Persuasive Argumentation and Epistemic Attitudes (preprint version)*

Carlo Proietti[1,2] and Antonio Yuste-Ginel[3]

[1] Lund University, Sweden
[2] ILLC, University of Amsterdam, The Netherlands
[3] University of Málaga, Spain

Abstract. This paper studies the relation between persuasive argumentation and the speaker's epistemic attitude. Dung-style abstract argumentation and dynamic epistemic logic provide the necessary tools to characterize the notion of persuasion. Within abstract argumentation, persuasive argumentation has been previously studied from a game-theoretic perspective. These approaches are blind to the fact that, in real-life situations, the epistemic attitude of the speaker determines which set of arguments will be disclosed by her in the context of a persuasive dialogue. This work is a first step to fill this gap. For this purpose we extend one of the logics of Schwarzentruber et al. with dynamic operators, designed to capture communicative phenomena. A complete axiomatization for the new logic via reduction axioms is provided. Within the new framework, a distinction between actual persuasion and persuasion from the speaker's perspective is made. Finally, we explore the relationship between the two notions.

Keywords: Argumentation Frameworks · Dynamic Epistemic Logic · Persuasion and Argument Labellings

1 Introduction

Persuasion is probably the most relevant motivation of arguing and debating. Roughly defined, we could say that a piece of argumentation is persuasive whenever the speaker succeeds to align the hearer with her goal. Within the traditional divide of liberal arts of the trivium between *logic*, *grammar* and *rhetoric*, persuasive communication has been mostly the object of rhetoric (see [12]), with minor interest for its formal aspects. The main objective of this paper consists in studying persuasive argumentation using a well-known formal tool: abstract argumentation frameworks [10]. Abstract argumentation frameworks place ourselves in a dialectical perspective, where the strength of an argument is measured

* The research activity of Antonio Yuste-Ginel is supported by MECD-FPU 2016/04113. Carlo Proietti gratefully acknowledges funding received from the European Commission (Marie Skodowska-Curie Individual Fellowship 2016, 748421) and Sveriges Riksbanken (P16-0596:1) for his research.

in terms of its relation to other arguments.[4] Besides, the notion of persuasion we are going to argue for can be read in dialectical terms too. Concretely, persuasion will be defined in terms of how the hearer assesses her arguments after the communication has taken place – and not in terms of her disposition to act in a certain way. For this purpose, we make use of the notion of *justification status* of an argument proposed by [24]. Hence, persuasion is going to be understood as a change in the hearer's justification status of a specific argument, the *issue* of the exchange, that fits the speaker communicative goal.

In addition to this dialectical flavour, our main conceptual contribution consists in taking into account the epistemic attitudes of the involved agents in order to provide a realistic notion of persuasive communication. An important antecedent in the study of persuasive communication from a formal perspective is [23], where dialogical logic is used to define different kinds of persuasive dialogues. Closer to our approach, previous works in abstract argumentation have investigated the question of persuasive argumentation from a game-theoretic point of view. Such approaches are mostly based on debate-game scenarios where the participants disclose arguments in turns, under the presupposition that they either know everything about the arguments available to their opponent [17] or that they don't consider them as relevant for their strategy [18]. These assumptions seem to be too strong when applied to real-life argumentative scenarios. First, communication does not necessarily unfolds as an exchange of single arguments in rigid turns. More importantly, it looks that it is precisely what the speaker thinks of the hearer's argumentative situation that makes her decide which arguments to disclose to persuade her. Indeed, there are cases in which, at an initial situation, the speaker can find a persuasive piece of argumentation but, due to her misinformation about the hearer, she ends up in a new situation in which the original goal of communication is not reachable any more.

Dynamic Epistemic Logic (DEL) is the mathematical study of informational changes [7], where information is understood in terms of the epistemic attitudes (knowledge and belief) of a set of agents. Previous works have studied relations between argumentation frameworks and epistemic logic [20,21,13,19]. For instance, [20,21] have focused on explaining the way that beliefs emerge from the set of arguments owned by an agent. Here, we introduce another problem: how the beliefs of the speaker determine the set of arguments that will be disclosed in the context of a discussion. For this aim, we propose a dynamic extension of one of the logics presented in [19], strongly related to the work of [2], that allows to reason about epistemic attitudes concerning the argumentative knowledge-base of other agents and about the effects of argumentative communication. In line with DEL, disclosure is encoded as a specific update of the model after an argument is announced. For our purposes we assume that the hearer is somehow "credulous": she will accept all the information that the speaker sends. As far as we know this is the first approach that combines tools imported both from

[4] See [14] for this notion of dialectics and [1] for its relation to argumentation frameworks.

AFs and DEL.[5] Its main contribution consists of a complete axiomatization of a dynamic logic of argumentative disclosure $\mathcal{L}^{!+}(A)$ and its doxastic/epistemic extensions.

The rest of this paper is organized as follows. Section 2.1 presents the preliminary concepts. In Section 2.2 we define communication within multi-agent frameworks and then introduce a notion of persuasion, understood as a change of the hearer's justification status of a given argument according to the speaker's goal. We also show how, in real-life situations, the information of the speaker about the hearer's information is crucial in order to perform persuasive argumentation. In Section 3 we extend one of the epistemic-argumentation logics of [19] with two dynamic modalities, meant to capture communicative phenomena. We provide a complete axiomatization for such extension via reduction axioms. In Section 4, we combine the AF tools with DEL to define the concept of epistemic-based persuasive arguments. We argue that this notion is more realistic than plain persuasion when applied to real-life argumentative scenarios. Finally, we discuss some relations between epistemic-based persuasion and plain persuasion. We close the paper by pointing out some future possible directions (Section 5).

2 Persuasive Communication

In this section we first recall the concepts that are needed for our notion of persuasion, namely, that of *argumentation framework*, its *multi-agent* version, *argument (complete) labelling* and *justification status*. Later on, we discuss how argumentative communication and persuasive argumentation can be modelled with the mentioned list of notions.

2.1 Preliminaries

Definition 1 (Argumentation Framework [10]). *An* argumentation framework *(AF, for short) is a pair* $(A, \rightsquigarrow)$ *where* $A \neq \emptyset$ *is a set of* arguments *and* $\rightsquigarrow \subseteq A \times A$ *is called the* attack *relation.*

For the purposes of this work we restrict our attention to *finite* AFs. Among others, AFs have been applied in several contexts to model multi-agent argumentative scenarios [4,18,16,19,17] where the agents' (partial) information is represented as a subgraph of a larger argumentative pool. Such scenarios are captured by the following definition.

Definition 2 (Multi-agent Argumentation Framework). *A* multi-agent argumentation framework *(MAF) for a non-empty and finite set of agents* Ag *is a triple* $(A, \rightsquigarrow, \{A_i\}_{i \in \mathsf{Ag}})$ *such that* $(A, \rightsquigarrow)$ *is an AF (called the* Universal

[5] Nevertheless, some papers [8,9] have combined tools from propositional dynamic logic and abstract argumentation. We will come back to them in the conclusions (Section 5).

Argumentation Framework *(UAF)) and* $A_i \subseteq A$*. Given a MAF and an agent* $i \in \mathsf{Ag}$*,* agent i's subgraph *is defined as* $(A_i, \rightsquigarrow_i)$ *where* $\rightsquigarrow_i = \rightsquigarrow \cap (A_i \times A_i)$*.*

Remark 1. Note that, given a MAF and an agent $i \in \mathsf{Ag}$, we have that $(A_i, \rightsquigarrow_i)$ is an AF. Note, also, that the way in which agents' subgraphs are defined amounts to assume, following [19], that agents share the same logical framework. That is to say, given two arguments $a, b \in A$, if the agent is aware of them, then she thinks that a attacks b if and only if it is actually the case that a attacks b. We add that this way of defining an agent's subgraph amounts also to assume that agents are somehow *ideal reasoners*: they cannot fail to see a conflict between arguments, e.g. an undercut or a rebuttal, where there is one.

In what follows, we restrict our attention to 2-agents AFs (2-AF for short), i.e. we assume that $\mathsf{Ag} = \{1, 2\}$ and denote 2-AFs as $(A, \rightsquigarrow, A_1, A_2)$. Given a 2-AF, *a pointed 2-AF* is a tuple $(A, \rightsquigarrow, A_1, A_2, a)$ where $a \in A_1 \cap A_2$. In a pointed 2-AF $(A, \rightsquigarrow, A_1, A_2, a)$, the UAF $(A, \rightsquigarrow)$ is intended to represent all the relevant arguments about the issue a while each A_i is intended to represent the arguments that agent i is aware of.

