Discussion Games for Preferred Semantics of Abstract Dialectical Frameworks *

Atefeh Keshavarzi Zafarghandi, Rineke Verbrugge, and Bart Verheij

Department of Artificial Intelligence,
Bernoulli Institute of Mathematics, Computer Science and Artificial Intelligence,
University of Groningen, Groningen, The Netherlands
{A.Keshavarzi.Zafarghandi, L.C.Verbrugge, Bart.Verheij}@rug.nl

Abstract. Abstract dialectical frameworks (ADFs) are introduced as a general formalism for modeling and evaluating argumentation. However, the role of discussion in reasoning in ADFs has not been clarified well so far. The current work provides a discussion game as a proof method for preferred semantics of ADFs to cover this gap. We show that an argument is credulously acceptable (deniable) by an ADF under preferred semantics iff there exists a discussion game that can defend the acceptance (denial) of the argument in question.

Keywords: argumentation, abstract dialectical frameworks, decision theory, game theory, structural discussion

1 Introduction

Abstract Dialectical frameworks (ADFs), first introduced in [7] and have been further refined in [5, 6], are expressive generalizations of Dung's widely used argumentation frameworks (AFs) [15]. ADFs are formalisms that abstract away from the content of arguments but are expressive enough to model different types of relations among arguments. The applications of ADFs have been presented in legal reasoning [1, 2] and text exploration [8].

Basically, the term 'dialectical method' refers to a discussion among two or more people who have different points of view about a subject but are willing to find out the truth by argumentation. That is, in classical philosophy, dialectic is a method of reasoning based on arguments and counter-arguments [20, 22].

In ADFs, dialectical methods have a role in picking the truth-value of arguments under principles governed by several types of semantics, defined mainly based on three-valued interpretations, a form of labelings. Thus, in ADFs, beyond an argument being *acceptable* (the same as *defended* in AFs) there is a symmetric notion of *deniable*. One of the most common argumentation semantics are the *admissible* semantics, which in ADFs come in the form of interpretations that do not contain unjustifiable information. The other semantics of ADFs fulfil the

* Supported by the Center of Data Science & Systems Complexity (DSSC) Doctoral Programme, at the University of Groningen.

admissibility property. Maximal admissible interpretations are called *preferred* interpretations. Preferred semantics have a higher computational complexity than other semantics in ADFs [25]. That is, answering the decision problems of preferred semantics is more complicated than answering the same problem of other semantics in a given ADF. Therefore, having a structural discussion to investigate whether a decision problem is fulfilled under preferred semantics in a given ADF has a crucial importance.

There exists a number of works in which the relation between semantics of AFs and structural discussions are studied [9, 16, 17, 19, 23, 24]. As far as we know, the relation between semantics of ADFs and dialectical methods in the sense of discussion among agents has not been studied yet [3]. We aim to investigate whether semantics of ADFs are expressible in terms of discussion games.

In this paper we introduce the first existing discussion game for ADFs. We focus on preferred semantics and we show that for an argument being credulously accepted (denied) under preferred semantics in a given ADF there is a discussion game successfully defending the argument. Given the unique structure of ADFs, standard existing approaches known from the AFs setting could not be straightforwardly reused [11,12,27,28]. We thus propose a new approach based on interpretations that can be revised by evaluating the truth values of parents of the argument in question. The current methodology can be reused in other formalisms that can be represented in ADFs, such as AFs.

In the following, we first recall the relevant background of ADFs. Then, in Section 3, we present the *preferred discussion game*, which is a game with perfect information, that can capture the notion of preferred semantics. We show that there exists a proof strategy for arguments that are credulously acceptable (deniable) under preferred semantics in a given ADF and vice versa. Further, we show soundness and completeness of the method.

2 Background: Abstract Dialectical Frameworks

The basic definitions in this section are derived from those given in [5–7].

Definition 1. *An abstract dialectical framework (ADF) is a tuple* $F = (A, L, C)$ *where:*

- A *is a finite set of arguments (statements, positions);*
- $L \subseteq A \times A$ *is a set of links among arguments;*
- $C = \{\varphi_a\}_{a \in A}$ *is a collection of propositional formulas over arguments, called acceptance conditions.*

