HAL
open science

Abduction and Dialogical Proof in Argumentation and Logic Programming

Richard Booth, Dov M. Gabbay, Souhila Kaci, Tjitze Rienstra, Leendert van Der Torre

To cite this version:

Richard Booth, Dov M. Gabbay, Souhila Kaci, Tjitze Rienstra, Leendert van Der Torre. Abduction and Dialogical Proof in Argumentation and Logic Programming. ECAI: European Conference on Artificial Intelligence, Aug 2014, Prague, Czech Republic. pp.117-122, 10.3233/978-1-61499-419-0-117 . lirmm-01082080

HAL Id: lirmm-01082080

https://hal-lirmm.ccsd.cnrs.fr/lirmm-01082080v1

Submitted on 29 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Abduction and Dialogical Proof in Argumentation and Logic Programming

Richard Booth[1] and **Dov Gabbay**[1,2] and **Souhila Kaci**[3] and **Tjitze Rienstra**[1,3] and **Leendert van der Torre**[1]

arXiv:1407.3896v1 [cs.AI] 15 Jul 2014

Abstract. We develop a model of abduction in abstract argumentation, where changes to an argumentation framework act as hypotheses to explain the support of an observation. We present dialogical proof theories for the main decision problems (i.e., finding hypotheses that explain skeptical/credulous support) and we show that our model can be instantiated on the basis of abductive logic programs.

1 Introduction

In the context of abstract argumentation [12], abduction can be seen as the problem of finding *changes* to an argumentation framework (or *AF* for short) with the goal of explaining observations that can be justified by making arguments accepted. The general problem of whether and how an AF can be changed with the goal of changing the status of arguments has been studied by Baumann and Brewka [3], who called it the *enforcing* problem, as well as Bisquert et al. [4], Perotti et al. [5] and Kontarinis et al. [15]. None of these works, however, made any explicit link with abduction. Sakama [20], on the other hand, explicitly focused on abduction, and presented a model in which additions as well as removals of arguments from an abstract AF act as explanations for the observation that an argument is accepted or rejected.

While Sakama did address computation in his framework, his method was based on translating abstract AFs into logic programs. Proof theories in argumentation are, however, often formulated as *dialogical* proof theories, which aim at relating the problem they address with stereotypical patterns found in real world dialogue. For example, proof theories for skeptical/credulous acceptance have been modelled as dialogues in which a proponent persuades an opponent to accept the necessity/possibility of an argument [17], while credulous acceptance has also been related to Socratic style dialogue [9]. Thus, the question of how decision problems in abduction in argumentation can similarly be modelled as dialogues remains open.

Furthermore, argumentation is often used as an abstract model for non-monotonic reasoning formalisms. For example, an *instantiated* AF can be generated on the basis of a logic program. Consequences can then be computed by looking at the extensions of the instantiated AF [12]. In the context of abduction, one may ask whether a model of abduction in argumentation can similarly be seen as an abstraction of *abductive* logic programming. Sakama, however, did not explore the instantiation of his model, meaning that this question too remains open.

[1] Computer Science and Communication, University of Luxembourg (richard.booth@uni.lu, tjitze.rienstra@uni.lu, leon.vandertorre@uni.lu)
[2] Dept. Computer Science, King's College London (dov.gabbay@kcl.ac.uk)
[3] LIRMM, University of Montpellier 2 (souhila.kaci@lirmm.fr)

This brings us to the contribution of this paper. We first present a model of abduction in abstract argumentation, based on the notion of an AAF (abductive argumentation framework) that encodes different possible changes to an AF, each of which may act as a hypothesis to explain an observation that can be justified by making an argument accepted. We then do two things:

1. We present sound and complete dialogical proof procedures for the main decision problems, i.e., finding hypotheses that explain skeptical/credulous acceptance of arguments in support of an observation. These proof procedures show that the problem of abduction is related to an extended form of persuasion, where the proponent uses *hypothetical* moves to persuade the opponent.
2. We show that AAFs can be instantiated by ALPs (abductive logic programs) in such a way that the hypotheses generated for an observation by the ALP can be computed by translating the ALP into an AAF. The type of ALPs we focus on are based on Sakama and Inoue's model of *extended* abduction [13, 14], in which hypotheses have a positive as well as a negative element (i.e., facts added to the logic program as well as facts removed from it).

In sum, our contribution is a model of abduction in argumentation with dialogical proof theories for the main decision problems, which can be seen as an abstraction of abduction in logic programming.

The overview of this paper is as follows. After introducing the necessary preliminaries in section 2 we present in section 3 our model of abduction in argumentation. In section 4 we present dialogical proof procedures for the main decision problems (explaining skeptical/credulous acceptance). In section 5 we show that our model of abduction can be used to instantiate abduction in logic programming. We discuss related work in section 6 and conclude in section 7.

2 Preliminaries

An argumentation framework consists of a set A of arguments and a binary *attack* relation $\rightsquigarrow$ over A [12]. We assume in this paper that A is a finite subset of a fixed set $\mathcal{U}$ called the *universe of arguments*.

Definition 1. *Given a countably infinite set $\mathcal{U}$ called the* universe of arguments, *an* argumentation framework *(*AF*, for short) is a pair $F = (A, \rightsquigarrow)$ where A is a finite subset of $\mathcal{U}$ and $\rightsquigarrow$ a binary relation over A. If $a \rightsquigarrow b$ we say that a* attacks *b. $\mathcal{F}$ denotes the set of all AFs.*

Extensions are sets of arguments that represent different viewpoints on the acceptance of the arguments of an AF. A *semantics* is a method to select extensions that qualify as somehow justifiable. We focus on one of the most basic ones, namely the *complete* semantics [12].

