A Dialectical Approach for Argument-Based Judgment Aggregation

Martin CAMINADA and Richard BOOTH
Cardiff University, UK

Abstract. The current paper provides a dialectical interpretation of the argumentation-based judgment aggregation operators of Caminada and Pigozzi. In particular, we define discussion-based proof procedures for the foundational concepts of *down-admissible* and *up-complete*. We then show how these proof procedures can be used as the basis of dialectical proof procedures for the *sceptical*, *credulous* and *super credulous* judgment aggregation operators.

Keywords. judgment aggregation, proof procedures, discussion games

1. Introduction

Given an argumentation framework, there can be more than one reasonable position on which arguments to accept and which arguments to reject [4], and different agents can take different positions. How to aggregate the agents' individual positions to form a group position has been studied by Caminada and Pigozzi [11]. For this, three different operators have been formulated: the *sceptical operator*, the *credulous operator* and the *super credulous operator*. These operators are such that, when each individual position is an admissible labelling, the collective outcome will also be an admissible labelling.[1]

Various follow-up research has been done based on the work of Caminada and Pigozzi. Podlaszewski [19], Caminada et al. [12] and Awad et al. [1,3] have examined issues of Pareto optimality and strategy proofness of the three judgment aggregation operators. Awad et al. [1,2] have examined the empirical acceptance of their outcomes, and Booth et al. [6] have recently provided a generalised theory and have shown how the operators of Caminada and Pigozzi fit in.

In the current paper, we examine how the three judgment aggregation operators of Caminada and Pigozzi can be given a dialectical interpretation. This is in line with recent work on argumentation-based discussion games [17,13,14,8,9,15]. However, instead of applying discussion games as proof procedures for argumentation semantics, we apply discussion games as proof procedures for the judgment aggregation operators. That is, we introduce argument games for the sceptical,

[1] In some cases, stronger results apply. For instance, when each agent's position is a complete labelling, applying the sceptical operator will yield a complete labelling. We refer to [11] for details.

credulous and super credulous operator, such that the ability to win the game coincides with the argument being accepted by the respective judgment aggregation operator. This is done by defining discussion games for two fundamental concepts used by the judgment aggregation operators: the *down-admissible* and *up-complete* labellings.

The remaining part of this paper is structured as follows. First, in Section 2 we briefly revisit Caminada and Pigozzi's work on argumentation based judgment aggregation. Then, in Section 3 we introduce the down-admissible game, as well as the discussion games for the sceptical and credulous operators based on it. In Section 4 we then introduce the up-complete game, as well as the discussion game for the super credulous operator based on it. We then round off with a discussion of the obtained results in Section 5.

2. Formal Preliminaries

For current purposes, we apply the labelling-based version of argumentation semantics [7,10,4]. In line with [11], we restrict ourselves to finite argumentation frameworks.

Definition 1. *An* argumentation framework *is a pair* $(Ar, \rightharpoonup)$ *where* Ar *is a finite set of arguments*[2] *and* $\rightharpoonup \subseteq Ar \times Ar$.

Definition 2. *Let* $(Ar, \rightharpoonup)$ *be an argumentation framework. A* labelling *is a total function* $\mathcal{L}ab : Ar \rightarrow \{\mathtt{in}, \mathtt{out}, \mathtt{undec}\}$. *We write* $\mathtt{in}(\mathcal{L}ab)$ *for* $\{A \mid \mathcal{L}ab(A) = \mathtt{in}\}$, $\mathtt{out}(\mathcal{L}ab)$ *for* $\{A \mid \mathcal{L}ab(A) = \mathtt{out}\}$ *and* $\mathtt{undec}(\mathcal{L}ab)$ *for* $\{A \mid \mathcal{L}ab(A) = \mathtt{undec}\}$. *We define a relation* $\sqsubseteq$ *between labellings s.t.* $\mathcal{L}ab_1 \sqsubseteq \mathcal{L}ab_2$ *iff* $\mathtt{in}(\mathcal{L}ab_1) \subseteq \mathtt{in}(\mathcal{L}ab_2)$ *and* $\mathtt{out}(\mathcal{L}ab_1) \subseteq \mathtt{out}(\mathcal{L}ab_2)$. *We define a function* Γ *such that* $\Gamma(\mathcal{L}ab)$ *is a labelling with* $\mathtt{in}(\Gamma(\mathcal{L}ab)) = \{A \mid \mathcal{L}ab(B) = \mathtt{out}$ *for each* $B \rightharpoonup A\}$ *and* $\mathtt{out}(\Gamma(\mathcal{L}ab)) = \{A \mid \mathcal{L}ab(B) = \mathtt{in}$ *for some* $B \rightharpoonup A\}$. *A labelling* $\mathcal{L}ab$ *is called* admissible *iff* $\mathcal{L}ab \sqsubseteq \Gamma(\mathcal{L}ab)$. *A labelling is called* complete *iff* $\mathcal{L}ab = \Gamma(\mathcal{L}ab)$.

We will sometimes write a labelling as a triple $(\mathtt{in}(\mathcal{L}ab), \mathtt{out}(\mathcal{L}ab), \mathtt{undec}(\mathcal{L}ab))$. We proceed to define the concepts of down-admissible and up-complete.