The semantics of an AF $(A, \rightsquigarrow)$ is presented in terms of its *extensions* [10]. An extension of $(A, \rightsquigarrow)$ is a set of arguments $B \subseteq A$ that meets certain conditions to be an "acceptable" opinion. Typically, the minimal conditions for B are *conflict-freeness* (no $a, b \in B$ attack each other) and *defense* of its own arguments (for any c that attacks $b \in B$ there is some $b' \in B$ that attacks c). Any set that has these two properties is said to be *admissible*. Any admissible set that is equal to the set of arguments it defends is said to be *complete*. In [4,6] an alternative but equivalent approach to Dung's semantics is offered in terms of *labellings*. For our present purposes we only introduce the notion of a complete labelling,

Definition 3 ((Complete) Argument Labelling). *Let* $(A, \rightsquigarrow)$ *be an AF. An argument labelling is a total function* $\mathcal{L} : A \to \{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$*. Furthermore, a labelling* $\mathcal{L}$ *for* $(A, \rightsquigarrow)$ *is said to be* complete *iff for all* $a \in A$ *it holds that:*

- $\mathcal{L}(a) = \mathsf{in}$ *iff for all* $c \in A$ *s.t.* $c \rightsquigarrow a$*:* $\mathcal{L}(c) = \mathsf{out}$
- $\mathcal{L}(a) = \mathsf{out}$ *iff there is a* $c \in A$ *s.t* $c \rightsquigarrow a$ *and* $\mathcal{L}(c) = \mathsf{in}$

It is not difficult to show that, for any complete labelling, the set of arguments that are labelled in forms a complete extension and, viceversa, from any complete extension B we easily obtain a complete labelling where all and only the arguments in B are labelled in.

An AF $(A, \rightsquigarrow)$ may contain more than one complete extension. Based on this, [24] defines the justification status of an argument relative to an AF in terms of its membership to complete labellings.

Definition 4 (Justification Status). *Let* $(A, \rightsquigarrow)$ *be an AF and let* $a \in A$*. The* justification status of a *is the outcome yielded by the function* $\mathcal{JS} : A \to \wp(\{\mathsf{in}, \mathsf{out}, \mathsf{undec}\})$ *defined as:*

$$\mathcal{JS}(a) := \{\mathcal{L}(a) \mid \mathcal{L} \quad \text{is a complete labelling of} \quad (A, \rightsquigarrow)\}$$

As noted in [24], two of the eight possible outcomes of $\mathcal{JS}$ are excluded. First, $\emptyset$ is never a possible outcome of $\mathcal{JS}$, since, as proved in [10], there is always at least a complete extension. Second, the value $\{\mathsf{in}, \mathsf{out}\}$ is also excluded from the range of $\mathcal{JS}$, since it can be proven that if $\mathsf{in}, \mathsf{out} \in \mathcal{JS}(a)$, then $\mathsf{undec} \in \mathcal{JS}(a)$. Let us denote by JS^* the set of the six possible justification status of an argument.

Notation. Let $(A, \rightsquigarrow, A_1, A_2, a)$ be a 2-AF, let $B \subseteq A$, we use $\mathcal{JS}^B(a)$ to refer to the justification status of a w.r.t. $(B, \rightsquigarrow \cap (B \times B))$. Note that $\mathcal{JS}^{A_i}(a)$ denotes the justification status of a for agent i. For the sake of readability, we sometimes write $\mathcal{JS}_i(a)$ instead of $\mathcal{JS}^{A_i}(a)$.

The authors of [24] give the following names to the possible outcomes of $\mathcal{JS}$: $\{\mathsf{in}\}$ is called *strong accept*; $\{\mathsf{in}, \mathsf{undec}\}$ is called *weak accept*; $\{\mathsf{undec}\}$ is called *determined borderline*; $\{\mathsf{in}, \mathsf{out}, \mathsf{undec}\}$ is called *undetermined borderline*, $\{\mathsf{out}, \mathsf{undec}\}$ is called *weak reject* and $\{\mathsf{out}\}$ is called *strong reject*. The following total pre-order defines an acceptance hierarchy of an argument with respect to an AF:

$$\text{strong accept} > \text{weak accept} > \text{determined borderline} = \text{undetermined borderline} > \text{weak reject} > \text{strong reject}$$

2.2 Communication and Persuasion

For the sake of simplicity, we assume that 1 is the speaker (or sender), i.e., the one that is trying to persuade while 2 is the hearer (or receiver).

Definition 5 (Communication Step). *Given a pointed 2-AF* $\mathcal{G} = (A, \rightsquigarrow, A_1, A_2, a)$ *and* $B \subseteq A_1$*, a* communication step *is a triple* $(\mathcal{G}, B, \mathcal{G}^B)$ *where* $\mathcal{G}^B$ *is called the* resulting pointed 2-AF *and it is defined as* $\mathcal{G}^B := (A, \rightsquigarrow, A_1, A_2^B, a)$ *where* $A_2^B = A_2 \cup B$*.*

Remark 2. First, agent 2 (the hearer) is assumed to always perform an *open update* [16], i.e., 2 always trusts 1 and incorporates all the received information into her subgraph. Second, the *disclosure policy* of 1 is left undecided, since our objective is to investigate how should 1 discloses her available information in order to persuade 2.

Note that $\mathcal{JS}^{A_i \cup B}(a)$ (see notation above) denotes the justification status of a for agent i after B has been communicated. For the sake of readability, we sometimes write $\mathcal{JS}_i^B(a)$ instead of $\mathcal{JS}^{A_i \cup B}(a)$. With these ingredients in mind, we are able to define a persuasive communication as one where the speaker reaches to align the hearer's justification status with her intended goal.

Definition 6 (Persuasive Communication). *Given a communication step* $(\mathcal{G}, B, \mathcal{G}^B)$ *and a goal of communication for the speaker* $\mathsf{goal} \in \mathsf{JS}^*$*. The communication step is said to be* persuasive *iff* $\mathcal{JS}_2^B(a) = \mathsf{goal}$*. Furthermore, a* goal *is said to be* achievable *through* $B \subseteq A_1$ *iff* $(\mathcal{G}, B, \mathcal{G}^B)$ *is persuasive. In general, a* goal *is said to be achievable iff there is some* B *through which* goal *is achievable.*

Remark 3. Given a pointed 2-AF $(A, \rightsquigarrow, A_1, A_2, a)$ and a $\mathsf{goal} \in \mathsf{JS}^*$, we have that goal is achievable w.r.t. $(A, \rightsquigarrow, A_1, A_2, a)$ iff there is an $A_2' \subseteq A$ such that: $A_2 \subseteq A_2'$, $A_2' \setminus A_2 \subseteq A_1$ and $\mathcal{JS}^{A_2'}(a) = \mathsf{goal}$.

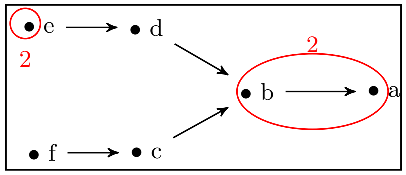
Note that persuasion is not always possible. In particular if $A_2 = A$ (2 has access to all relevant information) and $\mathsf{goal} \neq \mathcal{JS}_2(a)$ (the goal is not *trivial*), then we have that goal is not achievable: there is no $B \subseteq A_1$ s.t. $\mathcal{JS}_2^B(a) = \mathsf{goal}$. Besides, goal is not achievable when $\rightsquigarrow$ is well-founded in $(A, \rightsquigarrow, A_1, A_2, a)$ and goal is not $\{\mathsf{in}\}$ nor $\{\mathsf{out}\}$. As shown by [10], in such case there is only one complete extension (which is grounded, preferred and stable) in $(A, \rightsquigarrow)$ and in any of its possible subgraphs (since the relation $\rightsquigarrow$ is *a fortiori* well-founded there).

Our main objective can be reformulated now in more precise terms: given a pointed 2-AF $(A, \rightsquigarrow, A_1, A_2, a)$ and a $\mathsf{goal} \in \mathsf{JS}^*$, how can 1 select a set of arguments $B \subseteq A_1$ such that $(\mathcal{G}, B, \mathcal{G}^B)$ is persuasive? In order to define a persuasive disclosure policy for agent 1 (the speaker), we could adopt an external view and take into account all the relevant information (A, A_1 and A_2 together with $\rightsquigarrow$). If we adopt such a perspective, then a persuasive policy consists in selecting any set that produces persuasion. Nevertheless, this does not capture what agents do when they are actually trying to persuade another. In these cases, the speaker's information about the hearer's information is crucial to adopt a successful strategy. Let us make this point clear through the following example.

