A Dialectical Characterisation of Argument Game Proof Theories for Classical Logic Argumentation

Federico Castagna[1]

[1]*King's College London, Bush House, 30 Aldwych, London, WC2B 4BG, United Kingdom*

Abstract

Argument game-based proof theories provide procedural structures capable of determining the status of an argument. Given an argumentation framework, argument games identify the membership of an argument to a specific extension simulating a dialogue between two opposing contenders. The semantics meant to be captured dictates the rules of the played game, which serve to describe how the players can achieve victory. Dialectical Classical logic Argumentation (Dialectical Cl-Arg) is a recent approach that provides real-world dialectical characterisations of Cl-Arg arguments by resource-bounded agents while preserving the rational criteria established by the rationality postulates. This paper combines both subjects and introduces argument games for Dialectical Cl-Arg, highlighting the properties and strengths enjoyed by these games in comparison with the standard ones. The result will be a proof theory capable of approximating real-world non-monotonic single-agent reasoning processes.

Keywords

Argumentation, Argument game, Classical logic, Dialectic, Bounded reasoning

1. Introduction

Human reasoning evolved to produce and evaluate arguments [1]. Trying to consolidate possessed information by formulating reasons (arguments) that challenge or defend it is an ordinary procedure in which humans engage. This process is not only common but even necessary: how could it be possible, otherwise, to decide what to believe or trust without being misled by a non-reliable source of information? This 'scaffolding' (as defined in [2]) role of dialogues and arguments can be seen in social and lone thinking practices where the reasoner(s) evaluates the possessed information by constructing counter-arguments that assesses its acceptability. That is to say, every reasoning process entails a dialogue, and every dialogue entails arguments irrespective of the type of interacting agents: be they humans and AIs, among themselves and with humans. Thanks to its important role, argumentation has been developed as a theory able to characterize the essence of non-monotonic reasoning through the dialectical interplay of arguments [3]. Intuitively, in order to determine if a piece of information is acceptable, it will suffice to prove that the argument (in which the specific information is embedded) is justified under one of Dung's semantics. A way of doing this is to show the membership of the argument to a winning strategy of an argument game (as described, for example, in [4], [5] and [6]).

Although a plethora of works has successfully shown various instantiations of Dung's abstract argumentation framework (AF) and reached different goals, none of these approaches managed to provide a full rational account of real-world resource-bounded agents. Undoubtedly, the introduction of the rationality postulates [7][8] has allowed eschewing the arising of counter-intuitive results in AFs instantiations. However, the rationality postulates require a consumption of resources that far exceed the availability of real-world agents.

1.1. Contribution

The main contribution of this research is the development of argument games for Dialectical Classical Logic Argumentation (Dialectical Cl-Arg [9]), a recent approach that provides real-world dialectical characterisations of AFs by resource-bounded agents. This approach satisfies the rationality postulates (under minimal requirements) and revolves around the core notion of *dialectical defeats*. Such defeats enable argumentative interactions more aligned with the dialectical reasoning of real-world resource-bounded agents. Thus, their presence requires the implementation of *dialectical argument game* proof theories capable of conveying the same idea as single-agent reasoning processes.

2. Background

Argumentation has been developed as a theory of reasoning able to characterize the essence of non-monotonic logic via the dialectical interplay of arguments. According to Dung's seminal paper [3], an Argumentation Framework (AF) is composed of a set of arguments 'AR' and a binary *'attacks'* relation, which denotes the conflicts existing among the arguments in AR, i.e., AF = ⟨AR, *attacks*⟩. Various semantics *s* have also been presented and each of them specifies the status of (*sceptically* or *credulously*) justified (i.e., acceptable) arguments. Several works stemmed from [3], some of which introduced different ways of structuring arguments and instantiating Dung's abstract AF [10][11][12]. For example, Classical Logic Argumentation (Cl-arg) [13][14] is one of such instantiation that builds AFs using classical logic as its underlying language.

2.1. Dialectical Cl-Arg

Unlike standard formalisation of Cl-Arg, real-world agents behave pragmatically and do not need to: (i) always construct every argument defined by a base of wff and (ii) enforce consistency and subset minimality checks on their arguments (nor they have enough computational power to do these checks, given their limited resources). Dialectical Cl-Arg provides a formalisation of real-world modes of dialectical reasoning from resource-bounded agents whilst satisfying the rationality postulates [7][8].

Definition 1. [Dialectical Arguments] $X = (\Delta, \Gamma, \alpha)$ is a dialectical argument *defined by a base $\mathscr{B}$ of classical wff, if* $(\Delta \cup \Gamma) \subseteq \mathscr{B}$, $\Delta \cap \Gamma = \emptyset$, *and* $\Delta \cup \Gamma \vdash_c \alpha$ [1]. *If $\alpha = \wedge$, then X is said to be a*

---

[1]The symbol '$\vdash_c$' denotes the consequence relation of classical logic.

falsum *argument. If* Γ = ∅, *then X is said to be* unconditional; *else X is* conditional. *Finally, if* Δ = ∅, *then X is said to be* unassailable.

Δ, Γ *and* α *are respectively referred to as the* premises (*Prem*(X)), suppositions (*Supp*(X)) *and* conclusion (*Con*(X)) *of* X = (Δ, Γ, α).

Attacks and defeats for Dialectical Cl-Arg work differently than their respective counterparts for Cl-Arg. The reason is the presence of suppositions embedded in the internal structure of the arguments. This epistemic distinction between information considered true (i.e., the premises of an argument) and information supposed true for the sake of the argument (i.e., the supposition of the argument) proves useful also in solving the so-called 'foreign commitment problem' [2].

Definition 2. [Attacks and Defeats][*9*] *Let AR be a set of dialectical arguments defined by a base* $\mathscr{B}$. *The attack relation 'attacks'* ⊆ AR × AR *is defined as follows. For any* X = (Δ, Γ, α), Y = (Π, Σ, β) ∈ AR: attacks(X, Y) *iff:*

- *if* α ≠ ∧ *then* $\bar{α}$ ∈ Π (X *attacks Y on* $\bar{α}$, *equivalently on* Y' = ({$\bar{α}$}, ∅, $\bar{α}$));
- *if* α = ∧ (X *attacks Y on any* φ ∈ Γ ∩ Π, *equivalently on any* Y' = ({φ}, ∅, φ)).

Let ≺ *be a strict partial ordering over AR. Then, for every X, Y such that attacks*(X, Y), *defeats*(X, Y) *iff exactly one of the following holds:*

- *either X is an argument of the form* (∅, Γ, ∧);
- *else,* ∃ψ ∈ Π *such that attacks*(X, Y) *on* ψ, *and* X ⊀ ({ψ}, ∅, ψ).