Definition 3 ([11]). *Let* $\mathcal{L}ab$ *be a labelling of argumentation framework* $(Ar, \rightharpoonup)$. *The down-admissible labelling of* $\mathcal{L}ab$ *(written as* $\downharpoonleft \mathcal{L}ab$*) is the unique biggest (w.r.t.* $\sqsubseteq$*) admissible labelling of* $(Ar, \rightharpoonup)$ *that is smaller or equal (w.r.t.* $\sqsubseteq$*) to* $\mathcal{L}ab$.

Definition 4 ([11]). *Let* $\mathcal{L}ab$ *be an admissible labelling of argumentation framework* $(Ar, \rightharpoonup)$. *The up-complete labelling of* $\mathcal{L}ab$ *(written as* $\upharpoonleft \mathcal{L}ab$*) is the unique smallest (w.r.t.* $\sqsubseteq$*) complete labelling that is bigger or equal (w.r.t.* $\sqsubseteq$*) to* $\mathcal{L}ab$.

[2] For current purposes, we keep the internal structure of the arguments abstract, although we emphasize that our theory is compatible with instantiated argumentation theories like ASPIC+ [18], ABA [20] and logic-based argumentation [16].

Definition 5 ([11]). *Given labellings $\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n$ ($n \geqslant 1$) of argumentation framework $(Ar, \rightharpoonup)$, we define $\sqcap(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n)$ as the labelling $(\{A \mid \forall_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) = \texttt{in}\}, \{A \mid \forall_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) = \texttt{out}\}, \{A \mid \exists_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) \neq \texttt{in} \wedge \exists_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) \neq \texttt{out}\})$ and $\sqcup(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n)$ as the labelling $(\{A \mid \exists_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) = \texttt{in} \wedge \neg\exists_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) = \texttt{out}\}, \{A \mid \exists_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) = \texttt{out} \wedge \neg\exists_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) = \texttt{in}\}, \{A \mid \forall_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) = \texttt{undec} \vee (\exists_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) = \texttt{in} \wedge \exists_{i\in\{1\ldots n\}}\mathcal{L}ab_i(A) = \texttt{out})\})$.*[3]

Given the above defined concepts, we proceed to formally state the three judgment aggregation operators of [11].

Definition 6 ([11]). *Let $\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n$ ($n \geqslant 1$) be admissible labellings of argumentation framework $(Ar, \rightharpoonup)$. We define:*

- *the* sceptical outcome *as* $\downharpoonright \sqcap(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n)$
- *the* credulous outcome *as* $\downharpoonright \sqcup(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n)$
- *the* super credulous outcome *as* $\upharpoonleft\!\downharpoonright \sqcup(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n)$

We sometimes refer to $\downharpoonright \sqcap$ as the sceptical operator, $\downharpoonright \sqcup$ as the credulous operator and $\upharpoonleft\!\downharpoonright \sqcup$ as the super credulous operator. We refer to [11,12,19,1,3] for the formal properties of these operators.

3. Dialectical Proof Procedures for the Sceptical and Credulous Operators

Using the formal preliminaries stated above, we now turn our attention to specifying dialectical proof procedures for the three judgment aggregation operators. We start with the sceptical and credulous operators. As these are both based on the down-admissible labelling, we first define a discussion game for the down-admissible, based on the admissible game (for preferred semantics) of [15,9].

Definition 7. *Let $\mathcal{L}ab$ be a labelling of argumentation framework $(Ar, \rightharpoonup)$. A* down-admissible discussion *for $A \in Ar$ in $\mathcal{L}ab$ is a sequence of moves $[M_1, \ldots, M_m]$ ($m \geqslant 1$) such that:*

- *$M_1 = \texttt{in}(A)$*
- *each move M_j ($1 \leqslant j \leqslant m$) where j is odd (called a* proponent move*) is of the form $\texttt{in}(B)$ with $B \in Ar$*
- *each move M_j ($1 \leqslant j \leqslant m$) where j is even (called an* opponent move*) is of the form $\texttt{out}(B)$ with $B \in Ar$*
- *for each opponent move $M_j = \texttt{out}(B)$ ($2 \leqslant j \leqslant m$) there exists a proponent move $M_k = \texttt{in}(C)$ ($k < j$) such that $B \rightharpoonup C$*
- *for each proponent move $M_j = \texttt{in}(B)$ except the first one ($3 \leqslant j \leqslant m$) it holds that M_{j-1} is of the form $\texttt{out}(C)$ such that $B \rightharpoonup C$*
- *there exist no two opponent moves M_j and M_k ($j \neq k$) such that $M_j = M_k$*

A down-admissible discussion $[M_1, \ldots, M_m]$ is called terminated *iff*

[3] $\sqcap(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n)$ is called the *sceptical initial labelling* in [11] and $\sqcup(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n)$ is called the *credulous initial labelling* in [11].

