Arguing about Complex Formulas: Generalizing Abstract Dialectical Frameworks

Jesse Heyninck[1,2], **Matthias Thimm**[3], **Gabriele Kern-Isberner**[1], **Tjitze Rienstra**[3], **Kenneth Skiba**[3]

[1]Technische Universität Dortmund, Dortmund, Germany
[2]University of Cape Town and CAIR, South-Africa
[3]University of Koblenz-Landau, Koblenz, Germany

Abstract

Abstract dialectical frameworks (in short, ADFs) are a unifying model of formal argumentation, where argumentative relations between arguments are represented by assigning acceptance conditions to atomic arguments. This idea is generalized by letting acceptance conditions being assigned to complex formulas, resulting in conditional abstract dialectical frameworks (in short, cADFs). We define the semantics of cADFs in terms of a non-truth-functional four-valued logic, and study the semantics in-depth, by showing existence results and proving that all semantics are generalizations of the corresponding semantics for ADFs.

1 Introduction

Formal argumentation is one of the major approaches to knowledge representation. In the seminal paper (Dung 1995), *abstract argumentation frameworks* were conceived of as directed graphs where nodes represent arguments and edges between these nodes represent attacks. So-called *argumentation semantics* determine which sets of arguments can be reasonably upheld together given such an argumentation graph. Various authors have remarked that other relations between arguments are worth consideration. For example, in (Cayrol and Lagasquie-Schiex 2005), *bipolar argumentation frameworks* are developed, where arguments can support as well as attack each other.

The last decades saw a proliferation of such extensions of the original formalism of (Dung 1995), and it has often proven hard to compare the resulting different dialects of the argumentation formalisms. To cope with the resulting multiplicity, (Brewka et al. 2013) introduced *abstract dialectical argumentation* that aims to unify these different dialects (Polberg 2016). Just like in (Dung 1995), *abstract dialectical frameworks* (in short, ADFs) are directed graphs. In difference to abstract argumentation frameworks, however, in ADFs, edges between nodes do not necessarily represent attacks but can encode *any* relationship between arguments. Such a generality is achieved by associating an *acceptance condition* with each argument, which is a Boolean formula in terms of the parents of the argument that expresses the conditions under which an argument can be accepted. This results in an ADF being defined as a triple (At, L, C) where At represents a set of atoms or arguments, $L \subseteq \mathsf{At} \times \mathsf{At}$ represents a set of argumentative relations between the atoms and C is a set of acceptance conditions C_s for every s. As such, ADFs are able to capture all of the major semantics of abstract argumentation and offer a general framework for argumentation-based inference. Furthermore, ADFs were shown to be able to capture a number of non-argumentative formalisms such as logic programming (Brewka et al. 2013). Recently, first attempts were made to translate non-monotonic conditional logics in ADFs (Heyninck et al. 2019).

However, there are limits to the representative capabilities of ADFs, both on a conceptual as well as a more technical level. On the conceptual level, acceptance conditions are assigned to atoms, which means that, e. g., an attack on a set of arguments cannot be captured by ADFs. For example, to state that the set $\{p, q\}$ is attacked by r we would have to be able to set the acceptance condition of $p \wedge q$ to $\neg r$, which is not possible in ADFs. Likewise, it is not immediately obvious how to represent more complicated logic programming languages in ADFs, such as disjunctive logic programming. Such limitations are, not unsurprisingly, also reflected on a more technical level. For example, a (polynomial) translation of disjunctive logic programming into ADFs is impossible in view of complexity results on disjunctive logic programming and ADFs. Finally, in (Heyninck et al. 2019) shows that only a fragment of the full language of conditional logics can be translated in ADFs in view of their limited syntax.

In this paper, we generalize ADFs as to allow for the assignment of acceptance conditions to complex formulas. This results in *conditional abstract dialectical frameworks* (in short, cADFs) which are sets of acceptance pairs of the form $\phi \lhd \psi$ with arbitrary formulas ϕ and ψ, interpreted as a defeasible version of "ϕ is the case if and only if ψ is the case". The semantics of cADFs are formulated as a generalization of the semantics of ADFs, with the Γ-function, on its turn based on a non-truth-functional four-valued logic, as a central component. Some of the main results include existence results for all the major semantics, as well as the definition of the so-called *grounded state*, a single-state semantics which can be iteratively constructed and represents the minimal information entailed by a given cADF.

Outline of this Paper: We first state all the necessary preliminaries in Section 2 on propositional logic (Section

2.1), and abstract dialectical argumentation (Section 2.2). The syntax of *conditional abstract dialectical frameworks* cADFs is introduced in Section 3. In Section 4, a four-valued logic, which will form the basis of the semantics of cADFs, is defined and studied. In Section 5, we then define and study the admissible, complete, preferred and grounded semantics for cADFs. A unique, iteratively constructible analogue to the grounded extension, called the *grounded state*, is introduced in Section 6. Related work is discussed in Section 7 and a conclusion is drawn in Section 8.

2 Preliminaries

In the following, we briefly recall some general preliminaries on propositional logic, as well as technical details on conditional logic and ADFs (Brewka et al. 2013).

2.1 Propositional Logic

For a set At of atoms let $\mathcal{L}(\mathsf{At})$ be the corresponding propositional language constructed using the usual connectives $\wedge$ (*and*), $\vee$ (*or*), $\neg$ (*negation*) and $\rightarrow$ (*material implication*). A (classical) *interpretation* (also called *possible world*) ω for a propositional language $\mathcal{L}(\mathsf{At})$ is a function $\omega : \mathsf{At} \rightarrow \{\mathsf{T}, \mathsf{F}\}$. Let $\mathcal{V}^2(\mathsf{At})$ denote the set of all interpretations for At. We simply write $\mathcal{V}^2$ if the set of atoms is implicitly given. An interpretation ω *satisfies* (or is a *model* of) an atom $a \in \mathsf{At}$, denoted by $\omega \models a$, if and only if $\omega(a) = \mathsf{T}$. The satisfaction relation $\models$ is extended to formulas as usual. For $\Phi \subseteq \mathcal{L}(\mathsf{At})$ we also define $\omega \models \Phi$ if and only if $\omega \models \phi$ for every $\phi \in \Phi$. Define the set of models $\mathsf{Mod}^2(X) = \{\omega \in \mathcal{V}^2(\mathsf{At}) \mid \omega \models X\}$ for every formula or set of formulas X. A formula or set of formulas X_1 *entails* another formula or set of formulas X_2, denoted by $X_1 \vdash X_2$, if $\mathsf{Mod}^2(X_1) \subseteq \mathsf{Mod}^2(X_2)$. A formula ϕ is a tautology if $\mathsf{Mod}^2(\phi) = \mathcal{V}^2(\mathsf{At})$ and a falsity if $\mathsf{Mod}^2(\phi) = \emptyset$.

2.2 Abstract Dialectical Frameworks

We briefly recall some technical details on ADFs following loosely the notation from (Brewka et al. 2013). An ADF D is a tuple $D = (\mathsf{At}, L, C)$ where At is a finite set of atoms, $L \subseteq \mathsf{At} \times \mathsf{At}$ is a set of links, and $C = \{C_s\}_{s \in \mathsf{At}}$ is a set of total functions $C_s : 2^{par_D(\mathsf{At})} \rightarrow \{\top, \bot\}$ for each $s \in \mathsf{At}$ with $par_D(s) = \{s' \in \mathsf{At} \mid (s', s) \in L\}$ (also called acceptance functions). An acceptance function C_s defines the cases when the statement s can be accepted (truth value $\top$), depending on the acceptance status of its parents in D. By abuse of notation, we will often identify an acceptance function C_s by its equivalent *acceptance condition* which models the acceptable cases as a propositional formula.

Example 1. We consider the following ADF $D_1 = (\{a, b, c, d\}, L, C)$ with $L = \{(a, b), (b, a), (a, c), (b, c)\}$ and $C_a = \neg b$, $C_b = \neg a$, and $C_c = \neg a \vee \neg b$.
Informally, the acceptance conditions can be read as "a is accepted if b is not accepted", "b is accepted if a is not accepted" and "c is accepted if a is not accepted or b is not accepted".