X ⇒ Y *will stand for "defeats*(X, Y)*", and* X ⇏ Y *will stand for "¬defeats*(X, Y)*".*

Cl-Arg assumes instantiation of an AF by all arguments defined by a base $\mathscr{B}$ of classical wff, a task that proves to be unfeasible for agents with limited resources. As such, dialectical arguments (Definition 1) along with the described defeat relation (Definition 2) allow us to introduce a *dialectical AF* as an argumentation framework ⟨AR, *defeats*⟩ where AR is any subset of the dialectical arguments defined by a base $\mathscr{B}$.

The difference between a defeat and a dialectical defeat for dialectical AFs is the reference to a set $\mathscr{S}$ of arguments. The general idea is that, when challenging the acceptability of an argument with respect to a set $\mathscr{S}$, the defeating argument can suppose premises from all the arguments in $\mathscr{S}$. Whereas, the argument that defends $\mathscr{S}$ can only suppose the premises of the defeating argument. This new kind of defeat forces us to adjust the standard semantics accordingly.

Definition 3. [Dialectical defeats and semantics for dialectical AFs][*9*]
Let ⟨AR, *defeats*⟩ *be a dialectical AF,* $\mathscr{S}$ ⊆ AR *and* X, Y ∈ AR. *Then:*

1) X *dialectically defeats Y with respect to* $\mathscr{S}$, *denoted* X ⇒$_{\mathscr{S}}$ Y, *if defeats*(X, Y) *and Supp*(X) ⊆ *Prem*($\mathscr{S}$ ∪{Y}).

---

[2]As extensively explained in [15], the foreign commitment problem is the issue that arises in dialogical applications when agents are forced to commit to the premises of their interlocutors in order to challenge their arguments.

2) $\mathcal{S}$ is conflict free *if* $\forall Z, Y \in \mathcal{S}, Z \not\Rightarrow_{\mathcal{S}} Y$.

3) Y is acceptable *with respect to $\mathcal{S}$ if $\forall X$ such that $X \Rightarrow_{\mathcal{S}} Y, \exists Z \in \mathcal{S}$ such that $Z \Rightarrow_{\{X\}} X$.*

4) *Let $\mathcal{S}$ be conflict free. Then $\mathcal{S}$ is: an* admissible *extension iff $X \in \mathcal{S}$ implies X is acceptable with respect to $\mathcal{S}$; a* complete *extension iff $\mathcal{S}$ is admissible and if X is acceptable with respect to $\mathcal{S}$ then $X \in \mathcal{S}$; a* preferred *extension iff it is a set inclusion maximal complete extension; the* grounded *extension iff it is the set inclusion minimal complete extension.*

Dialectical AFs may enjoy some specific properties, as explained in [9]. Here we are going to outline four of them (P1, P2, P3, P4), which will prove useful in the next sections.

Remark 1. *Given a dialectical AF = $\langle AR, defeats \rangle$:*

(P1) $\forall X \in AR$: $\alpha \in Prem(X)$ *implies that* $(\{\alpha\}, \emptyset, \alpha) \in AR$;

(P2) $\forall X \in AR$: *if X' is the logically equivalent argument of X (i.e., the only difference between X and X' is the different distribution of premises and supposition), then $X' \in AR$;*

(P3) *If $(\Delta, \emptyset, \alpha) \in AR$ and $(\Gamma, \emptyset, \bar{\alpha}) \in AR$, then either $(\Delta, \emptyset, \wedge) \in AR$ or $(\Gamma, \emptyset, \wedge) \in AR$ or $(\Delta \cup \Gamma, \emptyset, \wedge) \in AR$;*

(P4) *If $(\Gamma, \emptyset, \alpha) \in AR$, $\Delta \subseteq \Gamma$, $\Delta \neq \emptyset$ and Δ is such that it does not share any propositional atomic formulae with $(\Gamma \setminus \Delta) \cup \{\alpha\}$ [3], then either $(\Delta, \emptyset, \wedge) \in AR$ or $(\Gamma \setminus \Delta, \emptyset, \alpha) \in AR$.*

2.2. Standard Argument Games

Finally, let us also review the fundamental notions of the standard argument games as described in [4]. Notice that these definitions have been modified to accommodate dialectical AFs (which is a fair straightforward adaptation). However, recall that the main contributions of this paper concern the development of argument games for Dialectical Cl-Arg that involves *dialectical defeats* (Definition 3): this entails a non-trivial modification of the standard games.

In a nutshell, an argument game is played by two players, PRO (for *proponent*) and OPP (for *opponent*), each of which is referred to as the other's 'counterpart'. PRO starts the game by moving an initial argument X that it wants to test. After that, both players take turns in moving arguments against their counterpart's moves. This generates disputes:

Definition 4. [Dispute] *A sequence of moves in which each player moves against its counterpart's argument is referred to as a* dispute. *Formally,* d = X—Y—Z—⋯ *is a dispute, and X—Y denotes a player moving argument Y against an argument X played by its counterpart (similarly, Y—Z). A sub-dispute d' of a dispute d is any sub-sequence of d that starts with the same initial argument. For example, if d = X—Y—Z, then $d' = X—Y$ would be a sub-dispute of d.*

Notice that, to avoid ambiguity, each argument of a dispute will be labelled with either P or O (that stands for either one of the two players, PRO or OPP). Hence, d = (P)X—(O)Y—(P)Z is a dispute where PRO moves the argument X, followed by Y played by OPP and countered by another move from PRO, Z.

---

[3]Alternatively, if we are dealing with first order logic, Δ is such that it does not share functions and predicates with $(\Gamma \setminus \Delta) \cup \{\alpha\}$.

We can now introduce the notion of the (unique) dispute tree, which represents the 'playing field' of the standard argument games. In other words, the dispute tree is the data structure that contains all the potential moves (and sequences of moves) available to the players.

Definition 5. [Dispute Tree] *Let AF = ⟨AR, defeats⟩ be a finite dialectical argumentation framework, and let A ∈ AR. The* dispute tree *induced by A in the AF is the tree $\mathcal{T}$ of arguments, such that $\mathcal{T}$'s root node is A, every branch of the tree is a different dispute, and ∀X, Y ∈ AR: X is a child of Y in $\mathcal{T}$ iff* defeats(X, Y).

From here on, we are going to write PRO(∗) and OPP(∗) to denote the sets of all PRO and OPP arguments in ∗, where ∗ can be replaced with d, $\mathcal{T}$ or any other tree that will be introduced in the remainder of the paper. Also, LAST(d) will identify the last argument played in a dispute d.

