ϕ is BD-valid, *written* $\mathsf{BD} \models \phi$, *iff $\forall v : v(\phi) \in \{\mathbf{T},\mathbf{B}\}$, and ϕ is* BD-satisfiable *iff $\exists v : v(\phi) \in \{\mathbf{T},\mathbf{B}\}$. A set of formulas Γ* entails *χ, written $\Gamma \models_{\mathsf{BD}} \chi$, iff*

$$\forall v\, [(\forall\phi \in \Gamma\ v(\phi) \in \{\mathbf{T},\mathbf{B}\}) \Rightarrow v(\chi) \in \{\mathbf{T},\mathbf{B}\}]$$

We will henceforth use $\phi \simeq \chi$ (ϕ and χ are weakly equivalent*) to denote that ϕ and χ entail one another and $\phi \equiv \chi$ (ϕ and χ are* strongly equivalent*) to denote that ϕ and χ have the same Belnapian value in every* BD *valuation.*

Let us briefly discuss the intuitive interpretation of $\circ$ and $\bullet$. Since we do not model sources and multiple agents, 'the information concerning p is reliable' ($\circ p$) can be construed as an assumption on the part of the reasoning agent that the available information concerning p is trustworthy, meaning that if we have the information that p holds, then we cannot have evidence for its negation, not p (and likewise, if we have information that p doesn't hold, there is no evidence for p holding). The negation, $\bullet p$, which is termed 'unreliable' but may be more precisely phrased as 'unavailable or unreliable', then covers two cases: the absence of information to support or refute p (so, neither p nor its negation can be inferred) or the presence of contradictory information (both the statement and its negation follow).

It is important to note that there is no BD-valid $\phi \in \mathcal{L}_{\mathsf{BD}}$ since $\mathbf{N}$ is preserved by $\neg$, $\wedge$, and $\vee$. In $\mathsf{BD}_\circ$ and $\mathsf{BD}_\triangle$, however, we can define

$$\top_{\mathsf{BD}_\circ} \coloneqq \circ\circ p \qquad \bot_{\mathsf{BD}_\circ} \coloneqq \bullet\circ p$$
$$\top_{\mathcal{L}_\triangle} \coloneqq \triangle p \vee \neg\triangle p \qquad \bot_{\mathcal{L}_\triangle} = \triangle p \wedge \neg\triangle p$$

and check that for every valuation v,

$$v(\top_{\mathsf{BD}_\circ}) = v(\top_{\mathsf{BD}_\triangle}) = \mathbf{T} \qquad v(\bot_{\mathsf{BD}_\circ}) = v(\bot_{\mathsf{BD}_\triangle}) = \mathbf{F}$$

One can also notice that $\mathsf{BD}_\circ$ is less expressive than $\mathsf{BD}_\triangle$. Indeed, $\circ\phi \equiv (\triangle\phi \wedge \neg\triangle\neg\phi) \vee (\triangle\neg\phi \wedge \neg\triangle\phi)$, while on the other hand, $\triangle$ cannot be defined via $\circ$ (Omori and Sano 2015, Corollaries 6.1 and 6.24). It is also easy to check that distributive and De Morgan laws hold w.r.t. $\neg$, $\wedge$, and $\vee$, and that the following equivalences hold for $\triangle$:

$$\triangle\triangle\phi \equiv \triangle\phi \qquad \neg\triangle\neg\triangle\phi \equiv \triangle\phi$$
$$\triangle(\phi \wedge \chi) \equiv \triangle\phi \wedge \triangle\chi \qquad \triangle(\phi \vee \chi) \equiv \triangle\phi \vee \triangle\chi \tag{1}$$

Thus, every $\phi \in \mathcal{L}_\triangle$ can be transformed into a strongly equivalent formula $\mathsf{NNF}(\phi)$ in *negation normal form*, i.e., built from literals of the form p, $\neg p$, $\triangle p$, $\neg\triangle p$, $\triangle\neg p$, and $\neg\triangle\neg p$ using $\wedge$ and $\vee$.

One can also show that in $\mathsf{BD}_\circ$ contraposition holds w.r.t. $\neg$ while in $\mathsf{BD}_\triangle$ w.r.t. $\neg\triangle$ and that the deduction theorem can be recovered using $\triangle$.

Proposition 1. *Let $\phi, \chi \in \mathcal{L}_\circ$ and $\varrho, \sigma, \tau \in \mathcal{L}_\triangle$. Then the following statements hold.*

1. $\phi \models_{\mathsf{BD}} \chi$ *iff* $\neg\chi \models_{\mathsf{BD}} \neg\phi$.
2. $\varrho, \sigma \models_{\mathsf{BD}} \tau$ *iff* $\varrho, \neg\triangle\tau \models_{\mathsf{BD}} \neg\triangle\sigma$.
3. $\varrho, \sigma \models_{\mathsf{BD}} \tau$ *iff* $\varrho \models_{\mathsf{BD}} \neg\triangle\sigma \vee \tau$.

Note that Proposition 1 entails that $\phi \simeq \chi$ iff $\phi \equiv \chi$ in $\mathsf{BD}_\circ$. On the other hand, this is not the case in $\mathsf{BD}_\triangle$: $p \simeq \triangle p$ but p is not strongly equivalent to $\triangle p$. Still, we can show that some $\triangle$'s can be removed from formulas while preserving weak equivalence.

Definition 2. *Let $\phi \in \mathcal{L}_{\circ,\triangle}$. We use $\phi^\flat$ to denote the result of removing $\triangle$'s that are* not in the scope of $\neg$ *from* $\mathsf{NNF}(\phi)$.

Proposition 2. *Let $\phi \in \mathcal{L}_{\circ,\triangle}$. Then, $\phi \simeq \phi^\flat$.*

We finish the section by establishing faithful embeddings of CPL into $\mathsf{BD}_\circ$, $\mathsf{BD}_\triangle$, and BD.

Proposition 3. *Let $\phi \in \mathcal{L}_{\mathsf{BD}}$. We let $\phi^\triangle$ denote the result of replacing each occurrence of each variable p in ϕ with $\triangle p$ and ϕ° denote the result of replacing p's with $\circ p$'s. Then ϕ is* CPL*-valid iff $\phi^\triangle$ is* BD*-valid iff ϕ° is* BD*-valid.*

Finally, CPL can be embedded even in BD itself.

Proposition 4. *Let $\phi, \chi \in \mathcal{L}_{\mathsf{BD}}$ and $\mathsf{Prop}[\{\phi, \chi\}] = \{p_1, \ldots, p_n\}$. Then $\phi \models_{\mathsf{CPL}} \chi$ iff*

$$\phi \wedge \bigwedge_{i=1}^{n} (p_i \vee \neg p_i) \models_{\mathsf{BD}} \chi \vee \bigvee_{i=1}^{n} (p_i \wedge \neg p_i).$$

Notice that the embeddings of CPL into $\mathsf{BD}_\circ$ and $\mathsf{BD}_\triangle$ increase the size of ϕ only linearly. Thus, since CPL-validity is coNP-complete, BD-validity of $\mathcal{L}_\circ$ and $\mathcal{L}_\triangle$ formulas is also coNP-complete. Likewise, BD entailment is coNP-complete even for $\mathcal{L}_{\mathsf{BD}}$-formulas since the embedding of $\models_{\mathsf{CPL}}$ into $\models_{\mathsf{BD}}$ increases the size of formulas only linearly.[1]

3 Abduction in $\mathsf{BD}_\triangle$ and $\mathsf{BD}_\circ$

We begin the presentation of abduction in $\mathsf{BD}_\triangle$ and $\mathsf{BD}_\circ$ with definitions of literals, terms, and clauses in $\mathcal{L}_\circ$ and $\mathcal{L}_\triangle$. Note that since $\circ$ and $\bullet$ do not distribute over $\wedge$ and $\vee$, we cannot assume that in $\mathcal{L}_\circ$, literals do not contain binary connectives, terms are $\vee$-free and clauses are $\wedge$-free *if we want every formula to be representable as a conjunction of clauses or a disjunction of terms*.

Definition 3 (Literals, terms, and clauses)**.**

- Propositional literal *is a variable p or its negation $\neg p$.*
- $\mathcal{L}_\triangle$-literal *has one of the following forms: p, $\neg p$, $\triangle p$, $\neg\triangle p$, $\triangle\neg p$, $\neg\triangle\neg p$ ($p \in \mathsf{Prop}$).* $\mathcal{L}_\triangle$-clause *is a disjunction of literals;* $\mathcal{L}_\triangle$-term *is a conjunction of literals.*
- $\mathcal{L}_\circ$-literal *has one of the following forms: p, $\neg p$, $\circ\phi$, $\bullet\phi$ ($\phi \in \mathcal{L}_\circ$). Clauses and terms are defined as above.*

The next statement is immediate.

Proposition 5. *Let $\phi \in \mathcal{L}_\circ$ and $\chi \in \mathcal{L}_\triangle$. Then there are conjunctions of $\mathcal{L}_\circ$- and $\mathcal{L}_\triangle$-clauses $\mathsf{DNF}(\phi)$ and $\mathsf{DNF}(\chi)$, and disjunctions of $\mathcal{L}_\circ$- and $\mathcal{L}_\triangle$-terms $\mathsf{CNF}(\phi)$ and $\mathsf{CNF}(\chi)$ s.t. $\phi \equiv \mathsf{DNF}(\phi) \equiv \mathsf{CNF}(\phi)$ and $\chi \equiv \mathsf{DNF}(\chi) \equiv \mathsf{CNF}(\chi)$.*

Note, however, that even though the definition of clauses and terms we gave above allows for a strongly equivalent representation of every $\mathcal{L}_\circ$-formula, the clauses and terms may be difficult to interpret in natural language. Indeed, while $\circ p$ ($\bullet p$) can be understood as 'information concerning p is (un)reliable', a formula such as $\bullet(p \wedge \circ(\neg q \vee \bullet(r \wedge \neg s)))$ does not have any obvious natural-language interpretation. It thus makes sense to consider *atomic* $\mathcal{L}_\circ$-literals.