1. *there exists no* M_{m+1} *such that* $[M_1, \ldots, M_m, M_{m+1}]$ *is a down-admissible discussion, or*
2. *there exists a proponent move* $\mathtt{in}(B)$ *and an opponent move* $\mathtt{out}(B)$ *for the same argument* B*, or*
3. *there exists a proponent move* $\mathtt{in}(B)$ *s.t.* $\mathcal{L}ab(B) \neq \mathtt{in}$*, or*
4. *there exists an opponent move* $\mathtt{out}(B)$ *s.t.* $\mathcal{L}ab(B) \neq \mathtt{out}$

and no subsequence $[M_1, \ldots M_l]$ *(*$l \leqslant m$*) is terminated. A terminated down-admissible discussion is won by the opponent if*

1. *there exists no* M_{m+1} *such that* $[M_1, \ldots, M_m, M_{m+1}]$ *is a down-admissible discussion and* M_m *is an opponent move, or*
2. *there exists a proponent move* $\mathtt{in}(B)$ *and an opponent move* $\mathtt{out}(B)$ *for the same argument* B*, or*
3. *there exists a proponent move* $\mathtt{in}(B)$ *s.t.* $\mathcal{L}ab(B) \neq \mathtt{in}$*, or*
4. *there exists an opponent move* $\mathtt{out}(B)$ *s.t.* $\mathcal{L}ab(B) \neq \mathtt{out}$

Otherwise, the terminated down-admissible discussion is won by the proponent.

We observe that the above discussion game is essentially the admissibility game of [15,9], with additional clauses 3 and 4 in both the termination criterion and the winning criterion. These additional clauses essentially state that for the proponent to win, the game has to stay "inside" the initial labelling $\mathcal{L}ab$.

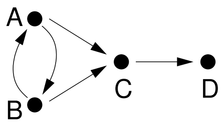

Figure 1. An argumentation framework

As an example of how the down-admissible discussion game works, consider the argumentation framework of Figure 1 and labelling $\mathcal{L}ab = (\{D\}, \{C\}, \{A, B\})$. Here, the discussion $[\mathtt{in}(D), \mathtt{out}(C), \mathtt{in}(A)]$ is terminated and won by the opponent (since $\mathcal{L}ab(A) \neq \mathtt{in}$), as is the discussion $[\mathtt{in}(D), \mathtt{out}(C), \mathtt{in}(B)]$ (since $\mathcal{L}ab(B) \neq \mathtt{in}$). We observe that $D \notin \mathtt{in}(\downharpoonleft \mathcal{L}ab)$ as $\downharpoonleft \mathcal{L}ab$ is the all-$\mathtt{undec}$ labelling $(\varnothing, \varnothing, \{A, B, C, D\})$.

We are now ready to formally state soundness and completeness of the down-admissible discussion game.

Theorem 1. *Let* $\mathcal{L}ab$ *be a labelling of argumentation framework* $(Ar, \rightharpoonup)$ *and let* $A \in Ar$*.* A *is labelled* $\mathtt{in}$ *by* $\downharpoonleft \mathcal{L}ab$ *iff the proponent has a winning strategy*[4] *in the down-admissible discussion game for* A *in* $\mathcal{L}ab$*.*

Proof. "$\Leftarrow$": Suppose the proponent has a winning strategy for A in $\mathcal{L}ab$. Then, from the definition of a winning strategy, it follows that there exists at least one discussion game for A in $\mathcal{L}ab$ that is won by the proponent. Now consider the labelling $\mathcal{L}ab'$ with $\mathtt{in}(\mathcal{L}ab')$ consisting of all proponent moves and $\mathtt{out}(\mathcal{L}ab')$

[4]We use the term winning strategy in the sense of [9].

consisting of all opponent moves of this discussion.[5] From the fact that the discussion is won by the proponent (together with winning condition 2) it follows that there is no argument B with both $B \in \texttt{in}(\mathcal{L}ab')$ and $B \in \texttt{out}(\mathcal{L}ab')$. This means that $\mathcal{L}ab'$ is a well-defined argument labelling. From the fact that the discussion is won by the proponent, it also follows (termination condition 1 and winning condition 1) that the last move (M_m) is a proponent move. This means that each opponent move in the discussion has been replied to. That is, for each opponent move $\texttt{out}(B)$ there exists a proponent move $\texttt{in}(B)$. Hence we obtain that (i) for each $B \in \texttt{out}(\mathcal{L}ab')$ there exists a $C \in \texttt{in}(\mathcal{L}ab')$ such that C attacks B. From the fact that the discussion is terminated with the last move (M_m) being a proponent move, it also follows that the opponent cannot make a move M_{m+1} anymore. This means that there is no attacker to any of the proponent's moves that hasn't already been moved. This implies that (ii) for each $B \in \texttt{in}(\mathcal{L}ab')$ it holds that each attacker C of B has $C \in \texttt{out}(\mathcal{L}ab')$. From conditions (i) and (ii) it follows that $\mathcal{L}ab'$ is an admissible labelling. From winning conditions 3 and 4 it follows that $\texttt{in}(\mathcal{L}ab') \subseteq \texttt{in}(\mathcal{L}ab)$ and $\texttt{out}(\mathcal{L}ab') \subseteq \texttt{out}(\mathcal{L}ab)$. That is, $\mathcal{L}ab'$ is an admissible labelling with $\mathcal{L}ab' \sqsubseteq \mathcal{L}ab$. As the down-admissible labelling $\downharpoonleft \mathcal{L}ab$ is the unique biggest (w.r.t. $\sqsubseteq$) admissible labelling with $\downharpoonleft \mathcal{L}ab \sqsubseteq \mathcal{L}ab$ it follows that $\mathcal{L}ab' \sqsubseteq \downharpoonleft \mathcal{L}ab$. From the fact that $A \in \texttt{in}(\mathcal{L}ab')$ it directly follows that $A \in \texttt{in}(\downharpoonleft \mathcal{L}ab)$.
"$\Rightarrow$": Let A be labelled in by $\downharpoonleft \mathcal{L}ab$. Now consider a discussion that starts with the proponent moving $\texttt{in}(A)$. As long as each proponent move is labelled in by $\downharpoonleft \mathcal{L}ab$ (as is the case with the first move) the opponent can only move arguments that are labelled out by $\downharpoonleft \mathcal{L}ab$ (this is because $\downharpoonleft \mathcal{L}ab$ is an admissible labelling). Moreover, when each opponent move is labelled out by $\downharpoonleft \mathcal{L}ab$, it is always possible for the proponent to reply with an argument that is labelled in by $\downharpoonleft \mathcal{L}ab$ (this is again because $\downharpoonleft \mathcal{L}ab$ is an admissible labelling). Suppose the proponent follows such a strategy (of choosing only moves that are labelled in by $\downharpoonleft \mathcal{L}ab$). As the opponent cannot repeat his moves, the discussion will terminate in a finite number of steps. Termination cannot be due to termination condition 2 (since the fact that $\downharpoonleft \mathcal{L}ab$ is a well defined labelling implies that there exists no argument B with both $B \in \texttt{in}(\downharpoonleft \mathcal{L}ab)$ and $B \in \texttt{out}(\downharpoonleft \mathcal{L}ab)$). Also, termination cannot be due to termination conditions 3 or 4, as the proponent's strategy ensures that (as we have observed) for each proponent move $\texttt{in}(B)$ it holds that $B \in \texttt{in}(\downharpoonleft \mathcal{L}ab)$ and for each opponent move $\texttt{out}(B)$ it holds that $B \in \texttt{out}(\downharpoonleft \mathcal{L}ab)$. This means that termination must be due to termination condition 1 (meaning no next move is possible). As the proponent's strategy ensures that the proponent can *always* move after an opponent move (as the fact that $\downharpoonleft \mathcal{L}ab$ is an admissible labelling means that for each $B \in \texttt{out}(\downharpoonleft \mathcal{L}ab)$ there exists a $C \in \texttt{in}(\downharpoonleft \mathcal{L}ab)$ such that C attacks B) this means that the last move cannot be an opponent move. That is, winning condition 1 cannot be applicable (nor can winning conditions 2, 3 and 4 be applicable). It then directly follows that the proponent wins the discussion. □