An argument game is said to be won by the proponent only if it has a winning strategy. That is to say, only if it can successfully defend the argument it wants to test (i.e., the root of $\mathcal{T}$) against any possible arguments moved by the opponent. PRO loses otherwise.

Definition 6. [Winning Strategy] *Let $\mathcal{T}$ be the dispute tree induced by A in a finite dialectical AF = ⟨AR, defeats⟩. Let also d be a dispute in $\mathcal{T}$. Then, a* winning strategy $\mathcal{T}'$ *for A is the dispute tree $\mathcal{T}$ pruned in a way such that:*

(6.1) *The set $\mathcal{T}'_D$ of disputes in $\mathcal{T}'$ is a non-empty finite set such that each dispute $d \in \mathcal{T}'_D$ is finite and is won by PRO (i.e., LAST(d) ∈ PRO($\mathcal{T}$));*

(6.2) *∀d ∈ $\mathcal{T}'_D$, ∀d' such that d' is some sub-dispute of d, LAST(d') = X and X ∈ PRO($\mathcal{T}$), then ∀Y ∈ OPP($\mathcal{T}$) such that Y ⇒ X, there is a d'' ∈ $\mathcal{T}'_D$ such that d' − Y is a sub-dispute of d''.*

Informally, the previous definition states that a winning strategy is the dispute tree $\mathcal{T}$ pruned in a way such that (6.1) $\mathcal{T}'_D$ is a non-empty finite set, its disputes are finite, end with a PRO argument and (6.2) are such that OPP has moved exhaustively (i.e., all the moves that OPP could have played, had been played) and also PRO has countered every defeating argument moved by OPP.

3. Developing Dialectical Argument Games

In the following sections, we are going to develop argument games for Dialectical Cl-Arg that accommodate the dialectical defeats and semantics introduced in Definition 3. The resulting proof theory will present some specific features that will distinguish it from the standard argument games, although the general structure remains similar. Intuitively, winning a dialectical game for an argument *A* means having a specific 'dialectical way' (depending on the semantics that the proof theory is meant to capture) for defending the information contained in *A*, hence showing the admissibility of the encoded data.

The main difference between a dispute tree $\mathcal{T}$ and a *dialectical dispute tree* $\mathcal{D}$ can be identified with the additional reference to a subset $\mathcal{S} \subseteq$ PRO($\mathcal{T}$). That is to say, $\mathcal{S}$ represents a candidate admissible set of PRO arguments such that PRO commits to their premises. Recall once again that, when challenging the acceptability of an argument with respect to a set $\mathcal{S}$, the defeating

argument can suppose premises from all the arguments in $\mathcal{S}$. Whereas, the argument that defends $\mathcal{S}$ can only suppose the premises of the defeating argument. Another important difference between standard and dialectical games is that the latter handles *partially instantiated dialectical AFs* (*pdAFs*), i.e. dialectical AFs that satisfy properties P1, P2, P3 and P4 [4].

As a consequence, each dialectical game enjoys specific properties that encapsulate the dialectical uses of arguments by real-world resource-bounded agents, thus succeeding in better approximating a process capable of bridging formal (proof-theoretical) and informal (real-world exchange of arguments) single-agent reasoning.

We can now formally introduce the (unique) dialectical dispute tree induced by A wrt a set $\mathcal{S}$ [5]:

Definition 7. [Dialectical Dispute Tree] *Let $\mathcal{T}$ be the dispute tree induced by* A *in a finite pdAF = $\langle AR, defeats \rangle$. Let also $\mathcal{S} \subseteq PRO(\mathcal{T})$. Then, the dialectical dispute tree $\mathcal{D}$ induced by* A *with respect to $\mathcal{S}$ is the dispute tree $\mathcal{T}$ pruned in a way such that $\forall X, Y \in AR$: X is a child of Y in $\mathcal{D}$ iff defeats(X, Y) and:*

1. *If $X \in PRO(\mathcal{D})$ and $Y \in OPP(\mathcal{D})$, then $X \Rightarrow_{\{Y\}} Y$, i.e. X defeats Y and Supp$(X) \subseteq$ Prem(Y);*
2. *If $X \in OPP(\mathcal{D})$ and $Y \in PRO(\mathcal{D})$, then $X \Rightarrow_{\mathcal{S}} Y$, i.e. X defeats Y with respect to a set $\mathcal{S}$ and Supp$(X) \subseteq$ Prem$(\mathcal{S} \cup \{Y\})$.*

The 'playing field' of the dialectical argument games (i.e., the data structure on the basis of which the games are played) is still depicted by the dispute tree $\mathcal{T}$. Indeed, the relationship existing between the dispute tree $\mathcal{T}$ induced by A in a finite pdAF and the dialectical dispute tree $\mathcal{D}$ induced by A wrt $\mathcal{S}$ is such that $\mathcal{D}$ is 'contained' in $\mathcal{T}$ (since $\mathcal{D}$ is a pruned version of $\mathcal{T}$), as shown in the following example.

Example 1. *Figure 1 presents the (incomplete) dispute tree $\mathcal{T}$ induced by A_1 in a finite pdAF = $\langle AR, defeats \rangle$ and the corresponding (incomplete) dialectical dispute tree $\mathcal{D}$ induced by A_1 wrt $\mathcal{S} = \{A_1, G_2, O_1\}$ in the same pdAF. Both trees are incomplete since the purpose of the example is just to show the relationship existing between them. For the same reason, we also avoid listing all the arguments of the pdAF.*

Observe that, unlike $\mathcal{T}$, where no set is taken into consideration, the defeats in $\mathcal{D}$ are parametrized to the set $\mathcal{S}$. This implies that, when defeating PRO's arguments, the arguments moved by OPP can only suppose the premises of the arguments in the set $\mathcal{S}$. No such restrictions exist for $\mathcal{T}$. Notice that, even if we keep extending both trees, dispute $d = (P)A_1 - (O)F_2 - (P)G_2$ will never be part of $\mathcal{D}$. This is because, according to Definition 7 (which also emphasizes how dialectical defeats work), PRO can move G_2 only if Supp$(G_2) \subseteq$ Prem(F_2). However, this is never going to be the case since the formula $\neg a \lor \neg b \notin$ Prem(F_2). Therefore, even if the two trees were identical in every other branch, the absence of dispute d will still make $\mathcal{D}$ 'contained' in $\mathcal{T}$.

---

[4] Refer to Remark 1.