Example 1 (A bloody crime). A murder was committed in Amsterdam last night. The main suspect, called Mr. 1, is being interrogated by the famous detective Ms. 2. The suspect wants to persuade 2 that he is innocent (a), i.e., $\mathsf{goal} = \{\mathsf{in}\}$. Unfortunately for him, he was seen by several witnesses holding a bloody knife in his hands close to the crime scene (b). He does possess two potential alibis. First, 1 has a (well-known criminal) identical twin brother (c). Nevertheless, and unknown to Ms. 2, his twin brother was in Venice the night of the crime (f). Besides, 1 used to work in a butcher's nearby (d) but he was fired a week ago (e). Imagine that 2 owns e because she has done the proper interrogations before. Figure 1 represents a pointed 2-AF depicting the story. Note that $\mathcal{JS}_1(a) = \mathcal{JS}_2(a) = \{\mathsf{out}\}$ (both agents think that 1 is not innocent). Note also that 1 will succeed if he discloses c but he will fail if he discloses either d or $\{c, d\}$. Even more, if he discloses either d or $\{c, d\}$, then goal is not achievable in the resulting graph. This shows that 1's election is crucial in order to reach his goal (proving himself innocent). It also suggests that, within the art of persuasion, some mistakes turn out to be irreparable. In a real-life situation, 1 will chose between c and d according to his information about 2's information.

3 A Dynamic Epistemic Logic for Argumentation

A question arises naturally from Example 1: how could we represent 1's epistemic states (belief and knowledge) about 2's argumentative state? One simple way to do so, using tools imported from standard epistemic logic and awareness logic



a: "1 is innocent"
b: "1 was seen close to the crime scene"
c: "1 has a twin brother living in the city"
f: "1's twin brother was in Venice last night"
d: "1 works in a butcher's nearby"
e: "1 was fired from the butcher's a week ago"

Fig. 1. A pointed 2-AF for Example 1.

[11], is offered in [19] under the name of $\mathcal{L}_1$. In order to avoid confusion, let us just simply call it $\mathcal{L}$. In this section, we first recall the syntax and semantics of $\mathcal{L}$. After that, we propose a dynamic extension of $\mathcal{L}$, named $\mathcal{L}^+$. $\mathcal{L}^+$ is designed to capture the notion of communication step (see Definition 5) within a DEL framework. We close the section by offering a complete axiomatization for the dynamic extension via reduction axioms.

3.1 Syntax and Semantics of $\mathcal{L}$

Given an AF $(A, \rightsquigarrow)$ and a finite set of agents Ag, the language $\mathcal{L}(A, \mathsf{Ag})$ is generated by the following Backus-Naur form (BNF, in what follows):

$$\varphi ::= \mathsf{owns}_i(a) \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box_i \varphi \qquad a \in A \quad i \in \mathsf{Ag}$$

$\mathsf{owns}_i(a)$ is intended to mean "i is aware of argument a" and $\Box_i\varphi$ is intended to mean "i believes (knows) that φ". In what follows, we assume that $\mathsf{Ag} = \{1, 2\}$ and restrict our attention to $\mathcal{L}(A)$.

Definition 7 (Model, Truth and Validity). *An* $\mathcal{L}(A)$-model *(or simply, a* model*) is a triple* $M = (W, \mathcal{R}, \mathcal{D})$ *where* $W \neq \emptyset$ *is set of* possible worlds, $\mathcal{R} : \mathsf{Ag} \to \wp(W \times W)$ *assigns an* accessibility relation *to each agent, and* $\mathcal{D} : (\mathsf{Ag} \times W) \to \wp(A)$ *is an* awareness function, *intended to represent the set of arguments each agent is aware of in each world. For notational convenience, we abbreviate* $(w, v) \in \mathcal{R}(i)$ *as* $w\mathcal{R}_i v$ *and* $a \in \mathcal{D}(i, w)$ *as* $a \in \mathcal{D}_i(w)$*. Furthermore, we assume that for every* $i, j \in \mathsf{Ag}$ *and every* $w, u \in W$*, it holds that:*

1. *If* $w\mathcal{R}_i u$*, then* $\mathcal{D}_i(w) = \mathcal{D}_i(u)$
2. *If* $w\mathcal{R}_i u$*, then* $\mathcal{D}_j(u) \subseteq \mathcal{D}_i(w)$

Let us denote by $\mathcal{M}$ *the class of all models. Given a model* $M = (W, \mathcal{R}, \mathcal{D})$*, a pointed model is a pair* (M, w) *s.t.* $w \in W$*. Truth in pointed models is defined as usual for propositional connectives. We just make explicit the clauses for* owns_i *and* $\Box_i$*:*

$$M, w \vDash \mathsf{owns}_i(a) \textit{ iff } a \in \mathcal{D}_i(w)$$

$$M, w \vDash \Box_i \varphi \quad \textit{iff } M, w' \vDash \varphi \textit{ for all } w' \textit{ such that } w\mathcal{R}_i w'$$

A formula $\varphi \in \mathcal{L}(A)$ is said to be valid*, denoted by $\models \varphi$, iff it is true in all pointed models for $\mathcal{L}(A)$.*

Condition 1. means that awareness of arguments is fully introspective with respect to belief (knowledge). In other words, if an agent is aware of an argument, then she believes (knows) so and if she is not aware of an argument, then she believes (knows) so. Note that, if we assume that $\mathcal{R}_i$ is serial (which holds both for the standard notion of belief and knowledge), then Condition 1 also implies that $\models \Box_i \mathsf{owns}_i(a) \rightarrow \mathsf{owns}_i(a)$, i.e., if an agent believes (knows) that she is aware of an argument, then she is right. Condition 2, defined by the axiom $\neg\mathsf{owns}_i(a) \rightarrow \Box_i \neg\mathsf{owns}_j(a)$, captures the intuition according to which if an agent is not aware of an argument, then she thinks that no one else is aware of it. Since we want to discuss both knowledge and belief, no restriction on $\mathcal{R}_i$ is imposed in Definition 7. Given an accessibility relation $\mathcal{R}_i$, we say that $\mathcal{R}_i$ is an *epistemic relation* iff it is a pre-order (reflexive and transitive) and we say that it is a *doxastic relation* iff it is serial, transitive and euclidean.

Let $((W, \mathcal{R}, \mathcal{D}), w)$ be a pointed model for $\mathcal{L}(A)$, we have that $(A, \rightsquigarrow, \mathcal{D}_1(w), \mathcal{D}_2(w))$ is a 2-AF (see Definition 2); let us call $(A, \rightsquigarrow, \mathcal{D}_1(w), \mathcal{D}_2(w))$ *the* 2*-AF induced by* (M, w). Furthermore, let $(A, \rightsquigarrow, \mathcal{D}_1(w), \mathcal{D}_2(w))$ be the 2-AF induced by (M, w), a *pointed* 2*-AF induced by* (M, w) is just a tuple $(A, \rightsquigarrow, \mathcal{D}_1(w), \mathcal{D}_2(w), a)$ s.t. $a \in \mathcal{D}_1(w) \cap \mathcal{D}_2(w)$. Note that a pointed model and one of its induced pointed 2-AFs represent both the argumentative situation of each agent with respect to a debate about a and their epistemic attitudes with respect to the other agent's argumentative situation. Analogously, let $\mathcal{G} = (A, \rightsquigarrow, A_1, A_2)$ be a 2-AF, *a pointed model for* $\mathcal{G}$ is a pointed model $((W, \mathcal{R}, \mathcal{D}), w)$ such that $A_1 = \mathcal{D}_1(w)$ and $A_2 = \mathcal{D}_2(w)$.

Proposition 1. *Let $(A, \rightsquigarrow, A_1, A_2)$ be a 2-AF. If $A_i \not\subseteq A_j$, then there does not exist any pointed model for $(A, \rightsquigarrow, A_1, A_2)$ s.t. $\mathcal{R}_j$ is reflexive.*

In other words, if j is not at least as well informed as i (in a strong sense of informedness [5]), then j cannot have any knowledge about $(A, \rightsquigarrow, A_1, A_2)$.

Example 2 (A bloody crime, revisited). Recall the story of Example 1 and the 2-AF shown in Figure 1. A pointed model for it is shown in Figure 2. Note that, since the relation $\rightsquigarrow$ is the same in every possible world, we can dispense with its representation. Nevertheless, awareness sets do change from world to world, so they must be included in the figure. Note that both $\mathcal{R}_1$ and $\mathcal{R}_2$ are doxastic relations. It is also simple to see that $M, w_0 \models \Box_1(\neg\mathsf{owns}_2(e)) \wedge \mathsf{owns}_2(e)$ (1 believes (wrongly) that 2 is not aware of e (the counter-argument to one of 1's alibis).

3.2 A Dynamic Extension of $\mathcal{L}$

Communication steps (see Definition 5) can be now represented in DEL style, as model transformers.

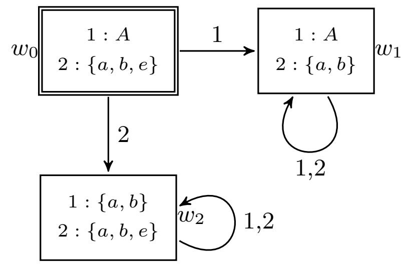


Fig. 2. Pointed model for the 2-AF of Figure 1

.