Using the down-admissible game, it becomes fairly straightforward to define a dialectical proof procedure for the sceptical and credulous operators. The eas-

[5] $\texttt{undec}(\mathcal{L}ab')$ then consists of all arguments that are neither proponent moves nor opponent moves.

iest way would be simply to start with $\mathcal{L}ab$ being either $\sqcap(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n)$ or $\sqcup(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n)$. Alternatively, it would be possible to define separate dialectical proof procedures for the sceptical and credulous operators, by slightly changing the rules of the down-admissible discussion game. For the sceptical game, we need to change clauses 3 and 4 regarding the termination and winning criterion to:

$3'$ there exists a proponent move $\texttt{in}(B)$ s.t. $\mathcal{L}ab_i(B) \neq \texttt{in}$ for some $i \in \{1 \ldots n\}$

$4'$ there exists an opponent move $\texttt{out}(B)$ s.t. $\mathcal{L}ab_i(B) \neq \texttt{out}$ for some $i \in \{1 \ldots n\}$

As an example of how the sceptical discussion game works, consider the argumentation framework of Figure 1 and labellings $\mathcal{L}ab_1 = (\{A, D\}, \{B, C\}, \emptyset)$ (of agent 1) and $\mathcal{L}ab_2 = (\{B, D\}, \{A, C\}, \emptyset)$ (of agent 2).

Proponent:	"We can all agree that D has to be accepted ($\texttt{in}(D)$)"
Opponent:	"But then we'd also all have to agree that D's attacker C has to be rejected ($\texttt{out}(C)$). Based on what grounds?"
Proponent:	"We can all agree that C has to be rejected because we can all agree that A has to be accepted ($\texttt{in}(A)$)"
Agent 2:	"Objection! I don't accept A myself."[6]

As the sceptical game is essentially the admissibility game of [15,9] with extra conditions $3'$ and $4'$, we can think of the sceptical game as the standard admissibility game with a twist: apart from participants proponent and opponent, there is now also a room full of potential hecklers (the agents whose labellings are being aggregated). If the discussion between the proponent and opponent touches an argument of which one of the agents in the room does not agree on its label, the agent shouts "Objection!" in which case the discussion ends and the proponent loses (regardless of whether it was a proponent or opponent move that was being objected to).

The discussion game for the credulous operator can be defined in a similar way. Again, the easiest way would be to simply start with $\mathcal{L}ab$ being $\sqcup(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n)$. Alternatively, rules 3 and 4 regarding the termination and winning criterion should be changed as follows:

$3''$ there exists a proponent move $\texttt{in}(B)$ s.t. $\mathcal{L}ab_i(B) = \texttt{out}$ for some $i \in \{1 \ldots n\}$ or $\mathcal{L}ab_i(B) \neq \texttt{in}$ for each $i \in \{1 \ldots n\}$

$4''$ there exists an opponent move $\texttt{out}(B)$ s.t. $\mathcal{L}ab_i(B) = \texttt{in}$ for some $i \in \{1 \ldots n\}$ or $\mathcal{L}ab_i(B) \neq \texttt{out}$ for each $i \in \{1 \ldots n\}$

As an example of how the credulous discussion game works, consider again the argumentation framework of Figure 1 and labellings $\mathcal{L}ab_1 = (\{A, D\}, \{B, C\}, \emptyset)$

[6] For the sake of the example, we have modelled the objection as a separate move that indicates termination condition $3'$.