[5] Notice, however, that there can be different dialectical dispute trees induced by A wrt to different sets $\mathcal{S}$. That is to say, a dialectical dispute tree $\mathcal{D}_1$ induced by A wrt $\mathcal{S}_1$ constitutes a different data structure than a dialectical dispute tree $\mathcal{D}_2$ induced by A wrt $\mathcal{S}_2$.
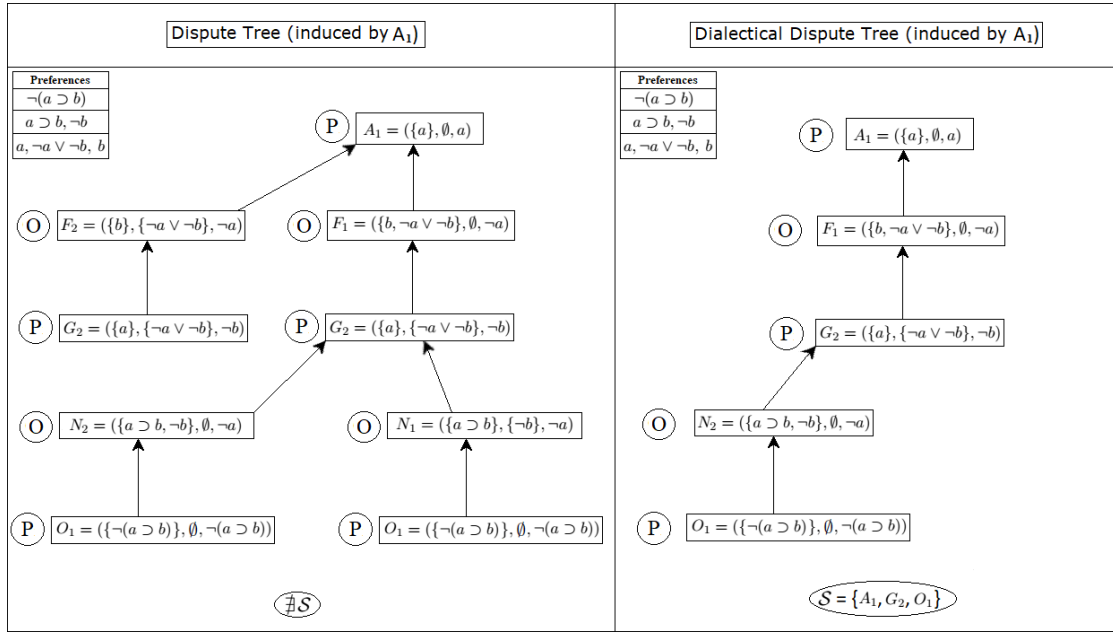
Figure 1: The (incomplete) dispute tree $\mathcal{T}$ (on the left) induced by A_1 in a finite pdAF = $\langle$AR, *defeats*$\rangle$ and the corresponding (incomplete) dialectical dispute tree $\mathcal{D}$ (on the right) induced by A_1 wrt $\mathcal{S} = \{A_1, G_2, O_1\}$ in the same pdAF = $\langle$AR, *defeats*$\rangle$.

Dialectical argument games share with the standard argument games the notion of winning strategy: in order to win the game for an argument A, PRO must have a winning strategy for it. It will lose otherwise. However, the two definitions slightly differ since a dialectical winning strategy has to take into account the set $\mathcal{S}$ targeted by the dialectical defeats:

Definition 8. [Dialectical Winning strategy] *Let $\mathcal{D}$ be the dialectical dispute tree induced by A wrt $\mathcal{S}$ in a finite pdAF = $\langle$AR, defeats$\rangle$ and let d be a dispute in $\mathcal{D}$. Then, a dialectical winning strategy $\mathcal{W}$ for A corresponds to the dialectical dispute tree $\mathcal{D}$ pruned in a way such that:*

(8.1) *The set $\mathcal{W}_D$ of disputes in $\mathcal{D}$ is a non-empty finite set such that each dispute $d \in \mathcal{W}_D$ is finite and is won by PRO (i.e., LAST(d) $\in$ PRO($\mathcal{D}$));*

(8.2) *$\forall d \in \mathcal{W}_D$, $\forall d'$ such that d' is some sub-dispute of d, LAST(d') = X and $X \in$ PRO($\mathcal{D}$), then $\forall Y \in$ OPP($\mathcal{D}$) such that $Y \Rightarrow_{\mathcal{S}} X$, there is a $d'' \in \mathcal{W}_D$ such that $d' - Y$ is a sub-dispute of d''.*

Similarly to Definition 6, the previous definition states that a dialectical winning strategy corresponds to the dialectical dispute tree $\mathcal{D}$ pruned in a way such that (8.1) $\mathcal{W}_D$ is a non-empty finite set, its disputes are finite, end with a PRO argument and are such that (8.2) OPP has moved exhaustively and also PRO has countered each defeating argument moved by OPP. The difference is in the dialectical defeats: the nodes are no more connected by means of the defeats relations among arguments, but through dialectical defeats among arguments that target the set $\mathcal{S}$.

We now have all the elements needed to formally introduce the protocol of the dialectical admissible/preferred game. Similar to a list of instructions, this protocol determines the legal moves that can be performed by the players. The game unfolds as a result of the legal arguments played and terminates when there are no more valid moves available. When this happens, the status of the root of the tree is evaluated. The presence of a winning strategy for such an argument assigns the victory to PRO. Strictly speaking, OPP never wins: its purpose is to counter each argument moved by the proponent in order to assist it in testing the admissibility of the root argument (indeed, argument games are formalisations of single-agent reasoning processes). Nevertheless, OPP can still prevent PRO's victory by invalidating its winning strategy.

4. Dialectical Admissible/Preferred Games

When we play a Φ-dialectical game we are increasingly building, starting from the root A and following the legal moves licensed by the protocol Φ, a dialectical dispute tree denoted as Φ-$\mathscr{D}^n$. Each node of such tree corresponds to an argument progressively played by either PRO or OPP that is labelled with a positive integer i (with $1 \leq i \leq n$). These additional labels allow identifying the order in which the arguments have been played, hence, also determining the current stage (i.e., the nth-stage) of the Φ-dialectical game. Recall that the dispute tree $\mathscr{T}$ induced by A represents the playing field of the games and every Φ-dialectical game for A is contained within its data structure (i.e., Φ-$\mathscr{D}^n$ is a 'pruned-version' of $\mathscr{T}$). Being a dialectical dispute tree, even Φ-$\mathscr{D}^n$ is constructed wrt a set $\mathscr{S} \subseteq \mathrm{PRO}(\mathscr{T})$, however, such $\mathscr{S}$ can gradually increase with each new move made by PRO during the game. Indeed, $\mathscr{S}$ is composed of the same arguments moved by PRO in Φ-$\mathscr{W}^n$ (i.e., a dialectical winning strategy for A of Φ-$\mathscr{D}^n$), which can be extended while the game proceeds [6]. As it will be shown, observe also that $\mathscr{S}$ is still a different set than $\mathrm{PRO}(\Phi$-$\mathscr{W}^n)$, meaning that it will modify its members according to the changes in $\mathrm{PRO}(\Phi$-$\mathscr{W}^n)$, but it will never be empty even if there is no winning strategy Φ-$\mathscr{W}^n$. Formally:

Definition 9. [Φ-$\mathscr{D}^n$ and Φ-$\mathscr{W}^n$] *A dialectical dispute tree Φ-$\mathscr{D}^n$ induced by* A *wrt* $\mathscr{S} \subseteq PRO(\mathscr{T})$ *(with* $\mathscr{S} \neq \emptyset$) *in a finite pdAF =* $\langle AR, defeats \rangle$ *is the tree that starts from the argument* A *and it is progressively built up to the nth-move of one of the two players, such that each node of the tree is labelled with a positive integer i (for* $1 \leq i \leq n$). *Moreover,* $\forall X, Y \in AR$: *X is a child of Y in* Φ-$\mathscr{D}^n$ *iff X is a legal move identified by* Φ, *defeats(X, Y) and:*

1. *If* $X \in PRO(\Phi$-$\mathscr{D}^n)$ *and* $Y \in OPP(\Phi$-$\mathscr{D}^n)$ [7], *then* $X \Rightarrow_{\{Y\}} Y$, *i.e. X defeats Y and Supp(X)* $\subseteq$ *Prem(Y);*

2. *If* $X \in OPP(\Phi$-$\mathscr{D}^n)$ *and* $Y \in PRO(\Phi$-$\mathscr{D}^n)$, *then* $X \Rightarrow_{\mathscr{S}} Y$, *i.e. X defeats Y with respect to a set* $\mathscr{S}$ *and Supp(X)* $\subseteq$ *Prem($\mathscr{S} \cup \{Y\}$).*

Finally, Φ-$\mathscr{W}^n$ *will denote a dialectical winning strategy for* A *of* Φ-$\mathscr{D}^n$.

---

[6] Although the set $\mathscr{S}$ can increase the number of its members while the game goes on, it can never exceed the size of $\mathrm{PRO}(\mathscr{T})$. Keep in mind that every Φ-dialectical game for A is contained in the dispute tree $\mathscr{T}$ induced by A (since $\mathscr{T}$ corresponds to the playing field of the game).

[7] Similarly to $\mathrm{PRO}(\mathscr{D})$ and $\mathrm{OPP}(\mathscr{D})$, writing $\mathrm{PRO}(\Phi$-$\mathscr{D}^n)$ and $\mathrm{OPP}(\Phi$-$\mathscr{D}^n)$ denote the sets of all PRO, respectively OPP, arguments in Φ-$\mathscr{D}^n$.

Every stage of a Φ-dialectical game can then be identified with a specific dialectical dispute tree $\Phi\text{-}\mathscr{D}^n$. As we are going to see, this notion is essential for a proper account of the dialectical defeats in the game protocol [8].

It is interesting to notice that, during a Φ-dialectical game, a dialectical defeat that occurred in an early stage of the game might not take place in a more advanced phase of the same game. This can be caused by an update of the current $\mathscr{S}$, the set parametrized by OPP for performing dialectical defeats. We denote this anomaly as 'disqualified defeats'.

Definition 10. [Disqualified dialectical defeats] *Let $\Phi\text{-}\mathscr{D}^n$ be the dialectical dispute tree of a Φ-dialectical game where X and Y denote arguments played respectively by OPP and PRO in $\Phi\text{-}\mathscr{D}^n$. Let also $X \Rightarrow_\mathscr{S} Y$ by supposing $\alpha \in Prem(\mathscr{S})$. If, after the game goes on, we will reach a stage $\Phi\text{-}\mathscr{D}^{n+k}$ (for $k > 0$) where $\alpha \notin Prem(\mathscr{S})$, then the defeat moved by X against Y will be invalidated and will be denoted as 'disqualified'. As such, X and all the arguments following it in the same dispute will be (temporary) pruned from the tree.*

Consider indeed that the status of disqualified defeats might be temporary and be updated again in a further stage of the game (when these defeats will become valid once more).
Definition 10 entails the following proposition:

Proposition 1. *Let $\Phi\text{-}\mathscr{D}^n$ be the dialectical dispute tree of a Φ-dialectical game:*

(I) *If the nth-move is an argument X played by OPP, then moving X cannot disqualify the dialectical defeat that X performs against a PRO argument.*

(II) *The presence of OPP arguments whose defeats have been disqualified will not affect the dialectical winning strategy.*

Proof 1.

(I) *Since X is the last argument (legally) played in $\Phi\text{-}\mathscr{D}^n$, it trivially does not comply with Definition 10.*

(II) *Even if some of the dialectical defeats moved by OPP arguments have been disqualified (hence are no more a threat for PRO), the defeat of the last argument played by OPP has not been disqualified due to (I). Notice also that every argument that follows X, whose defeat has been invalidated, will not be part of the dialectical winning strategy $\Phi\text{-}\mathscr{W}^n$ (including X). Otherwise, we would violate Definition 10.*

Notice that every dialectical game protocol Φ takes into account disqualified defeats, which are then also contemplated by the dialectical dispute tree $\Phi\text{-}\mathscr{D}^n$ (and dialectical winning strategy $\Phi\text{-}\mathscr{W}^n$) of Definition 9.

---

[8] Observe that it is possible for one (or more, depending on the protocol) dialectical winning strategy $\Phi\text{-}\mathscr{W}^n$ for A of $\Phi\text{-}\mathscr{D}^n$ to exist, although there is no dialectical winning strategy $\mathscr{W}$ for A of $\mathscr{D}$. This can happen, for example, when $\mathscr{D}$ is composed only by infinite disputes (recall that we need finite disputes to have winning strategies, as stated by Definition 8.1), whilst $\Phi\text{-}\mathscr{D}^n$ is composed by finite disputes, due to the restrictions imposed by the protocol Φ. In this situation, it is possible to identify in $\Phi\text{-}\mathscr{D}^n$ a winning strategy $\Phi\text{-}\mathscr{W}^n$. Such an example is illustrated in Figure 2(b).

4.1. Protocol

We can now formally introduce the protocol for the dialectical admissible/preferred game. As already stated, during the dialectical argument game, the players have to comply with a protocol Φ that identifies the legal moves allowed.