Definition 8 (Communication Model). *Given a pointed model* $(M, w) = ((W, \mathcal{R}, \mathcal{D}), w)$ *for* $\mathcal{L}(A)$*, the* communication pointed model $(M, w)^{+b} := ((W, \mathcal{R}, \mathcal{D}^{+b}), w)$ *shares domain, accessibility relations and distinguished world with* (M, w)*. The only difference is in the awareness function* $\mathcal{D}^{+b} : (\mathsf{Ag} \times W) \to \wp(A)$*, defined by cases for each* $i \in \mathsf{Ag}$ *and each* $v \in W$ *as follows:*

$$\begin{array}{ll} \mathcal{D}_i(v) \cup \{b\} & \text{if } b \in \mathcal{D}_1(w) \\ \mathcal{D}_i(v) & \text{otherwise} \end{array}$$

For notational convenience, we sometimes abbreviate $(M, w)^{+b}$ as M^{+b}, w.

Remark 4 ($(\cdot)^{+b}$ is a local update). This way of defining the updated model is somehow non-standard. Note that the function $(\cdot)^{+b}$ goes from *pointed models* to *pointed models*; and not from *models* to *models*, as it is the case in public announcement logics and logics with substitution operators [22,15]. Informally, the effect of an action changes if we move from one world to another, i.e., the action is *local*.

Example 3 (Communication Models). Figure 3 extends the model of Figure 2 and it shows the effects of the action $(\cdot)^{+d}$ when it is applied to different pointed models. Note that $(M, w_0)^{+d} \neq (M, w_0)$ but $(M, w_2)^{+d} = (M, w_2)$. Since the precondition of the action holds in w_0 ($d \in \mathcal{D}_1(w_0)$), we have that in the communication model $(M, w_0)^{+d}$ argument d has been added to the awareness sets of both agents in every world. However, if the action takes places in (M, w_2), where the precondition does not hold ($a \notin \mathcal{D}_1(w_2)$) the action $(\cdot)^{+d}$ has no effects.

We can easily go from the disclosure of single arguments to the disclosure of sets of arguments: given $B = \{b_1, ..., b_n\} \subseteq A$, define $(M, w)^{+B}$ as $(M, w)^{+b_1^{\cdot^{\cdot^{\cdot^{+b_n}}}}}$. In order to talk about the action $(\cdot)^{+a}$, we need to enrich the language with a dynamic modality. Define $\mathcal{L}^{+}(A)$ as the language generated by the following BNF:

$$\varphi ::= \mathsf{owns}_i(a) \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box_i \varphi \mid [+a]\varphi \qquad a \in A \quad i \in \mathsf{Ag}$$

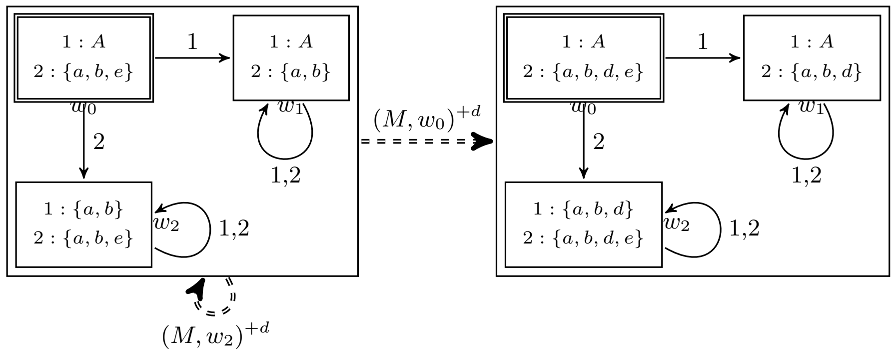


Fig. 3. Communication Model

For any $B = \{b_1, ..., b_n\} \subseteq A$, we define the shorthand $[+B]\varphi := [+b_1]...[+b_n]\varphi$. The notion of truth is extended to the new class of formulas as expected:

$$M, w \vDash [+b]\varphi \qquad \text{iff} \qquad M^{+b}, w \vDash \varphi$$

Remark 5. Definition 8 is enough to capture at least three of our intuitions. First, $M, w \vDash \mathsf{owns}_1(b)$ works as a precondition for the action to have any effect; intuitively, the speaker needs to be aware of what she communicates. Second, if $M, w \vDash \mathsf{owns}_1(b)$, then $M^{+b}, w \vDash \mathsf{owns}_2(b)$; i.e., if the precondition holds, then the hearer gets the communicated argument. Third, if $M, w \vDash \mathsf{owns}_1(b)$, then $M^{+b}, w \vDash \Box^k_{\mathsf{Ag}}(\mathsf{owns}_1(b) \land \mathsf{owns}_2(b))$ and any $k \in \mathbb{N}$; i.e. if the precondition holds, then the awareness of the communicated argument by both agents becomes common belief (knowledge). Note, however, that the communication model must be refined if we count with more than two agents and we want to model cases of private communication (for instance, by removing $\mathcal{R}$-arrows just for the speaker and the hearer).

3.3 Completeness

We provide an axiomatization for $\mathcal{L}^+(A)$. All systems are extensions of K, consisting of axioms $(Taut)$ and (K), and are closed under both inference rules.

- A extends K with axioms (PI) and (GNI).[6]
- BA extends A with axioms (D), (4) and (5).
- KA extends A with axioms (T), (4) and (5)

Besides, we denote by $\mathcal{M}^{\mathcal{K}}$ (resp. $\mathcal{M}^{\mathcal{B}}$) the class of all models where every $\mathcal{R}_i$ is an epistemic (resp. a doxastic) relation.

[6] PI stands for *positive introspection* and GNI for *generalized negative introspection*. The latter captures standard negative introspection as a special case where the indexes i and j are the same.

Theorem 1 (Soundness and Completeness for the static fragments). *The proof system* A *(respectively* AK*,* AB*) is sound and complete w.r.t the class of all models (resp. w.r.t.* $\mathcal{M}^{\mathcal{K}}$*,* $\mathcal{M}^{\mathcal{B}}$*).*

Proving soundness of all the mentioned systems is straightforward. Completeness proofs can be found in the Appendix.

Axioms			
All propositional tautologies	(Taut)	$\vdash \Box_i(\varphi \to \psi) \to (\Box_i\varphi \to \Box_i\psi)$	(K)
$\vdash \mathsf{owns}_i(a) \to \Box_i\mathsf{owns}_i(a)$	(PI)	$\vdash \neg\mathsf{owns}_i(a) \to \Box_i\neg\mathsf{owns}_j(a)$	(GNI)
$\vdash \neg\Box_i \bot$	(D)	$\vdash \Box_i\varphi \to \varphi$	(T)
$\vdash \Box_i\varphi \to \Box_i\Box_i\varphi$	(4)	$\vdash \neg\Box_i\varphi \to \Box_i\neg\Box_i\varphi$	(5)
Rules			
From $\varphi \to \psi$ and φ, infer ψ	MP	From φ infer $\Box_i\varphi$	NEC

Table 1. Axioms for the static fragments

Reduction Axioms for $[+a]$ Using Theorem 1, we obtain completeness results for the dynamic extensions (with $[+a]$ and its semantics) for all the three static logics (A, AK and AB). This is done by providing a set of reduction axioms for $[+a]$. Reduction axioms are valid formulas of the form $[a+]\varphi \leftrightarrow \psi$ s.t. the dynamic operator is "pushed inside" in the formula on the left side of the equivalence. A full set of reduction axioms plus the rule of substitution of equivalents enable us to find a provably equivalent formula in the static logic for every formula in the dynamic logic. The reader is referred to the Appendix and to [7,15] for further details.

In order to find the full set of reduction axioms for our new modality $[+a]$, we take a small detour. First, we show that $[+a]$ is definable in terms of another dynamic modality $[a!]$ (the *public announcement of argument* modality). Then, we offer the full set of reduction axioms for $[a!]$. We start by defining the new language. Given an AF $(A, \rightsquigarrow)$ and $\mathsf{Ag} = \{1, 2\}$, the language $\mathcal{L}^{!+}(A)$ is generated by the following BNF:

$$\varphi ::= \mathsf{owns}_i(a) \mid \neg\varphi \mid \varphi \wedge \varphi \mid [a!]\varphi \mid [+a]\varphi \qquad a \in A \quad i \in \mathsf{Ag}$$

The language $\mathcal{L}^{!}(A)$ is the result of removing the $[+a]\varphi$-clause.

