Building an Epistemic Logic for Argumentation [*]

François Schwarzentruber[*]   Srdjan Vesic[+]   Tjitze Rienstra[+]

[*] IRISA / ENS Cachan
francois.schwarzentruber@bretagne.ens-cachan.fr
[+] Computer Science and Communication
University of Luxembourg
{srdjan.vesic, tjitze.rienstra}@uni.lu

Abstract. In this paper, we study a multi-agent setting in which each agent is aware of a set of arguments. The agents can discuss and persuade each other by putting forward arguments and counter-arguments. In such a setting, what an agent will do, i.e. what argument she will utter, may depend on what she knows about the knowledge of other agents. For example, an agent does not want to put forward an argument that can easily be attacked, unless she believes that she is able to defend her argument against possible attackers. We propose a logical framework for reasoning about the sets of arguments owned by other agents, their knowledge about other agents' arguments, etc. We do this by defining an epistemic logic for representing their knowledge, which allows us to express a wide range of scenarios.

1   Introduction

Argumentation is the interdisciplinary study of how conclusions can be reached through logical reasoning. In the area of artificial intelligence, argumentation is usually seen as a reasoning approach based on construction and evaluation of arguments. The work of Pollock [10], Vreeswijk [16], and Simari and Loui [13] gave rise to other propositions on how to conceptualise this process. Nowadays, much research on the topic of argumentation is based on the argumentation theory proposed by Dung [4]. It allows to abstract from the origin and the structure of arguments, by representing an argumentation system as a directed graph, whose vertices correspond to arguments and arcs to attacks between them.

It is common that argumentation takes place between multiple agents, having different information and different goals. In such a setting, agents present arguments in order to persuade other agents. Their goal is often to make a certain argument accepted (or rejected). Some efforts were done in studying argumentation dialogues [11, 12] by applying game-theoretic notions. However, those approaches do not allow for reasoning about agents' knowledge, which is one of the essential factors in such a setting and influences agent's behaviour in a major way. For example, when deciding which argument

to utter, an agent may take into account his beliefs about whether another agent has an attacker of that argument. Moreover, an agent may want to reason about the knowledge of another agent. For example: what should I do if he knows that I know that he does not have an attacker of this argument?

In this paper, we define a logical framework for this setting. To do so, we use the epistemic modal logic. We define a logic which allows to formalise a broad spectrum of scenarios concerning the knowledge of agents in form of arguments (e.g. attacks between them) but also the knowledge of agents about the knowledge of other agents, and so on. We also provide a method to speak about the fact that an agent is aware of the existence of an argument.

The remainder of the paper is organised as follows. Section 2 introduces the setting and stresses the importance of the notion of awareness. Section 3 provides a logic to express beliefs about awareness. Section 4 extends this logic for expressing beliefs about the structure of the argumentation graph. Section 5 provides a solution for expressing beliefs about properties of a given argument. The last section concludes.

2 Setting

Since we represent the basic knowledge of agents in form of arguments, we first introduce the notion of an argumentation theory [4] which is used in our formalisation.

Definition 1 (Argumentation system). *An argumentation framework is a pair $\mathcal{A} = (A, \rightsquigarrow)$ where A is a set of arguments and $\rightsquigarrow \subseteq A \times A$ a binary relation. For each pair $(a, b) \in \rightsquigarrow$, we say that a attacks b. We will also sometimes use notation $a \rightsquigarrow b$ instead of $(a, b) \in \rightsquigarrow$.*

We model a situation where a set of agents $\{1, \ldots, n\}$ have different knowledge (in terms of arguments) and beliefs (about the knowledge of other agents). We can model this situation in abstract argumentation theory by representing the arguments and the attack relation between them by what we will call a *big argumentation framework*. We denote this framework by $\texttt{BAF} = (\mathcal{A}_\mathcal{B}, \rightsquigarrow_\mathcal{B})$. The big argumentation framework contains all arguments relevant to a particular discourse. Here we may imagine, for example, that $\texttt{BAF}$ is constructed from all available knowledge and beliefs on a subject such as nuclear energy, and the issue of whether or not we should build more nuclear power plants, or instead close them. The knowledge and opinions in $\texttt{BAF}$ may come, for example, from books, internet, scientific publications, but they may also be completely personal to an agent. Agents are resource bounded and are, in general, not aware of all arguments that belong to the $\texttt{BAF}$. An agent is aware of only those arguments that she has acquainted herself with, or that she has formed, in some way, on the basis of personal considerations or a priori knowledge.