Definition 11. [Dialectical Admissible Game legal moves] *Let $\Phi\text{-}\mathscr{D}^n$ and $\Phi\text{-}\mathscr{W}^n$ be defined as in Definition 9, let d be a dispute of $\Phi\text{-}\mathscr{D}^n$ and d' be a sub-dispute of d. Let also $(PL_n)X$ (for $n > 0$) denotes the argument X played by either one of the two players (P or O) as the (last) nth-move. Then Φ_P identifies legal moves in the following way:*

(11.0) *PRO moves the first argument.*

(11.1) *If $(PL_n)X$ and $n = 2k$ (for $k > 0$), then the next move $n{+}1$, say Y, is by PRO and it is such that:*

     (a) $Y \Rightarrow_{\{Z\}} Z$, where $Z \in OPP(\Phi\text{-}\mathscr{D}^n)$;

     (b) There exists a $\Phi\text{-}\mathscr{W}^{n+1}$ for A of $\Phi\text{-}\mathscr{D}^{n+1}$.

(11.2) *If $(PL_n)X$ and $n = 2k + 1$ (for $k \geq 0$), then the next move $n{+}1$, say Y, is by OPP and it is such that:*

     (a) $Y \Rightarrow_{\mathcal{S}} Z$, where $Z \in \mathcal{S}$ and $\mathcal{S} := PRO(\Phi\text{-}\mathscr{W}^n)$ [9];

     (b) If $d = d' - Z$, then $Y \notin OPP(d')$.

A Φ_P-dialectical game is said to be terminated when, during its turn, the corresponding player runs out of the legal moves identified by (11.1) or (11.2) of the protocol Φ_P. PRO wins only if it has a winning strategy once the game terminates. It loses otherwise.

     The previous protocol can be informally summarised as follows. PRO starts the game by playing the first argument [(11.0)] and, after that, OPP will make its move. Then, the two players alternate in playing only one argument at a time to reply to one of their counterpart's arguments. Observe that when $\mathcal{S}$ is initialized in the game and, subsequently, every time its arguments are updated by the changes in $PRO(\Phi\text{-}\mathscr{W}^n)$ [$(11.2(a))$], it is always the beginning of OPP's turn. This means that the condition for which $\mathcal{S} \neq \varnothing$ is continuously respected [10].

     Observe that the established protocol allows *backtracking* to other arguments. That is to say, when PRO moves it can either target the last argument played by OPP or another argument moved by OPP in the dialectical dispute tree generated thus far (i.e., an argument member of the set $OPP(\Phi\text{-}\mathscr{D}^n)$) [$(11.1(a))$]. Similarly, when OPP moves it can either target the last argument played by PRO or another argument moved by PRO in the current dialectical winning strategy (i.e., an argument member of the set $PRO(\Phi\text{-}\mathscr{W}^n)$) [$(11.2(a))$]. The *relevance conditions* [$(11.1(b))$ for PRO; $(11.2(a))$ for OPP] ensure that: after PRO has made its move, there will be a winning strategy $\Phi\text{-}\mathscr{W}^{n+1}$, hence providing the victory to PRO; after OPP has moved, instead, the previous winning strategy will cease to exist, thus preventing PRO from

---

[9]The symbol ':=' denotes a variable initialization rather than an equivalence relation. That is to say, at the beginning of each OPP's turn, the content of $\mathcal{S}$ is initialized to the current $PRO(\Phi\text{-}\mathscr{W}^n)$, i.e, the arguments member of $\mathcal{S}$ are the same as $PRO(\Phi\text{-}\mathscr{W}^n)$. This operation overwrites the previous contents of $\mathcal{S}$.

[10]That is because a situation in which $\mathcal{S} = PRO(\Phi\text{-}\mathscr{W}^n) = \varnothing$ never occurs at the beginning of OPP's turn.

winning. That is to say, PRO will be forced to generate a dialectical winning strategy during each of its turns, while OPP will have to invalidate such winning strategy during every one of its turns.

Finally, the restriction (11.2(*b*)) on the moves played by OPP is necessary (as also shown in the standard games of [4], [5] and [6]). Indeed, allowing OPP to repeat its arguments, since OPP is required to move exhaustively, could imply the generation of infinite disputes. To see why let us suppose that $(PL_n)X$ (for $n > 0$) identifies an argument X played by either one of the two players (denoted as P or O) as its nth move in a Φ-dialectical game. Then, there could be an infinite dispute d like the following:

$$d = (P_1)A - \cdots - (O_n)Y - (P_{n+1})Z - (O_{n+2})Y - (P_{n+3})Z - (O_{n+4})Y - \cdots$$

Intuitively, since Z is capable of defending itself by defeating Y, there is no need to further extend the dispute by repeating the same arguments: this is because Z has already shown its acceptability wrt $PRO(\Phi\text{-}\mathscr{W}^{n+1})$. Therefore, the only way for avoiding infinite disputes (and infinite dialectical admissible/preferred games) is to prevent OPP from repeating its arguments in the same disputes.

Remark 2. *Similarly to the standard argument games presented in [4], the protocol of the dialectical admissible games is identical to the protocol of the dialectical credulous preferred games. Indeed, it suffices to show the membership of an argument* A *in an admissible extension to show also that* A *is credulously justified under the preferred semantics as well. That is because every admissible extension of a dialectical AF is a subset of a preferred extension. This is a consequence of the Fundamental Lemma (FL) introduced by Dung [3] and adapted for Dialectical Cl-Arg in [9].*

As it has been defined, the admissible/preferred game satisfies the properties of soundness and completeness. This proves the equivalence existing between the victory of the Φ_P-dialectical game for an argument A and the membership of the same A to the admissible/preferred extensions of the corresponding finite pdAF.

Theorem 1. *Let* $\Phi_P\text{-}\mathscr{D}^n$ *identifies a terminated* Φ_P*-dialectical game. Then, there exists a dialectical winning strategy* $\Phi_P\text{-}\mathscr{W}^n$ *for* A*, such that the set* $PRO(\Phi_P\text{-}\mathscr{W}^n)$ *of arguments moved by PRO in* $\Phi_P\text{-}\mathscr{W}^n$ *is conflict-free, iff* A *is included in an admissible extension* Adm *of the pdAF.*