Definition 2. *Let $F = (A, \rightsquigarrow)$. An* extension *of F is a set $E \subseteq A$. An extension E is* conflict-free *iff for no $a, b \in E$ it holds that $a \rightsquigarrow b$. An argument $a \in A$ is* defended *by E iff for all b such that $b \rightsquigarrow a$ there is a $c \in E$ such that $c \rightsquigarrow b$. Given an extension E, we define $Def_F(E)$ by $Def_F(E) = \{a \in A \mid E \text{ defends } a\}$. An extension E is* admissible *iff E is conflict-free and $E \subseteq Def_F(E)$, and* complete *iff E is conflict-free and $E = Def_F(E)$. The set of complete extension of F will be denoted by $Co(F)$. Furthermore, the grounded extension (denoted by $Gr(F)$) is the unique minimal (w.r.t. $\subseteq$) complete extension of F.*

An argument is said to be *skeptically* (resp. *credulously*) accepted iff it is a member of all (resp. some) complete extensions. Note that the set of skeptically accepted arguments coincides with the grounded extension. Furthermore, an argument is a member of a complete extension iff it is a member of a *preferred* extension, which is a maximal (w.r.t. $\subseteq$) complete extension. Consequently, credulous acceptance under the preferred semantics (as studied e.g. in [17]) coincides with credulous acceptance under the complete semantics.

3 Abductive AFs

Abduction is a form of reasoning that goes from an observation to a hypothesis. We assume that an observation translates into a set $X \subseteq A$. Intuitively, X is a set of arguments that each individually support the observation. If at least one argument $x \in X$ is skeptically (resp. credulously) accepted, we say that the observation X is skeptically (resp. credulously) *supported*.

Definition 3. *Given an AF $F = (A, \rightsquigarrow)$, an observation $X \subseteq A$ is skeptically (resp. credulously) supported iff for all (resp. some) $E \in Co(F)$ it holds that $x \in E$ for some $x \in X$.*

The following proposition implies that checking whether an observation X is skeptically supported can be done by checking whether an individual argument $x \in X$ is in the grounded extension.

Proposition 1. *Let $F = (A, \rightsquigarrow)$ and $X \subseteq A$. It holds that F skeptically supports X iff $x \in Gr(F)$ for some $x \in X$.*

Proof of proposition 1. The if direction is immediate. For the only if direction, assume $F = (A, \rightsquigarrow)$ explains skeptical support for X. Then for every complete extension E of F, there is an $x \in X$ s.t. $x \in E$. Define G by $G = (A \cup \{a, b\}, \rightsquigarrow \cup \{(x, a) \mid x \in X\} \cup \{(a, b)\})$, where $a, b \notin A$. Then for every complete extension E of G it holds that $b \in E$, hence $b \in Gr(G)$. Thus $x \in Gr(G)$ for some $x \in X$. But $Gr(F) = Gr(G) \cap A$, hence $x \in Gr(F)$ for some $x \in X$. □

It may be that an AF F does not skeptically or credulously support an observation X. Abduction then amounts to finding a change to F so that X is supported. We use the following definition of an *AAF* (*Abductive AF*) to capture the changes w.r.t. F (each change represented by an AF G called an *abducible* AF) that an agent considers. We assume that F itself is also an abducible AF, namely one that captures the case where no change is necessary. Other abducible AFs may be formed by addition of arguments and attacks to F, removal of arguments and attacks from F, or a combination of both.

Definition 4. *An* abductive AF *is a pair $M = (F, I)$ where F is an AF and $I \subseteq \mathcal{F}$ a set of AFs called* abducible *such that $F \in I$.*

Given an AAF (F, I) and observation X, skeptical/credulous support for X can be explained by the change from F to some $G \in I$ that skeptically/credulously supports X. In this case we say that G *explains* skeptical/credulous support for X. The arguments/attacks added to and absent from G can be seen as the actual explanation.

Definition 5. *Let $M = (F, I)$ be an AAF. An abducible AF $G \in I$* explains *skeptical (resp. credulous) support for an observation X iff G skeptically (resp. credulously) supports X.*

One can focus on explanations satisfying additional criteria, such as minimality w.r.t. the added or removed arguments/attacks. We leave the formal treatment of such criteria for future work.

Example 1. *Let $M = (F, \{F, G_1, G_2, G_3\})$, where F, G_1, G_2 and G_3 are as defined in figure 1. Let $X = \{b\}$ be an observation. It holds that G_1 and G_3 both explain skeptical support for X, while G_2 only explains credulous support for X.*

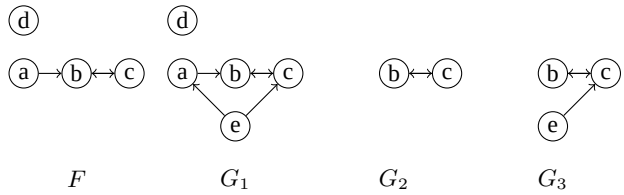

Figure 1. The AFs of the AAF $(F, \{F, G_1, G_2, G_3\})$.

Remark 1. *The main difference between Sakama's [20] model of abduction in abstract argumentation and the one presented here, is that he takes an explanation to be a set of independently selectable abducible arguments, while we take it to be a change to the AF that is applied as a whole. In section 5 we show that this is necessary when applying the abstract model in an instantiated setting.*

4 Explanation dialogues

In this section we present methods to determine, *given an AAF $M = (F, I)$ (for $F = (A, \rightsquigarrow)$) whether an abducible AF $G \in I$ explains credulous or skeptical support for an observation $X \subseteq A$.* We build on ideas behind the *grounded* and *preferred games*, which are dialogical procedures that determine skeptical or credulous acceptance of an argument [17]. To sketch the idea behind these games (for a detailed discussion cf. [17]): two imaginary players (PRO and OPP) take alternating turns in putting forward arguments according to a set of rules, PRO either as an initial claim or in defence against OPP's attacks, while OPP initiates different disputes by attacking the arguments put forward by PRO. Skeptical or credulous acceptance is proven if PRO can win the game by ending every dispute in its favour according to a "last-word" principle.