We can thus represent the knowledge of an agent i by a set $A_i \subseteq \mathcal{A}_\mathcal{B}$ of arguments. We assume, however, that all agents use the same logical language in order to understand each other and that they agree on the attack relation. That is, for every pair of arguments $a, b \in \mathcal{A}_\mathcal{B}$, all agents agree on whether or not a is a valid counter-argument to b (or whether a attacks b). So we have a model where all arguments of a particular discourse, and their attack relations, are represented by the big argumentation

framework $\texttt{BAF} = (\mathcal{A}_{\mathcal{B}}, \leadsto_{\mathcal{B}})$, and the knowledge of an agent i is represented by a set $A_i \subseteq \mathcal{A}_{\mathcal{B}}$. This induces, for an agent i, a framework $(A_i, \leadsto_i)$, with $\leadsto_i = \leadsto |_{A_i}$. Note that the formalisation we use, namely the hypothesis that there exists a big argumentation framework and that agents are aware of some arguments from this framework is already present in argumentation literature [11, 14]. In the rest of the paper, we develop logics for reasoning in this setting.

3 The First Attempt of an Epistemic Argumentation Logic

In this section, we propose a framework for representing the fact that different agents are aware of different arguments. Let $AGT = \{1, \ldots, n\}$ be a finite set of agents. The language of this logic, denoted by $\mathcal{L}_1$ is generated by the following BNF:

$$\varphi ::= owns(i, a) \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_i\varphi$$

where $i \in AGT$ is an agent, and $a \in \mathcal{A}_{\mathcal{B}}$ is an argument. For a finite set $S \subseteq \mathcal{A}_{\mathcal{B}}$, with $S = \{a_1, \ldots, a_k\}$, we define an abbreviation $owns(i, S) \stackrel{\text{def}}{=} owns(i, a_1) \wedge \ldots \wedge owns(i, a_k)$.

A formula $owns(i, a)$ means that agent i is aware of the argument a. The meaning of $\neg, \wedge$ and derived connectives $\vee, \rightarrow, \leftrightarrow$ are as usual. A formula $B_i\varphi$ means that agent i believes that φ holds. Some examples of statements that we can express are:

- $owns(1, \{a, b, c\}) \wedge B_1 owns(2, \{a, b\})$ (Agent 1 is aware of a, b and c and believes that agent 2 is aware of a and b.)
- $owns(1, \{a\}) \wedge B_1 B_2 \neg owns(1, \{a\})$ (Agent 1 is aware of a but believes that agent 2 believes he is not.)

The interpretation of the language is based on Kripke structures where states describe possible configurations of argument awareness for all agents. Formally, a state w and an agent i map to a set $D_i \subseteq A$, where D_i is the set of arguments that agent i is aware of in state w. For every agent i, the accessibility relation R_i captures the 'considers possible' relation. Formally:

Definition 2. *An $\mathcal{L}_1$-epistemic argumentation model is a Kripke structure $\mathcal{M} = (W, R, D)$ where:*

- *W is a non-empty set of states;*
- *R maps each agent i to an accessibility relation R_i over W;*
- *D maps each world w and each agent i to a set of arguments $D_i(w)$ such that:*
    1. *for all agents i, for all $w, u \in W$, wR_iu implies $D_i(u) = D_i(w)$.*
    2. *for all agents i, j, for all $w, u \in W$, wR_iu implies $D_j(u) \subseteq D_i(w)$.*

We use the familiar interpretation of belief by taking every R_i to be a KD45 relation [8]. That is, R_i is

- serial: $\forall s \in W, \exists t \in W$ s.t. $t \in R_i(s)$,
- transitive: $\forall s, t, u \in W, t \in R_i(s)$ and $u \in R_i(t)$ implies $u \in R_i(s)$,

– and Euclidean: $\forall s, t, u \in W$, $t \in R_i(s)$ and $u \in R_i(s)$ implies $t \in R_i(u)$,

The truth conditions are as follows:

– $\mathcal{M}, w \models owns(i, a)$ iff $a \in D_i(w)$;
– $\mathcal{M}, w \models B_i\varphi$ iff for all $u \in R_i(w)$, we have $\mathcal{M}, u \models \varphi$;
– $\mathcal{M}, w \models \varphi \wedge \psi$ iff $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$;
– $\mathcal{M}, w \models \neg\varphi$ iff it is not the case that $\mathcal{M}, w \models \varphi$

The two conditions from Definition 2 crucially capture our intuition behind awareness of arguments, as described in the previous section. The first condition says that in every world an agent considers possible, she is aware of the same set of arguments that she is aware of in the actual world. This condition corresponds to the following 'argument awareness introspection' axioms:

– $owns(i, a) \rightarrow B_i owns(i, a)$;
– $\neg owns(i, a) \rightarrow B_i \neg owns(i, a)$

The second condition stipulates that, if an agent is not aware of an argument, she believes that no agent is aware of that argument. This condition corresponds to the following axiom:

– $\neg owns(i, a) \rightarrow B_i \neg owns(j, a)$.