Proof 2. [Soundness] *It follows directly from 8.1 and 8.2*

Proof 3. [Completeness]*(Sketched). Assume that* A $\in$ Adm*. Since the pdAF is finite, then it is also finitary. We can build a dialectical winning strategy* $\Phi_P\text{-}\mathscr{W}^n$ *for* A *if, for each argument* Y *dialectically defeating* A *and moved by OPP, PRO chooses one argument* X *from* Adm *(even* A *itself) such that* $X \Rightarrow_{\{Y\}} Y$ *(Definition 11.2(b) prevents the generation of infinite disputes). This procedure can be repeated for every defeating argument moved by OPP.*
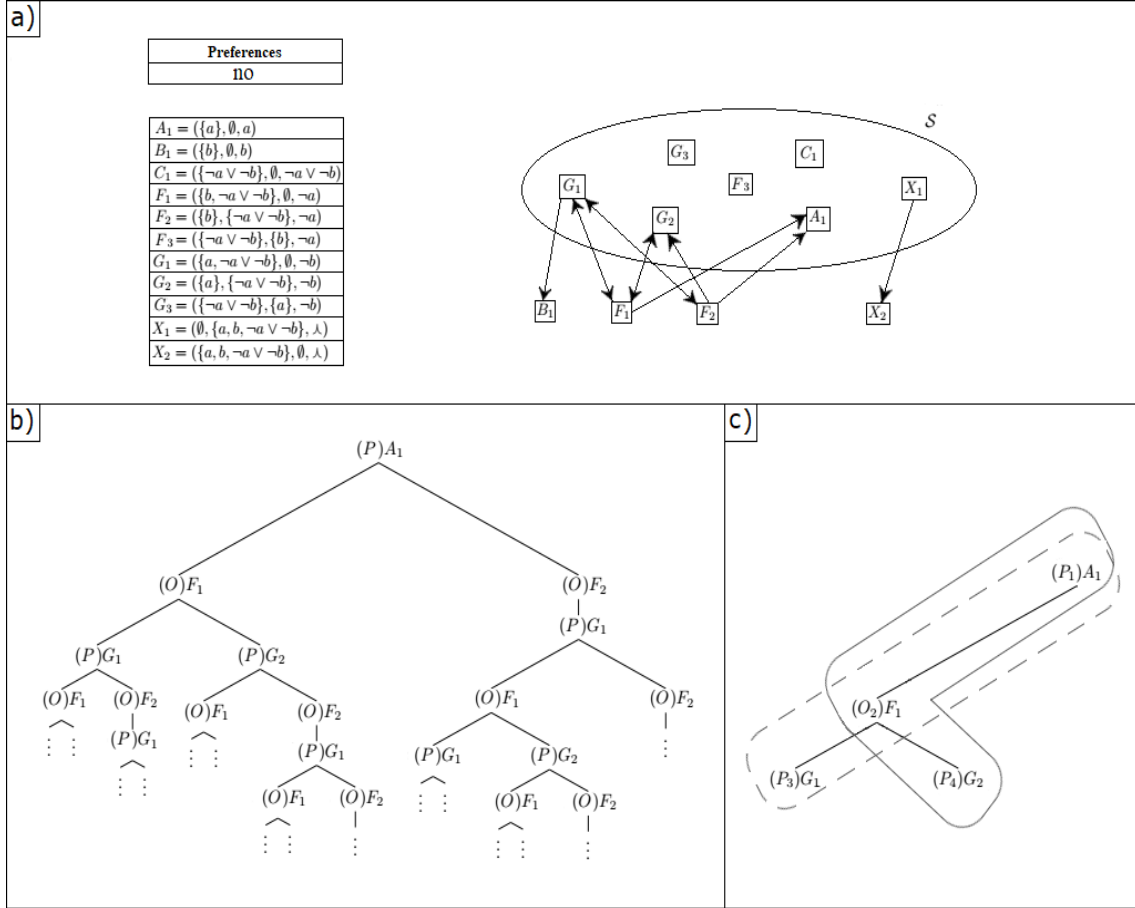
Figure 2: Figure a) illustrates a pdAF with a list of its arguments and the set $\mathscr{S}$ that is parametrized by the dialectical defeats. Consider also that X_2 is defeated by all the arguments of the pdAF, except A_1, B_1, C_1 (the arrows that should have highlighted such defeats have been omitted to avoid unnecessary graphical confusions). Figure b) displays the dialectical dispute tree $\mathscr{D}$ induced by A_1 wrt $\mathscr{S}$ in the pdAF of Figure a). Notice that $\mathscr{D}$ is composed of infinite disputes (the vertical dots represent the endless continuation of the disputes). A dialectical dispute tree $\Phi\text{-}\mathscr{D}^n$, with $n = 4$, is depicted in Figure c) and corresponds to a Φ-dialectical game played up to the nth-move. Observe that the number of each move (next to the label P or O) represents the order in which the arguments have been played in the game. In this example, we are assuming a protocol Φ that licenses legal moves where PRO can play more than one argument per turn, therefore, $\Phi\text{-}\mathscr{D}^n$ has two winning strategies (both of which are encircled in the figure).

4.2. Main Features of Dialectical Argument Games

Dialectical admissible/preferred argument games hold specific features that differentiate them from the standard argument games of [4][6][5] and depend upon their protocol and the properties possessed by each pdAF (especially P1, P2 and P3). In the following, we are going to list all of such features. Notice that the proofs of the results in this section are left out due to space constraints.

4.2.1. Features F1-F2

(F1) *The set of all the arguments moved by PRO in a dialectical winning strategy of a Φ_P-dialectical game (i.e., $PRO(\Phi_P\text{-}\mathscr{W}^n)$) is always conflict-free;*

(F2) *The relevance conditions, i.e., the conditions of the admissible/preferred protocol that compel both players to change the outcome of the game at the end of every turn, are essential to the unfolding of the Φ_P-dialectical games. This also justifies why the set $\mathscr{S}$ cannot be initialized with any set other than $PRO(\Phi_P\text{-}\mathscr{W}^n)$.*

4.2.2. Feature F3

Before the introduction of the third feature (F3) enjoyed by the dialectical admissible/preferred argument games, we need to formally define what is meant by the *uniqueness* of the dialectical winning strategy $\Phi\text{-}\mathscr{W}^n$.

Definition 12. [Uniqueness of the dialectical winning strategy $\Phi\text{-}\mathscr{W}^n$] *Let $\Phi\text{-}\mathscr{D}^n$ and let $\Phi\text{-}\mathscr{W}^n$ be defined as in Definition 9. Then $\Phi\text{-}\mathscr{W}^n$ is said to enjoy the uniqueness property if there is no other dialectical winning strategy for A wrt $\mathscr{S}$ simultaneously present in $\Phi\text{-}\mathscr{D}^n$.*

Let us consider the dialectical dispute tree $\Phi\text{-}\mathscr{D}^n$ of Figure 2(c). This tree has two winning strategies, say $\Phi\text{-}\mathscr{W}_1^n$ and $\Phi\text{-}\mathscr{W}_2^n$, each of which is composed of a single dispute. That is to say: $d_1 = (P_1)A_1-(O_2)F_1-(P_3)G_1$ and $d_2 = (P_1)A_1-(O_2)F_1-(P_4)G_2$, such that $\Phi\text{-}\mathscr{W}_1^n$ is composed of d_1, while $\Phi\text{-}\mathscr{W}_2^n$ is composed of d_2. Obviously, $\Phi\text{-}\mathscr{D}^n$ does not enjoy the uniqueness property. Indeed, both G_1 and G_2 defeat the same argument F_1, whereas only one of such defeats is actually needed. This implies that it suffices that either $\Phi\text{-}\mathscr{W}_1^n$ or $\Phi\text{-}\mathscr{W}_2^n$ is present for PRO to win the game. For the final outcome of the game, it is pointless to have both winning strategies simultaneously. It is also resources consuming, meaning that it does not comply well with the Dialectical Cl-Arg purpose of capturing resource-bounded real-world agents reasoning.