Definition 4. Atomic $\mathcal{L}_\circ$-literals *are formulas of a form p, $\neg p$, $\circ p$*[2]*, or $\bullet p$ with $p \in$* Prop. Atomic $\mathcal{L}_\circ$-clauses *are disjunctions of atomic literals and* atomic $\mathcal{L}_\circ$-terms *are conjunctions of atomic literals.*

Convention 1. *For a propositional literal l, we set $\overline{p} = \neg p$, $\overline{\neg p} = p$ and $\overline{\overline{l}} = l$. Given a formula ϕ, we use $\mathsf{Lit}(\phi)$, $\mathsf{Lit}_\circ(\phi)$, and $\mathsf{Lit}_\triangle(\phi)$ to denote the sets of propositional literals, atomic $\mathcal{L}_\circ$-literals, and $\mathcal{L}_\triangle$-literals occurring in ϕ.*

We can now define abduction problems and solutions.

[1] Membership in coNP is immediate since BD and the considered expansions have truth-table semantics.

[2] Note that since $\circ\neg\phi \equiv \circ\phi$, we do not need to consider literals $\circ\neg p$ and $\bullet\neg p$.

Definition 5 (Abduction problems and solutions)**.**

- *A* BD abduction problem *is a tuple* $\mathbb{P} = \langle\Gamma, \psi, \mathsf{H}\rangle$ *with* $\Gamma \cup \{\psi\} \subseteq \mathcal{L}_{\mathsf{BD}}$, $\Gamma \not\models_{\mathsf{BD}} \psi$, *and* H *is a finite set of* $\mathcal{L}_\triangle$*-literals or atomic* $\mathcal{L}_\circ$*-literals. If* H *is not restricted, we will omit it for brevity. We call* Γ *a* theory*, members of* H hypotheses*, and* χ *an* observation*.*
- *An* $\mathcal{L}_\circ$-solution *of* $\mathbb{P}$ *is an atomic* $\mathcal{L}_\circ$*-term* τ *composed from the literals in* H *s.t.* $\Gamma, \tau \models_{\mathsf{BD}} \psi$ *and* $\Gamma, \tau \not\models_{\mathsf{BD}} \bot$*.*
- *An* $\mathcal{L}_\triangle$-solution *is an* $\mathcal{L}_\triangle$*-term composed from the literals in* H *s.t.* $\Gamma, \tau \models_{\mathsf{BD}} \psi$ *and* $\Gamma, \tau \not\models_{\mathsf{BD}} \bot$*.*
- *A solution* τ *is* proper *if* $\tau \not\models_{\mathsf{BD}} \psi$*.*
- *A proper solution* τ *is* $\models_{\mathsf{BD}}$-minimal *if there is no proper solution* ϕ *s.t.* $\phi \not\simeq \tau$, $\tau \models_{\mathsf{BD}} \phi$*.*
- *A proper solution* τ *is* theory-minimal *if there is no proper solution* ϕ *s.t.* $\Gamma, \phi \not\models_{\mathsf{BD}} \tau$ *and* $\Gamma, \tau \models_{\mathsf{BD}} \phi$*.*

Convention 2. *Given an abduction problem* $\mathbb{P}$*, we will use* $\mathcal{S}(\mathbb{P})$, $\mathcal{S}^{\mathsf{p}}(\mathbb{P})$, $\mathcal{S}^{\mathsf{BD}}(\mathbb{P})$*, and* $\mathcal{S}^{\mathsf{Th}}(\mathbb{P})$ *to denote the sets of all solutions, all proper solutions, all* $\models_{\mathsf{BD}}$*-minimal solutions, and all theory-minimal solutions of* $\mathbb{P}$*, respectively.*

Convention 3. *To simplify the presentation of examples, we shall sometimes omit* H *when specifying abduction problems. In such cases, it is assumed that* H *contains all possible* $\mathcal{L}_\triangle$*-literals or atomic* $\mathcal{L}_\circ$*-literals (depending on which language we are considering).*

Observe that we have two notions of minimality. The first one ($\models_{\mathsf{BD}}$-minimality or, entailment-minimality) generalises *subset-minimality* by (Eiter and Gottlob 1995) to the BD setting. As we will see in Section 4.1, even though entailment between $\mathcal{L}_\triangle$- and atomic $\mathcal{L}_\circ$-terms is polynomially decidable, it is not equivalent to the containment of one term in the other, whence we need a more general criterion. This approach to minimality is also presented by (Aliseda 2006). Theory-minimal solutions are, essentially, *least specific* solutions in the terminology of (Stickel 1990; Sakama and Inoue 1995) or *least presumptive* solutions in the terminology of (Poole 1989). Theory-minimal solutions can also be seen as duals of *theory prime implicates* from (Marquis 1995).

In addition, it is easy to see that even though a theory-minimal solution is $\models_{\mathsf{BD}}$-minimal, the converse need not hold. Indeed, consider $\mathbb{P} = \langle\{p \vee q, r\}, q \wedge r\rangle$. In both $\mathcal{L}_\triangle$ and $\mathcal{L}_\circ$, there are two $\models_{\mathsf{BD}}$-minimal solutions: $\neg\triangle p$ and q in $\mathcal{L}_\triangle$, and $\neg p \wedge \circ p$ and q in $\mathcal{L}_\circ$. But $\neg\triangle p$ and $\neg p \wedge \circ p$ are *not theory-minimal*: $p \vee q, r, \neg\triangle p \models_{\mathsf{BD}} q$ and $p \vee q, r, \neg p \wedge \circ p \models_{\mathsf{BD}} q$.

One can also notice that allowing *any* $\mathcal{L}_\circ$-terms (rather than only atomic $\mathcal{L}_\circ$-terms) results in even weaker solutions. Consider for example the following problem $\mathbb{P} = \langle\Gamma, \chi, \mathsf{H}\rangle$ (for the sake of discussion, we locally abuse notation and terminology by admitting non-atomic $\mathcal{L}_\circ$-terms as solutions).

$$\Gamma = \{(p \wedge p') \vee (q \wedge q'), \neg(p \wedge p')\} \qquad \chi = q \wedge q'$$
$$\mathsf{H} = \{p, p', \circ p, \circ p', \circ(p \wedge p')\} \tag{2}$$

It is clear that $\circ(p \wedge p')$ and $\circ p \wedge \circ p'$ solve (2) and that $\circ p \wedge \circ p' \models_{\mathsf{BD}} \circ(p \wedge p')$. Moreover, one can see that $\circ(p \wedge p')$ is a theory-minimal solution. Furthermore, some abduction problems cannot be solved if we only allow atomic terms. Indeed, one can check that no atomic $\mathcal{L}_\circ$-term over p and q properly solves $\langle\{p \vee q\}, (p \vee \neg p) \wedge (q \vee \neg q)\rangle$. On the other hand, $\circ p \wedge \circ((p \vee \neg p) \wedge (q \vee \neg q))$ is a solution.

In what follows, we will illustrate the differences between abduction in $\mathsf{BD}_\circ$, $\mathsf{BD}_\triangle$, and classical logic. First, we can observe that some problems have abductive solutions both in $\mathcal{L}_\triangle$ and $\mathcal{L}_\circ$ (cf. Example 1). On the other hand, some problems can be solved only in $\mathcal{L}_\circ$. That is, there are no solutions in the form of $\mathcal{L}_\triangle$-terms[3] even though there are $\mathcal{L}_\circ$-terms that solve the problem.

Example 2. *As in Example 1, either Paula or Quinn is culpable, but now there is also evidence that implicates Paula —* p*. In this case, we want to justify that* Paula is innocent *—* $\neg p$*. There is, of course, no proper classical solution for* $\langle\{p \vee q, p\}, \neg p\rangle$*. Likewise, one can check that this problem admits no proper solutions in* $\mathcal{L}_\triangle$*, as* $\tau \models_{\mathsf{BD}} \neg p$ *for any* $\mathcal{L}_\triangle$*-term* τ *s.t.* $p \vee q, p, \tau \models_{\mathsf{BD}} \neg p$*.*

How can we solve the problem in $\mathsf{BD}_\circ$*? We can add the* $\mathcal{L}_\circ$*-term* $\bullet p$*, i.e., assume that the* evidence against Paula is unreliable*. This way, we have* $p \vee q, p, \bullet p \models_{\mathsf{BD}} \neg p$*, which is justified since one must not be convicted on unreliable evidence. It can be verified that* $\bullet p$ *is the unique* BD*-minimal proper solution for* $\langle\{p \vee q, p\}, \neg p\rangle$*.*

Conversely, some problems can be solved only in $\mathcal{L}_\triangle$.

Example 3. *Consider* $\mathbb{P} = \langle\{p \vee \neg p \vee q\}, q\rangle$*. To explain q, one must assume that p has value* $\mathbf{N}$*. Formally, this means that we assume* $\neg\triangle p \wedge \neg\triangle\neg p$*. One can check that this is a theory-minimal proper solution. On the other hand, it is easy to see that there is no atomic* $\mathcal{L}_\circ$*-term (and, in fact, no* $\mathcal{L}_\circ$*-formula at all) that can solve* $\langle\{p \vee \neg p \vee q\}, q\rangle$*.*

Since $\mathcal{L}_\circ$- and $\mathcal{L}_\triangle$-solutions to BD abduction problems are incomparable, it makes sense to ask which $\mathcal{L}_\circ$-solutions can be represented as $\mathcal{L}_\triangle$-solutions and vice versa. In the remainder of the section, we answer this question.

Definition 6.

- *A satisfiable* $\mathcal{L}_\triangle$*-term* τ *is* $\mathbf{N}$-free *if for every* $\neg\triangle l$ *occurring in* τ*,* $\triangle\bar{l}$ *or* $\bar{l}$ *also occurs in* τ*.*
- *A satisfiable atomic* $\mathcal{L}_\circ$*-term* τ *is* determined *if for every* p *s.t.* $\circ p$ *or* $\bullet p$ *occurs in* τ*,* p *or* $\neg p$ *also occurs in* τ*.*
- *An* $\mathcal{L}_\circ$*-solution (resp.* $\mathcal{L}_\triangle$*-solution)* ϱ *of a* BD *abduction problem* $\mathbb{P}$ *is* $\mathcal{L}_\triangle$*-representable (resp.* $\mathcal{L}_\circ$*-representable) if there exists an* $\mathcal{L}_\triangle$*-solution (resp.* $\mathcal{L}_\circ$*-solution)* σ *of* $\mathbb{P}$ *s.t.* $\varrho \simeq \sigma$*.*

Theorem 1.