(of agent 1) and $\mathcal{L}ab_2 = (\{B, D\}, \{A, C\}, \varnothing)$ (of agent 2).

Proponent:	"We can all agree that D has to be accepted ($\texttt{in}(D)$)"
Room:	"Aye" (Agent 1) "Aye" (Agent 2)
Opponent:	"But then we'd also all have to agree that D's attacker C has to be rejected ($\texttt{out}(C)$). Based on what grounds?"
Room:	"Aye" (Agent 1) "Aye" (Agent 2)
Proponent:	"We can all agree that C has to be rejected because we can all agree that A has to be accepted ($\texttt{in}(A)$)"
Room:	"Aye" (Agent 1) "Nay" (Agent 2)

As the credulous game is essentially the admissibility game of [15,9] with extra conditions $3''$ and $4''$, we can think of the credulous game as the standard admissibility game with a twist: after each move of the proponent and opponent, the agents in the room are asked for their opinion. Agents who agree with the label of the argument shout "Aye". Agents who have the opposite label[7] shout "Nay".[8] If there is at least one agent that shouts "Aye" and no agent that shouts "Nay" then the discussion continues. However, if there is no agent shouting "Aye" or at least one agent that shouts "Nay" then the discussion is terminated and the proponent loses (regardless of whether it was a proponent or opponent move that caused it).

4. A Dialectical Proof Procedure for the Super Credulous Operator

As the super credulous operator is based on the up-complete labelling, we first define an up-complete discussion game, based on the Grounded Discussion Game [8].

The Grounded Discussion Game is a sound and complete dialectical proof procedure to determine whether an argument is in the grounded extension.[9] It is based on a discussion between two participants (proponent and opponent) who use the following four kind of utterances.

$HTB(A)$ ("A has to be the case")
: With this move, the proponent claims that A has to be labelled `in`.

$CB(B)$ ("B can be the case, or at least cannot be ruled out")
: With this move, the opponent claims that B does not have to be labelled `out`.

$CONCEDE(A)$ ("I agree that A has to be the case")
: With this move, the opponent indicates that he now agrees with the proponent (who previously did a $HTB(A)$ move) that A has to be labelled `in`.

[7] with `in` being the opposite of `out`, and `out` being the opposite of `in`

[8] Our naming convention is inspired by the British parliament, where a similar procedure is used before holding a physical vote.

[9] The Grounded Discussion Game has a number of advantages compared to alternative dialectical proof procedures for grounded semantics like the Standard Grounded Game [17] and the Grounded Persuasion Game [13]. We refer to [8] for details.

RETRACT(B) ("I give up that B can be the case")

With this move, the opponent indicates that he no longer believes that B can be `in` or `undec`. That is, the opponent acknowledges that B has to be labelled `out`.

One of the key ideas of the game is that the proponent has burden of proof. That is, the proponent has to establish the acceptance of the main argument and make sure that the discussion does not go around in circles (meaning that arguments are not mentioned more than once).

Using the four moves of the Grounded Discussion Game, we proceed to define the up-complete discussion game.

Definition 8. *Let* $(Ar, \rightharpoonup)$ *be an argumentation framework. An* up-complete discussion game *is a sequence of discussion moves constructed by applying the following principles.*

BASIS (*HTB*) *If* $A \in Ar$ *then* $[HTB(A)]$ *is an up-complete discussion.*

STEP (*HTB*) *If* $[M_1, \dots, M_n]$ $(n \geqslant 1)$ *is an up-complete discussion without HTB-CB repeats,*[10] *and no CONCEDE or RETRACT move is applicable, and* $M_n = CB(A)$ *and* B *is an attacker of* A *then* $[M_1, \dots, M_n, HTB(B)]$ *is also an up-complete discussion.*

STEP (*CB*) *If* $[M_1, \dots, M_n]$ $(n \geqslant 1)$ *is an up-complete discussion without HTB-CB repeats, and no CONCEDE or RETRACT move is applicable, and* M_n *is not a CB move, and there is a move* $M_i = HTB(A)$ $(i \in \{1 \dots n\})$ *such that the discussion does not contain* $CONCEDE(A)$*, and for each move* $M_j = HTB(A')$ $(j > i)$ *the discussion contains a move* $CONCEDE(A')$*, and* B *is an attacker of* A *such that the discussion does not contain a move* $RETRACT(B)$*, then* $[M_1, \dots, M_n, CB(B)]$ *is an up-complete discussion.*

STEP (*CONCEDE*) *If* $[M_1, \dots, M_n]$ $(n \geqslant 1)$ *is an up-complete discussion without HTB-CB repeats, and* $CONCEDE(B)$ *is applicable then* $[M_1, \dots, M_n, CONCEDE(B)]$ *is an up-complete discussion.*

STEP (*RETRACT*) *If* $[M_1, \dots, M_n]$ $(n \geqslant 1)$ *is an up-complete discussion without HTB-CB repeats, and* $RETRACT(B)$ *is applicable then* $[M_1, \dots, M_n, RETRACT(B)]$ *is an up-complete discussion.*

A key issue in Definition 8 is when a *CONCEDE* or *RETRACT* move is applicable. In the original Grounded Discussion Game [8], a move $CONCEDE(B)$ is applicable iff the discussion contains a move $HTB(B)$, the discussion does not already contain a move $CONCEDE(B)$ and for every attacker A of B the discussion contains a move $RETRACT(A)$. Also, a move $RETRACT(B)$ is applicable iff the discussion contains a move $CB(B)$, the discussion does not already contain a move $RETRACT(B)$, and there is an attacker A of B such that the discussion contains a move $CONCEDE(A)$. For the up-complete discussion game, we need to slightly alter this condition as follows.