Figure 1 shows a model $\mathcal{M}$ where $\mathcal{M}, s \models owns(1, \{a, b, c\}) \wedge B_1 owns(2, \{a, b\})$. Notice that agent 1 has no belief as to whether or not agent 2 knows c, and agent 2 has no beliefs as to whether agent 1 knows a.
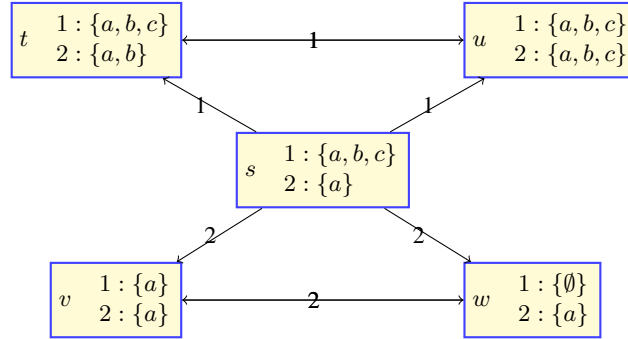

Fig. 1. An $\mathcal{L}_1$-epistemic argumentation logic model.

We say that φ is $\mathcal{L}_1$-satisfiable iff there exists an $\mathcal{L}_1$-argumentation epistemic model $\mathcal{M}$ and a world w such that $\mathcal{M}, w \models \varphi$. The satisfiability problem of a formula of $\mathcal{L}_1$ is the following decision problem:

– input: a formula $\varphi \in \mathcal{L}_1$;

– output: yes iff the formula φ is $\mathcal{L}_1$-satisfiable.

Having an algorithm to solve the satisfiability problem enables us to check consistency automatically. We now study the computational properties of the satisfiability problem of a formula of $\mathcal{L}_1$.

Theorem 1. *If there are more than 3 agents, the satisfiability problem of a formula of $\mathcal{L}_1$ is PSPACE-hard.*

Proof. When there is more than three agents, we can embed the satisfiability problem of $KD45_2$ into the satisfiability problem of a formula of $\mathcal{L}_1$. Let φ be a formula of $KD45_2$. Let i, j be two distinct agents such all agents appearing in φ are in $\{i, j\}$. Let k be a third distinct agent. We suppose the set of arguments to be the set of all propositions appearing in φ. We then define a polynomial translation tr from the $KD45_2$ language to $\mathcal{L}_1$ as follows:

– $tr(p) = owns(k, p)$;
– $tr(B_i\varphi) = B_i tr(\varphi)$;
– $tr(B_j\varphi) = B_j tr(\varphi)$.

We have that φ is $KD45_n$ satisfiable iff $tr(\varphi)$ is satisfiable in an epistemic argumentation model. Note that we must take care to verify the condition of Definition 2 in the sense 'left to right'. We need the extra agent k in order to be able to construct such a epistemic argumentation model. Technical details are left to the reader.

Theorem 2. *The satisfiability problem of a formula of $\mathcal{L}_1$ is in PSPACE.*

Proof. We can embed $\mathcal{L}_1$ into $KD45_n$ and call an optimal procedure for $KD45_n$ plus *distributed belief* which is in PSPACE [8]. Let φ be a formula of $\mathcal{L}_1$. Let m be the modal depth of the formula φ. The embedding works as follows. We add also an operator called *distributed belief* in the language, noted B_{dist}. This operator enables us to express the properties of Definition 3 up to the depth n with a formula of polynomial size in the length of φ. The semantics is defined as follows:

– $\mathcal{M}, w \models B_{dist}\psi$ iff for all $i \in AGT$, for all $u \in R_i(w)$ we have $\mathcal{M}, u \models \psi$.

We denote by $\boldsymbol{B_{dist}}\chi$ the formula $\chi \wedge B_{dist}\chi \wedge B_{dist}^2\chi \wedge \cdots \wedge B_{dist}^m\chi$. It corresponds to common knowledge up to level n, where n is the modal depth of φ. We then consider $tr(\varphi)$ as the conjunction of φ and the following formulas:

– $\boldsymbol{B_{dist}}(owns(i, a) \rightarrow B_i owns(i, a))$
– $\boldsymbol{B_{dist}}(\neg owns(i, a) \rightarrow B_i \neg owns(i, a))$
– $\boldsymbol{B_{dist}}(\neg owns(i, a) \rightarrow B_i \neg owns(j, a))$.

In that way, $tr(\varphi)$ imposes the $KD45_n$-model to satisfy the properties of Definition 3 up to level m. The formula φ is satisfiable in an epistemic argumentation model iff the formula $tr(\varphi)$ is satisfiable in $KD45_n$ plus distributed belief, where constructions of the form $owns(i, a)$ are considered as atomic propositions in $KD45_n$.

4 Expressing Belief About Properties of Arguments

The formalisation presented until now is only the first step towards describing an argument-based dialogue. There are still simple facts that cannot be expressed in the proposed logic. For example, imagine a non-expert person having an idea in some area. She can believe that her idea is interesting, and she is not aware of any attacker of her argument, but she also believes that there is an argument (from an expert) attacking her argument. The problem here is that to express a property about an argument one is not aware of. The next example formalises this consideration.