An ADF $D = (\mathsf{At}, L, C)$ is interpreted through 3-valued interpretations $\nu : \mathsf{At} \rightarrow \{\mathsf{T}, \mathsf{F}, \mathsf{U}\}$. We denote the set of all 3-valued interpretations over At by $\mathcal{V}^3(\mathsf{At})$. We define the information order $<_i$ over $\{\mathsf{T}, \mathsf{F}, \mathsf{U}\}$ by making U the minimal element: $\mathsf{U} <_i \mathsf{T}$ and $\mathsf{U} <_i \mathsf{F}$, and $\dagger \leq_i \ddagger$ iff $\dagger <_i \ddagger$ or $\dagger = \ddagger$ for any $\dagger, \ddagger \in \{\mathsf{T}, \mathsf{F}, \mathsf{U}\}$. This order is lifted point-wise as follows (given $\nu, \nu' \in \mathcal{V}^3(\mathsf{At})$): $\nu \leq_i \nu'$ iff $\nu(s) \leq_i \nu'(s)$ for every $s \in \mathsf{At}$. The set of two-valued interpretations extending a 3-valued interpretation v is defined as $[\nu]^2 = \{\omega \in \mathcal{V}^2(\mathsf{At}) \mid \nu \leq_i \omega\}$. Given a set of 3-valued interpretations $V \subseteq \mathcal{V}^3(\mathsf{At})$, $\sqcap_i V$ is the 3-valued interpretation defined via $\sqcap_i V(s) = \dagger$ if for every $\nu \in V$, $\nu(s) = \dagger$, for any $\dagger \in \{\mathsf{T}, \mathsf{F}, \mathsf{U}\}$, and $\sqcap_i V(s) = \mathsf{U}$ otherwise. Truth values based on a three-valued interpretations can now be assigned to complex formulas ϕ by taking $\sqcap_i [\nu]^2(\phi)$. All major semantics of ADFs single out three-valued interpretations in which the truth value of every atom $s \in \mathsf{At}$ is, in some sense, in alignment or agreement with the truth value of the corresponding condition C_s. The Γ-function enforces this intuition by mapping an interpretation ν to a new interpretation $\Gamma_D(\nu)$, which assigns to every atom s exactly the truth value assigned by ν to C_s, i.e.:

$$\Gamma_D(\nu) : \mathsf{At} \rightarrow \{\mathsf{T}, \mathsf{F}, \mathsf{U}\} \text{ where } s \rightarrow \sqcap_i \{\omega(C_s) \mid \omega \in [\nu]^2\}.$$

Definition 1. Let $D = (\mathsf{At}, L, C)$ be an ADF with $\nu \in \mathcal{V}(\mathsf{At})$ a 3-valued interpretation:

- ν is *admissible for* D iff $\nu \leq_i \Gamma_D(\nu)$.
- ν is *complete for* D iff $\nu = \Gamma_D(\nu)$.
- ν is *preferred for* D iff ν is $\leq_i$-maximal among all admissible interpretations.
- ν is *grounded for* D iff ν is $\leq_i$-minimal among all complete interpretations.

We denote by $\mathsf{admissible}$, $\mathsf{complete}(D)$, $\mathsf{prf}(D)$, and $\mathsf{grounded}(D)$ the sets of complete, preferred, and grounded interpretations of D, respectively.

Notice that ν is admissible iff $\nu(s) \leq_i \sqcap_i [\nu]^2(C_s)$ for every $s \in S$ and likewise, ν is complete iff $\nu(s) = \sqcap_i [\nu]^2(C_s)$ for every $s \in S$. It can thus be observed that the logic defined by $\sqcap_i [\nu]^2$ is, essentially, the logic underlying ADFs, in the sense that the evaluation of acceptance conditions under $\sqcap_i [\nu]^2$ is the fundamental operation underlying every semantical notion of ADFs. It should be furthermore noted that $\sqcap_i [\nu]^2$ does not give rise to a *truth-functional logic*. Recall that a truth-functional logic is a logic in which the truth value assigned to a complex formula is a function of the truth values of its component formulas. E.g. for a truth-functional logic, the truth value of $a \vee \neg b$ is determined completely by the truth value of a and $\neg b$. For example, given $\nu(a) = \mathsf{U}$ and $\nu(b) = \mathsf{U}$, $\sqcap_i [\nu]^2(a \vee \neg a) = \mathsf{T}$ whereas $\sqcap_i [\nu]^2(a \vee \neg b) = \mathsf{U}$.

Example 2 (Example 1 continued)**.** The ADF of Example 1 has three complete models ν_1, ν_2, ν_3 with:

$$\begin{array}{lll} \nu_1(a) = \mathsf{T} & \nu_1(b) = \mathsf{F} & \nu_1(c) = \mathsf{T} \\ \nu_2(a) = \mathsf{F} & \nu_2(b) = \mathsf{T} & \nu_2(c) = \mathsf{T} \\ \nu_3(a) = \mathsf{U} & \nu_3(b) = \mathsf{U} & \nu_3(c) = \mathsf{U} \end{array}$$

ν_3 is the grounded interpretation whereas ν_1 and ν_2 are both preferred.

3 Syntax of cADFs

The syntactical representation $D = (S, L, C)$ of an ADF contains some superfluous information. In particular, as there is a link between a statement s and s' iff s is mentioned in the acceptance condition of s', the set of links does not contain any information not already derivable from the set of acceptance conditions C. As such, given a set of atoms S, we can simply write an ADF as a set of statements $s \lhd C_s$ if C_s is the acceptance condition of s. So the ADF D_1 from Example 1 can be simply written as:

$$D_1 = \{a \lhd \neg b, b \lhd \neg a, c \lhd \neg a \vee \neg b\}$$

An ADF is determined by a set of propositional formulæ that, when evaluated to true, make a certain statement, which is a simple atom, true as well, and when evaluated to false, make the simple atom false as well. In other words, $\lhd$ can be read as a *approximate if and only if*: $s \lhd C_s$ means that the truth-values s and C_s should be aligned. $\lhd$ can truly be read as a *approximate iff*, since it might not always be possible to align the truth values of s and C_s in such a way that they take on exactly the same (determinate) truth value. To see this, consider, e. g., $a \lhd \neg a$. We generalise this framework by allowing these statements to be arbitrary propositional formulæ:

Definition 2. Given a set of atoms At, a *conditional abstract dialectical framework* cADF Π w.r.t. At is a finite set of *acceptance pairs* over At, where an *acceptance pair* is of the form:

$$\phi \lhd \psi$$

with ϕ and ψ being propositional formulæ over At.

In order to stick to ADF terminology we call ϕ the *statement* and ψ the *condition* of the acceptance pair $\phi \lhd \psi$. We omit the reference to the signature At when it is clear from context.

Example 3. Consider a cADF $\Pi_1 = \{c_1, c_2, c_3\}$ with

$$\begin{aligned} c_1 &: & p \vee s \vee q \lhd \top \\ c_2 &: & p \wedge s \lhd \neg q \\ c_3 &: & (p \wedge q) \vee (p \wedge s) \lhd t \end{aligned}$$

This cADF can be used to model an argument of a group of friends about making plans on a Sunday. They are discussing whether to go to a party (p), to the swimming pool (s) or go to a pub quiz (q). They want to do at least one of these three things (c_1). However, if they go to the quiz, they won't be able to still go to the pool *and* go to the party (represented by the attack of q on $p \wedge s$ in c_2). If everyone arrives on time (t), they would like to go to both the quiz and the party, or to both the pool and the party (c_3). We notice that without adding further atoms, an attack from q on the set $\{p, s\}$, as formalized by c_2, cannot be represented in ADFs.

We observe that this simple generalization w.r.t. ADFs results in the following additional points of expressiveness in comparison to ADFs:

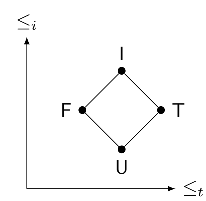
- cADFs allow for complex formulas as statements, as demonstrated by $(p \wedge q) \vee (p \wedge s) \lhd t$ in Example 3.
- cADFs allow for "incomplete" specifications, i.e. they do not force the user to formulate an acceptance condition for every atom, as demonstrated in Example 3, where t has no acceptance condition.
- cADFs allow for "overspecifications" or conflicting specifications, as demonstrated by the cADF $\{a \lhd b, \neg a \lhd b\}$ where both a and $\neg a$ have the acceptance condition b.
- cADFs allow for indeterminism, as demonstrated by the cADF $\{a \vee b \lhd \top\}$, where $a \vee b$ is required to be true, but no further information on which of the disjuncts is required to be true is given.

To cope with this higher expressiveness semantically, it will prove useful to move from three-valued interpretations to four-valued interpretations. To assign truth values to complex formulas on the basis of four-valued interpretations, we generalize the logic defined by $\sqcap_i[v]^2$ to a four-valued setting in Section 4. We then generalize the semantics of ADFs to cADFs on the basis of this four-valued logic in Section 5.

4 A Four-Valued Logic Based on Completions

We first define a four-valued logic 4CL which generalizes the idea of completions known from the logic underlying ADFs defined by $[\nu]^2$, which preserves classical tautologies and falsities. We first recall four-valued interpretations. A *four-valued interpretation* v : At $\rightarrow$ $\{\mathsf{T}, \mathsf{F}, \mathsf{I}, \mathsf{U}\}$ assigns to every atom a truth value T (true), F (false), U (undecided) or I (inconsistent). We will also write an interpretation $v \in \mathcal{V}^4(\{a_1, \ldots, a_n\})$ as $v(a_1) \ldots v(a_n)$, e. g., v over $\{p, q\}$ with $v(p) = \mathsf{T}$ and $v(q) = \mathsf{U}$ will be written as TU. We denote the set of four-valued interpretations over At by $\mathcal{V}^4(\mathsf{At})$. Notice that $\mathcal{V}^2(\mathsf{At}) \subseteq \mathcal{V}^3(\mathsf{At}) \subseteq \mathcal{V}^4(\mathsf{At})$. If it is clear that an interpretation is two- respectively three-valued, we will denote it by (a possibly indexed) ω respectively ν.

Two useful orders over these truth values are the *information order* $\leq_i$ and the *truth order* $\leq_t$, which form the following bilattice-structure (Fitting 2006):



Notice that $\mathcal{V}^4(\mathsf{At})$ also forms a bounded lattice under $\leq_i$ with v_U and v_I as least and greatest element respectively (where v_U is defined as the interpretation that sets $v_\mathsf{U}(a) = \mathsf{U}$ for every $a \in$ At and v_I is defined as $v_\mathsf{I}(a) = \mathsf{I}$ for every $a \in$ At).