Definition 9 (Announcement Model). *Given a model for* $\mathcal{L}(A)$*,* $M = (W, \mathcal{R}, \mathcal{D})$ *and an argument* $a \in A$*, let us define the* announcement model *as* $M^{a!} = (W, \mathcal{R}, \mathcal{D}^{a!})$ *that shares domain and accessibility relations with* M *and only varies in the awareness function, defined for every* $i \in \mathsf{Ag}$ *and every* $w \in W$ *as* $\mathcal{D}_i^{a!}(w) = \mathcal{D}_i(w) \cup \{a\}$*.*[7]

[7] The operation $(\cdot)^{a!}$ can be understood as a special case of public substitution [22,15]. From the perspective of dynamic awareness logic [2], the same action can be under-

Remark 6. Note that both operations, $(\cdot)^{a+}$ and $(\cdot)^{a!}$, always return a model inside the same class (no matter if we pick up an arbitrary one from $\mathcal{M}$, $\mathcal{M}^{\mathcal{K}}$ or $\mathcal{M}^{\mathcal{B}}$). Since none of the operations alters the accessibility relation, we just have to check that Conditions 1. and 2. from Definition 7 are satisfied by any output of $(\cdot)^{a!}$ and $(\cdot)^{a+}$. Details are left to the reader.

Remark 6 assures that both $\mathcal{L}^{!+}$ and $\mathcal{L}^{!}$ are interpreted on the same class of models as $\mathcal{L}$. Now, the truth clauses for the remaining formulas are as before. As for $[a!]\varphi$, we add the following clause to the definition of truth:

$$M, w \vDash [a!]\varphi \quad \text{iff} \quad M^{a!}, w \vDash \varphi$$

Table 2 shows the full list of reduction axioms for $[+a]$ and $[a!]$.[8] The following proposition is crucial to obtain soundness and completeness for the dynamic extensions of the logics mentioned so far.

Proposition 2. *Schemata shown in Table 2 are valid and the rule SE preserves validity.*

We consider the axiom system $\mathsf{A}^{!+}$ (resp. $\mathsf{AK}^{!+}$ and $\mathsf{AB}^{!+}$) that extends A (resp. AK, AB) with the axioms and rule of Table 2.

Theorem 2 (Completeness of $\mathcal{L}^{+!}$). *The axiom system $\mathsf{A}^{!+}$ (resp. $\mathsf{AK}^{!+}$ and $\mathsf{AB}^{!+}$) is sound and complete w.r.t. the class of models $\mathcal{M}$ (resp. $\mathcal{M}^{\mathcal{K}}$ and $\mathcal{M}^{\mathcal{B}}$).*

$\vdash [+a]\varphi \leftrightarrow (\mathsf{owns}_1(a) \to [a!]\varphi) \wedge (\neg\mathsf{owns}_1(a) \to \varphi)$	(Def+)
$\vdash [a!]\mathsf{owns}_i(a) \leftrightarrow \top$	(Atoms$^{=}$)
$\vdash [a!]\mathsf{owns}_i(b) \leftrightarrow \mathsf{owns}_i(b)$ where $a \neq b$	(Atoms$^{\neq}$)
$\vdash [a!]\neg\varphi \leftrightarrow \neg[a!]\varphi$	(Negation)
$\vdash [a!](\varphi \wedge \psi) \leftrightarrow ([a!]\varphi \wedge [a!]\psi)$	(Conjunction)
$\vdash [a!]\Box_i\varphi \leftrightarrow \Box_i[a!]\varphi$	(Box)
From $\varphi \leftrightarrow \psi$, infer $\delta \leftrightarrow \delta[\varphi/\psi]$	SE

Table 2. Reduction Axioms for $\mathcal{L}^{+!}(A)$

4 Epistemic-based Persuasive Arguments

In this section we combine the tools of Section 2 (AFs) and Section 3 (DEL) in order to define a notion of *persuasive argument based on the perspective of the speaker*. For this we need to take into account the speaker's epistemic situation.

stood in terms of the *consider* action, with the only difference that here arguments (and not formulas) are the content of announcements. A detailed comparison between the three operations is out of the scope of this paper.

[8] In Table 2 $\delta[\varphi/\psi]$ is the result of substituting one or more occurrences of ψ in δ by φ.

As our previous analysis suggests, "being persuasive from the speaker's perspective" does not guarantees actual persuasiveness (Definition 6), which heavily relies on external conditions. However, it is possible to provide sufficient conditions for the former to imply the latter. In other words, it is possible to isolate some epistemic conditions that are "safe" for the speaker, they guarantee that the communication is going to be persuasive. Let us first define persuasiveness from the speaker's perspective.

Definition 10 (Epistemic-based persuasive arguments). *Let (M, w) be a pointed model, let $(A, \rightsquigarrow, \mathcal{D}_1(w), \mathcal{D}_2(w), a)$ be a pointed 2-AF induced by (M, w) and let $\mathsf{goal} \in \mathsf{JS}^*$ be a goal of communication, we say that a set $B \subseteq \mathcal{D}_1(w)$ is* persuasive from 1's perspective *iff $\mathcal{JS}^{\mathcal{D}_2^{+B}(w')}(a) = \mathsf{goal}$ for all $w' \in W$ s.t. $w\mathcal{R}_1 w'$.*

Informally, a set of arguments is persuasive from the speaker's perspective iff it is persuasive in all her accessible AFs. An epistemic-based disclosure policy is just one that selects any set of arguments known (believed) to be persuasive by the speaker.

Definition 10 makes strong use of the notion of justification status. Consequently, persuasive sets cannot be described using $\mathcal{L}^+$. However, and due to the fact that we are working with finite AFs, we can get closer to this objective and express the "all accessible AFs" part of the definition. First, let us define the following shorthand in $\mathcal{L}(A)$:

$$\mathsf{2graph}(C) := \bigwedge_{a \in C} \mathsf{owns}_2(a) \wedge \bigwedge_{b \notin C} \neg \mathsf{owns}_2(b) \quad \text{for any} \quad C \subseteq A$$

The intuitive reading of $\mathsf{2graph}$ is "C is 2's subgraph" and, indeed, it holds that $M, w \models \mathsf{2graph}(C)$ iff $C = \mathcal{D}_2(w)$. Again, since A is finite, we can fix an enumeration of its subsets $\wp(A) = \{A_1, ..., A_n\}$. With these two tools in mind, we can obtain:

Proposition 3. *Given a pointed model (M, w) for $\mathcal{L}(A)$, an enumeration of the subsets of A, $\{A_1, ..., A_n\} = \wp(A)$, and a $\mathsf{goal} \in \mathsf{JS}^*$ we have that $B \subseteq A$ is persuasive from the speaker perspective iff there is an index m (with $1 \leq m \leq n$) s.t.:*

$$M, w \models \Box_1 [+B] \big(\bigvee_{1 \leq i \leq m} \mathsf{2graph}(C_i) \bigwedge_{m < i \leq n} \neg \mathsf{2graph}(C_i) \big)$$

and $\mathcal{JS}^{C_i}(a) = \mathsf{goal}$ for every $1 \leq i \leq m$.

It is immediate to show that if $\Box_1$ is not a factive attitude ($\mathcal{R}_1$ is not reflexive) then not any set that is persuasive from 1's perspective is ensured to be actually persuasive. More importantly, when the speaker lacks of knowledge about the hearer argumentative situation, she runs the risk of committing irreparable mistakes, as shown in the following example:

Example 4 (Wrong belief and persuasion). Recall the communication model depicted in Figure 3. Let us assume that w_0 is the real world and recall that

$\mathsf{goal} = \{\mathsf{in}\}$. Note that the set $\{d\}$ (the butcher's alibi) is persuasive from 1's perspective. Note, however, how she is wrong since in $(\mathcal{D}_2^{+d}(w_0), \rightsquigarrow)$ we have that $\mathcal{JS}_2(a) = \{\mathsf{out}\}$ (because 2 owns the counter-argument e). Even more, it can be easily shown that $\mathsf{goal} = \{\mathsf{in}\}$ is not achievable in $(A, \mathcal{D}_1^{+d}(w_0), \mathcal{D}_2^{+d}(w_0), \rightsquigarrow)$, i.e. if 1 discloses d, he will not longer be able to persuade 2 of his innocence

Remark 7 (Fully specific knowledge (belief) and epistemic-based persuasive sets). Note that $M, w \vDash \Box_1[+B]2\mathsf{graph}(C)$ can be read has "1 has fully specific knowledge (belief) about the effects of communicating B", i.e., the speaker is sure about what the hearer's subgraph will be after the communication. Note, also, that $M, w \vDash \Box_1[+B]2\mathsf{graph}(C)$ and $\mathcal{JS}^C(a) = \mathsf{goal}$ is a sufficient but not necessary condition for B to be persuasive from the speaker perspective in (M, w).