A move $CONCEDE(B)$ is applicable iff

[10] We say that there is a *HTB-CB* repeat iff $\exists i, j \in \{1 \dots n\} \exists A \in Ar : (M_i = HTB(A) \vee M_i = CB(A)) \wedge (M_j = HTB(A) \vee M_j = CB(A)) \wedge i \neq j$.

1. the discussion contains a previous move $HTB(B)$, and
2. the discussion does not already contain a move $CONCEDE(B)$, and
3. either

   a. for every attacker A of B the discussion contains a previous move $RETRACT(A)$, or
   b. B is labelled $\mathtt{in}$ by the initial labelling $\mathcal{L}ab$

A move $RETRACT(B)$ is applicable iff

1. the discussion contains a previous move $CB(B)$, and
2. the discussion does not already contain a move $RETRACT(B)$, and
3. either

   a. there exists an attacker A of B such that the discussion contains a previous move $CONCEDE(A)$, or
   b. B is labelled $\mathtt{out}$ by the initial labelling $\mathcal{L}ab$

The above definition of applicability of *CONCEDE* and *RETRACT* is almost the same as in the Grounded Discussion Game [8] (as is the rest of the up-complete game). The only difference is that the condition 3b has been added in regarding the applicability of *CONCEDE* and the applicability of *RETRACT*.

Just as in the Grounded Discussion Game, the proponent wins the up-complete game iff the opponent concedes the main argument (the argument the discussion started with).

Definition 9. *An up-complete discussion* $[M_1, \ldots, M_n]$ *is called* terminated *iff there exists no move* M_{n+1} *such that* $[M_1, \ldots, M_n, M_{n+1}]$ *is an up-complete discussion. A terminated up-complete discussion (with A being the main argument) is won by the proponent iff the discussion contains* $CONCEDE(A)$*, otherwise it is won by the opponent.*

As an example of how the up-complete discussion game works, consider the argumentation framework of Figure 1 and labelling $\mathcal{L}ab = (\{A\}, \{B\}, \{C, D\})$. The discussion $[HTB(D),\ CB(C),\ HTB(A), CONCEDE(A),\ RETRACT(C),\ CONCEDE(D)]$ is terminated and won by the proponent. We observe that $D \in \mathtt{in}(\upharpoonleft\mathcal{L}ab)$ as $\upharpoonleft\mathcal{L}ab = (\{A, D\}, \{B, C\}, \varnothing)$.

We are now ready to formally state soundness and completeness of the up-complete discussion game.

Theorem 2. *Let* $\mathcal{L}ab$ *be an admissible labelling (called the* initial labelling*) of argumentation framework* $(Ar, \rightharpoonup)$*. An argument* $A \in Ar$ *is labelled* $\mathtt{in}$ *by* $\upharpoonleft\mathcal{L}ab$ *iff the proponent has a winning strategy for A in the up-complete game.*

Proof. We first define the increasing sequence of labellings $L_0 \sqsubseteq \cdots \sqsubseteq L_u$ inductively by $L_0 = \mathcal{L}ab$ and $L_{i+1} = \Gamma(L_i)$ for $i \geqslant 0$, where u is minimal such that $L_u = L_{u+1}$. By results in [5] we know each L_i is admissible and $L_u = \upharpoonleft\mathcal{L}ab$.
"$\Leftarrow$": Suppose proponent has a winning strategy for A. Then in particular there is a terminated discussion $[M_1, \ldots, M_n]$ containing move $CONCEDE(A)$. For each $1 \leqslant i \leqslant n$ define labelling N_i by setting $N_i(B) = \mathcal{L}ab(B)$ if $B \in \mathtt{in}(\mathcal{L}ab) \cup \mathtt{out}(\mathcal{L}ab)$, $N_i(B) = \mathtt{in}$ if $B \in \mathtt{undec}(\mathcal{L}ab)$ and $M_j = CONCEDE(B)$ for