Our method adapts this idea so that the moves made by PRO are essentially *hypothetical* moves. That is, to defend the initial claim (i.e., to explain an observation) PRO can put forward, by way of hypothesis, any attack $x \rightsquigarrow y$ present in some $G \in I$. This marks a choice of PRO to focus only on those abducible AFs in which the attack $x \rightsquigarrow y$ is present. Similarly, PRO can reply to an attack $x \rightsquigarrow y$, put forward by OPP, with the claim that this attack is invalid, marking the choice of PRO to focus only on the abducible AFs in which the attack $x \rightsquigarrow y$ is *not* present. Thus, each move by PRO narrows down the set of abducible AFs in which all of PRO's moves are valid. The objective is to end the dialogue with a non-empty set of abducible AFs. Such a dialogue represents a proof that these abducible AFs explain skeptical or credulous support for the observation.

Alternatively, such dialogues can be seen as games that determine skeptical/credulous support of an observation by an AF that

are played simultaneously over all abducible AFs in the AAF. In this view, the objective is to end the dialogue in such a way that it represents a proof for at least one abducible AF. Indeed, in the case where $M = (F, \{F\})$, our method reduces simply to a proof theory for skeptical or credulous support of an observation by F.

Before we move on we need to introduce some notation.

Definition 6. *Given a set I of AFs we define:*
- $A_I = \cup\{A \mid (A, \rightsquigarrow) \in I\}$,
- $\rightsquigarrow_I = \cup\{\rightsquigarrow \mid (A, \rightsquigarrow) \in I\}$,
- $I_{x \rightsquigarrow y} = \{(A, \rightsquigarrow) \in I \mid x, y \in A, x \rightsquigarrow y\}$,
- $I_X = \{(A, \rightsquigarrow) \in I \mid X \subseteq A\}$.

We model dialogues as sequences of *moves*, each move being of a certain type, and made either by PRO or OPP.

Definition 7. *Let $M = (F, I)$ be an AAF. A* dialogue based on *M is a sequence $S = (m_1, \ldots, m_n)$, where each m_i is either:*
- *an* OPP attack *"**OPP:** $x \rightsquigarrow y$", where $x \rightsquigarrow_I y$,*
- *a* hypothetical PRO defence *"**PRO:** $y \rightsquigarrow^+ x$", where $y \rightsquigarrow_I x$,*
- *a* hypothetical PRO negation *"**PRO:** $y \rightsquigarrow^- x$", where $y \rightsquigarrow_I x$,*
- *a* conceding move *"**OPP: ok**",*
- *a* success claim move *"**PRO: win**".*

We denote by $S \cdot S'$ the concatenation of S and S'.

Intuitively, a move **OPP:** $y \rightsquigarrow x$ represents an attack by OPP on the argument x by putting forward the attacker y. A hypothetical PRO defence **PRO:** $y \rightsquigarrow^+ x$ represents a defence by PRO who puts forward y to attack the argument x put forward by OPP. A hypothetical PRO negation **PRO:** $y \rightsquigarrow^- x$, on the other hand, represents a claim by PRO that the attack $y \rightsquigarrow x$ is *not* a valid attack. The conceding move **OPP: ok** is made whenever OPP runs out of possibilities to attack a given argument, while the move **PRO: win** is made when PRO is able to claim success.

In the following sections we specify how dialogues are structured. Before doing so, we introduce some notation that we use to keep track of the abducible AFs on which PRO chooses to focus in a dialogue D. We call this set the *information state* of D after a given move. While it initially contains all abducible AFs in *M*, it is restricted when PRO makes a move **PRO:** $x \rightsquigarrow^+ y$ or **PRO:** $x \rightsquigarrow^- y$.

Definition 8. *Let $M = (F, I)$ be an AAF. Let $D = (m_1, \ldots, m_n)$ be a dialogue based on M. We denote the* information state in D after move i *by $J(D, i)$, which is defined recursively by:*

$$J(D,i) = \begin{cases} I & \text{if } i = 0, \\ J(D, i-1) \cap I_{x \rightsquigarrow y} & \text{if } m_i = \textbf{PRO: } x \rightsquigarrow^+ y, \\ J(D, i-1) \setminus I_{x \rightsquigarrow y} & \text{if } m_i = \textbf{PRO: } x \rightsquigarrow^- y, \\ J(D, i-1) & \text{otherwise.} \end{cases}$$

We denote by $J(D)$ the information state $J(D, n)$.

4.1 Skeptical explanation dialogues

We define the rules of a dialogue using a set of production rules that recursively define the set of sequences constituting dialogues. (The same methodology was used by Booth et al. [7] in defining a dialogical proof theory related to preference-based argumentation.) In a skeptical explanation dialogue for an observation X, an initial argument $x \in X$ is challenged by the opponent, who puts forward all possible attacks **OPP:** $y \rightsquigarrow x$ present in any of the abducible AFs present in the AAF, followed by **OPP: ok**. We call this a *skeptical OPP reply* to x. For each move **OPP:** $y \rightsquigarrow x$, PRO responds with a *skeptical PRO reply to $y \rightsquigarrow x$*, which is either a hypothetical defence **PRO:** $z \rightsquigarrow^+ y$ (in turn followed by a skeptical OPP reply to z) or a hypothetical negation **PRO:** $y \rightsquigarrow^- x$. Formally:

Definition 9 (Skeptical explanation dialogue)**.** *Let $F = (A, \rightsquigarrow)$, $M = (F, I)$ and $x \in A$.*
- *A* skeptical OPP reply to x *is a finite sequence* (***OPP:*** $y_1 \rightsquigarrow x$) $\cdot$ $S_1 \cdot \ldots \cdot$ (***OPP:*** $y_n \rightsquigarrow x$) $\cdot$ $S_n \cdot$ (***OPP: ok***) *where $\{y_1, \ldots, y_n\} = \{y \mid y \rightsquigarrow_I x\}$ and each S_i is a skeptical PRO reply to $y_i \rightsquigarrow x$.*
- *A* skeptical PRO reply to $y \rightsquigarrow x$ *is either: (1) A sequence* (***PRO:*** $z \rightsquigarrow^+ y$) $\cdot$ S *where $z \rightsquigarrow_I y$ and where S is a skeptical OPP reply to z, or (2) The sequence* (***PRO:*** $y \rightsquigarrow^- x$).

Given an observation $X \subseteq A$ we say that M generates the skeptical explanation dialogue *D for X iff $D = S \cdot$ (**PRO: win**), where S is a skeptical OPP reply to some $x \in X$.*

The following theorem establishes soundness and completeness.

Theorem 1. *Let $M = (F, I)$ be an AAF where $F = (A, \rightsquigarrow)$. Let $X \subseteq A$ and $G \in I$. It holds that G explains skeptical support for X iff M generates a skeptical explanation dialogue D for X such that $G \in J(D)$.*

Due to space constraints we only provide a sketch of the proof.

Sketch of proof. Let $M = ((A, \rightsquigarrow), I)$, $X \subseteq A$ and $G \in I$. *(Only if:)* Assume $x \in Gr(G)$ for some $x \in X$. By induction on the number of times the characteristic function [12] is applied so as to establish that $x \in Gr(G)$, it can be shown that a credulous OPP reply D to x exists (and hence a dialogue $D \cdot$ (**PRO: win**) for X) s.t. $G \in J(D \cdot$ (**PRO: win**)). *(If:)* Assume *M* generates a skeptical explanation dialogue D for X s.t. $G \in J(D)$. By induction on the structure of D it can be shown that $x \in Gr(G)$ for some $x \in X$. □

Example 2. *The listing below shows a skeptical explanation dialogue $D = (m_1, \ldots, m_8)$ for the observation $\{b\}$ that is generated by the AAF defined in example 1.*

i	m_i	$J(D,i)$
1	**OPP:** $c \rightsquigarrow b$	$\{F, G_1, G_2, G_3\}$
2	**PRO:** $e \rightsquigarrow^+ c$	$\{G_1, G_3\}$
3	**OPP: ok**	$\{G_1, G_3\}$
4	**OPP:** $a \rightsquigarrow b$	$\{G_1, G_3\}$
5	**PRO:** $e \rightsquigarrow^+ a$	$\{G_1\}$
6	**OPP: ok**	$\{G_1\}$
7	**OPP: ok**	$\{G_1\}$
8	**PRO: win**	$\{G_1\}$

The sequence $(m_1, \ldots, m_7)$ is a skeptical OPP reply to b, in which OPP puts forward the two attacks $c \rightsquigarrow b$ and $a \rightsquigarrow b$. PRO defends b from both c and a by putting forward the attacker e (move 2 and 5). This leads to the focus first on the abducible AFs G_1, G_3 (in which the attack $e \rightsquigarrow c$ exists) and then on G_1 (in which the attack $e \rightsquigarrow a$ exists). This proves that G_1 explains skeptical support for the observation $\{b\}$. Another dialogue is shown below.

i	m_i	$J(D,i)$
1	**OPP:** $c \rightsquigarrow b$	$\{F, G_1, G_2, G_3\}$
2	**PRO:** $e \rightsquigarrow^+ c$	$\{G_1, G_3\}$
3	**OPP: ok**	$\{G_1, G_3\}$
4	**OPP:** $a \rightsquigarrow b$	$\{G_1, G_3\}$
5	**PRO:** $a \rightsquigarrow^- b$	$\{G_3\}$
6	**OPP: ok**	$\{G_3\}$
7	**PRO: win**	$\{G_3\}$

Here, PRO defends b from c by using the argument e, but defends b from a by claiming that the attack $a \rightsquigarrow b$ is invalid. This leads to the focus first on the abducible AFs G_1, G_3 (in which the attack $e \rightsquigarrow c$ exists) and then on G_3 (in which the attack $a \rightsquigarrow b$ does not exist). This dialogue proves that G_3 explains skeptical support for $\{b\}$.

4.2 Credulous explanation dialogues

The definition of a credulous explanation dialogue is similar to that of a skeptical one. The difference lies in what constitutes an acceptable defence. To show that an argument x is skeptically accepted, x must be defended from its attackers by arguments other than x itself. For credulous acceptance, however, it suffices to show that x is a member of an admissible set, and hence x may be defended from its attackers by any argument, including x itself. To achieve this we need to keep track of the arguments that are, according to the moves made by PRO, accepted. Once an argument x is accepted, PRO does not need to defend x again, if this argument is put forward a second time.

Formally a *credulous OPP reply* to (x, Z) (for some $x \in A_I$ and set $Z \subseteq A_I$ used to keep track of accepted arguments) consists of all possible attacks **OPP:** $y \rightsquigarrow x$ on x, followed by **OPP: ok** when all attacks have been put forward. For each move **OPP:** $y \rightsquigarrow x$, PRO responds either by putting forward a hypothetical defence **PRO:** $z \rightsquigarrow^+ y$ which (this time *only if* $z \notin Z$) is followed by a credulous OPP reply to $(z, Z \cup \{z\})$, or by putting forward a hypothetical negation **PRO:** $y \rightsquigarrow^- x$. We call this response a *credulous PRO reply to* $(y \rightsquigarrow x, Z)$. A credulous explanation dialogue for a set X consists of a credulous OPP reply to $(x, \{x\})$ for some $x \in X$, followed by a success claim **PRO: win**.