On the other side, it is immediate to show that if $\mathcal{R}_1$ is reflexive ($\Box_1$ is a factive attitude) then any set that is persuasive from 1's perspective is actually persuasive. However, factivity of $\Box_1$ is not a necessary condition for this. The next proposition shows that it is sufficient for the speaker to have a "good enough" belief in order to guarantee persuasion:

Proposition 4. *Given a pointed model (M, w) for $\mathcal{L}(A)$, let $B \subseteq A$ be persuasive from the speaker's perspective. Let A_i be the set of all a_i such that $M, w \vDash \mathsf{owns}_2(a_i) \wedge \neg\Box_1\mathsf{owns}_2(a_i)$. If $A_i \not\rightsquigarrow \mathcal{D}_2^{+B}(w) \setminus A_i$ then B is persuasive.*[9]

In other words, if the arguments 2 is aware of unbeknownst to 1 (e.g. e in our example) are such that they don't conflict with the set of arguments 2 is expected by 1 to have after communicating B, then the communication turns out to be persuasive. This is a consequence of the fact that the justification status of an argument a only depends on the *upstream* of a [24].

5 Closing Words and Future Work

Summing up, we defined persuasion based on the notion of *justification status* of an argument [24]. We then provided a dynamic extension for the logic $\mathcal{L}_1$ of [19] in order to capture the communicative aspects of persuasive argumentation. Furthermore, we have shown that the new logic is axiomatizable via reduction axioms. Using this logic, we distinguished persuasive arguments from epistemic-based persuasive arguments. We argued that the second notion is more realistic to model real-life argumentative scenarios. Finally, we discussed some of the epistemic conditions for a set of arguments to be actually persuasive.

There are several open branches for future work; let us just mention two of them. First, one of the main assumptions of our framework is that hearers always perform an *open update* [16], i.e., they incorporate whatever the speaker says into their AFs. Nevertheless, this is not the case in many situations, where the

[9] Given an AF $(A, \rightsquigarrow)$ and two sets of arguments $B, C \subseteq A$, the attack relation is lifted from single arguments to sets of arguments as follows: $B \rightsquigarrow C$ iff $\exists b \in B, \exists c \in C (b \rightsquigarrow c)$

hearer might suspect that the speaker is acting dishonestly and, consequently, she might revise her AFs more prudently. Second, we recall that neither the notion of persuasion nor its epistemic-based counterpart are definable in the proposed logic. Augmenting the expressive power of the logic –using, for instance, similar techniques to those employed in [8,9]– in order to fully reason about persuasive argumentation inside the language looks a promising step for future work. Among other things, it might allow us to isolate not only sufficient epistemic conditions for a set of argument to be persuasive but also necessary ones.

References

1. Beirlaen, M., Heyninck, J., Pardo, P., Straßer, C.: Argument strength in formal argumentation. Journal of Applied Logics-IfCoLog Journal of Logics and their Applications **5**(3), 629–675 (2018)
2. van Benthem, J., Velázquez-Quesada, F.R.: The dynamics of awareness. Synthese **177**(1), 5–27 (2010)
3. Blackburn, P., De Rijke, M., Venema, Y.: Modal Logic: Graph. Darst, vol. 53. Cambridge University Press (2002)
4. Caminada, M.: On the issue of reinstatement in argumentation. In: European Workshop on Logics in Artificial Intelligence. pp. 111–123. Springer (2006)
5. Caminada, M., Sakama, C.: On the issue of argumentation and informedness. In: JSAI International Symposium on Artificial Intelligence. pp. 317–330. Springer (2015)
6. Caminada, M.W., Gabbay, D.M.: A logical account of formal argumentation. Studia Logica **93**(2-3), 109 (2009)
7. van Ditmarsch, H., van Der Hoek, W., Kooi, B.: Dynamic epistemic logic, vol. 337. Springer Science & Business Media (2007)
8. Doutre, S., Herzig, A., Perrussel, L.: A dynamic logic framework for abstract argumentation. In: Fourteenth International Conference on the Principles of Knowledge Representation and Reasoning (2014)
9. Doutre, S., Maffre, F., McBurney, P.: A dynamic logic framework for abstract argumentation: adding and removing arguments. In: International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. pp. 295–305. Springer (2017)
10. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence **77**(2), 321–357 (1995)
11. Fagin, R., Halpern, J.Y.: Belief, awareness, and limited reasoning. Artificial intelligence **34**(1), 39–76 (1987)
12. Groarke, L.: Informal logic. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, spring 2017 edn. (2017)
13. Grossi, D., van der Hoek, W.: Justified beliefs by justified arguments. In: KR (2014)
14. Johnson, R.H.: Manifest rationality: A pragmatic theory of argument. Routledge (2012)
15. Kooi, B.: Expressivity and completeness for public update logics via reduction axioms. Journal of Applied Non-Classical Logics **17**(2), 231–253 (2007)
16. Proietti, C.: The dynamics of group polarization. In: International Workshop on Logic, Rationality and Interaction. pp. 195–208. Springer (2017)

17. Rahwan, I., Larson, K.: Argumentation and game theory. In: Argumentation in Artificial Intelligence, pp. 321–339. Springer (2009)
18. Sakama, C.: Dishonest arguments in debate games. COMMA **75**, 177–184 (2012)
19. Schwarzentruber, F., Vesic, S., Rienstra, T.: Building an epistemic logic for argumentation. In: Logics in Artificial Intelligence, pp. 359–371. Springer (2012)
20. Shi, C., Smets, S., Velázquez-Quesada, F., et al.: Argument-based belief in topological structures. Electronic Proceedings in Theoretical Computer Science **251** (2017)
21. Shi, C., Smets, S., Velázquez-Quesada, F.R.: Beliefs supported by binary arguments. Journal of Applied Non-Classical Logics pp. 1–24 (2018)
22. Van Benthem, J., Van Eijck, J., Kooi, B.: Logics of communication and change. Information and computation **204**(11), 1620–1662 (2006)
23. Walton, D., Krabbe, E.C.: Commitment in dialogue: Basic concepts of interpersonal reasoning. SUNY press (1995)
24. Wu, Y., Caminada, M., Podlaszewski, M.: A labelling-based justification status of arguments. Studies in Logic **3**(4), 12–29 (2010)

Appendix: Notes and Proofs

Notes

Note 1. As mentioned, we have assumed that 1 is always the speaker while 2 is the hearer. Nevertheless, this logical setting can be easily generalized in order to model argumentative dialogues (where the role speaker/hearer changes in turns). Definition 8 should be extended to obtain a different operation $(\cdot)^a_i$ for each $i \in \mathsf{Ag}$ where the precondition is, consequently, $a \in \mathsf{owns}_i(w)$. Furthermore, the language must also be extended with the clause $[+_i a]\varphi$ for each $a \in A$, $i \in \mathsf{Ag}$. The completeness result presented below can be easily extended for such generalization.

Note 2 ($(\cdot)^{+b}$ as a global update). An alternative way of understanding the action $(\cdot)^{+b}$ (see Definition 7) such that the update becomes world-independent, i.e. $(\cdot)^{+b}$ goes from models to models, works as follows. First note that we still need a suitable notion of *actual world* so that we can express the precondition according to which the speaker has to be aware of b *in the actual world* for $(\cdot)^{+b}$ to have any effect. Hence, our former notion of *pointed model* is now simply called a *model*, i.e., a model is a tuple $M = (W, w, \mathcal{R}, \mathcal{D})$. A *pointed model* is now a pair $((W, w, \mathcal{R}, \mathcal{D}), v)$ where $v \in W$. Here w represents the actual world. The definition of $(\cdot)^{+b}$ stays the same (see Definition 8), but note that now the function $(\cdot)^{+b}$ goes from models to models. Finally, the notion of truth is redefined in *pointed models* as usual for the rest of the operators and it is the following one for $[+b]$:

$$(W, w, \mathcal{R}, \mathcal{D}), v \vDash [+b]\varphi \qquad \text{iff} \qquad (W, w, \mathcal{R}, \mathcal{D})^{+b}, v \vDash \varphi$$

Proofs

All the proofs follow standard methods. We just include some of them here for illustration.

Proposition 1

Proof. Let $\mathcal{G} = (A, \rightsquigarrow, A_i, A_j)$ be a 2-AF model, let (M, w) be a pointed model for $\mathcal{G}$, i.e., $A_i = \mathcal{D}_i(w)$ and $A_j = \mathcal{D}_j(w)$. Suppose, for the sake of contradiction, that $A_i \not\subseteq A_j$ but $\mathcal{R}_j$ is reflexive. We have that there is an argument $a \in A$ s.t. $a \in A_i$ but $a \notin A_j$ or equivalently $a \in \mathcal{D}_i(w)$ but $a \notin \mathcal{D}_j(w)$. But since $\mathcal{R}_j$ is reflexive, we know that $w\mathcal{R}_j w$. The last two assertions contradict Condition 2 of Definition 7.