Example 1. Let us consider the following BAF:
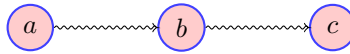
Imagine that agent 1 is not an expert and she has only argument b. The framework proposed in the previous section does not allow to represent a situation in which she has no beliefs about whether this argument is attacked or not. Namely, according to Definition 2, for every model $\mathcal{M}$, for every world w in $\mathcal{M}$ such that agent 1 has exactly the argument b, i.e. where $D_1(w) = \{b\}$, for every world u in $\mathcal{M}$ such that wR_1u, for every agent j, it holds that $a \notin D_j(u)$. That is, in every world of every possible model where agent 1 is aware, agent 1 believes that b is not attacked.

The previous example shows the formalism from the previous section is not expressive enough since it cannot represent the situation where an agent believes that there exists an attacker of one of her arguments, without being able to construct an attacker herself. We start by defining a new language which is richer and allows to speak about attacks between arguments. The solution we propose consists in mixing epistemic modal logic (that we proposed in the previous section) and a logical framework to speak about argumentation graphs, initially proposed by Grossi [7]. Let ATM be a countable set of atomic propositions. The new language is defined as a combination of those two languages:

$$\varphi ::= \langle U \rangle \psi \mid \neg \varphi \mid \varphi \wedge \varphi \mid B_i \varphi$$
$$\psi ::= p \mid \neg \psi \mid \psi \wedge \psi \mid isarg(a) \mid ownedby(i) \mid [attacks]\psi \mid [is\_attacked]\psi$$

where $p \in ATM$, a is an argument of the BAF and i is an agent. We define the language $\mathcal{L}_2$ as the set of formulas obtained with the rule φ. φ-formulas are epistemic modal logic formulas expressing beliefs about facts. The construction $\langle U \rangle \psi$ is read as 'there exists an argument verifying the property ψ'. Then a ψ-formula describes the property of a given argument. Propositions p are used to describe property of arguments, as for instance 'the current argument is about politics'. $isarg(a)$ states that the argument of which we speak now is argument a. $ownedby(i)$ means that the current argument is owned by i. The construction $[attacks]\psi$ means that all arguments that the current argument attacks verify the property ψ. The construction $[is\_attacked]\psi$ means that all arguments that the current argument is attacked by verify the property ψ. We define the following abbreviations: $\langle attacks \rangle \psi = \neg[attacks]\neg\psi$ and $\langle is\_attacked \rangle \psi = \neg[is\_attacked]\neg\psi$.

Example 2. Now, we can say that agent 1 does not have beliefs about whether argument b is attacked or not. We can write this as 'agent 1 does not believe that b is attacked and agent 1 does not believe that b is not attacked': $\neg B_1(\langle U \rangle(isarg(b) \wedge \langle is\_attacked \rangle \top)) \wedge \neg B_1(\langle U \rangle(isarg(b) \wedge \neg \langle is\_attacked \rangle \top))$. As another example, take the following formula which says that agent 1 believes that there exists an argument about global warming owned by the second agent: $B_1(\langle U \rangle(global\_warming \wedge ownedby(2)))$. We can also say that agent 1 does not have an attacker of b but agent 1 believes that agent 2 has an attacker of b on the subject of global warming. It is written as:

$\langle U \rangle(isarg(b) \wedge [is\_attacked]\neg ownedby(1)) \wedge B_1 isarg(b) \wedge \langle is\_attacked \rangle(ownedby(2) \wedge global\_warming)$.

We now define how to interpret formulas of language $\mathcal{L}_2$.

Definition 3. *A $\mathcal{L}_2-$epistemic argumentation model is a Kripke structure $\mathcal{M} = (W, R, \mathcal{A})$ based on a $\mathtt{BAF} = (\mathcal{A}_\mathcal{B}, \rightsquigarrow_\mathcal{B})$ where:*

- *W is a non-empty set of epistemic worlds;*
- *R maps each agent to a serial, transitive and Euclidean relation over W;*
- *$\mathcal{A}$ maps each world w to a labelled argumentation graph $\mathcal{A}_w = (A_w, \rightsquigarrow_w, L_w)$ where:*
  - *A_w is a finite subset of $\mathcal{A}_\mathcal{B} \cup \{?_0, ?_1, \dots\}$;*
  - *$\rightsquigarrow_w \subseteq A_w \times A_w$ is a binary relation such that for $a, b \in \mathcal{A}_\mathcal{B}$, $a\rightsquigarrow_\mathcal{B} b$ if and only if $a \rightsquigarrow_w b$;*
  - *L_w is a map from A_w to $2^{AGT \cup ATM}$.*

Furthermore, we impose:

1. *for all agents i, for all $w, u \in W$, $wR_i u$ implies that $\{a \in A_u \cap \mathcal{A}_\mathcal{B} \mid i \in L_u(a)\} = \{a \in A_w \cap \mathcal{A}_\mathcal{B} \mid i \in L_w(a)\}$;*
2. *for all agents i, for all $w, u \in W$, $wR_i u$ implies that $\{a \in A_u \cap \mathcal{A}_\mathcal{B}\} \subseteq \{a \in A_w \cap \mathcal{A}_\mathcal{B} \mid i \in L_w(a)\}$.*