1. *An* $\mathcal{L}_\circ$*-solution is* $\mathcal{L}_\triangle$*-representable iff it is determined.*
2. *An* $\mathcal{L}_\triangle$*-solution is* $\mathcal{L}_\circ$*-representable iff it is* $\mathbf{N}$*-free.*

Proof. For Statement 1, let τ be an $\mathcal{L}_\circ$-solution of some abduction problem. By definition, τ is satisfiable. Suppose that τ is determined. As $p \wedge \neg p \equiv \neg p \wedge \bullet p \equiv p \wedge \bullet p \equiv p \wedge \neg p \wedge \bullet p$, we can assume w.l.o.g. that τ has the following form:

$$\tau = \bigwedge_{i=1}^{m}(l_i \wedge \circ l_i) \wedge \bigwedge_{i'=1}^{m'}(l'_{i'} \wedge \bullet l'_{i'}) \wedge \bigwedge_{j=1}^{n} l''_j$$

[3] Of course, there are $\mathcal{L}_\triangle$*-formulas* that can solve the problem but we are interested in solutions in the form of terms.

where the l_i, $l'_{i'}$, and l''_j are propositional literals. Now, we observe that $l \wedge \circ l \equiv \triangle l \wedge \neg\triangle\bar{l}$. Thus, τ can be represented by the following $\mathbf{N}$-free term $\tau^{\circ\triangle}$:

$$\tau^{\circ\triangle} = \bigwedge_{i=1}^{m}(\triangle l_i \wedge \neg\triangle\overline{l_i}) \wedge \bigwedge_{i'=1}^{m'}(l'_{i'} \wedge \overline{l'_{i'}}) \wedge \bigwedge_{j=1}^{n} l''_j$$

For the converse, suppose τ is not determined, i.e., there is some $\circ p$ s.t. neither p nor $\neg p$ occur in τ or there is some $\bullet q$ s.t. neither q not $\neg q$ occur in τ. By examining the truth table semantics from Definition 1, we can see that there is no conjunction of $\mathcal{L}_\triangle$-literals that is weakly equivalent to $\circ p$ or $\bullet q$. Hence, τ is not $\mathcal{L}_\triangle$-representable.

For Statement 2, consider an $\mathcal{L}_\triangle$-solution τ. First, suppose that τ is an $\mathbf{N}$-free term, and let $\tau^\flat$ be as in Definition 2. We have that $\tau^\flat \simeq \tau$ with $\tau^\flat$ having the following form:

$$\tau^\flat = \bigwedge_{i=1}^{m}(l_i \wedge \neg\triangle\bar{l}_i) \wedge \bigwedge_{j=1}^{n} l'_j$$

It is clear that the following atomic $\mathcal{L}_\circ$-term represents $\tau^\flat$ (and hence, τ):

$$\tau^{\triangle\circ} = \bigwedge_{i=1}^{m}(l_i \wedge \circ l_i) \wedge \bigwedge_{j=1}^{n} l'_j$$

For the converse, suppose τ is *not* $\mathbf{N}$*-free*, and w.l.o.g. let $\neg\triangle p$ occur in $\tau^\flat$ but p not occur. From (Omori and Sano 2015, Propositions 6.23 and 6.23), we know that $\neg\triangle p$ (by itself, without p) is not definable in $\mathsf{BD}_\circ$. As τ is satisfiable (being a solution), this means τ is not $\mathcal{L}_\circ$-representable. □

We finish the section with a few observations. First, one can see that determined and $\mathbf{N}$-free terms can be recognised in polynomial time. Second, we note that none of the solutions in Examples 1–3 is representable in the other language. In Example 1, $\neg\triangle p$ is not $\mathcal{L}_\circ$-representable and $\circ p$ is not $\mathcal{L}_\triangle$-representable. In Example 2, $\bullet p$ is not $\mathcal{L}_\triangle$-representable. In Example 3, $\neg\triangle p \wedge \neg\triangle\neg p$ is not $\mathcal{L}_\circ$-representable. This shows that even though $\mathsf{BD}_\circ$ is less expressive than $\mathsf{BD}_\triangle$, their sets of solutions are incomparable. Finally, we remark that even if a problem has solutions in both languages, the solutions need not be (weakly or strongly) equivalent, especially if we consider BD- or theory-minimal solutions. Indeed, one can see that the solutions in Example 1 are theory-minimal, yet they are not equivalent. More than that, neither implies the other.

4 Complexity of Term Entailment

This section contains some technical results concerning entailment from $\mathcal{L}_\triangle$- and $\mathcal{L}_\circ$-terms that facilitate the proofs of complexity bounds in Section 5.

4.1 Entailment Between Terms

We begin with the complexity of entailment between terms. Recall from Definition 5 that to establish the $\models_{\mathsf{BD}}$-minimality of τ, we need to check whether there is another solution ϕ that is entailed by τ. In the next two theorems, we show that the entailment of atomic $\mathcal{L}_\circ$-terms and $\mathcal{L}_\triangle$-terms is recognisable in polynomial time.

Theorem 2. *Entailment between atomic $\mathcal{L}_\circ$-terms is decidable in deterministic polynomial time.*

Proof sketch. Let σ and σ' be atomic $\mathcal{L}_\circ$-terms. We begin by noting that σ is BD-unsatisfiable iff (i) p, $\neg p$, and $\circ p$ occur in σ; or (ii) $\circ p$ and $\bullet p$ occur in σ. The 'only if' direction is evident since $p \wedge \neg p \wedge \circ p$ and $\circ p \wedge \bullet p$ are unsatisfiable. For the 'if' direction, assume that there is no variable p s.t. p, $\neg p$, and $\circ p$ occur in σ, nor any variable q s.t. $\circ q$ and $\bullet q$ occur in σ. We construct a satisfying valuation v as follows:

- $v(r) = \mathbf{T}$ iff r occurs in σ but $\neg r$ and $\bullet r$ *do not*;
- $v(r) = \mathbf{F}$ iff $\neg r$ occurs in σ but r and $\bullet r$ do not;
- $v(r) = \mathbf{B}$ otherwise.

It is clear that $v(\sigma) \in \{\mathbf{T}, \mathbf{B}\}$. Indeed, it is easy to check that every *literal* occurring in σ has value $\mathbf{T}$ or $\mathbf{B}$ and that v is well defined.

It follows from this characterisation that the satisfiability of an atomic $\mathcal{L}_\circ$-term can be decided in polynomial time. In addition, observe that $\sigma \models_{\mathsf{BD}} \sigma'$ if σ is unsatisfiable and $\sigma \not\models_{\mathsf{BD}} \sigma'$ is σ is satisfiable but σ' is not. Finally, to show that entailment between *satisfiable* atomic $\mathcal{L}_\circ$-terms is decidable in polynomial time, we use the facts that $p \wedge \neg p \models_{\mathsf{BD}} \bullet p$ and $p \wedge \neg p \equiv p \wedge \bullet p \equiv \neg p \wedge \bullet p$. □

To establish a similar property of $\mathcal{L}_\triangle$-terms, we construct a faithful embedding of $\mathcal{L}_\triangle$-formulas into the language of CPL. The polynomial complexity evaluation of term entailment will follow since our embedding increases the size of formulas only *linearly* and since determining entailment of terms in CPL can also be done in polynomial time.

Definition 7. *Let $\phi \in \mathcal{L}_\triangle$ be in* NNF *and let $\sim$*[4] *denote the* classical negation*. We define ϕ^{cl} as follows.*

$$p^{\mathsf{cl}} = p^+ \qquad (\neg p)^{\mathsf{cl}} = p^-$$
$$(\triangle p)^{\mathsf{cl}} = p^+ \qquad (\triangle\neg p)^{\mathsf{cl}} = p^-$$
$$(\neg\triangle p)^{\mathsf{cl}} = {\sim} p^+ \qquad (\neg\triangle\neg p)^{\mathsf{cl}} = {\sim} p^-$$
$$(\chi \wedge \psi)^{\mathsf{cl}} = \chi^{\mathsf{cl}} \wedge \psi^{\mathsf{cl}} \qquad (\chi \vee \psi)^{\mathsf{cl}} = \chi^{\mathsf{cl}} \vee \psi^{\mathsf{cl}}$$

Lemma 1. *Let $\phi,\chi \in \mathcal{L}_\triangle$ be in* NNF*. Then $\phi \models_{\mathsf{BD}} \chi$ iff $\phi^{\mathsf{cl}} \models_{\mathsf{CPL}} \chi^{\mathsf{cl}}$.*

The next theorem follows immediately from Lemma 1.

Theorem 3. *Entailment between $\mathcal{L}_\triangle$-terms is decidable in deterministic polynomial time.*

To check that τ is a theory-minimal solution to $\langle\Gamma, \chi\rangle$, we need to establish that there is no other solution σ s.t. $\Gamma, \tau \models_{\mathsf{BD}} \sigma$ but $\Gamma, \sigma \not\models_{\mathsf{BD}} \tau$. The next results show that the complexity of checking term entailment w.r.t. a background theory differs depending on whether we consider $\mathcal{L}_\circ$ or $\mathcal{L}_\triangle$.

Theorem 4. *It is* coNP*-complete to decide whether $\Gamma, \varrho \models_{\mathsf{BD}} \sigma$, given $\Gamma \subseteq \mathcal{L}_{\mathsf{BD}}$ and atomic $\mathcal{L}_\circ$-terms ϱ and σ.*

[4] While we may assume that CPL is given over $\mathcal{L}_{\mathsf{BD}}$, we sometimes use $\sim$ rather than $\neg$ to make clear we are working in CPL.