We shall interpret the four truth values, at least for atoms, in the same way as (Belnap 2019): U (*undecided*) means that we have no explicit information for either the truth nor the falsity of an atom. T (*true*) respectively F (*false*) means that we have explicit information only for the truth respectively the falsity of the atom in question. Finally, I (*inconsistent*) means that we have explicit information for both the truth and the falsity of the atom in question. When it comes to complex formulas, we shall see that we take a somewhat hybrid position between the truth values expressing merely explicit information and the truth values standing for objective truth. In particular, the logic we will define here will allow for *logically contingent* formulas, i. e., formulas which are neither classical tautologies nor classical falsities, to be assigned any of the four truth values, whereas classical tautologies and classical falsities will always be assigned T respectively F by any interpretation. Intuitively, this means that even though the truth value of $s \in \mathsf{At}$ might be undetermined (U) or inconsistent (I), the logic will still evaluate $s \vee \neg s$ as true. This is in complete agreement with ADFs, where tautologies and logical falsities are always evaluated in agreement with classical logic by $\sqcap_i [v]^2$.

Semantically, we proceed as follows: we construct a set of sets of (two-valued) worlds on the basis of a four-valued interpretation v that represents the beliefs expressed by v. Just like in the logic underlying ADFs $\sqcap_i[\nu]^2$, a set of (two-valued) worlds will be used to represent a three-valued interpretation ν. The worlds in $[\nu]^2$ represent equally plausible candidates of the actual world in view of the beliefs expressed by the three-valued interpretation ν. Likewise, a set of three-valued interpretations $[v]^3$ will be used to represent the information expressed by a four-valued interpretation v. $[v]^3$ consists of the three-valued interpretations that *jointly* represent the information expressed by v. Notice the difference with $[\nu]^2$: $[\nu]^2$ consists of equally plausible candidates of the actual world in view of the information expressed by v, whereas $[v]^3$ contains interpretations that *taken together* represent the information expressed by v. We now develop this idea in more formal details.

Given a four-valued interpretation, we define the set of two-valued completions of v, $[v]^2$, in two steps. First, we construct $[v]^3$, which converts $v \in \mathcal{V}^4(\mathsf{At})$ to a set of three-valued interpretations $[v]^3 \subseteq \mathcal{V}^3(\mathsf{At})$. Then, we obtain $[v]^2 \subseteq \wp(\mathcal{V}^2(\mathsf{At}))$ by converting every three-valued interpretation $\nu \in [v]^3$ to a set of two-valued interpretations $[\nu]^2$.

Definition 3. Given a four-valued interpretation $v \in \mathcal{V}(\mathsf{At})$, $[v]^3 = \{\nu \in \mathcal{V}^3(\mathsf{At}) \mid \text{for every } s \in \mathsf{At} : \text{ if } v(s) = \mathsf{I} \text{ then } \nu(s) \in \{\mathsf{T},\mathsf{F}\}, \nu(s) = v(s) \text{ otherwise}\}$

In other words, $[v]^3$ is obtained by replacing every assignment of an atom s to I to an assignment of s to T or to F.

Notice that $[v]^3$ consists of the $\leq_i$-maximal three-valued interpretations that v extends:

Fact 1. For any $v \in \mathcal{V}^4(\mathsf{At})$, $[v]^3 = \max_{\leq_i}(\{\nu \in \mathcal{V}^3(\mathsf{At}) \mid \nu \leq_i v\})$.[1]

Example 4. Consider $v = \mathsf{TUI}$ over $\Sigma = abc$. Then $[v]^3 = \{\mathsf{TUT}, \mathsf{TUF}\}$.

We are now ready to define the *four-valued completions* $[v]^4$ of v:

Definition 4. Given some $v \in \mathcal{V}^4(\mathsf{At})$, the *four-valued completions of* v are defined as: $[v]^4 = \{[v']^2 \mid v' \in [v]^3\}$.

Thus, $[v]^4$ is obtained by first constructing $[v]^3$, and then taking for every $\nu \in [v]^3$ the set of two-valued completions of ν. The intuition behind this is as follows: $v(s) = \mathsf{I}$ means that we have information for both s being true and s being false. Thus, the interpretations where we set $\nu_1(s) = \mathsf{T}$ and $\nu_2(s) = \mathsf{F}$ are both (partial yet consistent) representations of the state of the world represented by v. Hence $[v]^3$ can be viewed as the set of three-valued interpretations that together form the representation of the state of the world represented by v. We then construct for every such representation a set of two-valued interpretations, which represent equally plausible candidates of the state of the world represented by $\nu \in [v]^3$. Altogether, $[v]^4$ contains a set of set of possible worlds, which together represent our knowledge about the actual state of the world.

It is useful to notice that for a three-valued interpretation $v \in \mathcal{V}^3(\mathsf{At})$, $[v]^4 = \{[v]^2\}$.

Example 5. Consider $v = \mathsf{TUI}$ over $\Sigma = \{abc\}$. Since $[v]^3 = \{\mathsf{TUT}, \mathsf{TUF}\}$, $[\mathsf{TUT}]^2 = \{\mathsf{TTT}, \mathsf{TFT}\}$ and $[\mathsf{TUF}]^2 = \{\mathsf{TTF}, \mathsf{TFF}\}$, we see that $[v]^4 = \{\{\mathsf{TTT}, \mathsf{TFT}\}, \{\mathsf{TTF}, \mathsf{TFF}\}\}$.

Notice that, in order to retain the four-valued structure of an interpretation v in its four-valued completion $[v]^4$, the two-step nature of the construction of $[v]^4$ and the resulting nested structure of $[v]^4$ is essential. Indeed, if $[v]^4$ would merely consist of possible worlds, we would somehow have to choose between letting the members $\omega \in [v]^4$ stand as equally plausible candidates of the actual world or partial descriptions of the information given by v, i. e., we would have to choose between U and I. Conceiving of $[v]^4$ as a set of sets of worlds avoids this choice: sets of worlds $\mathcal{V}' \in [v]^4$ represent partial descriptions of the information given by v, and members of these sets of worlds $\omega \in \mathcal{V}'$ represent equally plausible candidates of the information in $\mathcal{V}'$.

We can now define the assignment of truth values of complex formulas given an interpretation v based on our set of four-valued completions $[v]^4$:

Definition 5. Given a formula ϕ and an interpretation v, then:

$$v(\phi) = \begin{cases} \mathsf{T} & \text{if for every } \Omega' \in [v]^4, \sqcap_i \Omega'(\phi) = \mathsf{T} \\ \mathsf{F} & \text{if for every } \Omega' \in [v]^4, \sqcap_i \Omega'(\phi) = \mathsf{F} \\ \mathsf{I} & \text{if for some } \Omega_1 \in [v]^4, \sqcap_i \Omega_1(\phi) = \mathsf{T} \\ & \text{and for some } \Omega_2 \in [v]^4, \sqcap_i \Omega_2(\phi) = \mathsf{F} \\ \mathsf{U} & \text{otherwise} \end{cases}$$

Thus, a complex formula ϕ is assigned T (respectively F) relative to an interpretation v if every four-valued completion $\Omega' \in [v]^4$ of v, assigns T (respectively F) to ϕ. If there is disagreement among the four-valued completions of v on which determinate truth value ϕ should be assigned, $v(\phi) = \mathsf{I}$. Finally, if some of the four-valued completions of v do not assign any determinate truth value to ϕ, $v(\phi) = \mathsf{U}$.

This way of deriving a truth value for complex formulas on the basis of a four-valued interpretation is, to the best of our knowledge, completely new. It is perfectly in line with $\sqcap_i[v]^2$, the logic underlying ADFs, in the sense that for any three-valued interpretation $\nu \in \mathcal{V}^3(\mathsf{At})$ and any formula $\phi \in \mathcal{L}$, $\nu(\phi) = \sqcap_i[\nu]^2(\phi)$.

Fact 2. For any $\nu \in \mathcal{V}^3(\mathsf{At})$ and any $\phi \in \mathcal{L}(\mathsf{At})$, $\nu(\phi) = \sqcap_i[\nu]^2(\phi)$.

[1] Some proofs have been left out in view of spatial limitations.

Example 6. Consider $v = \mathsf{TUI}$ over $\Sigma = abc$. Observe that $[v]^4 = \{\{\mathsf{TTT}, \mathsf{TFT}\}, \{\mathsf{TTF}, \mathsf{TFF}\}\}$. Thus, we have the following assignments to complex formulas:

- $v(a \wedge c) = \mathsf{I}$, since $\sqcap_i\{\mathsf{TTT}, \mathsf{TFT}\}(a \wedge c) = \mathsf{T}$ and $\sqcap_i\{\mathsf{TTF}, \mathsf{TFF}\}(a \wedge c) = \mathsf{F}$;
- $v(b \wedge c) = \mathsf{U}$, since $\sqcap_i\{\mathsf{TTT}, \mathsf{TFT}\}(b \wedge c) = \mathsf{U}$ and $\sqcap_i\{\mathsf{TTF}, \mathsf{TFF}\}(b \wedge c) = \mathsf{F}$;
- $v(a \wedge \neg a) = \mathsf{F}$, since $\sqcap_i\{\mathsf{TTT}, \mathsf{TFT}\}(a \wedge \neg a) = \mathsf{F}$ and $\sqcap_i\{\mathsf{TTF}, \mathsf{TFF}\}(a \wedge \neg a) = \mathsf{F}$;

We first observe that 4CL preserves classical tautologies and falsities:

Proposition 1. If $\vdash \phi$ then for any $v \in \mathcal{V}^4$, $v(\phi) = \mathsf{T}$. Likewise, if $\vdash \neg\phi$ then for any $v \in \mathcal{V}^4$, $v(\phi) = \mathsf{F}$.