An example of a model is depicted in Figure 2. The model is still a Kripke model but now, each world w contains an argumentation graph $\mathcal{A}_w = (A_w, \rightsquigarrow_w)$. Each argument of A_w is either an argument from $\mathcal{A}_\mathcal{B}$, or an element of a set $\{?_0, ?_1, \dots\}$. A "question mark argument" denotes an argument the agent is not aware of, that is to say, they can not argue by using it, but they can imagine its existence. The idea is to represent facts like "there is an attacker of argument a" without being able to identify an actual attacker of this argument. $\rightsquigarrow_w$ is the attack relation in $\mathcal{A}_w$. $\rightsquigarrow_w$ is compatible with the attack relation of the $\mathtt{BAF}$ in the following sense: if two arguments a and b of the $\mathtt{BAF}$ appear in the argumentation graph $\mathcal{A}_w$ of the world w, then the attack relation in $\mathcal{A}_w$ is the same as in the $\mathtt{BAF}$. Intuitively, it says that there is no uncertainty concerning the attack of arguments of the $\mathtt{BAF}$ [11]. L is a valuation function that specifies the atomic properties of a (this can be anything, for example a subject of an argument) and the agents that own argument a, for all arguments $a \in A_w$. For example, in Figure 2, in worlds w and u, argument b, owned by agent 1 is about global warming. Condition (1) and condition (2) have the same meaning as condition (1) and (2) in Definition 2.

The truth conditions for φ-formulas are defined as follows:

- $\mathcal{M}, w \models B_i\varphi$ iff for all $u \in R_i(w)$, we have $\mathcal{M}, u \models \varphi$;
- $\mathcal{M}, w \models \langle U \rangle \psi$ iff there exists an argument $a \in A_w$ such that $\mathcal{A}_w, a \models \psi$.

The truth conditions for ψ-formulas are defined as follows:

- $\mathcal{A}_w, a \models p$ iff $p \in ATM$ and $p \in L_w(a)$;
- $\mathcal{A}_w, a \models isarg(b)$ iff $a = b$;
- $\mathcal{A}_w, a \models ownedby(i)$ iff $i \in AGT$ and $i \in L_w(a)$;
- $\mathcal{A}_w, a \models [attacks]\psi$ iff for all b such that $a \rightsquigarrow_w b$ we have $\mathcal{A}_w, b \models \psi$;
- $\mathcal{A}_w, a \models [is\_attacked]\psi$ iff for all b such that $b \rightsquigarrow_w a$ we have $\mathcal{A}_w, b \models \psi$;

Example 3 (Example 1 Cont.). The logic $\mathcal{L}_2$ is expressive enough to overcome problems of Example 1. Let $\alpha = \neg B_1(\langle U \rangle(isarg(b) \wedge \langle is\_attacked \rangle \top)) \wedge \neg B_1(\langle U \rangle(isarg(b) \wedge \neg \langle is\_attacked \rangle \top))$. This formula says that agent 1 does not have beliefs about whether argument b is attacked or not. Let $\mathcal{M}$ be the model from Figure 2, then $\mathcal{M}, w \models \alpha$. This means that in model $\mathcal{M}$ and world w, agent 1 does not have beliefs about whether argument b is attacked or not.
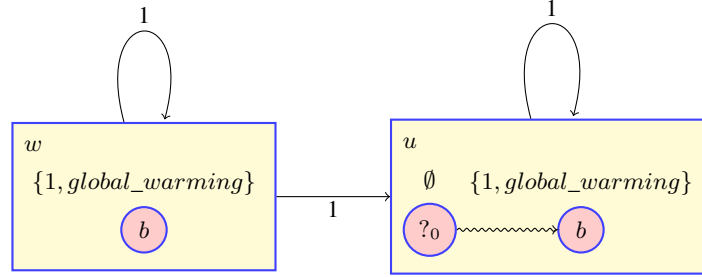

Fig. 2. An $\mathcal{L}_2$-epistemic argumentation logic model.

The language $\mathcal{L}_2$ is a conservative extension of the language $\mathcal{L}_1$. Indeed, we can embed $\mathcal{L}_1$ into $\mathcal{L}_2$ by preserving validities with the following translation:

- $tr(owns(i, a)) = \langle U \rangle(ownedby(i) \wedge isarg(a))$.

We define the notion of $\mathcal{L}_2$-satisfiable formula. A formula φ is $\mathcal{L}_2$-satisfiable iff there exists a big argumentation framework $\mathtt{BAF} = (\mathcal{A}_\mathcal{B}, \rightsquigarrow_\mathcal{B})$ and a $\mathcal{L}_2-$epistemic argumentation model $\mathcal{M} = (W, R, \mathcal{A})$ based on $\mathtt{BAF}$, and a world $\in W$, such that $\mathcal{M}, w \models \varphi$. In the same way, we define the satisfiability problem of $\mathcal{L}_2$.

Theorem 3. *Even if there are no occurrences of arguments in the formula we want to check, the satisfiability problem of $\mathcal{L}_2$ is EXPTIME-hard.*

Proof. The global satisfiability problem of modal logic K is defined as follows:

- input: two formulas φ, ψ where there is only one modal operator $\square$;

– output: yes iff there exists a pointed model $\mathcal{M}, w$ for logic K such that $\mathcal{M}, u \models \varphi$ for all $u \in W$ and $\mathcal{M}, w \models \psi$.