Decision problem	$\mathcal{L}_\triangle$	$\mathcal{L}_\circ$
$\tau \models_{\mathsf{BD}} \psi$?	in P	coNP
$\tau \models_{\mathsf{BD}} \sigma$?	in P	in P
$\Gamma, \tau \models_{\mathsf{BD}} \sigma$?	in P	coNP
$\Gamma, \tau \not\models_{\mathsf{BD}} \bot$?	in P	NP
$\Gamma, \tau \models_{\mathsf{BD}} \psi$?	coNP	coNP

Table 1: Complexity of entailment tasks, where $\Gamma \cup \{\psi\} \subseteq \mathcal{L}_{\mathsf{BD}}$, and σ and τ are terms in the considered language ($\mathcal{L}_\triangle$ or $\mathcal{L}_\circ$). Unless specified otherwise, all results are completeness results.

Proof. Membership is evident since $\mathsf{BD}_\circ$ has truth-table semantics. For hardness, observe that Γ is *classically unsatisfiable* iff $\Gamma, \bigwedge\limits_{p \in \mathsf{Prop}[\Gamma]} \circ p \models_{\mathsf{BD}} q$ with $q \notin \mathsf{Prop}[\Gamma]$. □

In $\mathcal{L}_\triangle$, however, this problem is tractable.

Theorem 5. *It can be decided in polynomial time whether* $\Gamma, \varrho \models_{\mathsf{BD}} \sigma$ *, given* $\Gamma \subseteq \mathcal{L}_{\mathsf{BD}}$ *and* $\mathcal{L}_\triangle$*-terms* ϱ *and* σ*.*

Proof. It suffices to show the result for the case where σ is an $\mathcal{L}_\triangle$-literal. Due to Propositions 1 and 2, $\Gamma, \varrho \models_{\mathsf{BD}} \sigma$ iff $\Gamma, \varrho, (\neg\triangle\sigma)^\flat \models_{\mathsf{BD}} \bot$, so we may focus on solving the latter task. First, we assume w.l.o.g. that all formulas in Γ are in NNF and that $\varrho \wedge (\neg\triangle\sigma)^\flat$ is BD-satisfiable (this can be checked in polynomial time by Theorem 3). By Lemma 1, we have that $\Gamma, \varrho, (\neg\triangle\sigma)^\flat \models_{\mathsf{BD}} \bot$ iff $\Gamma^{\mathsf{cl}}, \varrho^{\mathsf{cl}}, ((\neg\triangle\sigma)^\flat)^{\mathsf{cl}} \models_{\mathsf{CPL}} \bot$. Since $\Gamma \subseteq \mathcal{L}_{\mathsf{BD}}$, Γ^{cl} is $\sim$-free (cf. Definition 7). As $\varrho \wedge (\neg\triangle\sigma)^\flat$ is BD-satisfiable by assumption, then so is $\varrho^{\mathsf{cl}} \wedge ((\neg\triangle\sigma)^\flat)^{\mathsf{cl}}$ (by Lemma 1). Thus, there is no variable r s.t. r and $\sim r$ both occur in $\varrho^{\mathsf{cl}} \wedge ((\neg\triangle\sigma)^\flat)^{\mathsf{cl}}$. Now take Γ^{cl} and substitute every p that occurs in $\varrho^{\mathsf{cl}} \wedge ((\neg\triangle\sigma)^\flat)^{\mathsf{cl}}$ positively (resp. negatively) with $\top$ (resp. $\bot$). We then exhaustively apply the following CPL-equivalence-preserving transformations to all subformulas of $\bigwedge\limits_{\phi \in \Gamma^{\mathsf{cl}}} \phi$, denoting the result by $(\Gamma^{\mathsf{cl}})^\sharp$:

$$\top \wedge \psi \rightsquigarrow \psi \quad \top \vee \psi \rightsquigarrow \top \quad \bot \wedge \psi \rightsquigarrow \bot \quad \bot \vee \psi \rightsquigarrow \psi \tag{3}$$

Clearly, $(\Gamma^{\mathsf{cl}})^\sharp$ can be computed in polynomial time in the size of Γ^{cl}. Moreover, if $(\Gamma^{\mathsf{cl}})^\sharp = \bot$, then $\Gamma^{\mathsf{cl}}, \varrho^{\mathsf{cl}}, ((\neg\triangle\sigma)^\flat)^{\mathsf{cl}} \models_{\mathsf{CPL}} \bot$ holds. To complete the proof, we show that $\Gamma^{\mathsf{cl}}, \varrho^{\mathsf{cl}}, ((\neg\triangle\sigma)^\flat)^{\mathsf{cl}} \not\models_{\mathsf{CPL}} \bot$ when $(\Gamma^{\mathsf{cl}})^\sharp \neq \bot$. Let us assume $(\Gamma^{\mathsf{cl}})^\sharp \neq \bot$ and set $v(r) = \mathbf{T}$ for every $r \in \mathsf{Prop}((\Gamma^{\mathsf{cl}})^\sharp)$. It is clear that $v((\Gamma^{\mathsf{cl}})^\sharp) = \mathbf{T}$ since $(\Gamma^{\mathsf{cl}})^\sharp$ is $\sim$-free (cf. Definition 7). Now, we extend v to the variables occurring in ϱ^{cl} and $((\neg\triangle\sigma)^\flat)^{\mathsf{cl}}$ as expected: if s occurs, we set $v(s) = \mathbf{T}$, if $v(\sim s)$ occurs, we set $v(s) = \mathbf{F}$. As $\varrho^{\mathsf{cl}} \wedge ((\neg\triangle\sigma)^\flat)^{\mathsf{cl}}$ is CPL-satisfiable, we now have $v\left(\bigwedge\limits_{\phi \in \Gamma^{\mathsf{cl}}} \phi \wedge \varrho^{\mathsf{cl}} \wedge ((\neg\triangle\sigma)^\flat)^{\mathsf{cl}}\right) = \mathbf{T}$, as required. □

4.2 Entailment of Formulas From Terms

Let us now consider the complexity of entailment of $\mathcal{L}_{\mathsf{BD}}$-*formulas*[5] by $\mathcal{L}_\circ$-terms and $\mathcal{L}_\triangle$-terms. We begin with $\mathcal{L}_\circ$-terms. The following statement is straightforward.

Theorem 6. *It is is* coNP*-complete to decide whether* $\phi \models_{\mathsf{BD}} \chi$*, given an atomic* $\mathcal{L}_\circ$*-term* ϕ *and* $\chi \in \mathcal{L}_{\mathsf{BD}}$*.*

Proof. Membership is immediate as $\mathsf{BD}_\circ$-entailment is in coNP. To show coNP-hardness, observe that χ is CPL-valid iff $\bigwedge\limits_{p \in \mathsf{Prop}(\chi)} \circ p \models_{\mathsf{BD}} \chi$ since $\circ p$ ensures $v(p) \in \{\mathbf{T}, \mathbf{F}\}$ and $\wedge$, $\vee$, and $\neg$ behave classically on $\mathbf{T}$ and $\mathbf{F}$. □

On the other hand, if ϕ is an $\mathcal{L}_\triangle$-term, $\phi \models_{\mathsf{BD}} \chi$ can be decided in deterministic polynomial time.

Theorem 7. *It can be decided in polynomial time whether* $\phi \models_{\mathsf{BD}} \chi$*, given an* $\mathcal{L}_\triangle$*-term* ϕ *and* $\chi \in \mathcal{L}_{\mathsf{BD}}$*.*

Proof. Note first, that if $\mathsf{Prop}(\phi) \cap \mathsf{Prop}(\chi) = \varnothing$ and ϕ is satisfiable, then $\phi \not\models_{\mathsf{BD}} \chi$. Indeed, we can just evaluate all variables of χ as $\mathbf{N}$ which will make χ have value $\mathbf{N}$ as well (cf. Definition 1). By Theorem 3, it takes polynomial time to determine whether an $\mathcal{L}_\triangle$-term is satisfiable. Thus, we consider the case when ϕ is satisfiable.

We assume w.l.o.g. that χ is in NNF. By Lemma 1, we have that $\phi \models_{\mathsf{BD}} \chi$ iff $\phi^{\mathsf{cl}} \models_{\mathsf{CPL}} \chi^{\mathsf{cl}}$. Since $\chi \in \mathcal{L}_{\mathsf{BD}}$, χ^{cl} is $\sim$-free (cf. Definition 7). The rest of the proof is similar to that of Theorem 5. The only difference is that we will check whether χ^{cl} is reduced to $\top$ using (3). □

Finally, we observe that in the presence of Γ, the entailment of $\mathcal{L}_{\mathsf{BD}}$-formulas from terms becomes coNP-complete.

Theorem 8. *It is* coNP*-complete to decide whether* $\Gamma, \tau \models_{\mathsf{BD}} \psi$*, given* $\Gamma \cup \{\psi\} \in \mathcal{L}_{\mathsf{BD}}$ *and an* $\mathcal{L}_\triangle$*-term or an atomic* $\mathcal{L}_\circ$*-term* τ*.*

Proof. The proof is a straightforward reduction from the coNP-complete entailment problem for $\mathcal{L}_{\mathsf{BD}}$-formulas: $\phi \models_{\mathsf{BD}} \chi$ iff $\phi \vee p, \neg p \wedge \circ p \models_{\mathsf{BD}} \chi$ iff $\phi \vee p, \neg p \wedge \neg\triangle p \models_{\mathsf{BD}} \chi$. To see why the preceding holds, note that $(\phi \vee p) \wedge (\neg p \wedge \circ p) \equiv \phi$ and $\neg p \wedge \neg\triangle p \equiv \neg p \wedge \circ p$. □

We summarise the results of the section in Table 1. Note that the polynomial decidability of $\Gamma, \tau \not\models_{\mathsf{BD}} \bot$ for $\mathcal{L}_\triangle$-terms follows immediately from Theorem 5. Similarly, NP-completeness of $\Gamma, \tau \not\models_{\mathsf{BD}} \bot$ for atomic $\mathcal{L}_\circ$-terms is a corollary of Theorem 8. To see it, use a fresh p for ψ and set $\tau = \bigwedge\limits_{q \in \mathsf{Prop}[\Gamma]} \circ q$. It is immediate that $\Gamma, \bigwedge\limits_{q \in \mathsf{Prop}[\Gamma]} \circ q \models_{\mathsf{BD}} p$ iff Γ is classically unsatisfiable.

[5]Recall that in BD abduction problems, the theory is formulated in $\mathcal{L}_{\mathsf{BD}}$ while solutions are formulated in $\mathcal{L}_\circ$ or $\mathcal{L}_\triangle$.