Proof. This is so because for any $v \in \mathcal{V}^4$, $[v]^2(\phi) = \mathsf{T}[\mathsf{F}]$ for any tautology[falsity]. □

We can also define entailment in 4CL in the usual way. We set T and I as designated truth values in compliance with (Belnap 2019):

Definition 6. Given a set of formulas $\Psi \cup \{\phi\} \subseteq \mathcal{L}(\mathsf{At})$, $\mathsf{Mod}^4(\Psi) = \{v \in \mathcal{V}^4(\mathsf{At}) \mid v(\psi) \in \{\mathsf{T}, \mathsf{I}\}$ for every $\psi \in \Psi\}$ and $\Psi \models_{\mathsf{4CL}} \phi$ iff $\mathsf{Mod}^4(\Psi) \subseteq \mathsf{Mod}^4(\phi)$.

We now show that $\models_{\mathsf{4CL}}$ is paraconsistent:

Proposition 2. There exists a set of formulas $\Phi \subseteq \mathcal{L}(\mathsf{At})$ s.t. $\mathsf{Mod}(\Phi) = \emptyset$ yet $\mathsf{Mod}^4(\Phi) \neq \emptyset$.

Proof. Consider the signature $\mathsf{At} = \{p, q\}$, $\Phi = \{p, \neg p\}$ and $v \in \mathcal{V}^4(\mathsf{At})$ with $v(p) = \mathsf{I}$ and $v(q) = \mathsf{U}$. $[v]^2_4 = \{\{\mathsf{TT}, \mathsf{TF}\}, \{\mathsf{FT}, \mathsf{FF}\}\}$ and thus $v(\neg p) = v(p) = \mathsf{I}$ and $v(q) = \mathsf{U}$. □

We notice, though, that there might still be sets of formulas $\Phi \in \mathcal{L}(\mathsf{At})$ for which no $v \in \mathcal{V}^4(\mathsf{At})$ exists s.t. $v(\phi) \in \{\mathsf{T}, \mathsf{I}\}$ for every $\phi \in \Phi$. To see this, it suffices to observe that for any falsity ϕ and any interpretation $v \in \mathcal{V}^4(\mathsf{At})$, $v(\phi) = \mathsf{F}$. In other words, the logic defined above is still explosive for contradictions. But for inconsistent sets of formulas containing no contradictions, the logic is non-explosive.

Proposition 3. For every set of formulas $\Phi \subseteq \mathcal{L}(\mathsf{At})$ s.t. for every $\phi \in \Phi$, $\mathsf{Mod}(\phi) \neq \emptyset$, there is some $v \in \mathcal{V}^4(\mathsf{At})$, s.t. $v(\phi) \in \{\mathsf{I}, \mathsf{T}\}$ for every $\phi \in \Phi$.

Remark 1. Observe that the logic 4CL, like the logic defined by $\sqcap_i[v]^2$, is *not* truth-functional. To see this consider the interpretation v with $v(a) = \mathsf{U}$ and $v(b) = \mathsf{U}$. Then $v(a \vee \neg a) = \mathsf{T}$ yet $v(b \vee \neg a) = \mathsf{U}$. Thus, we see that 4CL is not truth-functional, as $v(a) = v(b) = \mathsf{U}$ yet $v(a \vee \neg a) \neq v(b \vee \neg a)$.

We finally notice the following useful property:

Proposition 4. Let $v_1, v_2 \in \mathcal{V}^4(\mathsf{At})$ and $\phi \in \mathcal{L}(\mathsf{At})$ be given. Then $v_1 \leq_i v_2$ implies $v_1(\phi) \leq_i v_2(\phi)$.

5 Semantics of cADFs

In this section, we define, motivate and study the semantics of cADFs. In more detail, in Section 5.1 we define the central Γ_Π-function and use it to define the main semantics for cADFs. In Section 5.2 we motivate the design choices made in generalizing the Γ-function from ADFs to cADFs. In Section 5.3 we show some central semantical properties of the semantics of cADFs.

5.1 The Γ_Π-Function and Resulting cADF-Semantics

A cADF Π over At is interpreted through 4-valued interpretations. Just like for ADFs, it is of crucial importance to construct a Γ-function that allows to characterize all semantics in terms of (post-)fixpoints of this function.

The Γ-function, conceptually, performs the following operation for ADFs: given an interpretation ν and an ADF D, $\Gamma_D(\nu)$ assigns to every atom s the truth value determined by ν and C_s. In other words, $\Gamma_D(\nu)(s)$ is the value s should take in view of the information expressed by $s \lhd C_s$ and ν. If (for every $s \in S$), this value is compatible (in terms of $\leq_i$) with the actual value $v(s)$, then v will be admissible or even complete. We generalize this idea to the case of cADFs, and take, intuitively, $\Gamma_\Pi(v)$ as the *set of interpretations* that evaluate ϕ in accordance with the information given by $\phi \lhd \psi \in \Pi$ and v. More formally, we define the Γ-function $\Gamma_\Pi : \mathcal{V}^4(\mathsf{At}) \to \wp(\mathcal{V}^4(\mathsf{At}))$ for a cADF Π and an interpretation $v \in \mathcal{V}^4(\mathsf{At})$ as follows:

$$\Gamma_\Pi(v) = \min_{\leq_i}\{v' \in \mathcal{V}^4 \mid \forall \phi \lhd \psi \in \Pi : v'(\phi) \geq_i v(\psi)\}$$

Example 7. Let $\Pi = \{p \vee s \lhd \top; \neg s \lhd p\}$ formulated over the signature $\Sigma = \{p, s\}$. We have the following interpretations and corresponding outcomes of the Γ_Π-function:

v	$\Gamma_\Pi(v)$	v	$\Gamma_\Pi(v)$
UU	{UT, TU}	FU	{UT}
UT	{UT, TU}	FT	{UT}
UF	{UT, TU}	FF	{UT}
UI	{UT, TU}	FI	{UT}
TU	{TF, FI}	IU	{TI, FI}
TT	{TF, FI}	IT	{TI, FI}
TF	{TF, FI}	IF	{TI, FI}
TI	{TF, FI}	II	{TI, FI}

We explain $\Gamma_\Pi(\mathsf{UU})$ as follows: in view of $p \vee s \lhd \top$ and $\mathsf{UU}(\top) = \mathsf{T}$, every interpretation $v' \in \Gamma_\Pi(\mathsf{UU})$ has to assign a truth value at least as informative as T to $p \vee s$, i.e. $v'(p \vee s) \geq_i \mathsf{T}$. Likewise, since $\mathsf{UU}(p) = \mathsf{U}$ and $\neg s \lhd p \in \Pi$, $v' \in \Gamma_\Pi(\mathsf{UU})$ has to set $v'(\neg s) \geq_i \mathsf{U}$, which is trivially the case. The two $\leq_i$-minimal interpretations that satisfy this constraint are: UT and TU.

As a second example, consider FF. Like with UU, every interpretation $v' \in \Gamma_\Pi(\mathsf{FF})$ has to assign $v'(p \vee s) \geq_i \mathsf{T}$. However, since $\mathsf{FF}(p) = \mathsf{F}$ and $\neg s \lhd p \in \Pi$, any $v' \in \Gamma_\Pi(\mathsf{FF})$ has to set $v'(\neg s) \geq_i \mathsf{F}$. UT is the unique $\leq_i$-minimal interpretation satisfying these constraints.

We first notice that Γ_Π is indeed a generalization of the Γ_D-function for ADFs. To show this in a more formally precise manner, we first define the cADF Π_D associated with an ADF D.

Definition 7. Given an ADF $D = (S, L, C)$, we define the *cADF* Π_D *associated with* D as $\Pi_D = \{s \lhd C_s \mid s \in S\}$.

We can now show that for any three-valued interpretation ν, $\Gamma_{\Pi_D}(\nu)$ coincides with $\Gamma_D(\nu)$, i.e. the Γ-function for ADFs coincides with the Γ-function for the associated cADFs for three-valued interpretations.

Proposition 5. For any ADF $D = (S, L, C)$ and any $\nu \in \mathcal{V}^3(S)$, $\Gamma_{\Pi_D}(\nu) = \{\Gamma_D(\nu)\}$.

Proof. Consider an ADF $D = (S, L, C)$ and some $\nu \in \mathcal{V}^3(S)$. $v \in \Gamma_{\Pi_D}$ iff v is among the $\leq_i$-minimal interpretations s.t. $v(s) \geq_i \nu(C_s)$ for every $s \in S$. With Fact 2, $\nu(C_s) = \sqcap_i[\nu]^2(C_s)$ for every $s \in S$. This means that $\Gamma_D(s) = \nu(C_s)$ and thus Γ_D is the unique $\leq_i$-minimal interpretation s.t. $v(s) \geq_i \nu(C_s)$. □

The above result shows that the Γ_Π-function is a direct generalization of the well-studied Γ_D-function known from ADFs. This allows us to define the main semantics of cADFs in terms of (post-)fixpoints of the Γ_Π-functions, just like in the case of ADFs.