It is EXPTIME-hard. We polynomially reduce the global satisfiability of logic K to the satisfiability of a formula of $\mathcal{L}_2$ in the following way: $\varphi \models_g \psi$ iff $\neg\langle U\rangle\neg tr(\varphi) \wedge \langle U\rangle tr(\psi)$ satisfiable where $tr(\Box\varphi) = [attacks]\varphi$. So the satisfiability problem of $\mathcal{L}_2$ is EXPTIME-hard. All of this works because of the presence of ?-arguments.

Now, concerning the satisfiability problem of $\mathcal{L}_2$, we have a tableau method decision procedure in [3] dealing with nominals (or arguments in our case), a K-operator $\Box$ (or $[attacks]$ in our case) universal modality. In [1], the author explains that the satisfiability problem of the hybrid logic where we add the converse operator and the universal operator is EXPTIME-complete. In our case, it means that given finite sets A, B of formulas of the form $\langle U\rangle\psi$, checking whether $\bigwedge_{\langle U\rangle\psi\in A}\langle U\rangle\psi \wedge \bigwedge_{\langle U\rangle\psi\in B}\neg\langle U\rangle\psi$ is satisfiable can be solved with an EXPTIME procedure. If we combine with a tableau procedure for $KD45_n$ [8] we obtain the following result.

Theorem 4. *The satisfiability problem of $\mathcal{L}_2$ is in EXPTIME.*

Proof. We give the idea of algorithm to solve the satisfiability problem of $\mathcal{L}_2$. Let us consider $\varphi \in \mathcal{L}_2$. Let $arg(\varphi)$ be the set of arguments that appear in the formula φ. We add also an operator called *distributed belief* to the language, denoted B_{dist}, in order to be able to express the properties of Definition 3. The semantics is defined as follows:

– $\mathcal{M}, w \models B_{dist}\varphi$ iff for all $i \in AGT$, for all $u \in R_i(w)$ we have $\mathcal{M}, u \models \varphi$.

We denote by $\boldsymbol{B_{dist}}\chi$ the formula $\chi \wedge B_{dist}\chi \wedge B_{dist}^2\chi \wedge \cdots \wedge B_{dist}^m\chi$. It corresponds to common knowledge up to level m, where m is the modal depth of φ. We denote by $\langle U\rangle SF(\varphi)$ the set of all subformulas of φ of the form $\langle U\rangle\varphi$. Let $Att \in 2^{arg(\varphi)^2}$. We define $T_{Att}(\varphi)$ as the following conjunction, which imposes the constraint of the Definition 3 up to depth m:

– φ;
– $\boldsymbol{B_{dist}}((\langle U\rangle isarg(a) \wedge ownedby(i)) \rightarrow B_i(\langle U\rangle isarg(a) \wedge ownedby(i)))$ for $i \in AGT$;
– $\boldsymbol{B_{dist}}((\langle U\rangle isarg(a) \wedge \neg ownedby(i)) \rightarrow B_i(\neg\langle U\rangle isarg(a)))$ for $i \in AGT$;
– $\boldsymbol{B_{dist}}([U](isarg(a) \rightarrow \neg isarg(b)))$ for all $a, b \in A$ such that $a \neq b$.
– $\boldsymbol{B_{dist}}(\langle U\rangle isarg(a) \wedge \langle U\rangle isarg(b)) \rightarrow \langle U\rangle(isarg(a) \rightarrow \langle attacks\rangle isarg(b))$ for all $(a, b) \in Att$;
– $\boldsymbol{B_{dist}}([U](isarg(a) \rightarrow [attacks]\neg isarg(b)))$ for all $(a, b) \notin Att$;

Now we define the algorithm.

```
for Att ∈ 2^(arg(φ)²)
      for A, B ⊆ ⟨U⟩SF(T_Att(φ))
            PROP[A, B] =
                satsolver_K^(U,converse)(⋀_(⟨U⟩φ∈A) ⟨U⟩ψ ∧ ⋀_(⟨U⟩φ∈B) ¬⟨U⟩φ))
      endFor
      if (modified_KD45_n_tableau_method(T_Att(φ), PROP))
            return sat
      endIf
endFor
return unsat
```

We loop on all possible attack relations Att over arguments that appear in the formula φ. Somehow, we browse all possible BAF. Our aim is then to check if $T_{Att}(\varphi)$ is satisfiable.

The first step consists in computing which subformulas of $T_{Att}(\varphi)$ of the form $\langle U \rangle \psi$ are consistent. Thus, $PROP[A, B]$ will contain 'sat' if $\bigwedge_{\langle U \rangle \psi \in A} \langle U \rangle \psi \wedge \bigwedge_{\langle U \rangle \psi \in B} \neg \langle U \rangle \psi$ is satisfiable and 'unsat' otherwise. The procedure $satsolver\_K^{U,converse}$ is an EXPTIME procedure to solve the satisfiability problem of K plus converse, plus universal modality and hybrid logic. Indeed, arguments are considered as nominals (it is impossible to have two different nodes labelled by the same arguments).