some $j \leqslant i$, $N_i(B) = \texttt{out}$ if $B \in \texttt{undec}(\mathcal{L}ab)$ and $M_j = \mathit{RETRACT}(B)$ for some $j \leqslant i$, $N_i(B) = \texttt{undec}$ otherwise. We note that N_i is well-defined due to there being no *HTB*-*CB* repeats. Then, by induction we can show that for each i we have $N_i \sqsubseteq L_j$ for some j (which depends on i). In particular $N_n \sqsubseteq L_j \sqsubseteq L_u = \upharpoonleft\mathcal{L}ab$ and so, since $N_n(A) = \texttt{in}$ we have A labelled $\texttt{in}$ by $\upharpoonleft\mathcal{L}ab$.
"⇒": Suppose A is labelled $\texttt{in}$ by $\upharpoonleft\mathcal{L}ab$. For any $B \in \texttt{in}(L_u) \cup \texttt{out}(L_u)$ let $r(B)$ be minimal such that $L_{r(B)}(B) = L_u(B)$. Note that, for *any* discussion starting with $\mathit{HTB}(A)$, $L_{r(A)}$ labels the arguments of any *HTB* and *CB* moves with $\texttt{in}$, $\texttt{out}$ respectively (this can be proved by induction on the length of the discussion). Then we can define a strategy for the proponent as follows: whenever opponent plays $\mathit{CB}(B)$ and no *CONCEDE* or *RETRACT* moves are applicable, play $\mathit{HTB}(C)$ where C is any argument such that $C \rightharpoonup B$ and $r(C) = r(B) - 1$. (C exists by the admissibility of $L_{r(B)}$ and the minimality of $r(B)$. Note if $r(B) = 0$ then B will be immediately *RETRACT*ed after $\mathit{CB}(B)$ is played.) To show this yields a *winning* strategy we claim that, for any *terminated* discussion following this strategy starting with $\mathit{HTB}(A)$ and for all arguments B, any move $\mathit{CB}(B)$ or $\mathit{HTB}(B)$ will eventually be followed by a $\mathit{RETRACT}(B)$ and $\mathit{CONCEDE}(B)$ move respectively. In particular $\mathit{HTB}(A)$ will eventually be followed by $\mathit{CONCEDE}(A)$, so the proponent wins. The claim is proved by induction on $r(B)$. If $r(B) = 0$ then the conclusion follows from applicability conditions 3b for *CONCEDE* and *RETRACT*. So suppose $r(B) = i > 0$ and the claim holds for all C such that $r(C) < i$. Suppose $\mathit{CB}(B)$ is played. If it is not immediately *RETRACT*ed then, following the strategy, the next move is $\mathit{HTB}(C)$ with $C \rightharpoonup B$ and $r(C) = r(B) - 1$. By induction, C is eventually *CONCEDE*ed, at which point B must also be *RETRACT*ed. If $\mathit{HTB}(B)$ is played and is not immediately *CONCEDE*ed then eventually all attackers of B must be played as *CB*. However for any such attacker C we have $r(C) < r(B)$ and so, by induction, every attacker must eventually be *RETRACT*ed. □

Given the up-complete game, it becomes possible to combine this with the down-admissible game to provide dialectical proof procedures for the super credulous operator. The idea is to embed the down-admissible game inside of the up-complete game. That is, to determine whether argument A is labelled $\texttt{in}$ by the super credulous labelling, we start with running the up-complete game for argument A (first move: $\mathit{HTB}(A)$). In the up-complete game defined above, the opponent has to move *CONCEDE* when the discussion hits an argument that is labelled $\texttt{in}$ by the initial labelling. However, what we are interested in is not so much whether an argument is labelled $\texttt{in}$ by the initial labelling, but whether the argument is labelled $\texttt{in}$ by the down-admissible of the initial labelling. This can be determined by running the down-admissible game. So whenever the proponent wants to do a *HTB* move for an argument (say B) he thinks is in the down-admissible of the initial labelling, instead of doing an $\mathit{HTB}(B)$ move, he starts the down-admissible game (first move: $\texttt{in}(B)$). If the proponent wins the down-admissible game, the entire game counts as an $\mathit{HTB}(B)$ move with B being in the down-admissible of the initial labelling. This means that (condition 3b) the opponent has to respond with a $CONCEDE(B)$ move.[11]

If one then substitutes the down-admissible game (inside of the up-complete game) by the credulous game (which after all is the down-admissible game based on the credulous initial labelling $\downharpoonright \sqcup(\mathcal{L}ab_1, \ldots, \mathcal{L}ab_n))$ one obtains a discussion game for the super-credulous operator.

5. Discussion

In the current paper, we have shown that it is possible to define sound and complete dialectical proof procedures for the down-admissible and up-complete labellings. These proof procedures were obtained by relatively minor changes to the dialectical proof procedures for preferred [9] and grounded [8] semantics. The proof procedure for down-admissible (the down-admissible game) is basically the admissibility game of [9] with the additional constraint that the discussion needs to stay "inside" of the initial labelling (that is, for the proponent to win the game, for each $\mathtt{in}(B)$ move it has to hold that $B \in \mathtt{in}(\mathcal{L}ab)$, and for each $\mathtt{out}(B)$ move it has to hold that $B \in \mathtt{out}(\mathcal{L}ab)$). The proof procedure for the up-complete labelling (the up-complete game) is basically the Grounded Discussion Game [8] with immediate $CONCEDE$ and $RETRACT$ moves whenever the discussion touches arguments in the initial labelling (that is, when uttering a move $HTB(B)$ with $B \in \mathtt{in}(\mathcal{L}ab)$ or $CB(B)$ with $B \in \mathtt{out}(\mathcal{L}ab)$). In the special case that the initial labelling is the all-`undec` labelling, the up-complete game coincides precisely with the Grounded Discussion Game, as the grounded labelling is the up-complete of the all-`undec` labelling.

Based on the proof procedures of the down-admissible and up-complete, we then outlined dialectical proof procedures for the sceptical, credulous and super credulous operator. The proof procedures for the sceptical and credulous operator are based on the down-admissible game with particular initial labellings. The proof procedure of the super credulous operator is based on the up-complete game with the down-admissible game embedded in it.