Completeness of the static logics. Definition of deduction from assumptions ($\Gamma \vdash_* \varphi$), consistent set ($\Gamma \nvdash_* \bot$) (where $* \in \{\mathsf{A}, \mathsf{AK}, \mathsf{AB}\}$) and maximal consistency are standard [3]. Let us denote by $\mathfrak{MC}^*$ the class of all maximal consistent sets in $* \in \{\mathsf{A}, \mathsf{AK}, \mathsf{AB}\}$. When the context is clear or irrelevant, we just write $\mathfrak{MC}$.
For the proof of the next two claims, the reader is referred to [3] (p. 199).

Proposition 5 (Properties of MC-sets). *Let $\Gamma \in \mathfrak{MC}$:*

- *For every $\varphi \in \mathcal{L}(A)$: $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$*
- *If $\Gamma \vdash \varphi$, then $\varphi \in \Gamma$*

Lemma 1 (Lindenbaum). *Let $\Gamma \subseteq \mathcal{L}(A)$, if $\Gamma \nvdash \bot$, then there is a $\Gamma' \in \mathfrak{MC}$ s.t. $\Gamma \subseteq \Gamma'$.*

Definition 11 (Canonical Model). *Given a language $\mathcal{L}(A)$, define the canonical model $M^c = (W^c, \mathcal{R}^c, \mathcal{D}^c)$ as:*

$$W^c := \{\Gamma \subseteq \mathcal{L}(A) \mid \Gamma \in \mathfrak{MC}\}$$

$$\Gamma \mathcal{R}_i^c \Delta \quad \textit{iff} \quad \{\varphi \in \mathcal{L}(A) \mid \Box_i \varphi \in \Gamma\} \subseteq \Delta$$

$$a \in \mathcal{D}_i^c(\Gamma) \quad \textit{iff} \quad \mathsf{owns}_i(a) \in \Gamma$$

Lemma 2 (Canonicity). *Given $\mathcal{L}(A)$, its canonical model M^c is a model.*

Proof. All we need to show is that conditions 1 and 2 of Definition 7 are satisfied by M^c. For Condition 1, let us suppose that $\Gamma \mathcal{R}_i^c \Delta$ (*). Suppose that $a \in \mathcal{D}_i^c(\Gamma)$, by definition this is equivalent to $\mathsf{owns}_i(a) \in \Gamma$. From this we obtain $\Gamma \vdash \mathsf{owns}_i(a)$ and note that, by monotonicity of $\vdash$ and Ax (PI) we have that $\Gamma \vdash \mathsf{owns}_i(a) \rightarrow \Box_i \mathsf{owns}_i(a)$. Applying modus ponens we get $\Gamma \vdash \Box_i \mathsf{owns}_i(a)$. By Lemma 5 we have that $\Box_i \mathsf{owns}_i(a) \in \Gamma$. This, together with (*) and the definition of $\mathcal{R}_i^c$ implies $\mathsf{owns}_i(a) \in \Delta$ which is equivalent by definition of $\mathcal{D}^c$ to $a \in \mathcal{D}_i^c(\Delta)$. We have proven the left to right inclusion, for the right to left, the proof is analogous to Condition 2, where (GNI) is applied with $i = j$
As for Condition 2, suppose $\Gamma \mathcal{R}_i^c \Delta$ (*). Suppose $a \in \mathcal{D}_j^c(\Delta)$. The latter is equivalent by definition to $\mathsf{owns}_j(a) \in \Delta$, which implies $\Delta \vdash \mathsf{owns}_j(a)$. Now, suppose, reasoning by contradiction, that $a \notin \mathcal{D}_i^c(\Gamma)$. This is equivalent by definition to $\mathsf{owns}_i(a) \notin \Gamma$. By Proposition 5 we have that $\neg\mathsf{owns}_i(a) \in \Gamma$ which implies $\Gamma \vdash \neg\mathsf{owns}_i(a)$. Using axiom (GNI) and MP we have that $\Gamma \vdash \Box_i \neg\mathsf{owns}_j(a)$. This, together with (*) and the definition of $\mathcal{R}^c$, implies $\neg\mathsf{owns}_j(a) \in \Delta$ and hence $\Delta \vdash \neg\mathsf{owns}_j(a)$ that contradicts the consistency of Δ. Therefore, $a \in \mathcal{D}_i^c(\Gamma)$.

Lemma 3 (Existence Lemma). *If $\Diamond_i\varphi \in \Gamma$, then there is a $\Delta \in W^c$ s.t. $\Gamma\mathcal{R}_i\Delta$ and $\varphi \in \Delta$.*

Since $\Box_i$ is a normal modal operator, the proof is completely standard. The reader is referred to [3] (pp. 200-201).

Lemma 4 (Truth Lemma). *For each $\Gamma \in W^c$ and and each $\varphi \in \mathcal{L}(A)$:*

$$\varphi \in \Gamma \quad \textit{iff} \quad M^c, \Gamma \vDash \varphi$$

Proof. The proof is by induction on the construction of φ (see [3] chapter 4).

Finally Theorem 1 follows from the Truth Lemma by the typical argument.

Completeness for the dynamic extensions For the completeness of $\mathsf{A}^{!+}$, $\mathsf{AK}^{!+}$ and $\mathsf{AB}^{!+}$, we apply the general method described in [15]. Let us show some of the details.

Proposition 2

Proof. Proving that SE preserves validity can be done by induction of the construction on φ. The proof is simple but long, so we leave it for the reader. Note that in standard awareness logic [11], this is not generally the case, since awareness sets do not need to be closed under logical equivalence. In $\mathcal{L}$, however, the range of awareness operators is restricted to arguments, and therefore soundness of SE is guaranteed.
As for the validity of the reduction axioms, let us just show two cases. Let (M, w) be a pointed model:

- $\vDash [+a]\varphi \leftrightarrow (\mathsf{owns}_1(a) \to [a!]\varphi) \wedge (\neg\mathsf{owns}_1(a) \to \varphi)$
  "$\to$". Suppose $M, w \vDash [+a]\varphi$. This is true iff $M^{+a}, w \vDash \varphi$. Let us reason by cases. If $a \notin \mathcal{D}_1(w)$, then the first conjunct is trivially true. For the second conjunct, note that if $a \notin \mathcal{D}_1(w)$, then $M^{+a}, w = M, w$. We can then substitute M^{+a}, w by M, w and obtain $M, w \vDash \varphi$ which implies $M, w \vDash \neg\mathsf{owns}_1(a) \to \varphi$. Now, if $a \in \mathcal{D}_1(a)$ then the second conjunct is trivially true. For the first conjunct we have that if $a \in \mathcal{D}_1(a)$, then $M^{+a}, w = M^{a!}, w$. By substituting equals in the hypothesis we obtain $M^{a!}, w \vDash \varphi$ which is equivalent by the semantic definition of $[a!]$ to $M, w \vDash [a!]\varphi$.
  "$\leftarrow$". This direction is analogous, each of the cases ($a \in \mathcal{D}_1(w)$ and $a \notin \mathcal{D}_1(w)$) makes one of the conjuncts trivially true and allows us to obtain the true consequent of the other. With that information is easy to deduce $M, w \vDash [+a]\varphi$.
- $\vDash [a!]\Box_i\varphi \leftrightarrow \Box_i[a!]\varphi$
  Suppose $M, w \vDash [a!]\Box_i\varphi$. This is true iff $M^{a!}, w \vDash \Box_i\varphi$ (Definition 9) iff $M^{a!}, w' \vDash \varphi$ for every w' s.t. $w\mathcal{R}_i w'$ (Definition 7) iff $M, w' \vDash [a!]\varphi$ for every w' s.t. $w\mathcal{R}_i w'$ (substituting equivalents of Definition 9 in the last assertion) iff $M, w \vDash \Box_i[a!]\varphi$ (Definition 7).

Definition 12 (Complexity measures).

+-depth *Define* $\overset{+}{n}$: $\mathcal{L}^{!+}(A) \to \mathbb{N}$ *that returns the number of nested* $[+a]$ *in* φ *for any* $a \in A$. *More detailed:* $\overset{+}{n}\ (\mathsf{owns}_i(a)) := 0$, $\overset{+}{n}\ (\star\varphi) := \overset{+}{n}\ (\varphi)$ *where* $\star \in \{\neg, \Box_i, [a!]\}$, $\overset{+}{n}\ (\varphi \wedge \psi) := max(\overset{+}{n}\ (\varphi), \overset{+}{n}\ (\psi))$ *and* $\overset{+}{n}\ ([+a]\varphi) := 1 + \overset{+}{n}\ (\varphi)$.