5 Complexity of BD Abduction

This section considers the complexity of the principal decision problems related to BD abduction, namely, solution recognition, solution existence, and relevance and necessity of hypotheses. Table 2 summarises the obtained results for BD abduction, alongside results for classical abduction.[6]

We begin with a useful technical lemma.

Definition 8. *We say that* Γ BD*-consistently entails* χ *(*$\Gamma \models^{\mathsf{cons}}_{\mathsf{BD}} \chi$*) iff* Γ *is* BD*-satisfiable and* $\Gamma \models_{\mathsf{BD}} \chi$.

Lemma 2. *Let* $\Gamma \cup \{\chi\} \subseteq \mathcal{L}_{\mathsf{BD}}$*,* σ *be an* $\mathcal{L}_{\triangle}$*-term, and* τ *be an atomic* $\mathcal{L}_{\circ}$*-term. Then*

1. *deciding whether* $\Gamma, \sigma \models^{\mathsf{cons}}_{\mathsf{BD}} \chi$ *is* coNP*-complete;*
2. *deciding whether* $\Gamma, \tau \models^{\mathsf{cons}}_{\mathsf{BD}} \chi$ *is* DP*-complete.*

Proof. We begin with Statement 1. coNP-membership is immediate since entailment is in coNP and verifying the consistency of Γ, σ can be done in polynomial time using Theorem 5. For hardness, we reduce BD-entailment to BD-consistent entailment as follows:

$$\phi \models_{\mathsf{BD}} \chi \text{ iff } \phi \vee p, q \models^{\mathsf{cons}}_{\mathsf{BD}} \chi \vee p \quad (p, q \notin \mathsf{Prop}[\{\phi, \chi\}], \sigma = q)$$

Let $\phi \not\models_{\mathsf{BD}} \chi$. We can thus find v such that $v(\phi) \in \{\mathbf{T}, \mathbf{B}\}$ and $v(\chi) \notin \{\mathbf{T}, \mathbf{B}\}$. Since p and q are fresh, we let $v(p) = \mathbf{F}$ and $v(q) = \mathbf{T}$ which falsifies the consistent entailment. Conversely, let $\phi \vee p, q \not\models^{\mathsf{cons}}_{\mathsf{BD}} \chi \vee p$. It is clear that $\{\phi \vee p, q\}$ is BD-satisfiable. Thus, there is a valuation s.t. $v(\phi \vee p) \in \{\mathbf{T}, \mathbf{B}\}$, $v(q) \in \{\mathbf{T}, \mathbf{B}\}$ but $v(\chi \vee p) \notin \{\mathbf{T}, \mathbf{B}\}$. Hence, $v(\phi) \in \{\mathbf{T}, \mathbf{B}\}$ and $v(\chi) \notin \{\mathbf{T}, \mathbf{B}\}$, as required.

For Statement 2, membership follows immediately from Table 1. To show hardness, we reduce the DP-complete Sat-UnSat problem for CPL to BD-consistent entailment. Let ϕ, χ be propositional formulas and assume w.l.o.g. that $\mathsf{Prop}(\phi) \cap \mathsf{Prop}(\chi) = \varnothing$. Set $\Xi = \mathsf{Prop}(\phi) \cup \mathsf{Prop}(\chi)$ and pick $p \notin \Xi$. We show that ϕ is CPL-satisfiable and χ is CPL-unsatisfiable iff $\phi, p \wedge \bigwedge\limits_{q \in \Xi} \circ q \models^{\mathsf{cons}}_{\mathsf{BD}} p \wedge \neg\chi$.

If ϕ is CPL-unsatisfiable, then $\phi \wedge p \wedge \bigwedge\limits_{q \in \Xi} \circ q$ is BD-unsatisfiable, so the consistent entailment fails. If χ is CPL-satisfiable, let v be a classical valuation s.t. $v(\chi) = \mathbf{T}$ (whence, $v(p \wedge \neg\chi) = \mathbf{F}$) and $v(\phi \wedge p) = \mathbf{T}$ (recall that $\mathsf{Prop}(\phi) \cap \mathsf{Prop}(\chi) = \varnothing$, so such valuation must exist unless ϕ is CPL-unsatisfiable). Again, the consistent entailment fails.

For the converse, let ϕ be CPL-satisfiable and χ is CPL-unsatisfiable. It is clear that $\phi, p \wedge \bigwedge\limits_{q \in \Xi} \circ q \models^{\mathsf{cons}}_{\mathsf{BD}} p \wedge \neg\chi$ because $p \wedge \bigwedge\limits_{q \in \Xi} \circ q \models_{\mathsf{BD}} p \wedge \neg\chi$ and $\phi \wedge p \wedge \bigwedge\limits_{q \in \Xi} \circ q$ is BD-satisfiable as $p \notin \mathsf{Prop}(\phi)$ and ϕ is CPL-satisfiable. □

5.1 Solution Recognition

We use the preceding lemma to establish the complexity of recognising arbitrary and proper solutions.

[6]Results for CPL come from (Eiter and Gottlob 1995) or can be obtained as corollaries of the latter paper or our own results.

Recognition and existence	$\mathcal{L}_{\triangle}$	$\mathcal{L}_{\circ}$	CPL
$\tau \in \mathcal{S}(\mathbb{P})$? / $\tau \in \mathcal{S}^{\mathsf{p}}(\mathbb{P})$?	coNP	DP	DP
$\tau \in \mathcal{S}^{\mathsf{BD}}(\mathbb{P})$?	DP	DP	DP
$\tau \in \mathcal{S}^{\mathsf{Th}}(\mathbb{P})$?	in Π^{P}_2	in Π^{P}_2	in Π^{P}_2
$\mathcal{S}(\mathbb{P}) \neq \varnothing$? / $\mathcal{S}^{\mathsf{p}}(\mathbb{P}) \neq \varnothing$?	Σ^{P}_2	Σ^{P}_2	Σ^{P}_2
Relevance	$\mathcal{L}_{\triangle}$	$\mathcal{L}_{\circ}$	CPL
w.r.t. $\mathcal{S}(\mathbb{P})$, $\mathcal{S}^{\mathsf{p}}(\mathbb{P})$, $\mathcal{S}^{\mathsf{BD}}(\mathbb{P})$	Σ^{P}_2	Σ^{P}_2	Σ^{P}_2
w.r.t. $\mathcal{S}^{\mathsf{Th}}(\mathbb{P})$	in Σ^{P}_3	in Σ^{P}_3	in Σ^{P}_3
Necessity	$\mathcal{L}_{\triangle}$	$\mathcal{L}_{\circ}$	CPL
w.r.t. $\mathcal{S}(\mathbb{P})$, $\mathcal{S}^{\mathsf{p}}(\mathbb{P})$, $\mathcal{S}^{\mathsf{BD}}(\mathbb{P})$	Π^{P}_2	Π^{P}_2	Π^{P}_2
w.r.t. $\mathcal{S}^{\mathsf{Th}}(\mathbb{P})$	in Π^{P}_3	in Π^{P}_3	in Π^{P}_3

Table 2: Complexity of abductive reasoning problems. Unless specified otherwise, all results are completeness results.

Theorem 9. *It is* coNP*-complete to decide, given a* BD *abduction problem* $\mathbb{P}$ *and an* $\mathcal{L}_{\triangle}$*-term* σ*, whether* σ *is a (proper) solution of* $\mathbb{P}$*.*

Proof. coNP-completeness of recognising $\sigma \in \mathcal{S}(\mathbb{P})$ follows immediately from Lemma 2 since σ is a solution of $\langle \Gamma, \chi, \mathsf{H} \rangle$ iff $\Gamma, \chi \models^{\mathsf{cons}}_{\mathsf{BD}} \chi$. For proper solutions, we observe that recognising an *arbitrary* solution is reducible to the recognition of *a proper* solution as follows: if we let $p \notin \mathsf{Prop}[\Gamma \cup \{\chi\} \cup \mathsf{H}]$, then σ is a solution of $\langle \Gamma, \chi, \mathsf{H} \rangle$ iff σ is a *proper solution* of $\langle \Gamma \cup \{p\}, p \wedge \chi, \mathsf{H} \cup \{p\} \rangle$. Indeed, $\Gamma, \sigma \models_{\mathsf{BD}} \chi$ holds iff $\Gamma, p, \sigma \models_{\mathsf{BD}} \chi \wedge p$ holds, and likewise, $\Gamma, p, \sigma \models_{\mathsf{BD}} \bot$ iff $\Gamma, \sigma \models_{\mathsf{BD}} \bot$. Moreover, since p does not occur in τ, it is clear that $\sigma \not\models_{\mathsf{BD}} \chi \wedge p$. This establishes coNP-hardness. For membership, we note that checking the properness condition ($\sigma \not\models_{\mathsf{BD}} \chi$) is in P since σ is an $\mathcal{L}_{\triangle}$-term (Theorem 7), so we remain in coNP. □

Theorem 10. *It is* DP*-complete to decide, given a* BD *abduction problem* $\mathbb{P}$ *and an atomic* $\mathcal{L}_{\circ}$*-term* σ*, whether* σ *is a (proper) solution of* $\mathbb{P}$*.*

Proof. The arguments are the same as in the proof of Theorem 9, but use the complexity results for atomic $\mathcal{L}_{\circ}$-terms from Table 1 and Lemma 2. □

We next show how to recognize $\models_{\mathsf{BD}}$-minimal solutions.

Theorem 11. *It is* DP*-complete to decide, given a* BD *abduction problem* $\mathbb{P}$ *and a term* σ*, whether* σ *is a* $\models_{\mathsf{BD}}$*-minimal (*$\mathcal{L}_{\circ}$*- or* $\mathcal{L}_{\triangle}$*-) solution of* $\mathbb{P}$*.*

Proof sketch. The upper bound exploits the fact that for every term σ, we can identify polynomially many terms $\sigma_1, \ldots, \sigma_k$ s.t. (i) $\sigma \models_{\mathsf{BD}} \sigma_i$ but $\sigma_i \not\models_{\mathsf{BD}} \sigma$ ($1 \leq i \leq k$), and (ii) if σ is a proper solution but not $\models_{\mathsf{BD}}$-minimal, then some σ_i is a (proper) solution. Thus, to verify $\models_{\mathsf{BD}}$-minimality, it suffices to check that none of the σ_i is a solution.