In addition, arguments put forward by PRO in defence of the observation may not conflict. Such a conflict occurs when OPP puts forward **OPP:** $x \rightsquigarrow y$ and **OPP:** $y \rightsquigarrow z$ (indicating that both y and z are accepted) while PRO does not put forward **PRO:** $y \rightsquigarrow^- z$. If this situation does not occur we say that the dialogue is *conflict-free*.

Definition 10 (Credulous explanation dialogue)**.** *Let $F = (A, \rightsquigarrow)$, $M = (F, I)$, $x \in A$ and $Z \subseteq A$.*

- *A* credulous OPP reply to (x, Z) *is a finite sequence* (***OPP:*** $y_1 \rightsquigarrow x$)$\cdot S_1 \cdot \ldots \cdot$(***OPP:*** $y_n \rightsquigarrow x$)$\cdot S_n \cdot$(***OPP: ok***) *where* $\{y_1, \ldots, y_n\} = \{y \mid y \rightsquigarrow_I x\}$ *and each S_i is a credulous PRO reply to* $(y_i \rightsquigarrow x, Z)$.
- *A* credulous PRO reply to $(y \rightsquigarrow x, Z)$ *is either: (1) a sequence* (***PRO:*** $z \rightsquigarrow^+ y$)$\cdot S$ *such that $z \rightsquigarrow_I y$, $z \notin Z$ and S is a credulous OPP reply to* $(z, Z \cup \{z\})$*, (2) a sequence* (***PRO:*** $z \rightsquigarrow^+ y$) *such that $z \rightsquigarrow_I y$ and $z \in Z$, or (3) the sequence* (***PRO:*** $y \rightsquigarrow^- x$).

Given a set $X \subseteq A$ we say that M generates the credulous explanation dialogue *D for X iff $D = S \cdot$ (**PRO: win**), where S is a credulous OPP reply to $(x, \{x\})$ for some $x \in X$. We say that D is* conflict-free *iff for all $x, y, z \in A_I$ it holds that if D contains the moves **OPP:** $x \rightsquigarrow y$ and **OPP:** $y \rightsquigarrow z$ then it contains the move **PRO:** $y \rightsquigarrow^- z$.*

The following theorem establishes soundness and completeness.

Theorem 2. *Let $M = (F, I)$ be an AAF where $F = (A, \rightsquigarrow)$. Let $X \subseteq A$ and $G \in I$. It holds that G explains credulous support for X iff M generates a conflict-free credulous explanation dialogue D for X such that $G \in J(D)$.*

Sketch of proof.. Let $M = ((A, \rightsquigarrow), I)$, $X \subseteq A$ and $G \in I$. *(Only if:)* Assume for some $x \in X$ and $E \in Co(G)$ that $x \in E$. Using the fact that $E \subseteq Def_G(E)$ one can recursively define a credulous OPP reply D to (x, Z) for some $Z \subseteq A$ and hence a credulous explanation dialogue $D \cdot$ (**PRO: win**). Conflict-freeness of E implies conflict-freeness of D. *(If:)* Assume *M* generates a credulous explanation dialogue $D \cdot$ (**PRO: win**) for X such that $G \in J(D)$. Then D is a credulous OPP reply to $(a, \{a\})$ for some $a \in X$. It can be shown that the set $E = \{a\} \cup \{x \mid$ **PRO:** $x \rightsquigarrow^+ z \in D\}$ satisfies $E \subseteq Def_G(E)$. Conflict-freeness of D implies conflict-freeness of E. Hence $a \in E$ for some $E \in Co(G)$. □

Example 3. *The listing below shows a conflict-free credulous explanation dialogue $D = (m_1, \ldots, m_6)$ for the observation $\{b\}$ generated by the AAF defined in example 1.*

i	m_i	$J(D, i)$
1	***OPP:*** $c \rightsquigarrow b$	$\{F, G_1, G_2, G_3\}$
2	***PRO:*** $b \rightsquigarrow^+ c$	$\{F, G_1, G_2, G_3\}$
3	***OPP:*** $a \rightsquigarrow b$	$\{F, G_1, G_2, G_3\}$
4	***PRO:*** $a \rightsquigarrow^- b$	$\{G_2, G_3\}$
5	***OPP: ok***	$\{G_2, G_3\}$
6	***PRO: win***	$\{G_2, G_3\}$

Here, the sequence $(m_1, \ldots, m_5)$ is a credulous OPP reply to $(b, \{b\})$. PRO defends b from OPP's attack $c \rightsquigarrow b$ by putting forward the attack $b \rightsquigarrow c$. Since b was already assumed to be accepted, this suffices. At move m_4, PRO defends itself from the attack $a \rightsquigarrow b$ by negating it. This restricts the focus on the abducible AFs G_2 and G_3. The dialogue proves that these two abducible AFs explain credulous support for the observation $\{b\}$. Finally, the skeptical explanation dialogues from example 2 are also credulous explanation dialogues.

5 Abduction in logic programming

In this section we show that AAFs can be instantiated with abductive logic programs, in the same way that regular AFs can be instantiated with regular logic programs. In sections 5.1 and 5.2 we recall the necessary basics of logic programming and the relevant results regarding logic programming as instantiated argumentation. In section 5.3 we present a model of abductive logic programming based on Sakama and Inoue's model of extended abduction [13, 14], and in section 5.2 we show how this model can be instantiated using AAFs.