With our generalized Γ_Π-function at hand, we can now define the main semantics for cADFs as straightforward generalizations of the ADF-semantics:

Definition 8. Let a cADF Π over At and an interpretation $v \in \mathcal{V}^4(\mathsf{At})$ be given, then:

- v is *admissible* for Π iff there is some $v' \in \Gamma_\Pi(v)$ s.t. $v \leq_i v'$.
- v is *complete* for Π iff $v \in \Gamma_\Pi(v)$.
- v is *preferred* for Π if it is a $\leq_i$-maximal among all admissible interpretation for Π;
- v is *grounded* for Π if it is a $\leq_i$-minimal among all complete interpretation for Π;
- v is a *two-valued model* for Π iff $v \in \mathcal{V}^2(\mathsf{At})$ and v is complete.

Example 8 (Example 7 ctd.)**.** We see that for Π from Example 7, there are two complete interpretations: TF and UT. This can be seen by observing that $\mathsf{TF} \in \Gamma_\Pi(\mathsf{TF})$ and $\mathsf{UT} \in \Gamma_\Pi(\mathsf{UT})$. Since these interpretations are $\leq_i$-incomparable, both interpretations are also grounded. The admissible interpretations are: UU, UT, TU and TF. Thus, UT and TF are also preferred.

Example 9. Let $\Pi = \{b \lhd p, f \lhd b, \neg f \lhd p\}$ formulated over $\Sigma = \{b, f, p\}$ be given. $v_{\mathsf{U}} = \mathsf{UUU}$ is the unique complete interpretation and thus also grounded. It is also the unique admissible interpretation.

Notice that e.g. TIT is *not* complete, since $\Gamma_\Pi(\mathsf{TIT}) = \{\mathsf{TIU}\}$. The reason for $\Gamma_\Pi(\mathsf{TIT})(p) = \mathsf{U}$ is since there is no acceptance pair $p \lhd \phi \in \Pi$. The intuition is that p is only accepted if we have good information to do so, but no such information is given by any $\phi \lhd \psi \in \Pi$.

It is interesting to note that for $\Pi' = \Pi \cup \{p \lhd p\}$, $\mathsf{TIT} \in \Gamma_{\Pi'}(\mathsf{TIT}) = \{\mathsf{TIU}, \mathsf{TIT}, \mathsf{TIF}\}$.

As can be seen in the example above, if an atom a occurs in no statement of ϕ of any acceptance pair $\phi \lhd \psi \in \Pi$, then $v(a) = \mathsf{U}$ for any admissible or complete interpretation v. However, should this be undesired, one can simply add the acceptance pair $a \lhd a$ for such an atom.

5.2 Design Choices in Γ_Π and Comparison with Γ_D

We now discuss the design choices that had to be made when generalizing the Γ-function from ADFs to cADFs. In particular, given the increase in syntactical expressiveness, we had to generalize Γ_Π as to adequately handle this increased expressiveness semantically.

A first generalization is caused by the fact that statements ϕ of acceptance pairs $\phi \lhd \psi$ are possibly non-atomic formulas. Since Γ_Π contains all interpretations v' that align, for any $\phi \lhd \psi \in \Pi$, the truth value of ϕ with $v(\psi)$, there might now be more than one interpretation v' which achieves this. As a case in point, consider the cADF $\Pi = \{p \vee q \lhd \top\}$, where acceptance of $p \vee q$ (which is required by any $v \in \mathcal{V}^4$, since $v(\top) = \mathsf{T}$ for any $v \in \mathcal{V}^4$) can be guaranteed by any interpretation that satisfies p or q. Therefore, the Γ-function might contain multiple interpretations which all do an equally good job of aligning the truth values of statements ϕ with their respective conditions ψ. Thus, Γ_Π is defined as a *non-deterministic operator* (Pelov and Truszczynski 2004; Heyninck and Arieli 2021), in the sense that a single interpretation v might give rise to a non-singleton set of interpretations $\{v_1, \ldots, v_n\} = \Gamma_\Pi(v)$. In the example above, we have e.g. $\Gamma_\Pi(v) = \{\mathsf{TU}, \mathsf{UT}\}$ for any $v \in \mathcal{V}^4(\{p, q\})$.

A second generalization w.r.t. the Γ-function for ADFs is the fact that alignment of statements ϕ with their corresponding condition ψ cannot always be done in an exact way. In more detail, for ADFs D, alignment by Γ_D of s is always exact, in the sense that $\Gamma_D(v)(s)$ *coincides* with the truth value assigned by $\sqcap_i[v]^2(C_s)$. This is not always possible for cADFs, since we might have conflicting specifications in a cADF. Take for example the cADF $\Pi = \{p \lhd \top; \neg p \lhd \top\}$. Clearly, for any $v \in \mathcal{V}^4(\mathsf{At})$, there exists no $v' \in \mathcal{V}^4(\mathsf{At})$ s.t. $v'(\phi) = v(\psi)$ for every $\phi \lhd \psi$. Indeed, this is one of the reasons we had to move to a four-valued logic, since now we can at least specify an interpretation v' which brings $v'(p)$ and $v'(\neg p)$ in alignment with $v(\top)$, in the sense that $v'(p)$ and $v'(\neg p)$ are at least as informative as $v(\top)$, i.e. $v'(p) \geq_i v(\top)$ and $v'(\neg p) \geq_i v(\top)$ (for any $v \in \mathcal{V}^4(\mathsf{At})$).

5.3 Semantical Properties of cADF-semantics

In this section, we show central semantical results on the semantics of cADFs. In particular, we show some relationships between the semantics, and we show under which conditions admissible, complete, grounded and preferred interpretations are guaranteed to exist.

We start by observing that, just like for ADFs, complete interpretations are admissible:

Proposition 6. Let a cADF Π and a complete interpretation v for Π be given. Then v is admissible.

Proof. Suppose v is complete for Π. Then $v \in \Gamma_\Pi(v)$ and thus $v \leq_i v'$ for some $v' \in \Gamma_\Pi(v)$. $\square$

For showing the existence of admissible and preferred interpretations, it will be useful to limit attention to what we will call *well-formed cADFs*. The main idea is that we want to avoid cADFs Π for which $\Gamma_\Pi(v) = \emptyset$ for some $v \in \mathcal{V}^4(\mathsf{At})$, as occurs in e.g. the following example:

Example 10. $\Pi = \{p \lhd \top, \neg p \lhd \top, p \vee \neg p \lhd p\}$.

v	$\Gamma_\Pi(v)$	v	$\Gamma_\Pi(v)$
T	$\{\mathsf{I}\}$	F	$\{\mathsf{I}\}$
U	$\{\mathsf{I}\}$	I	$\emptyset$

Notice that $\Gamma(\mathsf{I}) = \emptyset$.

Definition 9. A *well-formed cADF* is a cADF Π s.t. $\Gamma_\Pi(v) \neq \emptyset$ for any $v \in \mathcal{V}^4(\mathsf{At})$.

We observe that a syntactic sufficient condition for well-formedness of a cADF Π is to simply require that for every acceptance pair $\phi \lhd \psi \in \Pi$, the statement ϕ is a logically contingent formula. We call such cADFs *unconstrained*:

Definition 10. A cADF Π is *unconstrained* iff for every $\phi \lhd \psi \in \Pi$, ϕ is logically contingent.

We explain the term of *unconstrained* cADF as follows. Notice that an acceptance pair $\phi \lhd \psi$, where ϕ is a tautology or a falsity, can be seen as a constraint, in the sense that it forces ψ to be set to the value of ϕ (i.e. $v(\psi) = \mathsf{T}$ if ϕ is a tautology and $v(\psi) = \mathsf{F}$ if ψ is a falsity) for any complete extension. To see this, observe that $v(\phi) = \mathsf{T}[\mathsf{F}]$ for any $v \in \mathcal{V}^4$ if ϕ is a tautology[falsity]. In particular, for any $v' \in \Gamma_\Pi(v)$, it will hold that $v(\phi) = \mathsf{T}[\mathsf{F}]$. It is quite interesting that the framework naturally allows for the formulation of constraints, but for the development of the meta-theory, it will prove useful to restrict attention to well-formed cADFs. It is an interesting question for future work to see whether *constrained argumentation frameworks* (Coste-Marquis, Devred, and Marquis 2006) can be captured using such constraints.

Proposition 7. Any unconstrained cADF Π is well-formed.

Proof sketch. Suppose that Π is an unconstrained cADF. It can be shown that for every $\phi \lhd \psi \in \Pi$, $v_\mathsf{I}(\phi) = \mathsf{I}$. Thus, for every $v' \in \mathcal{V}^4(\mathsf{At})$ there is some $v \in \mathcal{V}^4(\mathsf{At})$ (namely $v = v_\mathsf{I}$) s.t. $v(\phi) \geq_i v'(\psi)$ for every $\phi \lhd \psi \in \Pi$. Since $\leq_i$ is well-founded and Π is finite, $\Gamma_\Pi(v') \neq \emptyset$ for any $v' \in \mathcal{V}^4(\mathsf{At})$. $\square$

However, there are well-formed cADFs that are not unconstrained:

Example 11. Consider $\Pi = \{a \vee \neg a \lhd a \vee \neg a\}$. Then clearly, for any $v \in \mathcal{V}^4(\mathsf{At})$, $\Gamma_\Pi(v) = \{\mathsf{T}\}$ (since $\mathsf{U}(a \vee \neg a) = \mathsf{T}$ with Lemma 1).