An ADF can be represented by a graph in which nodes indicate arguments and links show the relation among arguments. Each argument a in an ADF is attached by a propositional formula, called acceptance condition, φ_a over $par(a)$ such that, $par(a) = \{b \mid (b, a) \in R\}$. The acceptance condition of each argument clarifies under which condition the argument can be accepted [5–7]. Further, the acceptance conditions indicate the type of links. An *interpretation* v (for F) is a function $v : A \mapsto \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$, that maps arguments to one of the three truth values true ($\mathbf{t}$), false ($\mathbf{f}$), or undecided ($\mathbf{u}$). Truth values can be ordered via information ordering relation $<_i$ such that, $\mathbf{u} <_i \mathbf{t}$ and $\mathbf{u} <_i \mathbf{f}$ and no other pair of truth values are related by $<_i$. Relation $\leqslant_i$ is the reflexive and transitive closure of $<_i$. Interpretations can be ordered via $\leqslant_i$ with respect to their information content. It is said that an interpretation v is an *extension* of another interpretation w, if $w(a) \leqslant_i v(a)$ for each $a \in A$, denoted by $w \leqslant_i v$. Interpretations are incomparable if neither $w \nleqslant_i v$ nor $v \nleqslant_i w$, denoted by $w \not\sim v$.

Semantics for ADFs can be defined via the *characteristic operator* Γ_F which maps interpretations to interpretations. Given an interpretation v (for F), the partial valuation of φ_a by v, is $\varphi_a^v = \varphi_a[b/\top : v(b) = \mathbf{t}][b/\bot : v(b) = \mathbf{f}]$, for $b \in par(a)$. Applying Γ_F on v leads to v' such that for each $a \in A$, v' is as follows:

$$v'(a) = \begin{cases} \mathbf{t} & \text{if } \varphi_a^v \text{ is irrefutable (i.e., a tautology)}, \\ \mathbf{f} & \text{if } \varphi_a^v \text{ is unsatisfiable (i.e., } \varphi_a^v \text{ is a contradiction)}, \\ \mathbf{u} & \text{otherwise.} \end{cases}$$

From now on whenever there is no ambiguity, in order to make three-valued interpretations more readable, we rewrite them by the sequence of truth values, by choosing the lexicographic order on arguments. For instance, $v = \{a \mapsto \mathbf{t}, b \mapsto \mathbf{u}, c \mapsto \mathbf{f}\}$ can be represented by the sequence $\mathbf{tuf}$. The semantics of ADFs are defined via the characteristic operator as in Definition 2.

Definition 2. *Given an ADF F, an interpretation v is:*

- *admissible in F iff $v \leqslant_i \Gamma_F(v)$, denoted by adm;*
- *preferred in F iff v is $\leqslant_i$-maximal admissible, denoted by prf;*
- *a (two-valued) model of F iff v is two-valued and $\Gamma_F(v) = v$, denoted by mod.*

The notion of an argument being accepted and the symmetric notion of a argument being denied in an interpretation are as follows.

Definition 3. *Let $F = (A, L, C)$ be an ADF and let v be an interpretation of F.*

- *An argument $a \in A$ is called acceptable with respect to v if φ_a^v is irrefutable.*
- *An argument $a \in A$ is called deniable with respect to v if φ_a^v is unsatisfiable.*

One of the main decision problems of ADFs is whether an argument is credulously acceptable (deniable) under a particular semantics. Given an ADF $F = (A, L, C)$, an argument $a \in A$ and a semantics $\sigma \in \{adm, prf, mod\}$, argument a is *credulously acceptable (deniable)* under σ if there exists a σ interpretation v of F in which a is acceptable (a is deniable, respectively).

3 Discussion Game for Preferred Semantics

In this section, we present the structure of the discussion game for preferred semantics. The aim is to show that an argument is credulously accepted (denied) under preferred semantics in an ADF iff there exists a discussion game and a winning strategy for a player who starts the game.

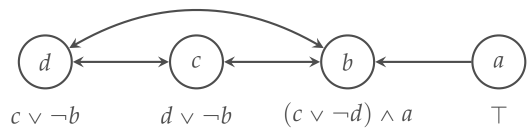

Fig. 1. ADF of Example 1

A *preferred discussion game*, which is similar to Socrates' form of reasoning [10, 29], is a (non-deterministic) two-player game of perfect information between defender (proponent) and challenger (opponent). So, all agents know all acceptance conditions. The game starts with a belief of proponent (P) about credulous acceptance (denial) of an argument under preferred semantics in a given ADF. Then opponent (O) challenges the proponent by investigating the consequences of P's belief and demanding reasons for those consequences. The game continues alternately: P has to convince O why consequences of the claim can be held. Till the time that there is a new claim by P or there is a new challenge by O and there is no contradiction, the game will be continued.

Since each preferred interpretation is an admissible interpretation, if we want to investigate whether an argument is credulously acceptable (deniable) under preferred semantics, we study whether the argument is credulously acceptable (deniable) under admissible semantics. The key advantage of the current method is that the credulous problem for preferred in an ADF F can be solved without enumeration of all admissible interpretations of F. In the following, Examples 1 and 2 represent preferred discussion games, in which there are winning strategies for P's belief.