For $\mathcal{L}_{\triangle}$-terms, we may assume (recall Proposition 2) that all $\triangle$'s in σ occur under $\neg$, i.e., $\sigma = \sigma^{\flat}$. In this case, $\sigma^{\flat} \models_{\mathsf{BD}} \sigma'^{\flat}$ iff $\mathsf{Lit}_{\triangle}(\sigma^{\flat}) \supseteq \mathsf{Lit}_{\triangle}(\sigma'^{\flat})$, so we need only to check each of the (linearly many) terms σ^{-l} obtained by deleting one $\mathcal{L}_\triangle$-literal from σ.

For atomic $\mathcal{L}_\circ$-terms, however, we cannot just remove literals because one term may BD-entail another even though they have no common atomic literals, as in $p \wedge \neg p \models_{\mathsf{BD}} \bullet p$. Nevertheless, we can still identify syntactically a polynomial number of candidate better terms, each obtained by picking a variable p occurring in σ and replacing the set of p-literals in σ by a $\models_{\mathsf{BD}}$-weaker set of p-literals. ☐

In the case of theory-minimal solutions, we establish a Π_2^{P} upper bound. We expect that this case is indeed harder than $\models_{\mathsf{BD}}$-minimality, intuitively because the presence of the theory means we cannot readily identify a polynomial number of candidate better solutions to check. We leave the search for a matching lower bound for future work and remark that, to the best of our knowledge, the complexity of the analogous problem in CPL is also unknown.

Theorem 12. *It is in Π_2^{P} to decide, given a* BD *abduction problem $\mathbb{P}$ and an (atomic $\mathcal{L}_\circ$- or $\mathcal{L}_\triangle$-)term σ, whether σ is a theory-minimal solution of $\mathbb{P}$.*

5.2 Solution Existence

We now turn to the fundamental task of determining whether an abduction problem has a solution. To establish the complexity of deciding whether $\mathcal{S}(\mathbb{P}) = \varnothing$, we provide reductions from classical abduction problems. We adapt the definition of *classical abduction problems* from (Eiter and Gottlob 1995; Creignou and Zanuttini 2006) to our notation.

Definition 9. *A* classical abduction problem *is a tuple $\mathbb{P} = \langle \Gamma, \chi, \mathsf{H}\rangle$ s.t. $\Gamma \cup \{\chi\} \subseteq \mathcal{L}_{\mathsf{BD}}$ and* H *is a set of propositional literals.*

- *A* solution *of $\mathbb{P}$ is a conjunction τ of literals from* H *such that $\Gamma, \tau \models_{\mathsf{CPL}} \psi$ and $\Gamma, \tau \not\models_{\mathsf{CPL}} \bot$.*
- *A solution τ is* proper *if $\tau \not\models_{\mathsf{CPL}} \psi$.*
- *A proper solution τ is* theory-minimal *if there is no proper solution ϕ s.t. $\Gamma, \tau \models_{\mathsf{CPL}} \phi$ and $\Gamma, \phi \not\models_{\mathsf{CPL}} \tau$.*
- *A proper solution τ is* $\models_{\mathsf{CPL}}$-minimal *if there is no proper solution ϕ s.t. $\tau \models_{\mathsf{CPL}} \phi$ and* CPL $\not\models \phi \leftrightarrow \tau$.

Theorem 13. *It is Σ_2^{P}-complete to decide whether a* BD *abduction problem has a (proper) $\mathcal{L}_\triangle$- or $\mathcal{L}_\circ$-solution.*

Proof. Membership follows immediately from Theorems 9 and 10. To show Σ_2^{P}-hardness, we reduce solution existence for classical abduction problems $\mathbb{P}_{\mathsf{cl}} = \langle \Gamma_{\mathsf{cl}}, \chi_{\mathsf{cl}}, \mathsf{H}\rangle$ of the following form:[7]

$$\begin{aligned}\Gamma_{\mathsf{cl}} &= \{\neg\phi \vee (p \wedge \tau), \neg p \vee \tau\} \cup \\ &\quad \{\neg r \leftrightarrow r' \mid r \in \mathsf{Prop}(\phi) \setminus \mathsf{Prop}(p \wedge \tau)\} \\ &\quad\quad (p \notin \mathsf{Prop}(\phi \wedge \tau), \tau \text{ is a term}) \\ \chi_{\mathsf{cl}} &= p \wedge \tau \\ \mathsf{H} &= \{r \mid r \in \mathsf{Prop}(\phi) \setminus \mathsf{Prop}(p \wedge \tau)\} \cup \{r' \mid \neg r \leftrightarrow r' \in \Gamma_{\mathsf{cl}}\}\end{aligned} \tag{4}$$

[7] We write $\neg r \leftrightarrow r'$ as a shorthand for $(r \wedge \neg r') \vee (\neg r \wedge r')$.

By (Eiter and Gottlob 1995, Theorem 4.2), determining the existence of classical solutions for these problems is Σ_2^{P}-hard. We reduce $\mathbb{P}_{\mathsf{cl}}$ to $\mathbb{P}^{\mathbf{4}} = \langle \Gamma^{\mathbf{4}}, \chi^{\mathbf{4}}, \mathsf{H}\rangle$, where:

$$\begin{aligned}\Gamma^{\mathbf{4}} &= \Gamma_{\mathsf{cl}} \cup \{q \vee \neg q \mid q \in \mathsf{Prop}[\Gamma_{\mathsf{cl}}]\} \\ \chi^{\mathbf{4}} &= \chi_{\mathsf{cl}} \vee \bigvee_{q \in \mathsf{Prop}[\Gamma_{\mathsf{cl}}]} (q \wedge \neg q)\end{aligned} \tag{5}$$

First let σ be a solution of $\mathbb{P}_{\mathsf{cl}}$. It is immediate from (4) that σ is a *proper* solution because we cannot use variables occurring in χ_{cl}. Moreover, by Proposition 4, we have that σ is a *proper* solution of $\mathbb{P}^{\mathbf{4}}$. And since $\sigma \in \mathcal{L}_{\mathsf{BD}}$, it is both an $\mathcal{L}_\triangle$- and $\mathcal{L}_\circ$-proper solution.

For the converse, let σ' be a solution of $\mathbb{P}^{\mathbf{4}}$. As H contains only positive literals and no variables from $\mathsf{Prop}(p \wedge \tau)$, it follows that $\Gamma_{\mathsf{cl}}, \sigma' \not\models_{\mathsf{CPL}} \bot$ and $\sigma' \not\models_{\mathsf{CPL}} p \wedge \tau$. Moreover, applying Proposition 4, we obtain that $\Gamma_{\mathsf{cl}}, \sigma' \models_{\mathsf{CPL}} \chi_{\mathsf{cl}}$. This shows that σ' is a proper solution of $\mathbb{P}_{\mathsf{cl}}$. ☐

5.3 Relevance and Necessity of Hypotheses

Two other natural reasoning tasks that arise in the context of abduction are the recognition of which hypotheses are *relevant*, in the sense that they belong to at least one (minimal) solution, and which are *necessary* (or *indispensable*), as they occur in every (minimal) solution. Both of these decision problems have been investigated in the case of CPL abduction, see (Eiter and Gottlob 1995).

The following theorem shows that the complexity of relevance and necessity w.r.t. (proper) solutions and $\models_{\mathsf{BD}}$-minimal solutions coincides with the complexity of the analogous problems for ($\subseteq$-minimal) solutions in CPL.

Theorem 14. *It is Σ_2^{P}-complete (resp. Π_2^{P}-complete) to decide, given a* BD *abduction problem $\mathbb{P} = \langle \Gamma, \chi, \mathsf{H}\rangle$ and $h \in \mathsf{H}$, whether h is relevant (resp. necessary) w.r.t. $\mathcal{S}(\mathbb{P})$. The same holds for relevance and necessity w.r.t. $\mathcal{S}^{\mathsf{p}}(\mathbb{P})$ and $\mathcal{S}^{\mathsf{BD}}(\mathbb{P})$.*

Proof. Membership is straightforward since solution recognition is coNP-complete for proper $\mathcal{L}_\triangle$-solutions and DP-complete for proper $\mathcal{L}_\circ$-solutions and it suffices to guess a proper solution containing (or omitting) h and verify it.

For the hardness results for $\models_{\mathsf{BD}}$-minimal solutions, we construct a reduction from the class of BD abduction problems presented in (5) and adapt the approach from (Eiter and Gottlob 1995). Namely, we let $\mathbb{P}^{\mathbf{4}} = \langle \Gamma^{\mathbf{4}}, \chi^{\mathbf{4}}, \mathsf{H}\rangle$ be as in (5) and pick fresh variables r, r', and r''. Now set $\Xi = \mathsf{Prop}[\Gamma^{\mathbf{4}}] \cup \{r, r', r''\}$ and define $\mathbb{P}^{\mathsf{rd}} = \langle \Gamma^{\mathsf{rd}}, \chi^{\mathsf{rd}}, \mathsf{H}^{\mathsf{rd}}\rangle$ as follows.

$$\begin{aligned}\Gamma^{\mathsf{rd}} &= \{\neg r \vee \psi \mid \psi \in \Gamma^{\mathbf{4}}\} \cup \{s \vee \neg s \mid s \in \Xi\} \cup \\ &\quad \{\neg r' \vee (p \wedge \tau), \neg r \vee \neg r', \neg(r \vee r') \vee r''\} \\ \chi^{\mathsf{rd}} &= (p \wedge r'' \wedge \tau) \vee \bigvee_{s \in \Xi} (s \wedge \neg s) \\ \mathsf{H}^{\mathsf{rd}} &= \mathsf{H}^{\mathbf{4}} \cup \{r, r'\}\end{aligned} \tag{6}$$

Now let $\mathcal{S}^{\mathsf{p}}(\mathbb{P}^{\mathbf{4}})$ be the set of all proper solutions of $\mathbb{P}^{\mathbf{4}}$ and recall that $\mathcal{S}(\mathbb{P}^{\mathbf{4}}) = \mathcal{S}^{\mathsf{p}}(\mathbb{P}^{\mathbf{4}})$. It is clear that

$\mathcal{S}^{\mathsf{p}}(\mathbb{P}^{\mathsf{rd}}) = \mathcal{S}(\mathbb{P}^{\mathsf{rd}})$

$$\mathcal{S}^{\mathsf{p}}(\mathbb{P}^{\mathsf{rd}}) = \{\varrho\wedge r \mid \varrho\in\mathcal{S}(\mathbb{P}^{4})\}\cup\left\{\varrho'\wedge r' \mid \exists \mathsf{H}'\subseteq\mathsf{H} : \varrho'=\bigwedge_{l\in\mathsf{H}} l\right\}$$

and that $\mathbb{P}^{4}$ has (proper) solutions iff r is relevant and r' is not necessary.