Depth *Define* $d : \mathcal{L}^{!}(A) \to \mathbb{N}$ *as* $d(\mathsf{owns}_i(a)) = 0$, $d(\star\varphi) = 1 + d(\varphi)$ *where* $\star \in \{\neg, \Box_i, [a!]\}$ *and* $d(\varphi \wedge \psi) = max(d(\varphi), d(\psi))$.

O-depth *Define* $O : \mathcal{L}^{!}(A) \to \mathbb{N}$ *that returns the number of nested operators in* φ. *More detailed:* $Od(\mathsf{owns}_i(a)) = 0$, $Od(\neg\varphi) = Od(\Box_i\varphi) := Od(\varphi)$, $Od(\varphi \wedge \psi) := max(Od(\varphi), Od(\psi))$, *and* $Od([a!]\varphi) = 1 + Od(\varphi)$.

Ord *Define* $Ord : \mathcal{L}^{!}(A) \to \mathbb{N}$ *that returns the depth of the outermost occurrence of* $[a!]$. *More detailed:* $Ord(\mathsf{owns}_i(a)) := 0$, $Ord(\neg\varphi) = Ord(\Box_i\varphi) := Ord(\varphi)$, $Ord(\varphi \wedge \psi) := max(Ord(\varphi), Ord(\psi))$, $Ord([a!]\varphi) = 1 + d(\varphi)$.

Lemma 5 (From $\mathcal{L}^{!+}(A)$ to $\mathcal{L}^{!}(A)$). *For every* $\varphi \in \mathcal{L}^{!+}(A)$, *there is a* $\psi \in \mathcal{L}^{!}(A)$ *s.t.* $\vdash_{\mathsf{A}^{!+}} \varphi \leftrightarrow \psi$.

Proof. By induction on $\overset{+}{n}\ (\varphi)$. If $\overset{+}{n}\ (\varphi) = 0$, we have that $\varphi \in \mathcal{L}^{!}(A)$, and by (Taut) we have that $\vdash \varphi \leftrightarrow \varphi$, so we are done.

Assume as induction hypothesis that for every $\varphi \in \mathcal{L}^{!+}(A)$ s.t. $\overset{+}{n}\ (\varphi) \leq k$, there is a $\psi \in \mathcal{L}^{!}(A)$ s.t.: $\vdash_{\mathsf{A}^{+!}} \varphi \leftrightarrow \psi$. Suppose $\overset{+}{n}\ (\varphi) = k + 1$. Take every $\delta_i \in sub(\varphi)$ s.t. $\overset{+}{n}\ (\delta_i) \leq k$. Note that by induction hypothesis we have that there is a $\delta'_i \in \mathcal{L}^{!}(A)$ s.t.: $\vdash_{\mathsf{A}^{+!}} \delta_i \leftrightarrow \delta'_i$. By SE we have that $\vdash \varphi \leftrightarrow \varphi[\delta_i/\delta'_i]$. Note that $\overset{+}{n}\ (\varphi[\delta_i/\delta'_i]) = 1$. It is easy to see that (Def+) and SE assures the existence of a formula $\psi \in \mathcal{L}^{!}(A)$ for every φ s.t $\overset{+}{n}\ (\varphi) = 1$ satisfying $\vdash_{\mathsf{A}^{+!}} \varphi \leftrightarrow \psi$. In particular, we have that there is a $\psi \in \mathcal{L}^{!}(A)$ s.t.: $\vdash_{\mathsf{A}^{+!}} \varphi[\delta_i/\delta'_i] \leftrightarrow \psi$. By transitivity of $\leftrightarrow$ we have that $\vdash_{\mathsf{A}^{+!}} \varphi \leftrightarrow \psi$.

Remark 8. For every $\varphi \leftrightarrow \psi \in \{(Atoms^{=}) - (Box)\}$ it holds that $Ord(\varphi) > Ord(\psi)$.

Lemma 6. *For every* $\varphi \in \mathcal{L}^{!}(A)$ *s.t.* $Od(\varphi) = 1$ *there is a* $\psi \in \mathcal{L}(A)$ *s.t.* $\vdash_{\mathsf{A}^{+!}} \varphi \leftrightarrow \psi$.

Proof. Suppose $Od(\varphi) = 1$, the rest of the proof is by induction on $Ord(\varphi)$.

For the basic case, suppose $Ord(\varphi) = 0$, then $\varphi \in \mathcal{L}(A)$ and $\vdash \varphi \leftrightarrow \varphi$, so we are done.

Suppose, as induction hypothesis, that for every $\varphi \in \mathcal{L}^{!}(A)$ s.t. $Ord(\varphi) \leq k$ there is a $\psi \in \mathcal{L}(A)$ s.t. $\vdash_{\mathsf{A}^{+!}} \varphi \leftrightarrow \psi$. Now, suppose $Ord(\varphi) = k + 1$. We have, by definition of Ord, that there is $[a!]\delta \in sub(\varphi)$ s.t. $Ord([a!]\delta) = k + 1$. Note that, since $Od(\varphi) = 1$, then $Od([a!]\delta) = 1$ (there are no nested announcements in $[a!]\delta$) and therefore there is an axiom in Table 2 of the form $\vdash_{\mathsf{A}^{+!}} [a!]\delta \leftrightarrow \delta'$. By Remark 8, $Ord([a!]\delta) > Ord(\delta')$. By the induction hypothesis we have that there is a $\sigma \in \mathcal{L}(A)$ s.t. $\vdash_{\mathsf{A}^{+!}} \delta' \leftrightarrow \sigma$. By transitivity of $\leftrightarrow$ we have that $\vdash_{\mathsf{A}^{+!}} [a!]\delta \leftrightarrow \sigma$ and, by SE, $\vdash_{\mathsf{A}^{+!}} \varphi \leftrightarrow \varphi[[a!]\delta/\sigma]$. We can repeat the same argument for every $\delta_i \in sub(\varphi)$ s.t. $Ord(\delta_i) = k + 1$. It is clear than the resulting formula ψ is in $\mathcal{L}(A)$ and that $\vdash_{\mathsf{A}^{+!}} \varphi \leftrightarrow \psi$.

Lemma 7 (From $\mathcal{L}^!(A)$ to $\mathcal{L}(A)$). *For every $\varphi \in \mathcal{L}^!(A)$ there is a $\psi \in \mathcal{L}(A)$ s.t. $\vdash_{\mathsf{A}+!} \varphi \leftrightarrow \psi$*

Proof. By induction on $Od(\varphi)$. The atomic case is straightforward since, if $Od(\varphi) = 0$, then $\varphi \in \mathcal{L}(A)$ and we are done. As for the inductive step, suppose as induction hypothesis that for every $\varphi \in \mathcal{L}^!(A)$ s.t $Od(\varphi) \leq k$ there is a $\psi \in \mathcal{L}(A)$ s.t. $\vdash_{\mathsf{A}+!} \varphi \leftrightarrow \psi$. Suppose $Od(\varphi) = k+1$. Then, there is a $\delta \in sub(\varphi)$ s.t. $Od(\delta) \leq k$. By the induction hypothesis we have that there is a $\delta' \in \mathcal{L}(A)$ s.t. $\vdash_{\mathsf{A}+!} \delta \leftrightarrow \delta'$ and by SE it holds that $\vdash_{\mathsf{A}+!} \varphi \leftrightarrow \varphi[\delta/\delta']$. We can repeat the same argument for every $\delta_i \in sub(\varphi)$ s.t. $Od(\delta_i) \leq k$. Note that, since $Od(\delta'_i) = 0$, we have that $Od(\varphi[\delta_i/\delta'_i]) = 1$ and by SE $\vdash_{\mathsf{A}+!} \varphi \leftrightarrow \varphi[\delta_i/\delta'_i]$. By Lemma 6 we have that there is a $\psi \in \mathcal{L}(A)$ s.t. $\vdash_{\mathsf{A}+!} \varphi[\delta_i/\delta'_i] \leftrightarrow \psi$ and by transitivity of $\leftrightarrow$ it holds that $\vdash_{\mathsf{A}+!} \varphi \leftrightarrow \psi$.

Theorem 2

Proof. We prove completeness for $\mathsf{A}^{!+}$ w.r.t. $\mathcal{M}$, the other two cases are completely analogous. Let $\varphi \in \mathcal{L}^{!+}(A)$, suppose $\vDash \varphi$. By lemmas 5 and 7 and transitivity of $\leftrightarrow$ we have that there is a $\psi \in \mathcal{L}(A)$ s.t. $\vdash_{\mathsf{A}+!} \varphi \leftrightarrow \psi$. From soundness of $\mathsf{A}^{+!}$ and the initial hypothesis it follows that $\vDash \psi$ and, by completeness of A we have that $\vdash_{\mathsf{A}} \psi$. Since $\mathsf{A}^{+!}$ is an extension of A, we have that $\vdash_{\mathsf{A}+!} \psi$. By SE we obtain $\vdash_{\mathsf{A}+!} \varphi$.