5.1 Logic programs and partial stable semantics

A logic program P is a finite set of rules, each rule being of the form $C \leftarrow A_1, \ldots, A_n, \sim B_1, \ldots, \sim B_m$ where $C, A_1, \ldots, A_n, B_1, \ldots, B_m$ are *atoms*. If $m = 0$ then the rule is called *definite*. If both $n = 0$ and $m = 0$ then the rule is called a *fact* and we identify it with the atom C. We assume that logic programs are ground. Alternatively, P can be regarded as the set of ground instances of a set of non-ground rules. We denote by At_P the set of all (ground) atoms occurring in P. The logic programming semantics we focus on can be defined using *3-valued interpretations* [19]:

Definition 11. *A 3-valued interpretation I of a logic program P is a pair $I = (T, F)$ where $T, F \subseteq At_P$ and $T \cap F = \emptyset$. An atom $A \in At_P$ is* true *(resp.* false, undecided*) in I iff $A \in T$ (resp. $A \in F$, $A \in At_P \setminus (T \cup F)$).*

The following definition of a *partial stable model* is due to Przymusinski [19]. Given a logic program P and 3-valued interpretation I of P, the *GL-transformation* $\frac{P}{I}$ is a logic program obtained by replacing in every rule in P every premise $\sim B$ such that B is true (resp. undecided, false) in I by the atoms 0 (resp. $\frac{1}{2}$, 1), where 0 (resp. $\frac{1}{2}$,

1) are defined to be false (resp. undecided, true) in every interpretation. It holds that for all 3-valued interpretations I of P, $\frac{P}{I}$ is definite (i.e., consists only of definite rules). This means that $\frac{P}{I}$ has a unique *least* 3-valued interpretation (T, F) with minimal T and maximal F that satisfies all rules. That is, for all rules $C \leftarrow A_1, \ldots, A_n$, in $\frac{P}{I}$, C is true (resp. *not* false) in (T, F) if for all $i \in \{1, \ldots, n\}$, A_i is true (resp. *not* false) in (T, F). Given a 3-valued interpretation I, the least 3-valued interpretation of $\frac{P}{I}$ is denoted by $\Gamma(I)$. This leads to the following definition of a *partial stable model* of a logic program, along with the associated notions of consequence.

Definition 12. *[19] Let P be a logic program. A 3-valued interpretation I is a* partial stable model *of P iff $I = \Gamma(I)$. We say that an atom C is a skeptical (resp. credulous) consequence of P iff C is true in all (resp. some) partial stable models of P.*

It has been shown that the above defined notion of skeptical consequence coincides with the *well-founded* semantics [19].

5.2 Logic programming as argumentation

Wu et al. [22] have shown that a logic program P can be transformed into an AF F in such a way that the consequences of P under the partial stable semantics can be computed by looking at the complete extensions of F. The idea is that an argument consists of a conclusion $C \in At_P$, a set of rules $R \subseteq P$ used to derive C and a set $N \subseteq At_P$ of atoms that must be underivable in order for the argument to be acceptable. The argument is attacked by another argument with a conclusion C' iff $C' \in N$. The following definition, apart from notation, is due to Wu et al. [22].

Definition 13. *Let P be a logic program. An instantiated argument is a triple (C, R, N), where $C \in At_P$, $R \subseteq P$ and $N \subseteq At_P$. We say that P generates (C, R, N) iff either:*

- *$r = C \leftarrow \sim B_1, \ldots, \sim B_m$ is a rule in P, $R = \{r\}$ and $N = \{B_1, \ldots, B_m\}$.*
- *(1) $r = C \leftarrow A_1, \ldots, A_n, \sim B_1, \ldots, \sim B_m$ is a rule in P, (2) P generates, for each $i \in \{1, \ldots, n]$ an argument (A_i, R_i, N_i) such that $r \notin R_i$, and (3) $R = \{r\} \cup R_1 \cup \ldots \cup R_n$ and $N = \{B_1, \ldots, B_m\} \cup N_1 \cup \ldots \cup N_n$.*

We denote the set of arguments generated by P by A_P. Furthermore, the attack relation generated by P is denoted by $\rightsquigarrow_P$ and is defined by $(C, R, N) \rightsquigarrow_P (C', R', N')$ iff $C \in N'$.

The following theorem states that skeptical (resp. credulous) acceptance in $(A_P, \rightsquigarrow_P)$ corresponds with skeptical (resp. credulous) consequences in P as defined in definition 12. It follows from theorems 15 and 16 due to Wu et al. [22].

Theorem 3. *Let P be a logic program. An atom $C \in At_P$ is a skeptical (resp. credulous) consequence of P iff some $(C, R, N) \in A_P$ is skeptically (resp. credulously) accepted in $(A_P, \rightsquigarrow_P)$.*

5.3 Abduction in logic programming

The model of abduction in logic programming that we use is based on the model of *extended* abduction studied by Inoue and Sakama [13, 14]. They define an abductive logic program (ALP) to consist of a logic program and a set of atoms called *abducibles*.

Definition 14. *An abductive logic program is a pair (P, U) where P is a logic program and $U \subseteq At_P$ a set of facts called abducibles.*

Note that, as before, the set U consists of ground facts of the form $C \leftarrow$ (identified with the atom C) and can alternatively be regarded as the set of ground instances of a set of non-ground facts. A hypothesis, according to Inoue and Sakama's model, consists of both a positive element (i.e., abducibles added to P) and a negative element (i.e., abducibles removed from P).