Example 1. Given an ADF $F = (\{a, b, c, d\}, \{\varphi_a : \top, \varphi_b : (c \vee \neg d) \wedge a, \varphi_c : d \vee \neg b, \varphi_d : c \vee \neg b\})$, depicted in Figure 1.

- Assume that P claims that d is credulously acceptable under preferred semantics. The knowledge of P consists of information about the truth value of d, and there is no further information about the truth values of other arguments. This initial knowledge of P can be shown by the interpretation $v_0 = \mathbf{uuut}$.
- O checks the consequence of P's belief. O says that, based on the acceptance condition of d, argument d is acceptable in a preferred interpretation iff either c is accepted or b is denied in that interpretation. That is, O revises the information of v_0 to two interpretations; $v_1 = \mathbf{uutt}$ and $v_1' = \mathbf{ufut}$, and challenges P by asking, 'Why does either b have to be assigned to $\mathbf{f}$ or c have to be assigned to $\mathbf{t}$, if d is assigned to $\mathbf{t}$ in a preferred interpretation?'
- In both v_1 and v_1' there exists a new challenge, then the dialogue between players can be continued on any of them. P attempts to defeat the challenge by convincing O about the truth value of the arguments which are challenged by O in the preceding step.

  P chooses to work on v_1 in which the only new challenged argument is c. P checks under which condition c can be accepted in a preferred interpretation. Based on, $\varphi_c = d \vee \neg b$, c is assigned to $\mathbf{t}$ if and only if either d is assigned to $\mathbf{t}$ or b is assigned to $\mathbf{f}$. That is, the new information of P about the truth values of arguments can be represented by $v_2 = \mathbf{uutt}$ and $v'_2 = \mathbf{uftt}$. In the former one there is no new claim, that is, the dialogue v_0, v_1 and v_2 cannot be continued by O anymore. Further, in v_2 P answers the question of O (why is c assigned to $\mathbf{t}$), with no contradiction. Thus, P wins this dialogue. Since P can defend the initial claim via this dialogue, P wins the game and there is no need of continuing the game.

Definitions 4– 6 are needed to define the systematic method of computation of moves of each player in Definition 8. In the following, w and v are interpretations such that $w \leqslant_i v$.

Definition 4. *An argument a is **recently presented** in interpretation v with respect to w if $w(a) = \mathbf{u}$ and $v(a) \neq \mathbf{u}$.*

In contrast with standard interpretations in ADFs, in Definition 5 we define so-called minimal interpretations that give only value to argument a and $par(a)$. In the following the notations of $v(b)$ and $w_a(b)$ are used to indicate the truth value of argument b in v and w_a, respectively.

Definition 5. *Let v be an interpretation of an ADF F, in which $a \mapsto \mathbf{t}/\mathbf{f}$ and $par(a) \neq \emptyset$. An interpretation w_a, which is defined over $(par(a) \cup \{a\})$, is called a **minimal interpretation around** a **in** F, if $\Gamma_F(w_a)(a) = v(a)$, and there exists no $w' <_i w_a$ such that $\Gamma_F(w')(a) = v(a)$. In contrast, when $par(a) = \emptyset$ then w_a assigns a to $\Gamma_F(v)(a)$.*

Since the acceptance condition of each argument is indicated by a propositional formula, argument a may have more than one minimal interpretation around a in F. The set of all minimal interpretations around a in F is denoted by W_a.

Definition 6. *Let $A' = \{a_1, \ldots, a_n\}$ be the set of arguments recently presented in v w.r.t. w and choose $W_{A'} = \{w_{a_1}, \ldots, w_{a_n}\}$ s.t. $w_{a_i} \in W_{a_i}$, for $1 \leqslant i \leqslant n$. The output of the binary function $\delta(v, W_{A'})$ is called an **evaluation of the parents of arguments in** A' **w.r.t.** v **and** $W_{A'}$, defined as follows:*

$$\delta(v, W_{A'})(b) = \begin{cases} v(b) & \nexists i \text{ s.t. } b \mapsto \mathbf{t}/\mathbf{f} \in w_{a_i}, \\ \mathbf{u} & \exists i \text{ s.t. } b \mapsto \mathbf{t}/\mathbf{f} \in w_{a_i}, b \mapsto \mathbf{t}/\mathbf{f} \in v \text{ and } w_{a_i}(b) \neq v(b), \\ \mathbf{u} & \exists i, j \text{ s.t. } b \mapsto \mathbf{t} \in w_{a_i} \text{ and } b \mapsto \mathbf{f} \in w_{a_j}, \\ \mathbf{t} & \exists i \text{ s.t. } b \mapsto \mathbf{t} \in w_{a_i}, b \mapsto \mathbf{t}/\mathbf{u} \in v \text{ and } \nexists j \text{ s.t. } b \mapsto \mathbf{f} \in w_{a_j}, \\ \mathbf{f} & \exists i \text{ s.t. } b \mapsto \mathbf{f} \in w_{a_i}, b \mapsto \mathbf{f}/\mathbf{u} \in v \text{ and } \nexists j \text{ s.t. } b \mapsto \mathbf{t} \in w_{a_j}, \end{cases}$$