To show hardness w.r.t. $\mathcal{S}^{\mathsf{BD}}(\mathbb{P})$, it suffices to observe that r is relevant to $\mathbb{P}^{\mathsf{rd}}$ iff it is relevant w.r.t. $\models_{\mathsf{BD}}$-minimal solutions. Similarly, r' is (not) necessary in $\mathbb{P}^{\mathsf{rd}}$ iff it is (not) necessary w.r.t. $\models_{\mathsf{BD}}$-minimal solutions. □

Finally, we present the following upper bound for the recognition of relevant and necessary hypotheses w.r.t. theory-minimal solutions. The proof is an easy consequence of Theorem 12. To the best of our knowledge, no analogous problem has been considered for CPL abduction problems.

Theorem 15. *It is in Σ_3^{P} (resp. Π_3^{P}) to decide, given a* BD *abduction problem $\mathbb{P} = \langle\Gamma,\chi,\mathsf{H}\rangle$ and $h\in\mathsf{H}$, whether h is relevant (resp. necessary) w.r.t. $\mathcal{S}^{\mathsf{Th}}(\mathbb{P})$.*

6 Generating Solutions to BD Abduction Problems by Reduction to CPL

In this section, we show how to apply classical consequence-finding procedures to generate solutions for BD abduction problems, by reducing BD abduction to CPL abduction.

We first observe that Lemma 1 allows us to faithfully translate abduction problems with $\mathcal{L}_\triangle$-solutions into CPL.

Theorem 16. *Let $\mathbb{P} = \langle\Gamma,\chi,\mathsf{H}\rangle$ be a* BD *abduction problem. Then ϕ is a ($\models_{\mathsf{BD}}$-, theory-minimal, proper) $\mathcal{L}_\triangle$-solution of $\mathbb{P}$ iff ϕ^{cl} is a ($\models_{\mathsf{CPL}}$-, theory-minimal, proper) solution of $\mathbb{P}^{\mathsf{cl}}_\triangle = \langle\Gamma^{\mathsf{cl}},\chi^{\mathsf{cl}},\mathsf{H}^{\mathsf{cl}}_\triangle\rangle$ with $\mathsf{H}^{\mathsf{cl}}_\triangle = \{p^+ \mid p\in\mathsf{H}\}\cup\{p^- \mid \neg p\in\mathsf{H}\}\cup\{\sim l \mid \neg\triangle l\in\mathsf{H}\}$.*

The translation of $\mathsf{BD}_\circ$ abduction into CPL abduction is, however, more complicated.

Definition 10. *Let $\phi = \bigwedge\limits_{i=1}^{m} p_i \wedge \bigwedge\limits_{i'=1}^{m'} \neg p_{i'} \wedge \bigwedge\limits_{j=1}^{n} \circ p_j \wedge \bigwedge\limits_{j'=1}^{n'} \bullet p_{j'}$ be an atomic $\mathcal{L}_\circ$-term, and $X \supseteq \mathsf{Prop}(\phi)$ be a finite set of propositional variables. The* classical counterpart *of ϕ relative to X, denoted ϕ°_X, is defined as follows:*

$$\phi^{\sim} = \bigwedge_{i=1}^{m} p_i^+ \wedge \bigwedge_{i'=1}^{m'} p_{i'}^- \wedge \bigwedge_{j=1}^{n} p_j^\circ \wedge \bigwedge_{j'=1}^{n'} \sim p_{j'}^\circ$$

$$\phi_X^{\leftrightarrow} = \bigwedge_{q\in X}(\sim q^\circ \leftrightarrow (q^+ \leftrightarrow q^-))$$

$$\phi_X^\circ = \phi^{\sim} \wedge \phi_X^{\leftrightarrow} \tag{7}$$

where the p_i^+, p_i^-, and q°s are fresh variables (not in X).

Lemma 3. *Let $\chi,\psi\in\mathcal{L}_{\mathsf{BD}}$, $\Xi = \mathsf{Prop}(\chi)\cup\mathsf{Prop}(\psi)$, and ϕ be an* atomic *$\mathcal{L}_\circ$-term s.t. $\mathsf{Prop}(\phi)\subseteq\Xi$. Then*

$$\phi,\chi \models_{\mathsf{BD}} \psi \textit{ iff } \phi^\circ_\Xi,\chi^{\mathsf{cl}} \models_{\mathsf{CPL}} \psi^{\mathsf{cl}}$$

We can now use Lemma 3 to construct a faithful embedding of BD abduction problems with $\mathcal{L}_\circ$-solutions into classical abduction problems. Note that the size of ϕ°_X is linear in the cardinality of X. This, however, is only possible because ϕ is an *atomic term*. Furthermore, since atomic $\mathcal{L}_\circ$-terms are not always translated into conjunctions of literals, we need to modify the statement of Theorem 16. Moreover, the resulting embedding preserves only theory-minimality since $\phi^{\sim}$ does not govern the interaction between p°s, p^+s, and p^-s.

Theorem 17. *Let $\mathbb{P} = \langle\Gamma,\chi,\mathsf{H}\rangle$ be a* BD *abduction problem and $\Xi = \mathsf{Prop}[\Gamma\cup\{\chi\}]$. Then ϕ is a (theory-minimal, proper) $\mathcal{L}_\circ$-solution of $\mathbb{P}$ iff $\phi^{\sim}$ is a (theory-minimal, proper) solution of $\mathbb{P}^{\mathsf{cl}}_\circ = \left\langle\Gamma^{\mathsf{cl}}\cup\{\phi^{\leftrightarrow}_\Xi\}, \phi^{\leftrightarrow}_\Xi \rightarrow \chi^{\mathsf{cl}}, \mathsf{H}^{\mathsf{cl}}_\circ\right\rangle$ with $\mathsf{H}^{\mathsf{cl}}_\circ = \{p^+ \mid p\in\mathsf{H}\}\cup\{p^- \mid \neg p\in\mathsf{H}\}\cup\{p^\circ \mid \circ p\in\mathsf{H}\}\cup\{\sim p^\circ \mid \bullet p\in\mathsf{H}\}$.*

Theorems 16 and 17 show that we can use classical techniques of abductive reasoning (such as the ones based upon consequence finding, cf. (Inoue 1992; del Val 2000; Marquis 2000; Inoue 2002)) to solve BD abductive problems. Let us now illustrate how our translations work.

Example 4. *Recall Example 1. We consider two* BD *problems: $\mathbb{P}_\triangle = \langle\Gamma,\chi,\mathsf{H}_\triangle\rangle$ and $\mathbb{P}_\circ = \langle\Gamma,\chi,\mathsf{H}_\circ\rangle$, where:*

$$\Gamma = \{p\vee q, \neg p, \neg q\} \quad \chi = q \quad \mathsf{H}_\triangle = \{p, \neg p, \neg\triangle p, \neg\triangle\neg p\}$$
$$\mathsf{H}_\circ = \{p, \neg p, \circ p, \bullet p\}$$

Applying Theorems 16 and 17, we obtain the following classical problems $\mathbb{P}^{\mathsf{cl}}_\triangle = \langle\Gamma^{\mathsf{cl}},\chi^{\mathsf{cl}},\mathsf{H}^{\mathsf{cl}}_\triangle\rangle$ and $\mathbb{P}^{\mathsf{cl}}_\circ = \langle\Gamma^\circ,\chi^{\mathsf{cl}},\mathsf{H}^{\mathsf{cl}}_\circ\rangle$, where:

$$\Gamma^{\mathsf{cl}} = \{p^+\vee q^+, p^-, q^-\} \qquad \chi^{\mathsf{cl}} = q^+$$
$$\mathsf{H}^{\mathsf{cl}}_\triangle = \{p^+, p^-, \sim p^+, \sim p^-\}$$
$$\Gamma^\circ = \Gamma^{\mathsf{cl}}\cup\left\{(\sim r^\circ \leftrightarrow (r^+ \leftrightarrow r^-)) \mid r\in\{p,q\}\right\}$$
$$\mathsf{H}^{\mathsf{cl}}_\circ = \{p^+, p^-, p^\circ, \sim p^\circ\}$$

Now if we apply classical consequence-finding procedures, we need to look for clauses entailed by $\Gamma^{\mathsf{cl}}_\triangle\cup\{\sim\chi^{\mathsf{cl}}\}$ and $\Gamma^{\mathsf{cl}}_\circ\cup\{\sim\chi^{\mathsf{cl}}\}$. For $\mathcal{L}_\triangle$-solutions, we can take (the negation of) any clause. *For $\mathcal{L}_\circ$-solutions, the clauses must contain* only negative *occurrences of p^+ and p^-.*

It is easy to check that $\sim p^+$ and p° are theory-minimal classical *solutions of $\mathbb{P}^{\mathsf{cl}}_\triangle$ and $\mathbb{P}^{\mathsf{cl}}_\circ$. They correspond to $\neg\triangle p$ and $\circ p$, respectively, which are (as expected) theory-minimal $\mathcal{L}_\triangle$- and $\mathcal{L}_\circ$- solutions of $\mathbb{P}^\triangle$ and $\mathbb{P}^\circ$.*

We finish the section by noting that in general, $\mathcal{L}_\triangle$-solutions are not uniquely generated from classical clauses (cf. $p^{\mathsf{cl}} = (\triangle p)^{\mathsf{cl}} = p^+$) and that by Proposition 2, $\phi \simeq \phi^\flat$. Thus, for practical purposes, it makes sense to convert classical clauses to $\mathcal{L}_\triangle$ solutions τ s.t. $\tau = \tau^\flat$.