We now show the first existence result, which states that any well-formed cADF admits admissible interpretations:

Proposition 8. For any well-formed cADF, there exists an admissible interpretation.

Proof. For any well-formed cADF Π, $\Gamma_\Pi(v_\mathsf{U}) \neq \emptyset$. Since $v_\mathsf{U} \leq_i v$ for any $v \in \mathcal{V}^4(\mathsf{At})$, v_U is admissible. $\square$

We immediately obtain an existence result for preferred interpretations:

Corollary 1. For any well-formed cADF, there exists a preferred interpretation.

We now show an existence result for the complete and grounded interpretations. This is done by first showing that Γ_Π satisfies monotonicity under the *Smyth-order* (Smyth 1976). The *Smyth-order* $\preceq^S_i \subseteq \wp(\mathcal{V}^4) \times \wp(\mathcal{V}^4)$ is defined as follows: $\mathcal{V}_1 \preceq^S_i \mathcal{V}_2$ iff for every $v_2 \in \mathcal{V}_2$ there is some $v_1 \in \mathcal{V}_1$ s.t. $v_1 \leq_i v_2$.

Remark 2. Notice that $\preceq^S_i$ is a transitive and reflexive relation over $\wp(\mathcal{V}^4(\mathsf{At}))$. Furthermore, $\preceq^S_i$ is a partial order over the set of $\leq_i$-minimal subsets $\mathcal{V}^4$ (i.e. $\preceq^S_i$ is transitive, reflexive and anti-symmetric over $\wp_{\leq_i}(\mathcal{V}^4(\mathsf{At})) = \{\mathcal{V}' \subseteq \mathcal{V}^4(\mathsf{At}) \mid \mathcal{V}' = \min_{\leq_i}(\mathcal{V}')\}$).

Proposition 9. For any well-formed cADF Π, Γ_Π is $\preceq^S_i$-monotonic.

Proof. First observe that for any $v_1 \leq_i v_2$ and any $\phi \lhd \psi \in \Pi$, $v_1(\psi) \leq_i v_2(\psi)$. Suppose now that $v' \in \mathcal{V}^4$ s.t. $v'(\phi) \geq_i v_2(\psi)$ for every $\phi \lhd \psi \in \Pi$. Then $v'(\phi) \geq_i v_1(\psi)$ for every $\phi \lhd \psi \in \Pi$. Thus, there is some $v \in \Gamma_\Pi(v_1)$ s.t. $v \leq_i v'$. In particular, this means that for every $v' \in \Gamma_\Pi(v_2)$, there is some $v \in \Gamma_\Pi(v_1)$ s.t. $v \leq_i v'$. $\square$

Proposition 10. For any well-formed cADF Π, there exists a complete interpretation.

Proof. Notice that since $v_\mathsf{I} \geq_i v$ for every $v \in \mathcal{V}^4(\mathsf{At})$, $v_\mathsf{I} \geq_i v_1$ for any $v_1 \in \Gamma_\Pi(v_\mathsf{I})$ (notice that since Π is well-formed, $\Gamma_\Pi(v_\mathsf{I}) \neq \emptyset$). Since Γ_Π is $\preceq^S_i$-monotonic with Proposition 9, $\Gamma_\Pi(v_1) \preceq^S_i \Gamma_\Pi(v_\mathsf{I})$ for any $v_1 \in \Gamma_\Pi(v_\mathsf{I})$. Thus, for any $v_1 \in \Gamma_\Pi(v_\mathsf{I})$, there is some $v_2 \in \Gamma_\Pi(v_1)$ s.t. $v_2 \leq_i v_1$. We can use the above line of argument to construct a chain of interpretations $\ldots \leq_i v_n \leq_i v_{n-1} \leq_i \ldots v_2 \leq_i v_1 \leq_i v_0 = v_\mathsf{I}$ s.t. for every $1 \leq i < n$, $v_i \in \Gamma_\Pi(v_{i-1})$ and $\Gamma_\Pi(v_i) \preceq^S_i \Gamma_\Pi(v_{i-1})$. Since $\mathcal{V}^4(\mathsf{At})$ is finite, this chain ends, i.e. there some $i \in \mathbb{N}$ s.t. $v_i = v_{i+1}$. Since $v_{i+1} \in \Gamma_\Pi(v_i) = \Gamma_\Pi(v_{i+1})$, v_i is a complete interpretation (notice that $\Gamma_\Pi(v_i) = \Gamma_\Pi(v_{i+1})$ follows from the anti-symmetry of $\preceq^S_i$ over $\wp_{\leq_i}(\mathcal{V}^4(\mathsf{At}))$ (Remark 2) and the fact that $\Gamma_\Pi(v) \in \wp_{\leq_i}(\mathcal{V}^4(\mathsf{At}))$ for any $v \in \mathcal{V}^4(\mathsf{At})$). $\square$

We immediately obtain an existence result for the grounded interpretation as well:

Corollary 2. For every well-formed cADF Π, there exists a grounded interpretation.

Another useful order on $\wp(\mathcal{V}^4) \times \wp(\mathcal{V}^4)$ is the *Hoare-order* $\preceq^H_i$ defined as: $\mathcal{V}_1 \preceq^S_i \mathcal{V}_2$ iff for every $v_1 \in \mathcal{V}_1$ there is some $v_2 \in \mathcal{V}_2$ s.t. $v_1 \leq_i v_2$.

Proposition 11. For every well-formed cADF Π s.t. Γ_Π is $\preceq^H_i$-monotonic, if v is preferred then it is complete.

Property	Condition on Π	Result
$\exists$ of admissible int.	well-formed	Prop. 8
$\exists$ of preferred int.	well-formed	Cor. 1
$\exists$ of complete int.	well-formed	Prop. 10
$\exists$ of grounded int.	well-formed	Cor. 2
preferred $\subseteq$ complete	well-formed & $\preceq_i^H$-monotonic	Prop. 11

Table 1: Summary of results from Section 5.3

Proof. Let a well-formed cADF Π s.t. Γ_Π is $\preceq_i^H$-monotonic be given and consider a preferred interpretation $v \in \mathcal{V}^4(\mathsf{At})$. Suppose towards a contradiction that $v \notin \Gamma_\Pi(v)$. Since v is preferred, it is admissible and thus there is some $v' \in \Gamma_\Pi(v)$ s.t. $v \leq_i v'$. Since $v \notin \Gamma_\Pi(v)$, $v <_i v'$. With $\preceq_i^H$-monotonicity of Γ_Π, we obtain that $\Gamma(v) \preceq_i^H \Gamma(v')$ and thus there is some $v'' \in \Gamma(v')$ s.t. $v' \leq_i v''$. But then v' is admissible, contradicting v being preferred. $\square$

We observe, however, that not every cADF has a $\preceq_i^H$-monotonic Γ_Π function:

Example 12. Let $\Pi = \{p \vee (q \wedge s) \lhd s, p \wedge q \lhd s\}$ over the signature $\{p, q, s\}$. Then $\Gamma_\Pi(\mathsf{UUT}) = \{\mathsf{UTT}, \mathsf{TUU}\}$ and $\Gamma_\Pi(\mathsf{TUT}) = \{\mathsf{TTU}\}$. Since $\mathsf{UUT} \leq_i \mathsf{TUT}$, yet there is no $v \in \Gamma_\Pi(\mathsf{TUT})$ s.t. $\mathsf{UTT} \leq_i v$, we see that Γ_Π is not $\preceq_i^H$-monotonic.

We summarize our results in Table 1.

6 Grounded Interpretations and the Grounded State

One of the crucial properties of ADFs is that a unique grounded interpretation is guaranteed to exist. This property does not generalize to the grounded semantics of cADFs, in view of the indeterminism that cADFs allow to express. As a case in point consider $\Pi = \{p \vee q \lhd \top\}$, which has two $\leq_i$-minimal complete interpretations: v_1 and v_2 with:

$$v_1(p) = \mathsf{T} \quad v_1(q) = \mathsf{U} \quad \text{and} \quad v_2(p) = \mathsf{U} \quad v_2(q) = \mathsf{T}$$

Thus, there might be cADFs that do not have a unique grounded interpretation. This might be seen as problematic, since the grounded interpretation for ADFs can be calculated efficiently and straightforwardly by iterating Γ_D starting from v_{U}. Since the grounded interpretation v_g is $\leq_i$-minimally complete and unique for ADFs, it approximates any other complete interpretation of the ADF in question (in the sense that $v_g \leq_i v$ for any complete interpretation v). We are now interested in defining a similar concept for cADFs, that is, a unique representation of the $\leq_i$-minimal information expressed by a cADF that can be unambigously obtained by application of Γ_Π and approximates any complete interpretation. This can be done by looking at a set of interpretations instead of a single interpretation. We note that this idea is not new. For example, many well-founded semantics for disjunctive logic programming take up this idea, resulting in a well-founded state (Baral, Lobo, and Minker 1992; Alcântara, Damásio, and Pereira 2005).[2] Accordingly, we will be interested in a *grounded state* $\mathcal{V}' \subseteq \mathcal{V}^4(\mathsf{At})$ that represents the minimal knowledge entailed by a cADF. This grounded state can be defined as the $\preceq_i^S$-minimal fixpoint of Γ'_Π, a generalization of Γ_Π to sets of interpretations. Γ'_Π is obtained as follows:

Definition 11. Given a cADF Π and $\mathcal{V}' \subseteq \mathcal{V}^4(\mathsf{At})$:

$$\Gamma'_\Pi(\mathcal{V}') = \min_{\leq_i} \bigcup_{v \in \mathcal{V}'} \Gamma_\Pi(v)$$

The following fact gives an equivalent characterization of Γ'_Π, which avoids the superfluous $\leq_i$-minimization in $\Gamma_\Pi(v)$:

Fact 3. Given a cADF Π and $\mathcal{V}' \subseteq \mathcal{V}^4(\mathsf{At})$, $\Gamma'_\Pi(\mathcal{V}') =$

$$\min_{\leq_i}\{v' \in \mathcal{V}^4 \mid \exists v \in \mathcal{V}' : \forall \phi \lhd \psi \in \Pi : v'(\phi) \leq v(\psi)\}$$

Proof. Let some cADF Π and $\mathcal{V}' \subseteq \mathcal{V}^4(\mathsf{At})$ be given. Then $\Gamma'_\Pi(\mathcal{V}') = \min_{\leq_i} \bigcup_{v \in \mathcal{V}'} \Gamma_\Pi(v)$ by definition. By definition of Γ_Π, this means that $\Gamma'_\Pi(\mathcal{V}') = \min_{\leq_i} \bigcup_{v \in \mathcal{V}'} \min_{\leq_i}\{v' \in \mathcal{V}^4 \mid v'(\phi) \geq_i v(\psi)$ for every $\phi \lhd \psi \in \Pi\}$. But then $\Gamma'_\Pi(\mathcal{V}') = \min_{\leq_i}\{v' \in \mathcal{V}^4 \mid \exists v \in \mathcal{V}' : \forall \phi \lhd \psi \in \Pi : v'(\phi) \leq v(\psi)\}$. $\square$

Definition 12. Let a cADF Π be given. Then we say $\mathcal{V}' \subseteq \mathcal{V}^4(\mathsf{At})$ is:

- a *complete state* (for Π) iff $\mathcal{V}' = \Gamma'_\Pi(\mathcal{V}')$.
- a *grounded state* (for Π) iff $\mathcal{V}'$ is a $\preceq_i^S$-minimally complete state (for Π).

Proposition 12. Let a cADF Π be given. Then there exists a unique grounded state which can be obtained by iterating Γ'_Π, starting with v_{U}.

Proof. We now show that Γ'_Π is a $\preceq_i^S$-monotonic operator over $\wp_{\leq_i}(\mathcal{V}^4(\mathsf{At}))$. For this, define $G_\Pi(v) = \{v' \in \mathcal{V}^4(\mathsf{At}) \mid v'(\phi) \geq_i v(\psi)$ for every $\psi \lhd \phi\}$. We first show that $\mathcal{V}_1 \preceq_i^S \mathcal{V}_2$ implies $\bigcup_{v \in \mathcal{V}_1} G_\Pi(v) \preceq_i^S \bigcup_{v \in \mathcal{V}_2} G_\Pi(v)$. Indeed, consider some $v_2 \in \bigcup_{v \in \mathcal{V}_2} G_\Pi(v)$. This means that $v_2(\psi) \geq_i v(\phi)$ for some $v \in \mathcal{V}_2$ and every $\phi \lhd \psi \in \Pi$. Since $\mathcal{V}_1 \preceq_i^S \mathcal{V}_2$, there is some $v' \in \mathcal{V}_1$ s.t. $v' \leq_i v$. Thus, $v'(\psi) \leq_i v(\psi)$ for every $\phi \lhd \psi \in \Pi$ (Proposition 4). Thus, $v_2(\psi) \geq_i v'(\phi)$ for every $\phi \lhd \psi \in \Pi$ and $v_2 \in \bigcup_{v \in \mathcal{V}_2} G_\Pi(v)$. Since $\Gamma'_\Pi(\mathcal{V}_2) \subseteq \bigcup_{v \in \mathcal{V}_2} G_\Pi(v)$, we derive that $\bigcup_{v \in \mathcal{V}_1} G_\Pi(v) \preceq_i^S \Gamma'_\Pi(\mathcal{V}_2)$. In other words, for every $v_2 \in \Gamma'_\Pi(\mathcal{V}_2)$ there is some $v_1 \in \bigcup_{v \in \mathcal{V}_1} G_\Pi(v)$ s.t. $v_1 \leq_i v_2$. Since $\Gamma'_\Pi(\mathcal{V}_1) = \min_{\leq_i} \bigcup_{v \in \mathcal{V}_1} G_\Pi(v)$, it follows that $\Gamma'_\Pi(\mathcal{V}_1) \preceq_i^S \Gamma'_\Pi(\mathcal{V}_2)$.

We now show Γ'_Π admits a $\preceq_i^S$-minimal fixpoint. This fixpoint is constructed by applying Γ'_Π iteratively, starting with v_{U} (recall $v_{\mathsf{U}}(a) = \mathsf{U}$ for every $a \in \mathsf{At}$). Since $v_{\mathsf{U}} \leq_i v$ for any $v \in \mathcal{V}^4(\mathsf{At})$, $v_{\mathsf{U}} \preceq_i^S \Gamma_\Pi(v_{\mathsf{U}})$. By the

[2]Some semantics explicitly use the idea of a set of interpretations (Alcântara, Damásio, and Pereira 2005), whereas other semantics are phrased syntactically, resulting in a set of disjunctions (Baral, Lobo, and Minker 1992), which is clearly equivalent to a set of interpretations (see also (Seipel, Minker, and Ruiz 1997).

$\preceq_i^S$-monotonicity of Γ_Π, $\Gamma_\Pi^\alpha(v_{\mathsf{U}}) \preceq_i^S \Gamma_\Pi^\beta(v_{\mathsf{U}})$ for any ordinals α, β with $\alpha \leq \beta$. Since $\mathsf{At}(\Pi)$ is finite, this chain reaches an endpoint, i.e. for some ordinal γ, $\Gamma_\Pi^\gamma(v_{\mathsf{U}}) = \Gamma^{\gamma+1}\Pi(v_{\mathsf{U}})$. Thus, we have shown that $\Gamma_\Pi(v_{\mathsf{U}})$ admits a fixpoint. To show that this fixpoint is the $\preceq_i^S$-minimal fixpoint, consider some $\mathcal{V}' \subseteq \mathcal{V}^4(\mathsf{At})$ s.t. $\Gamma_\Pi(\mathcal{V}') = \mathcal{V}'$. Since $\mathcal{V}' = \Gamma_\Pi(\mathcal{V}') = \min_{\leq_i}(\bigcup_{v \in \mathcal{V}'} G_\Pi(v))$, $\mathcal{V}' \in \wp_{\leq_i}(\mathcal{V}(\mathsf{At}))$. Notice that $v_{\mathsf{U}} \preceq_i^S \mathcal{V}'$. Since Γ_Π is $\preceq_i^S$-monotonic, for any ordinal α, $\Gamma_\Pi^\alpha(v_{\mathsf{U}}) \preceq_i^S \Gamma^\alpha(\mathcal{V}')$. Since $\mathcal{V}^4$ is a fixpoint of Γ_Π, this means $\Gamma_\Pi^\alpha(v_{\mathsf{U}}) \preceq_i^S \mathcal{V}'$ for any ordinal α. In particular this holds for the ordinal γ for which $\Gamma_\Pi^\gamma(v_{\mathsf{U}}) = \Gamma^{\gamma+1}\Pi(v_{\mathsf{U}})$. With the anti-symmetry of $\preceq_i^S$, $\mathcal{V}' = \Gamma_\Pi^\gamma(v_{\mathsf{U}})$ or $\Gamma_\Pi^\gamma(v_{\mathsf{U}}) \prec_i^S \mathcal{V}'$. □

The grounded state is a generalization of the grounded interpretation for ADFs:

Proposition 13. For any ADF D, the grounded state coincides with $\{v\}$, where v is the grounded model of D.

Furthermore, the grounded state approximates any complete interpretation:

Proposition 14. For any cADF Π, where $\mathcal{V}'$ is the grounded state for Π and v is a complete interpretation of Π, we have that: $\mathcal{V}' \preceq_i^S \{v\}$.

We illustrate the construction of the grounded state with an example:

Example 13. Let $\Pi = \{p \vee q \lhd \top, s \lhd p, s \lhd q\}$ over the signature $\{p, q, s\}$. Then we can obtain the grounded state for Π by the following calculation:

- The first iteration is obtained as follows:
$$\Gamma'_\Pi(\{v_{\mathsf{U}}\}) = \{\mathsf{TUU}, \mathsf{UTU}\}.$$
- As a second step we calculate
$$\begin{aligned}\Gamma'_\Pi(\Gamma'_\Pi(v_{\mathsf{U}})) &= \min_{\leq_i}(\Gamma_\Pi(\mathsf{TUU}) \cup \Gamma_\Pi(\mathsf{UTU})) \\ &= \min_{\leq_i}(\{\mathsf{TUT}, \mathsf{UTT}\}) \\ &= \{\mathsf{TUT}, \mathsf{UTT}\}.\end{aligned}$$
- As a third step, we calculate
$$\begin{aligned}\Gamma'_\Pi(\Gamma'_\Pi(\Gamma'_\Pi(v_{\mathsf{U}}))) &= \min_{\leq_i}(\Gamma_\Pi(\mathsf{TUT}) \cup \Gamma_\Pi(\mathsf{UTT})) \\ &= \{\mathsf{TUT}, \mathsf{UTT}\}\end{aligned}$$

Since in the third step a fixed point was reached, we see that the grounded state of Π is $\{\mathsf{TUT}, \mathsf{UTT}\}$. We see that the grounded state consists of two interpretations, which both make s true, and either make p or q true.