*The set of all possible evaluations of parents of arguments in A' is called **all evaluations of parents of** A', and denoted by $\delta_{A'}(v)$ such that:*

$$\delta_{A'}(v) = \{\delta(v, W_{A'}) \mid W_{A'} = \{w_{a_1}, \ldots, w_{a_n}\} \text{ s.t. } w_{a_i} \in W_{a_i} \text{ for } 1 \leqslant i \leqslant n\}$$

Note that when A' contains only one argument a, we address an evaluation of parents of a with $\delta(v, w_a)$ in which w_a is a minimal interpretation around a, and we denote the set of all evaluations of A' with $\delta_a(v)$.

In Example 1, it is assumed that d is credulously accepted, $v_0 = \mathbf{uuut}$. In comparison to interpretation $v_{\mathbf{u}} = \mathbf{uuuu}$, argument d is recently presented in v_0. Based on the acceptance condition of d, namely $\varphi_d : c \vee \neg b$, interpretations $w_d = \{b \mapsto \mathbf{u}, c \mapsto \mathbf{t}, d \mapsto \mathbf{t}\}$ and $w'_d = \{b \mapsto \mathbf{f}, c \mapsto \mathbf{u}, d \mapsto \mathbf{t}\}$ are minimal interpretations around d in F. As a consequence, the evaluation of the parents of the argument in question may lead to more than one interpretation. For instance, the evaluation of the parents of d with respect to v_0 and w_d is $\delta(v_0, w_d) = \mathbf{uutt}$, and with respect to v_0 and w'_d it is $\delta(v_0, w'_d) = \mathbf{ufut}$. Therefore, the set of evaluations of parents of d is $\delta_d(v_0) = \{\mathbf{uutt}, \mathbf{ufut}\}$.

Now we are going to define moves of each player based on the evaluation of the parents of the recently presented arguments, proposed in Definition 6. The information of each player in games can be represented by an interpretation. In the first claim of P there exists only information about the truth value of the argument which is claimed.

Definition 7. *The first claim of P about credulous acceptance (denial) of an argument is named **initial claim**, denoted by interpretation v_0, in which the argument in question is assigned to* **t** *(***f***, respectively) and all other argument are assigned to* **u**.

After each claim move of P, presented by interpretation v, O checks the conditions under which the claim of P can be valid. That is, O evaluates the truth values of the parents of arguments in A', recently presented by P in v with $\delta_{A'}(v)$. Then, O demands P to propose logical reasons for those results with the hope of leading to a contradiction. The game continues alternately: P has to convince O why at least one consequence of the claim can be held.

Definition 8. *Given interpretations v and w, such that $v \leqslant_i w$. Let A' be a set of arguments, recently presented in w. 1. If w is given by P, it is named that $a \in A'$ is **claimed** by P in w and $\delta_{A'}(w)$ is named **challenge move**. 2. If w is given by O, it is named that $a \in A'$ is **challenged** by O in w and $\delta_{A'}(w)$ is named **claim move**.*

Specifically, the initial claim is a claim move in comparison to the interpretation that assigns all arguments to **u**. Actually, a preferred discussion game can be represented as a labeled rooted tree in which the root is labeled by the initial claim, v_0. The nodes of depth $i > 0$ are labeled by all $\delta(v, W_{A'})$ such that v is the label of the directly preceding node of the tree with depth $i - 1$, and $W_{A'} = \{w_a \mid \text{ s.t. } a \in A'\}$ in which A' is a set of arguments that are recently presented in v with respect to the label of the directly preceding node of v. A part of the tree of Example 1, including a winning strategy for P, is depicted in Figure 2.