The second step is now to run a tableau method for $KD45_n$ logic on the formula $T_{Att}(\varphi)$. For that, we use the tableau method described in [8]. This tableau method runs as usual for the Boolean connectives and beliefs operators but it considers formulas of the form $\langle U \rangle \psi$ as atoms. We extend this tableau method with a new rule applied on a node w, when all the other rules have already been applied:

– Let A be the set of $\langle U \rangle \psi$ formula written in the node w. Let B be the set of $\langle U \rangle \psi$ such that $\neg \langle U \rangle \psi$. If $PROP[A, B]$ is unsat, we close the current tableau branch.

This modified version of the tableau method runs in PSPACE $\subseteq$ EXPTIME. So the global algorithm runs in EXPTIME. The proof of completeness and soundness of this algorithm are classical.

5 Expressing Properties of Arguments Containing Belief

The logical language developed in Section 4 is powerful enough to enable to speak about arguments without specifying them, but some facts about the framework we study in this paper still cannot be expressed in it. Consider the following example.

Example 4. Agent i owns a and believes that there exists an argument attacked by a which is owned by agent k, but agent j believes that this argument is not owned by k.

Languages $\mathcal{L}_1$ and $\mathcal{L}_2$ do not allow to express statement from the previous example.

We now present a preliminary but promising approach for expressing properties over arguments but also beliefs about those properties. The approach we present here is inspired by other applied logics mixing:

- knowledge and time [5], [9]: the author speaks about moments in time and knowledge about properties of the moment;
- time and space [2]: the author speaks about evolution of objects in the time.

The language $\mathcal{L}_3$ is defined by the following rule:

$$\varphi ::= p \mid isarg(a) \mid ownedby(i) \mid B_i\varphi \mid \langle U \rangle\varphi \mid [attacks]\varphi \mid [is\_attacked]\varphi$$

where $p \in ATM$, $a \in \mathcal{A}_\mathcal{B}$ and $i \in \{1, \dots, n\}$ is an agent. In $\mathcal{L}_3$, we can mix doxastic operator B_i and speaking about arguments. The reading of $B_i\varphi$ is now 'agent i believes φ about the current argument'.

Example 5. The formula $\langle U \rangle(isarg(a) \land ownedby(i) \land B_i\langle attacks \rangle(ownedby(k) \land B_j\neg ownedby(k)))$ is in $\mathcal{L}_3$ and captures the sentence of Example 4.

Definition 4. *$\mathcal{F} = (W, R)$ is a Kripke epistemic frame if and only if W is a non-empty set of possible worlds and $\mathcal{R}$ is a function mapping each agent i to a serial, transitive and Euclidean relation R_i over W. $\mathcal{MA} = (\mathcal{M}, \mathcal{A}, L)$ is a world/argument model if and only if:*

- *$\mathcal{M} = (W, R)$, is a Kripke epistemic frame;*
- *$\mathcal{A} = (A, \rightsquigarrow)$ is an argumentation graph such that $A = \mathcal{A}_\mathcal{B} \cup \{?_0, ?_1, \dots\}$ such that for all $a, b \in \mathcal{A}_\mathcal{B}$, $a \rightsquigarrow_\mathcal{B} b$ iff $a \rightsquigarrow b$;*
- *L maps all couples $(w, a) \in W \times A$ to elements of $2^{AGT \cup ATM}$.*

The truth conditions are:

- $\mathcal{MA}, (w, a) \models p$ iff $p \in ATM$ and $p \in L(w, a)$;
- $\mathcal{MA}, (w, a) \models isarg(b)$ iff $a = b$;
- $\mathcal{MA}, (w, a) \models ownedby(i)$ iff $i \in \{1, \dots, n\}$ and $i \in L(w, a)$;
- $\mathcal{MA}, (w, a) \models B_i\varphi$ iff for all $u \in R_i(w)$, we have $\mathcal{MA}, (u, a) \models \varphi$;
- $\mathcal{MA}, (w, a) \models \langle U \rangle\varphi$ iff there exists $b \in A$, we have $\mathcal{MA}, (w, b) \models \varphi$;
- $\mathcal{MA}, (w, a) \models [attacks]\varphi$ iff for all $b \in A$, $a \rightsquigarrow b$ implies $\mathcal{MA}, (w, b) \models \varphi$;
- $\mathcal{MA}, (w, a) \models [is\_attacked]\varphi$ iff for all $b \in A$, $b \rightsquigarrow a$ implies $\mathcal{MA}, (w, b) \models \varphi$.

Example 6. Figure 3 shows a model for the formula from Examples 4 and 5, namely: $\langle U \rangle(isarg(a) \land ownedby(i) \land B_i\langle attacks \rangle(ownedby(k) \land B_j\neg ownedby(k)))$. We consider a model (on the right) built from the epistemic frame on the left and the BAF in the middle.