(F3) *Any dialectical winning strategy $\Phi_P\text{-}\mathscr{W}^n$ enjoys the uniqueness property.*

4.2.3. Advantages of F1-F3

Each of these features proves to be useful in some specific way. *F1* allow eschewing consistency checks on the premises of arguments moved by PRO in any dialectical winning strategy $\Phi_P\text{-}\mathscr{W}^n$. In addition, it is of particular importance in providing the games with a wide range of efficiency improvements. Without the need to perform any extra check or to enforce further restrictions in the protocol (unlike in [4]), this property allows every dialectical game to prevent the proponent from: playing self-defeating arguments, playing arguments already moved by the opponent and playing arguments that defeat (or are defeated by) other arguments already moved by the proponent. *F2* revolves around the relevance conditions (11.1(b) and 11.2(a) of Definition 11) that can be summarised as the conditions that force the two players to change the outcome of the game at the end of every turn. These requirements are fundamental for real-world agents that reason with limited availability of resources. Indeed, it would be nonsensical to allow such agents to play arguments irrelevant for the outcome of the game: this would mean wasting valuable resources. The last feature, *F3*, concerns the uniqueness property. Uniqueness, enforced

on a dialectical winning strategy $\Phi_P\text{-}\mathscr{W}^n$ by the protocol of the dialectical admissible/preferred argument game, is certainly a desirable feature since allows for shorter and simpler games. This means that it is possible to evaluate the status of the dialectical dispute tree root faster than the standard games.

5. Conclusions and Future Work

The main aspects of the real-world uses of argumentation by resource-bounded agents include: (i) showing arguments inconsistencies by supposing the premises of the opponent's arguments; (ii) handling only finite subsets of the arguments of the AFs; (iii) reducing resources consumption by employing dialectical means (while still satisfying the rationality postulates) [9]. These features would constitute the hallmarks of an argument game based on Dialectical Cl-Arg, thus capable of better approximating non-monotonic single-agent real-world reasoning processes than the standard argument games. In this paper, we have achieved some important results. We have developed argument game proof theories (denoted as *dialectical argument games*) for the admissible and preferred semantics of Dialectical Cl-Arg. The limited space prevented us from entirely appreciating the extent of the formalism that would have also included a specific protocol for Dung's grounded semantics and a fully-fledged list of efficiency improvements for the introduced games. Incorporating dialectical defeats in the standard structure of the argument games proved to be a non-trivial process which yielded the discovery of interesting properties that differentiate dialectical games from the standard ones. That is to say, dialectical games enjoy (a) specific relevance conditions that identify their protocols and yield (b) the uniqueness of their winning strategies, whilst property *F1* ensures (c) the conflict-freeness of the set of arguments moved by the proponent in the winning strategy. The latter is of particular importance since it provides the games with a various range of efficiency improvements.

Extending further the study commenced in [16], we plan to increase the range of dialectical argument games protocols investigating the stable [6], semi-stable [17] and ideal semantics [18] [19]. Another research direction that will be pursued involves generalising the developed dialectical argument games to dialogues, following the guidelines of the already existing literature in the field (mainly [20]). This would have the interesting consequence of allowing to move from non-monotonic single-agent inference to distributed non-monotonic reasoning.

References

[1] H. Mercier, D. Sperber, Why do humans reason? arguments for an argumentative theory, Behavioral and brain sciences 34 (2011) 57–74.

[2] S. Modgil, Dialogical scaffolding for human and artificial agent reasoning., in: AIC, 2017, pp. 58–71.

[3] P. M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, Artificial intelligence 77 (1995) 321–357.

[4] S. Modgil, M. Caminada, Proof theories and algorithms for abstract argumentation frameworks., Argumentation in artificial intelligence 105129 (2009).

[5] G. A. Vreeswik, H. Prakken, Credulous and sceptical argument games for preferred semantics, in: European Workshop on Logics in Artificial Intelligence, Springer, 2000, pp. 239–253.

[6] M. Caminada, Y. Wu, An argument game for stable semantics, Logic Journal of IGPL 17 (2009) 77–90.

[7] M. Caminada, L. Amgoud, On the evaluation of argumentation formalisms, Artificial Intelligence 171 (2007) 286–310.

[8] M. Caminada, W. Carnielli, P. Dunne, Semi-stable semantics, Journal of Logic and Computation 22 (2012) 1207–1254.

[9] M. D'Agostino, S. Modgil, Classical logic, argument and dialectic, Artificial Intelligence 262 (2018) 15–51.

[10] P. M. Dung, R. A. Kowalski, F. Toni, Assumption-based argumentation, in: Argumentation in artificial intelligence, Springer, 2009, pp. 199–218.

[11] H. Prakken, An abstract framework for argumentation with structured arguments, Argument and Computation 1 (2010) 93–124.

[12] S. Modgil, H. Prakken, A general account of argumentation with preferences, Artificial Intelligence 195 (2013) 361–397.

[13] N. Gorogiannis, A. Hunter, Instantiating abstract argumentation with classical logic arguments: Postulates and properties, Artificial Intelligence 175 (2011) 1479–1497.

[14] P. Besnard, A. Hunter, Elements of argumentation, volume 47, MIT press Cambridge, 2008.

[15] M. Caminada, S. Modgil, N. Oren, Preferences and unrestricted rebut, Computational Models of Argument (2014).

[16] F. Castagna, Argument games for dialectical classical logic argumentation, Online Handbook of Argumentation for AI (2020) 2–6.

[17] M. Caminada, An algorithm for computing semi-stable semantics, in: European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Springer, 2007, pp. 222–234.

[18] P. M. Dung, P. Mancarella, F. Toni, Computing ideal sceptical argumentation, Artificial Intelligence 171 (2007) 642–674.

[19] M. Caminada, A labelling approach for ideal and stage semantics, Argument and Computation 2 (2011) 1–21.

[20] H. Prakken, Coherence and flexibility in dialogue games for argumentation, Journal of logic and computation 15 (2005) 1009–1040.