Definition 9. *A **dialogue** is the sequence of labels of a branch of the tree corresponding to the game which is started by an initial claim, and continued by applying $\delta(v_i, W_{A'})$, for $i \geqslant 0$ s.t. $a \in A'$ is recently presented in v_i.*

We say that there is a *contradiction* in a dialogue if the dialogue consists of interpretations v_i and v_{i+1} that are incomparable. For instance, the dialogue

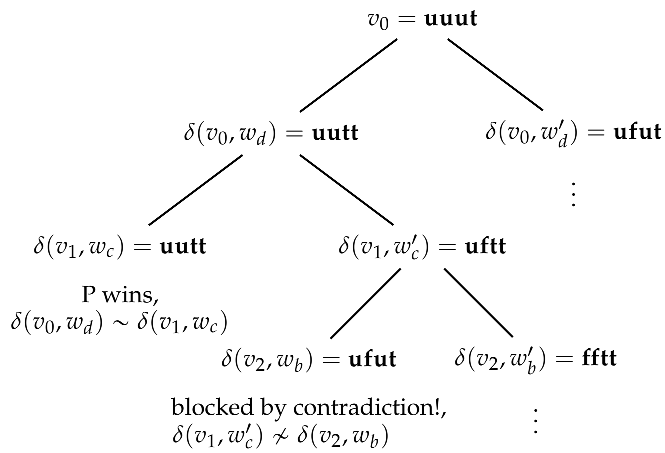

Fig. 2. Associated tree of the game in Example 1

$[v_0, \delta(v_0, w_d), \delta(v_1, w'_c), \delta(v_2, w_b)]$ in Fig. 2 leads to a contradiction. Definition 10–11 explain under which conditions a dialogue can be continued or halted.

Definition 10. *Let $[v_0, \ldots, v_n]$ be a dialogue with no contradiction.* ***The dialogue is continued*** *on v_n 1. by O if an argument is claimed in v_n by P; or 2. by P if an argument is challenged in v_n by O.*

Definition 11. *Let $[v_0, \ldots, v_n]$ be a dialogue. It is said that* ***the dialogue is blocked*** *on v_n when: 1. a is challenged in v_{n-1} by O, and $v_{n-1} \sim v_n$. We say that the game is blocked by P in this step. Or, 2. a is claimed in v_{n-1} by P, and $v_{n-1} \sim v_n$. We say that the game is blocked by O in this step. Or 3. there is a contradiction, that is, $v_{n-1} \nsim v_n$.*

In Example 1, dialogue $[v_0, \delta(v_0, w_{0d}), \delta(v_{1_0}, w_{0c})]$ is blocked by P. If a dialogue is blocked by P, it means that P could defeat a challenge of O without making a new claim. Thus, there is no further move for O. Therefore, P won the dialogue. Since P can defend the initial claim via this dialogue, P wins the game, as well. Thus, after this dialogue there is no need of continuing the game.

– *P wins the dialogue if the dialogue is blocked by P.*

Example 2 investigates the other condition under which P wins the dialogue.

Example 2. Let F be the ADF given in Example 1.

– P believes that d can be denied in a preferred interpretation in F, $v_0 = \mathbf{uuuf}$.
– The challenge move of O on d leads to $v_1 = \delta(v_0, w_d) = \mathbf{utff}$.
– The recently challenged arguments are b and c. The minimal interpretations around b are $w_b = \{a \mapsto \mathbf{t}, c \mapsto \mathbf{t}\}$ and $w'_b = \{a \mapsto \mathbf{t}, d \mapsto \mathbf{f}\}$, and the minimal interpretation around c is $w_c = \{b \mapsto \mathbf{t}, d \mapsto \mathbf{f}\}$. Thus, $v_2 = \delta(v_1, W_{bc}) = \mathbf{ttuf}$ and $v'_2 = \delta(v_1, W'_{bc}) = \mathbf{ttff}$.

– Since $v_1 \not\sim v_2$, O cannot continue this dialogue. However, $v_1 <_i v'_2$ and the challenge move on v'_2 is $\delta(v'_2, w_a) = v'_2$. Thus, the game is blocked by O.

If a dialogue is blocked by O, it means that O cannot find a contradiction between P's claim and O's challenging, which is done by O in an element of the claim move, and O cannot make a new challenge for P. Thus, P wins the dialogue and the initial claim of P is proved via this dialogue.

– *P wins the dialogue if the dialogue is blocked by O.*

The ADF of Example 1 can also be used as an example in which there is a winning strategy for O, explained in Example 3.

Example 3. Given ADF F of Example 1.

– P believes that b can be denied in a preferred interpretation in F, $v_0 = \mathbf{ufuu}$.
– There are three different dialogues based on this initial claim; 1. $[v_0 = \mathbf{ufuu}, v_1 = \mathbf{ufft}, v_2 = \mathbf{uufu}]$, 2. $[v_0 = \mathbf{ufuu}, v_1 = \mathbf{ufft}, v'_2 = \mathbf{uuuu}]$, 3. $[v_0 = \mathbf{ufuu}, v'_1 = \mathbf{ffuu}, v''_2 = \mathbf{ufuu}]$.