Definition 15. *Let ALP $= (P, U)$ be an abductive logic program. A hypothesis is a pair (Δ^+, Δ^-) such that $\Delta^+, \Delta^- \subseteq U$ and $\Delta^+ \cap \Delta^- = \emptyset$. A hypothesis (Δ^+, Δ^-) skeptically (resp. credulously) explains a query $Q \in At_P$ if and only if Q is a skeptical (resp. credulous) consequence of $(P \cup \Delta^+) \setminus \Delta^-$.*

Note that Sakama and Inoue focus on computation of explanations under the stable model semantics of P, and require P to be acyclic to ensure that a stable model of P exists and is unique [14]. We, however, define explanation in terms of the consequences according to the partial stable models of P, which always exist even if P is not acyclic [19], so that we do not need this requirement.

The following example demonstrates the previous two definitions.

Example 4. *Let ALP $= (P, U)$ where $P = \{(p \leftarrow \sim s, r), (p \leftarrow \sim s, \sim q), (q \leftarrow \sim p), r\}$ and $U = \{r, s\}$. The hypothesis $(\{s\}, \emptyset)$ skeptically explains q, witnessed by the unique model $I = (\{r, s, q\}, \{p\})$ satisfying $I = \Gamma(I)$. Similarly, $(\{s\}, \{r\}))$ skeptically explains q and $(\emptyset, \{r\}))$ credulously explains q.*

5.4 Instantiated abduction in argumentation

In this section we show that an AAF (F, I) can be instantiated on the basis of an abductive logic program (P, U). The idea is that every possible hypothesis (Δ^+, Δ^-) maps to an abducible AF generated by the logic program $(P \cup \Delta^+) \setminus \Delta^-$. The hypotheses for a query Q then correspond to the abducible AFs that explain the observation X consisting of all arguments with conclusion Q. The construction of (F, I) on the basis of (P, U) is defined as follows.

Definition 16. *Let ALP $= (P, U)$ be an abductive logic program. Given a hypothesis (Δ^+, Δ^-), we denote by $F_{(\Delta^+, \Delta^-)}$ the AF $(A_{(P \cup \Delta^+) \setminus \Delta^-}, \rightsquigarrow_{(P \cup \Delta^+) \setminus \Delta^-})$. The AAF* generated by *ALP is denoted by M_{ALP} and defined by $M_{ALP} = ((A_P, \rightsquigarrow_P), I_{ALP})$, where $I_{ALP} = \{F_{(\Delta^+, \Delta^-)} \mid \Delta^+, \Delta^- \subseteq U, \Delta^+ \cap \Delta^- = \emptyset\}$.*

The following theorem states the correspondence between the explanations of a query Q in an abductive logic program *ALP* and the explanations of an observation X in the AAF M_{ALP}.

Theorem 4. *Let ALP $= (P, U)$ be an abductive logic program, $Q \in At_P$ a query and (Δ^+, Δ^-) a hypothesis. Let $M_{ALP} = (F, I)$. We denote by X_Q the set $\{(C, R, N) \in A_P \mid C = Q\}$. It holds that (Δ^+, Δ^-) skeptically (resp. credulously) explains Q iff $F_{(\Delta^+, \Delta^-)}$ skeptically (resp. credulously) explains X_Q.*

Proof of theorem 4. Via theorem 3 and definitions 15 and 16. □

This theorem shows that our model of abduction in argumentation can indeed be seen as an abstraction of abductive logic programming.

Example 5. *Let ALP $= (P, U)$ be the ALP as defined in example 4. All arguments generated by ALP are:*

$$\begin{array}{lllllll} a & = & (p, \{(p \leftarrow \sim s, r), r\}, \{s\}) & & d & = & (r, \{r\}, \emptyset) \\ b & = & (q, \{(q \leftarrow \sim p)\}, \{p\}) & & e & = & (s, \{s\}, \emptyset) \\ c & = & (p, \{(p \leftarrow \sim s, \sim q)\}, \{s, q\}) & & & & \end{array}$$

Given these definitions, the AAF in example 1 is equivalent to M_{ALP}. In example 4 we saw that q is skeptically explained by $(\{s\}, \emptyset)$ and $(\{s\}, \{r\})$, while $(\emptyset, \{r\})$ only credulously explains it. Indeed, looking again at example 1, we see that $G_1 = F_{(\{s\},\emptyset)}$ and $G_3 = F_{(\{s\},\{r\})}$ explain skeptical support for the observation $\{b\} = X_q$, while $G_2 = F_{(\emptyset,\{r\})}$ only explains credulous support.

Remark 2. *This method of instantiation shows that, on the abstract level, hypotheses cannot be represented by independently selectable abducible arguments. The running example shows e.g. that a and d cannot be added or removed independently. (Cf. remark 1.)*

6 Related work

We already discussed Sakama's [20] model of abduction in argumentation and mentioned some differences. Our approach is more general because we consider a hypothesis to be a change to the AF that is applied as a whole, instead of a set of independently selectable abducible arguments. On the other hand, Sakama's method supports a larger range semantics, including (semi-)stable and skeptical preferred semantics. Furthermore, Sakama also considers observations leading to rejection of arguments, which we do not.

Some of the ideas we applied also appear in work by Wakaki et al. [21]. In their model, ALPs generate instantiated AFs and hypotheses yield a division into active/inactive arguments.

Kontarinis et al. [15] use term rewriting logic to compute changes to an abstract AF with the goal of changing the status of an argument. Two similarities to our work are: (1) our production rules to generate dialogues can be seen as a kind of term rewriting rules. (2) their approach amounts to rewriting goals into statements to the effect that certain attacks in the AF are enabled or disabled. These statements resemble the moves **PRO:** $x \rightsquigarrow^+ y$ and **PRO:** $x \rightsquigarrow^- y$ in our system. However, they treat attacks as entities that can be enabled or disabled independently. As discussed, different arguments (or in this case attacks associated with arguments) cannot be regarded as independent entities, if the abstract model is instantiated.