We have shown that our discussion games can be given an intuitive interpretation. For the sceptical and credulous games, one could envision a panel discussion between a proponent and an opponent, with the audience consisting of the agents whose opinions (labellings) are being aggregated, and who are able to actively interfere ("heckle") with the panel discussion.[12] By providing such an intuitive interpretation, the game goes beyond the purely computational function of traditional proof procedures. This is in line with what should be arguably one of the main aims of formal argumentation theory: to bridge the gap between computer-based reasoning and human reasoning.

[11] One could ask whether a similar down-admissible game is necessary when the opponent wants to do a CB move for an argument that is labelled `out` by the down-admissible of the initial labelling. The answer is negative, as carrying on for one more step in the up-complete game will yield an (HTB) argument that is labelled `in` by the down-admissible of the initial labelling. On this argument we then run the embedded down-admissible game as described above.

[12] For the super credulous game, audience participation is only possible during the credulous subgame.

References

[1] E. Awad. *Collective Judgement in Contested Domains: The Case of Conflicting Arguments*. PhD thesis, Masdar Institute, 2015.

[2] E. Awad, J. F. Bonnefon, M. W. A. Caminada, Th. Malone, and I. Rahwan. Experimental assessment of aggregation rules in argumentation-enabled collective intelligence. Technical report, arXiv:1604.00681, 2016.

[3] E. Awad, M. W. A. Caminada, G. Pigozzi, M. Podlaszewski, and I Rahwan. Pareto optimality and strategy proofness in group argumentation evaluation. Technical report, arXiv:1604.00693, 2016.

[4] P. Baroni, M.W.A. Caminada, and M. Giacomin. An introduction to argumentation semantics. *Knowledge Engineering Review*, 26(4):365–410, 2011.

[5] R. Booth. Judgment aggregation in abstract dialectical frameworks. In *Advances in Knowledge Representation, Logic Programming, and Abstract Argumentation - Essays Dedicated to Gerhard Brewka on the Occasion of His 60th Birthday*, volume 9060 of *Lecture Notes in Computer Science*, pages 296–308. Springer, 2015.

[6] R. Booth, E. Awad, and I. Rahwan. Interval methods for judgment aggregation in abstract argumentation. In *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR 2014)*, pages 594–597, 2014.

[7] M.W.A. Caminada. On the issue of reinstatement in argumentation. In M. Fischer, W. van der Hoek, B. Konev, and A. Lisitsa, editors, *Logics in Artificial Intelligence; 10th European Conference, JELIA 2006*, pages 111–123. Springer, 2006. LNAI 4160.

[8] M.W.A. Caminada. A discussion game for grounded semantics. In Elizabeth Black, Sanjay Modgil, and Nir Oren, editors, *Theory and Applications of Formal Argumentation (proceedings TAFA 2015)*, pages 59–73. Springer, 2015.

[9] M.W.A. Caminada, W. Dvořák, and S. Vesic. Preferred semantics as socratic discussion. *Journal of Logic and Computation*, 2014. (in print).

[10] M.W.A. Caminada and D.M. Gabbay. A logical account of formal argumentation. *Studia Logica*, 93(2-3):109–145, 2009. Special issue: new ideas in argumentation theory.

[11] M.W.A. Caminada and G. Pigozzi. On judgment aggregation in abstract argumentation. *Autonomous Agents and Multi-Agent Systems*, 22(1):64–102, 2011.

[12] M.W.A. Caminada, G. Pigozzi, and M. Podlaszewski. Manipulation in group argument evaluation. In Toby Walsh, editor, *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, pages 121–126, 2011.

[13] M.W.A. Caminada and M. Podlaszewski. Grounded semantics as persuasion dialogue. In Bart Verheij, Stefan Szeider, and Stefan Woltran, editors, *Computational Models of Argument - Proceedings of COMMA 2012*, pages 478–485, 2012.

[14] M.W.A. Caminada and M. Podlaszewski. User-computer persuasion dialogue for grounded semantics. In Jos W.H.M. Uiterwijk, Nico Roos, and Mark H.M. Winands, editors, *Proceedings of BNAIC 2012; The 24th Benelux Conference on Artificial Intelligence*, pages 343–344, 2012.

[15] M.W.A. Caminada and Y. Wu. An argument game of stable semantics. *Logic Journal of IGPL*, 17(1):77–90, 2009.

[16] N. Gorogiannis and A. Hunter. Instantiating abstract argumentation with classical logic arguments: Postulates and properties. *Artificial Intelligence*, 175(9-10):1479–1497, 2011.

[17] S. Modgil and M.W.A. Caminada. Proof theories and algorithms for abstract argumentation frameworks. In I. Rahwan and G.R. Simari, editors, *Argumentation in Artificial Intelligence*, pages 105–129. Springer, 2009.

[18] S. Modgil and H. Prakken. The ASPIC+ framework for structured argumentation: a tutorial. *Argument & Computation*, 5:31–62, 2014. Special Issue: Tutorials on Structured Argumentation.

[19] M. Podlaszewski. *Poles Apart: Navigating the Space of Opinions in Argumentation*. PhD thesis, Université du Luxembourg, 2015.

[20] F. Toni. A tutorial on assumption-based argumentation. *Argument & Computation*, 5:89–117, 2014. Special Issue: Tutorials on Structured Argumentation.