Each of the dialogues of this game is blocked by contradictions. That is, in each dialogue P cannot defeat the challenge of O. On the other hand, O defeats P in all the ways that P attempts to prove the initial claim, by finding contradictions. That is, P cannot make any reasonable discussion to defend the initial claim. Thus, O wins all dialogues and wins the game in consequence.

– *O wins the dialogue, when O can block the dialogue by contradiction.*

The examples which were studied above illustrate that each player only has to consider the arguments which are recently presented by the competitor in the directly preceding move. The discussion game that can decide credulous acceptance (denial) problem in ADFs under preferred semantics is called, *preferred discussion game*, introduced in Definition 12.

Definition 12. *Given an ADF $F = (A, L, C)$. A preferred discussion game for credulous acceptance (denial) of an argument of A is a sequence $[\Delta_0, \ldots, \Delta_n](n \geqslant 0)$ such that all the following conditions hold:*

– *Δ_0 consists of an initial claim;*
– *for $i \geqslant 1$, $\Delta_i = \bigcup_v \delta_{A'}(v)$, for each $v \in \Delta_{i-1}$ s.t. set of arguments of A' are recently presented in v;*
– *each $[v_0, \ldots, v_m]$ such that $v_i \in \Delta_i$ is a dialogue of the game, for $1 \leqslant m \leqslant n$, when: $v_i = \delta(v_{i-1}, W_{A'})$, s.t. the set A' is recently presented in v_{i-1};*
– *the game is finished in Δ_n if at least a dialogue of the game is blocked by P or O, or if all the dialogues lead to contradictions.*

In Definition 12, 1. if i is odd, for each $v \in \Delta_{i-1}$, Δ_i consists of all challenge moves $\delta_{A'}(v)$ such that $a \in A'$ is claimed in v; and 2. if $i \geqslant 2$ is even, for each $v \in \Delta_{i-1}$, Δ_i consists of all claim moves $\delta_{A'}(v)$ such that $a \in A'$ is challenged in v. The winning strategy of each player is explained in Definition 13.

Definition 13. *Let F be a given ADF. Let $[\Delta_0, \ldots, \Delta_n]$ be a preferred discussion game for credulous acceptance (denial) of an argument.*

- *P has a winning strategy in the game if P wins a dialogue of the game.*
- *O has a winning strategy in the game if O wins all dialogues of the game.*

Let F be an ADF and let $[\Delta_0, \ldots, \Delta_n]$ be a preferred discussion game of an initial claim of F. The *length* of the preferred discussion game is the length of the sequence $[\Delta_0, \ldots, \Delta_n]$, which is the number of elements of the sequence.

Proposition 1. *Let $F = (A, L, C)$ be an ADF and $|A| = n$. The length of each preferred discussion game of F is at most $n + 1$.*

Proof. Toward a contradiction, assume that that there exists a preferred discussion game $[\Delta_0, \ldots, \Delta_m]$ of F such that $m > n$. On the other hand, each dialogue $[v_0, \ldots, v_i]$ of the game is continued in v_i if $v_{i-1} <_i v_i$. This can be done by indicating the truth value of an argument in v_i that is not indicated before. Since the number of arguments of F is n, the longest dialogue contains interpretations such that $v_0 < \cdots < v_{n-1}$, and in the next step, the parents of arguments of claimed or challenged items in v_{n-1} will be evaluated. That is, the longest dialogue can be a sequence of $n + 1$ interpretations. Thus, the length of each game cannot be more that $n + 1$.

Since we assumed in the definition of ADFs that each ADF is finite, the immediate result of Proposition 1 is that each preferred discussion game halts and there exists a winning strategy either for O or P.

Theorem 1. *Let an ADF $F = (A, L, C)$ be given.*

- *Soundness: if there exists a winning strategy in a preferred discussion game with initial claim of accepting (denying) an argument a, then a is credulously acceptable (deniable) under preferred semantics in F.*
- *Completeness: if an argument a is credulously acceptable (deniable) under preferred semantics in F, then there is a preferred discussion game with a winning strategy for the initial claim of accepting (denying) of a.*

Proof. Soundness: assume that there is winning strategy for P in a preferred discussion game $[\Delta_0, \ldots, \Delta_n]$, for accepting (denying) of an argument a. Therefore, there is a winning dialogue $[v_0, \ldots, v_m]$ for P, for $0 < m \leqslant n$. To show the soundness it is enough to investigate whether v_m is an admissible interpretation. Towards a contradiction, assume that v_m is not an admissible interpretation, that is, $v_m \nleqslant_i \Gamma_F(v_m)$. Thus, there exists an argument b s.t. $b \mapsto \mathbf{t}/\mathbf{f} \in v_m$, however, the valuation of the acceptance condition of b under v_m is not the same as v_m; we prove the case that $b \mapsto \mathbf{t} \in v_m$.The proof method for the case in which $b \mapsto \mathbf{f} \in v_m$ is analogous.