Goal oriented change of AFs is also studied by Baumann [2], Baumann and Brewka [3], Bisquert et al. [4] and Perotti et al. [5]. Furthermore, Booth et al. [8] and Coste-Marquis et al. [11] frame it as a problem of *belief revision*. Other studies in which changes to AFs are considered include [6, 10, 16, 18].

7 Conclusions and Future work

We developed a model of abduction in abstract argumentation, in which changes to an AF act as explanations for skeptical/credulous support for observations. We presented sound and complete dialogical proof procedures for the main decision problems, i.e., finding explanations for skeptical/credulous support. In addition, we showed that our model of abduction in abstract argumentation can be seen as an abstract form of abduction in logic programming.

As a possible direction for future work, we consider the incorporation of additional criteria for the selection of good explanations, such as minimality with respect to the added and removed arguments/attacks, as well as the use of arbitrary preferences over different abducible AFs. An interesting question is whether the proof theory can be adapted so as to yield only the preferred explanations.

8 Acknowledgements

Richard Booth is supported by the Fonds National de la Recherche, Luxembourg (DYNGBaT project).

REFERENCES

[1] *Scalable Uncertainty Management - 7th International Conference, SUM 2013, Washington, DC, USA, September 16-18, 2013. Proceedings*, 2013.

[2] Ringo Baumann, 'Normal and strong expansion equivalence for argumentation frameworks', *Artif. Intell.*, **193**, 18–44, (2012).

[3] Ringo Baumann and Gerhard Brewka, 'Expanding argumentation frameworks: Enforcing and monotonicity results', in *Proc. COMMA*, pp. 75–86, (2010).

[4] Pierre Bisquert, Claudette Cayrol, Florence Dupin de Saint-Cyr, and Marie-Christine Lagasquie-Schiex, 'Enforcement in argumentation is a kind of update', In *SUM* [1], pp. 30–43.

[5] Guido Boella, Dov M. Gabbay, Alan Perotti, Leon van der Torre, and Serena Villata, 'Conditional labelling for abstract argumentation', in *TAFA*, pp. 232–248, (2011).

[6] Guido Boella, Souhila Kaci, and Leendert van der Torre, 'Dynamics in argumentation with single extensions: Attack refinement and the grounded extension (extended version)', in *ArgMAS*, pp. 150–159, (2009).

[7] Richard Booth, Souhila Kaci, and Tjitze Rienstra, 'Property-based preferences in abstract argumentation', in *ADT*, pp. 86–100, (2013).

[8] Richard Booth, Souhila Kaci, Tjitze Rienstra, and Leon van der Torre, 'A logical theory about dynamics in abstract argumentation', In *SUM* [1], pp. 148–161.

[9] Martin Caminada, 'Preferred semantics as socratic discussion', in *Proceedings of the 11th AI* IA Symposium on Artificial Intelligence*, pp. 209–216, (2010).

[10] Claudette Cayrol, Florence Dupin de Saint-Cyr, and Marie-Christine Lagasquie-Schiex, 'Change in abstract argumentation frameworks: Adding an argument', *Journal of Artificial Intelligence Research*, **38**(1), 49–84, (2010).

[11] Sylvie Coste-Marquis, Sébastien Konieczny, Jean-Guy Mailly, and Pierre Marquis, 'On the revision of argumentation systems: Minimal change of arguments status', *Proc. TAFA*, (2013).

[12] Phan Minh Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *Artif. Intell.*, **77**(2), 321–358, (1995).

[13] Katsumi Inoue and Chiaki Sakama, 'Abductive framework for nonmonotonic theory change', in *IJCAI*, pp. 204–210. Morgan Kaufmann, (1995).

[14] Katsumi Inoue and Chiaki Sakama, 'Computing extended abduction through transaction programs', *Ann. Math. Artif. Intell.*, **25**(3-4), 339–367, (1999).

[15] Dionysios Kontarinis, Elise Bonzon, Nicolas Maudet, Alan Perotti, Leon van der Torre, and Serena Villata, 'Rewriting rules for the computation of goal-oriented changes in an argumentation system', in *Computational Logic in Multi-Agent Systems*, 51–68, Springer, (2013).

[16] Beishui Liao, Li Jin, and Robert C. Koons, 'Dynamics of argumentation systems: A division-based method', *Artif. Intell.*, **175**(11), 1790–1814, (2011).

[17] Sanjay Modgil and Martin Caminada, 'Proof theories and algorithms for abstract argumentation frameworks', in *Argumentation in Artificial Intelligence*, 105–129, (2009).

[18] Emilia Oikarinen and Stefan Woltran, 'Characterizing strong equivalence for argumentation frameworks', *Artificial intelligence*, **175**(14-15), 1985–2009, (2011).

[19] Teodor C. Przymusinski, 'The well-founded semantics coincides with the three-valued stable semantics', *Fundam. Inform.*, **13**(4), 445–463, (1990).

[20] Chiaki Sakama, 'Abduction in argumentation frameworks and its use in debate games', in *Proceedings of the 1st International Workshop on Argument for Agreement and Assurance (AAA)*, (2013).

[21] Toshiko Wakaki, Katsumi Nitta, and Hajime Sawamura, 'Computing abductive argumentation in answer set programming', in *Proc. ArgMAS*, pp. 195–215, (2009).

[22] Yining Wu, Martin Caminada, and Dov M. Gabbay, 'Complete extensions in argumentation coincide with 3-valued stable models in logic programming', *Studia Logica*, **93**(2-3), 383–403, (2009).

View publication stats