Remark 3. All semantics defined in this paper have been implemented in Java using the `Tweety`-library. The implementation can be found online.

7 Related Work

To the best of our knowledge, no generalizations of ADFs as we have suggested here have been proposed before. However, *epistemic graphs* (Hunter, Polberg, and Thimm 2020) can be regarded as an orthogonal approach to extend the expressivity of ADFs. There, general propositional formulas are interpreted through a probabilistic semantics (that is not related to ADF semantics), thus yielding an expressive probabilistic and argumentative formalism. Instead, we have a purely qualitative formalism that generalises the original ADF semantics directly.

ADFs have been generalized in other works, in particular as to allow for the handling of weights (Brewka et al. 2018; Bogaerts 2019). As our semantics, they allow for an extension of the set of truth-values $\{\mathsf{T}, \mathsf{F}, \mathsf{U}\}$ with other values. In fact, in (Brewka et al. 2018) an instantiation of weighted ADFs using Belnap's four-valued logic is discussed. However, in (Brewka et al. 2018) this results in five truth-values, since in weighted ADFs, the truth-values are always supplemented with an information-theoretic minimum U that is not part of the original set of truth-values. Furthermore, this instantiation uses Belnap's four-valued logic to evaluate complex formulas, which means that tautologies can be both assigned Belnap's inconsistent and incomplete truth-values (but never the external U-value). Finally, weighted ADFs have the same syntax as ADFs, and thus, the syntax of cADFs also generalize the syntax of weighted ADFs.

As a side effect of the semantics of cADFs, we obtain also a four-valued semantics of ADFs and argumentation frameworks. Four-valued semantics for abstract argumentation frameworks have been suggested in (Baroni, Giacomin, and Liao 2015) and studied in (Arieli 2012). In (Arieli 2012) argumentation labellings that map arguments to four truth values, in, out, none and both, are defined. Adjusting notation to our setting by letting T stand for in, F for out, U for none and I for both, we see that such argumentation labellings are nothing less and nothing more than four-valued interpretations over the set of arguments. However, using the translation of argumentation frameworks in ADFs from (Brewka et al. 2013), we do not get an equivalence between p-admissible labellings and admissible interpretations of the translated argumentation frameworks.

Example 14. Consider the argumentation framework $\mathsf{AF} = (\{A, B, C\}, (A, B), (C, B), (C, C), (B, B)\}$. Then the corresponding cADF is given by $\Pi = \{A \lhd \neg B \wedge \neg C; C \lhd \neg C; B \lhd \neg B\}$. It can be checked that FIU is p-admissible[3] for AF, but FIU is not admissible for Π, as $\Gamma_\Pi(\mathsf{FIU}) = \{\mathsf{UUI}\}$.

It remains a question for future work whether the translation of argumentation frameworks in cADFs can be adjusted to avoid this discrepancy.

The logic 4CL we designed as a generalization of the logic $\sqcap_i[v]^2$ underlying ADFs has not been suggested in the literature on many-valued logics, to the best of our knowledge. The semantics of 4CL bears some similarities to that of *generalized possibilistic logic* (Dubois 2012), where a pair of sets of possible worlds is used to represent the information given by a four-valued interpretation. However, the crucial difference is that $[v]^4$ might consist of more than two sets of possible worlds, and thus the logics behave quite differently. For example, in generalized possibilistic logic, there exists

[3] We refer to (Arieli 2012) for definitions of p-admissible labellings.

no model that assigns to p, q and $\neg p \vee \neg q$ a designated truth value, whereas, in 4CL, $v_{\mathsf{I}}(p) = v_{\mathsf{I}}(q) = v_{\mathsf{I}}(\neg p \vee \neg q) = \mathsf{I}$.

8 Conclusion

In this paper, we have defined and studied cADFs, which generalize ADFs and allow for indeterminism, over- and underspecfications. Semantics for cADFs are defined in terms of a Γ-function mapping four-valued interpretations to sets of four-valued interpretations. There remains still a lot of work to be done on cADFs. As a first next step, there are still some semantics that need to be generalized form ADFs to cADFs, in particular the stable semantics. Thereafter, we plan to study the computational complexity and realizability (in the style of (Pührer 2020)) of cADFs. On the basis of these steps, we will then have a clear view of which formalisms can be captured by cADFs. Among the most interesting candidates for such representational results, we have our eyes on disjunctive and propositional logic programming (Minker and Seipel 2002; Ferraris 2005) and logics for nonmonotonic conditionals (Kraus, Lehmann, and Magidor 1990).

Acknowledgements

The research reported here was supported by the Deutsche Forschungsgemeinschaft under grant KE 1413/11-1. We thank the reviewers for helpful comments on a previous version of this paper.

References

Alcântara, J.; Damásio, C. V.; and Pereira, L. M. 2005. A well-founded semantics with disjunction. In *International Conference on Logic Programming*, 341–355. Springer.

Arieli, O. 2012. Conflict-tolerant semantics for argumentation frameworks. In *European Workshop on Logics in Artificial Intelligence*, 28–40. Springer.

Baral, C.; Lobo, J.; and Minker, J. 1992. Generalized disjunctive well-founded semantics for logic programs. *Annals of Mathematics and Artificial Intelligence* 5(2):89–131.

Baroni, P.; Giacomin, M.; and Liao, B. 2015. I don't care, i don't know… i know too much! on incompleteness and undecidedness in abstract argumentation. In *Advances in Knowledge Representation, Logic Programming, and Abstract Argumentation*. Springer. 265–280.

Belnap, N. D. 2019. How a computer should think. In *New Essays on Belnap-Dunn Logic*. Springer. 35–53.

Bogaerts, B. 2019. Weighted abstract dialectical frameworks through the lens of approximation fixpoint theory. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 2686–2693.

Brass, S., and Dix, J. 1995. Disjunctive semantics based upon partial and bottom-up evaluation. In *ICLP*, 199–213.

Brewka, G.; Strass, H.; Ellmauthaler, S.; Wallner, J. P.; and Woltran, S. 2013. Abstract dialectical frameworks revisited. In *Twenty-Third International Joint Conference on Artificial Intelligence*.

Brewka, G.; Strass, H.; Wallner, J. P.; and Woltran, S. 2018. Weighted abstract dialectical frameworks. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Cayrol, C., and Lagasquie-Schiex, M.-C. 2005. On the acceptability of arguments in bipolar argumentation frameworks. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, 378–389. Springer.

Coste-Marquis, S.; Devred, C.; and Marquis, P. 2006. Constrained argumentation frameworks. *International Conference on Principles of Knowledge Representation and Reasoning* 6:112–122.

Dubois, D. 2012. Reasoning about ignorance and contradiction: many-valued logics versus epistemic logic. *Soft Computing* 16(11):1817–1831.

Dung, P. M. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *AI* 77:321–358.

Ferraris, P. 2005. Answer sets for propositional theories. In *International Conference on Logic Programming and Nonmonotonic Reasoning*, 119–131. Springer.

Fitting, M. 2006. Bilattices are nice things. *Conference on Self-reference* 53–77.

Heyninck, J., and Arieli, O. 2021. Approximation fixpoint theory for non-deterministic operators and its application in disjunctive logic programming. In *International Conference on Principles of Knowledge Representation and Reasoning*.

Heyninck, J.; Kern-Isberner, G.; Skiba, K.; and Thimm, M. 2019. Interpreting conditionals in argumentative environments. In *NMR 2020 Workshop Notes*, 73.

Hunter, A.; Polberg, S.; and Thimm, M. 2020. Epistemic graphs for representing and reasoning with positive and negative influences of arguments. *Artificial Intelligence* 281:103236.

Kraus, S.; Lehmann, D.; and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44(1-2):167–207.

Minker, J., and Seipel, D. 2002. Disjunctive logic programming: A survey and assessment. In *Computational logic: logic programming and beyond*. Springer. 472–511.

Pelov, N., and Truszczynski, M. 2004. Semantics of disjunctive programs with monotone aggregates: an operator-based approach. In *Proceedings of NMR'04*, 327–334.

Polberg, S. 2016. Understanding the abstract dialectical framework. In *European Conference on Logics in Artificial Intelligence*, 430–446. Springer.

Pührer, J. 2020. Realizability of three-valued semantics for abstract dialectical frameworks. *Artificial Intelligence* 278:103198.

Seipel, D.; Minker, J.; and Ruiz, C. 1997. Model generation and state generation for disjunctive logic programs. *The Journal of Logic Programming* 32(1):49–69.

Smyth, M. B. 1976. Powerdomains. In *International Symposium on Mathematical Foundations of Computer Science*, 537–543. Springer.