$b \mapsto \mathbf{t} \in v_m$ means that either P claims this assignment in an interpretation v_i, $0 \leqslant i < m$, or O challenges it in an interpretation v_i, $0 < i < m$. Assume that this is claimed by P in v_i, $0 \leqslant i < m$. An element of the challenge move of O on v_i is v_{i+1}. That is, O presents the truth values of $par(b)$ in v_{i+1}. Since there is a winning strategy for P in this dialogue, $v_{m-1} \sim v_m$. That is, $\varphi_b^{v_m} \equiv \top$, since v_m consists of the truth values of $par(b)$ presented in v_{i+1}. Thus, $\Gamma_F(v_m)(b) = \mathbf{t}$. Therefore, the assumption that v_m is not an admissible interpretation is rejected. The proof method for a challenge move is analogous.

Completeness: assume that an argument a is credulously accepted under preferred semantics in F (the proof method in case a is credulously denied is analogous). Then, there is a preferred interpretation v of F in which a is accepted. We construct the corresponding preferred discussion game as follows. Let v_0, the initial claim, be an interpretation in which a is assigned to $\mathbf{t}$ and all other arguments of A are assigned to $\mathbf{u}$. Extend v_0 to v_1 by changing the truth values of the parents of a in v_0 by their truth values in v. Continue this method and construct v_{i+1} by changing the truth value of the parents of arguments which are recently presented in v_i, by the ones which are in v, for $i > 0$. Since the number of arguments is finite, this procedure will end in some v_n. To construct v_{i+1} only the truth values of the arguments which are assigned to $\mathbf{u}$ in v_i can be changed, then $v_i < v_{i+1}$, for $0 \leqslant i < n$. Let $v_{n+1} = v_n$. The sequence $[v_0, \dots, v_{n+1}]$ is a dialogue of the preferred discussion game $[\Delta_0, \dots, \Delta_{n+1}]$ of F, in which $v_0 \in \Delta_0$. Further, this dialogue is a winning strategy for P in this game.

4 Conclusion and Future Work

In this paper, preferred discussion games between two agents, proponent and opponent, are considered as a proof method to investigate credulous acceptance (denial) of arguments in an ADF under preferred semantics. Some notable results of the current work are: 1. The method is sound and complete. 2. The presented methodology can be reused in AFs and generalizations of AFs that can be represented as subclasses of ADFs, namely set argumentation frameworks [21] and bipolar argumentation framework [13]. 3. Winning *one* dialogue of the game by P is sufficient to show that there exists a preferred interpretation in which the argument in question is assigned to the truth value which is claimed. In contrast, for AFs [23, 26, 27], P has a winning strategy if P can address *all* O's challenges. 4. In each move each player has to study the truth value of arguments that are recently presented in the directly preceding move. In contrast, in [9], O has to check all past moves of P to find a contradiction. 5. To investigate the credulous decision problem of ADFs under preferred semantics, there is no need to enumerate all preferred interpretations of an ADF. 6. Preferred semantics of an ADF corresponds to a preferred discussion game with winning strategy for P. 7. In [14] it is shown that in the class of acyclie ADFs all semantics coincide. Thus, in acyclic ADFs the presented game can be used to decide the credulous problem on other semantics. As future work, we could investigate structural discussion games for other semantics of ADFs. In addition, we could study discussion games for other decision problems of ADFs. Further, we could investigate whether the presented method is more effective than the methods used in current ADF-solvers, e.g. [4, 18]. This study may lead to new ADF-solvers that work locally on an argument to answer decision problems.