7 Discussion and Future Work

We have studied abductive reasoning in the four-valued paraconsistent logic BD, motivating and comparing $\mathcal{L}_\triangle$- and $\mathcal{L}_\circ$-solutions. Our complexity analysis (Table 2) provides an almost complete picture of the complexity of the main decision problems related to abduction. In particular, we established that the complexity of solution existence in $\mathsf{BD}_\triangle$ and $\mathsf{BD}_\circ$ is not higher than in the classical case (Eiter and Gottlob 1995; Creignou and Zanuttini 2006; Pichler and Woltran 2010; Pfandler, Pichler, and Woltran 2015). Moreover, by exhibiting reductions of abduction in $\mathsf{BD}_{\triangle}$ and $\mathsf{BD}_{\circ}$ to abduction in CPL, we have shown that existing procedures for generating abductive solutions in classical logic can be employed for paraconsistent abduction.

A few questions remain open. First, we do not know the exact complexity of theory-minimal solution recognition and relevance. One way to approach this would be to establish the complexity of closely related notion of theory prime implicants in CPL (Marquis 1995). There is also the question of how to embed BD abduction problems with $\mathcal{L}_{\circ}$-solutions into classical problems while preserving $\models_{\mathsf{BD}}$-minimal solutions (recall that Theorem 17 only preserves theory-minimal solutions). Also, since some BD abduction problems can be solved by arbitrary $\mathcal{L}_{\circ}$-terms but not atomic ones, it would be interesting to explore the computational properties of $\mathcal{L}_{\circ}$-solutions based upon non-atomic $\mathcal{L}_{\circ}$-terms.

A more general direction for future work is to consider abduction in expansions of BD. Of particular interest are *functionally complete* expansions of BD, e.g., the bi-lattice language expansion or an expansion with a 'quarter-turn' connective from (Ruet 1996) (cf. (Omori and Sano 2015) for more details). An important technical question would be whether all $\mathcal{L}_{\triangle}$- and $\mathcal{L}_{\circ}$-solutions can be represented as terms in functionally complete languages. A further challenge is to come up with an intuitive natural-language interpretation of literals and terms in such languages. Another option is to consider theories containing implicative formulas (cf. (Omori and Wansing 2017) for details). This would allow us to define Horn-like fragments of the languages in question. As solution existence in *classical* Horn abduction problems is NP-complete (Creignou and Zanuttini 2006), it makes sense to check whether abduction in Horn BD is also simpler than in the general case.

Additionally, we plan to consider *modal* expansions of BD. Abduction in classical modal logic is well-researched. In particular, (Levesque 1989) and (Sakama and Inoue 2016) study abduction in classical epistemic and doxastic logics; (Mayer and Pirri 1995) apply tableaux procedures to solution generation in $\mathbf{K}$, $\mathbf{D}$, $\mathbf{T}$, and $\mathbf{S4}$. (Bienvenu 2009) compares different definitions of prime implicates closely related to abductive solutions and provides complexity results and algorithms for prime implicate recognition and generation in multimodal $\mathbf{K}_n$; (Nepomuceno-Fernández, Soler-Toscano, and Velázquez-Quesada 2017) consider abduction in dynamic epistemic logic. Modal expansions of BD are also well known (cf. (Priest 2008) and (Drobyshevich 2020)). There are also public announcement (Rivieccio 2014) and dynamic (Sedlár 2016) BD logics. Thus, it is natural to consider abductive reasoning in *modal paraconsistent* framework and see whether classical decision procedures and complexity results can be transferred there.

Acknowledgements

The authors were supported by the ANR AI Chair INTENDED (ANR-19-CHIA-0014) (1st and 3rd), JSPS KAKENHI (JP21H04905) and the JST CREST (JP-MJCR22D3) (2nd), and an NII MOU grant (3rd).

References

Aliseda, A. 2006. *Abductive Reasoning*, volume 330 of *Synthese Library*. Springer.

Belnap, N. 1977a. A Useful Four-Valued Logic. In *Modern Uses of Multiple-Valued Logic*, 5–37. Dordrecht: Springer Netherlands.

Belnap, N. 1977b. How a Computer Should Think. In *Contemporary Aspects of Philosophy*. Oriel Press.

Bienvenu, M.; Inoue, K.; and Kozhemiachenko, D. 2024. Abductive Reasoning in a Paraconsistent Framework. Technical report — arXiv:2408.07287.

Bienvenu, M. 2009. Prime Implicates and Prime Implicants: From Propositional to Modal Logic. *Journal of Artificial Intelligence Research* 36:71–128.

Bueno-Soler, J.; Carnielli, W.; Coniglio, M.; and Rodrigues Filho, A. 2017. Formal (In)consistency, Abduction and Modalities. In *Springer Handbook of Model-Based Science*. Cham: Springer International Publishing. 315–335.

Carnielli, W. 2006. Surviving Abduction. *Logic Journal of the IGPL* 14(2):237–256.

Chlebowski, S.; Gajda, A.; and Urbański, M. 2022. An Abductive Question-Answer System for the Minimal Logic of Formal Inconsistency $\mathbf{mbC}$. *Studia Logica* 110(2):479–509.

Creignou, N., and Zanuttini, B. 2006. A Complete Classification of the Complexity of Propositional Abduction. *SIAM Journal of Computing* 36(1):207–229.

del Val, A. 2000. The Complexity of Restricted Consequence Finding and Abduction. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI)*, 337–342.

Drobyshevich, S. 2020. A General Framework for FDE-Based Modal Logics. *Studia Logica* 108(6):1281–1306.

Du, J.; Wang, K.; and Shen, Y.-D. 2015. Towards Tractable and Practical ABox Abduction Over Inconsistent Description Logic Ontologies. *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence* 29(1).

Dunn, J. 1976. Intuitive Semantics for First-Degree Entailments and 'Coupled Trees'. *Philosophical Studies* 29(3):149–168.

Eiter, T., and Gottlob, G. 1995. The Complexity of Logic-Based Abduction. *Journal of the Association for Computing Machinery* 42(1):3–42.

Inoue, K. 1992. Linear Resolution for Consequence Finding. *Artificial Intelligence* 56(2-3):301–353.

Inoue, K. 2002. Automated Abduction. In *Computational Logic: Logic Programming and Beyond*, volume 2408 of *Lecture notes in computer science*. Berlin, Heidelberg: Springer Berlin Heidelberg. 311–341.

Levesque, H. 1989. A Knowledge-Level Account of Abduction. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence (IJCAI)*, volume 2, 1061–1067.

Marquis, P. 1995. Knowledge Compilation Using Theory Prime Implicates. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI)*, volume I, 837–845.

Marquis, P. 2000. Consequence Finding Algorithms. In *Handbook of Defeasible Reasoning and Uncertainty Management Systems: Algorithms for Uncertainty and Defeasible Reasoning*. Springer. 41–145.

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

Nepomuceno-Fernández, A.; Soler-Toscano, F.; and Velázquez-Quesada, F. 2017. Abductive Reasoning in Dynamic Epistemic Logic. In *Springer Handbook of Model-Based Science*. Springer. 269–293.

Omori, H., and Sano, K. 2014. Da Costa Meets Belnap and Nelson. In *Recent Trends in Philosophical Logic*, volume 41 of *Trends in Logic*. Cham: Springer International Publishing. 145–166.

Omori, H., and Sano, K. 2015. Generalizing Functional Completeness in Belnap-Dunn Logic. *Studia Logica* 103(5):883–917.

Omori, H., and Wansing, H. 2017. 40 years of FDE: An Introductory Overview. *Studia Logica* 105(6):1021–1049.

Omori, H., and Waragai, T. 2011. Some Observations on the Systems LFI1 and LFI1*. In *22nd International Workshop on Database and Expert Systems Applications*, 320–324.

Pfandler, A.; Pichler, R.; and Woltran, S. 2015. The Complexity of Handling Minimal Solutions in Logic-Based Abduction. *Journal of Logic and Computation* 25(3):805–825.

Pichler, R., and Woltran, S. 2010. The Complexity of Handling minimal Solutions in Logic-Based Abduction. In *Proceedings of 19th European Conference on Artificial Intelligence (ECAI)*. 895–900.

Poole, D. 1989. Explanation and Prediction: An Architecture for Default and Abductive Reasoning. *Computational Intelligence* 5(2):97–110.

Priest, G. 2008. Many-Valued Modal Logics: A Simple Approach. *The Review of Symbolic Logic* 1(2):190–203.

Rivieccio, U. 2014. Bilattice Public Announcement Logic. In *Proceedings of the Tenth Conference on Advances in Modal Logic (AiML)*, 459–477.

Rodrigues, A.; Coniglio, M.; Antunes, H.; Bueno-Soler, J.; and Carnielli, W. 2023. Paraconsistency, Evidence, and Abduction. In *Handbook of Abductive Cognition*. Cham: Springer International Publishing. 313–350.

Ruet, P. 1996. Complete Sets of Connectives and Complete Sequent Calculus for Belnap's Logic. In *Logic Colloquium 96*. Ecole Normale Supérieure (Paris). Laboratoire d'Informatique (technical report).

Sakama, C., and Inoue, K. 1995. The Effect of Partial Deduction in Abductive Reasoning. In *Proceedings of the Twelth International Conference on Logic Programming (ICLP)*.

Sakama, C., and Inoue, K. 2016. Abduction, Conversational Implicature and Misleading in Human Dialogues. *Logic Journal of the IGPL* 24(4):526–541.

Sano, K., and Omori, H. 2014. An Expansion of First-Order Belnap–Dunn logic. *Logic Journal of IGPL* 22(3):458–481.

Sedlár, I. 2016. Propositional Dynamic Logic With Belnapian Truth Values. In *Proceedings of the Eleventh Conference on Advances in Modal Logic (AiML)*, 503–519.

Stickel, M. 1990. Rationale and Methods for Abductive Reasoning in Natural-Language Interpretation. In *Natural Language and Logic*, volume 459 of *Lecture Notes in Artificial Intelligence*, 233–252. Springer.