References

1. Al-Abdulkarim, L., Atkinson, K., Bench-Capon, T.J.M.: Abstract dialectical frameworks for legal reasoning. In: Legal Knowledge and Information Systems JURIX. Frontiers in Artificial Intelligence and Applications, vol. 271, pp. 61–70. IOS Press (2014)
2. Al-Abdulkarim, L., Atkinson, K., Bench-Capon, T.J.M.: A methodology for designing systems to reason with legal cases using abstract dialectical frameworks. Artificial Intelligence and Law **24**(1), 1–49 (2016)
3. Barth, E.M., Krabbe, E.C.: From Axiom to Dialogue: A Philosophical Study of Logics and Argumentation. Walter de Gruyter, Berlin-New York (1982)
4. Brewka, G., Diller, M., Heissenberger, G., Linsbichler, T., Woltran, S.: Solving advanced argumentation problems with answer-set programming. In: Conference on Artificial Intelligence AAAI. pp. 1077–1083. AAAI Press (2017)
5. Brewka, G., Ellmauthaler, S., Strass, H., Wallner, J.P., Woltran, S.: Abstract dialectical frameworks. An overview. IFCoLog Journal of Logics and their Applications (FLAP) **4**(8) (2017)
6. Brewka, G., Strass, H., Ellmauthaler, S., Wallner, J.P., Woltran, S.: Abstract dialectical frameworks revisited. In: Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence (IJCAI 2013). pp. 803–809 (2013)
7. Brewka, G., Woltran, S.: Abstract dialectical frameworks. In: Proceedings of the Twelfth International Conference on the Principles of Knowledge Representation and Reasoning (KR 2010). pp. 102–111 (2010)
8. Cabrio, E., Villata, S.: Abstract dialectical frameworks for text exploration. In: Proceedings of the 8th International Conference on Agents and Artificial Intelligence (ICAART (2). pp. 85–95. SciTePress (2016)
9. Caminada, M.: Argumentation semantics as formal discussion. Handbook of Formal Argumentation **1**, 487–518 (2017)
10. Caminada, M.W.: A formal account of Socratic-style argumentation. Journal of Applied Logic **6**(1), 109–132 (2008)
11. Caminada, M.W., Dvořák, W., Vesic, S.: Preferred semantics as Socratic discussion. Journal of Logic and Computation **26**(4), 1257–1292 (2014)
12. Cayrol, C., Doutre, S., Mengin, J.: On decision problems related to the preferred semantics for argumentation frameworks. Journal of Logic and Computation **13**(3), 377–403 (2003)
13. Cayrol, C., Lagasquie-Schiex, M.: Bipolarity in argumentation graphs: Towards a better understanding. International Journal of Approximate Reasoning **54**(7), 876–899 (2013)
14. Diller, M., Keshavarzi Zafarghandi, A., Linsbichler, T., Woltran, S.: Investigating subclasses of abstract dialectical frameworks. In: Proceedings of Computational Models of Argument COMMA 2018. pp. 61–72. Amsterdam (2018)
15. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence **77**, 321–357 (1995)
16. Dung, P.M., Thang, P.M.: A sound and complete dialectical proof procedure for sceptical preferred argumentation. In: Proc. of the LPNMR-Workshop on Argumentation and Nonmonotonic Reasoning (ArgNMR07). pp. 49–63 (2007)
17. van Eemeren, F.H., Garssen, B., Krabbe, E.C.W., Henkemans, A.F.S., Verheij, B., Wagemans, J.H.M. (eds.): Handbook of Argumentation Theory. Springer, New York (2014)

18. Ellmauthaler, S., Strass, H.: The DIAMOND system for computing with abstract dialectical frameworks. In: Proceedings of Computational Models of Argument COMMA. vol. 266, pp. 233–240. IOS Press (2014)
19. Jakobovits, H., Vermeir, D.: Dialectic semantics for argumentation frameworks. In: Proceedings of the 7th International Conference on Artificial Intelligence and Law. pp. 53–62. ACM Press (1999)
20. Krabbe, E.C.: Dialogue logic. In: Handbook of the History of Logic, vol. 7, pp. 665–704. Elsevier (2006)
21. Linsbichler, T., Pührer, J., Strass, H.: A uniform account of realizability in abstract argumentation. In: 22nd European Conference on Artificial Intelligence ECAI. vol. 285, pp. 252–260. IOS Press (2016)
22. Macoubrie, J.: Logical argument structures in decision-making. Argumentation **17**(3), 291–313 (2003)
23. Modgil, S., Caminada, M.: Proof theories and algorithms for abstract argumentation frameworks. In: Argumentation in Artificial Intelligence, pp. 105–129. Springer (2009)
24. Prakken, H., Sartor, G.: Argument-based extended logic programming with defeasible priorities. Journal of Applied Non-classical Logics **7**(1-2), 25–75 (1997)
25. Strass, H., Wallner, J.P.: Analyzing the computational complexity of abstract dialectical frameworks via approximation fixpoint theory. Artificial Intelligence **226**, 34–74 (2015)
26. Thang, P.M., Dung, P.M., Hung, N.D.: Towards a common framework for dialectical proof procedures in abstract argumentation. Journal of Logic and Computation **19**(6), 1071–1109 (2009)
27. Verheij, B.: A labeling approach to the computation of credulous acceptance in argumentation. In: Proceedings of the 20th International Joint Conference on Artificial Intelligence IJCAI. pp. 623–628. IJCAI (2007)
28. Vreeswijk, G., Prakken, H.: Credulous and sceptical argument games for preferred semantics. In: Proceeding of the 7th European Workshop on Logic for Artificial Intelligence JELIA. vol. 1919, pp. 239–253. Springer (2000)
29. Walton, D., Krabbe, E.: Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning. State University of New York Press (1995)