In previous sentence, the context is a couple (w, a) where w is a possible epistemic situation and a is an argument. Table 1 illustrates how to follow the nodes of a graph in Figure 3 when analysing the formula from our running example.

Note that this framework contains $KD45 \times K$. The satisfiability problem of $S5 \times K$ is NEXPTIME-complete [6, page 339]. We conjecture that the satisfiability of the logic $KD45 \times K$ is also NEXPTIME-complete so the satisfiability problem in our setting may be NEXPTIME-hard.
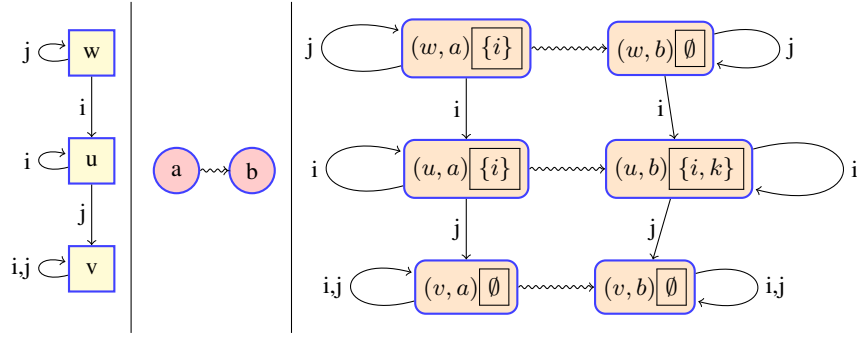
Fig. 3. A world / argument model.

Part of the sentence (syntax)	context world/argument (semantics)
i owns a and	(w, a)
i believes that	all (t, a) such that wR_it, in our case: only (u, a)
there exists an argument x attacked by a...	there exists (u, x) such that $a \rightsquigarrow x$, here: (u, b)
agent j believes that ...	all (t, b) such that uR_jt, in our case: only (v, b)

Table 1. A world / argument model.

6 Summary

In this paper, we provide three languages to deal with argumentation and beliefs. The first one ($\mathcal{L}_1$) enables us to speak about beliefs about awareness of arguments. The second one ($\mathcal{L}_2$) is a conservative extension of $\mathcal{L}_1$ and enables to speak about beliefs about the structure of the argumentation graph. The third logic ($\mathcal{L}_3$) enables to speak about beliefs about a specific argument. The third logic has many promising features, but is not a conservative extension of $\mathcal{L}_2$. A part of our future work will be to investigate whether it is possible to slightly change $\mathcal{L}_3$ in order to make it a conservative extension of $\mathcal{L}_2$.

This paper presents a landscape of incremental logics, in the sense that every logic is more expressive than the previous one. As expected, the complexity of consistency checking of a formula increases. For $\mathcal{L}_1$ it is PSPACE-complete, for $\mathcal{L}_2$, it is EXPTIME-complete and for $\mathcal{L}_3$ we conjecture it to be NEXPTIME-hard. As this complexity is high, a part of our future work will be to study their syntactic fragments.

The paper presents the first attempt to formalise agents' beliefs in a multi-agent argumentation setting. We are inspired by the work of Grossi [7]. That paper shows that an argumentation framework can be seen as a Kripke structure. In our paper, we "import" those ideas on argumentation level (inside of each possible world), and develop a framework for reasoning about agents' beliefs. The solution presented in this paper is the first attempt to model awareness about arguments. An urgent extension of our work

is to make a detailed comparison with the logic of awareness about propositional facts [15].

Growing interest in game theoretic investigations of argument-based dialogues [11, 12, 14] shows that a logical framework is needed to represent knowledge and beliefs of agents in such a setting. We believe that our work is the first step in that direction.

References

1. C. Areces and M. de Rijke. From description to hybrid logics, and back. *Advances in Modal Logic*, 3:17–36, 2001.
2. B. Bennett, A. Cohn, F. Wolter, and M. Zakharyaschev. Multi-dimensional modal logic as a framework for spatio-temporal reasoning. *Applied Intelligence*, 17(3):239–251, 2002.
3. T. Bolander and T. Braüner. Tableau-based decision procedures for hybrid logic. *Journal of Logic and Computation*, 16(6):737–763, 2006.
4. P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence Journal*, 77:321–357, 1995.
5. R. Fagin, Y. Moses, J. Halpern, and M. Vardi. *Reasoning about knowledge*. The MIT Press, 2003.
6. D. Gabbay. *Many-dimensional modal logics: theory and applications*, volume 148. North Holland, 2003.
7. D. Grossi. On the logic of argumentation theory. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'10)*, pages 409–416. IFAAMAS, 2010.
8. J. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial intelligence*, 54(3):319–379, 1992.
9. J. Halpern and M. Vardi. The complexity of reasoning about knowledge and time. i. lower bounds. *Journal of Computer and System Sciences*, 38(1):195–237, 1989.
10. J. Pollock. How to reason defeasibly. *Artificial Intelligence Journal*, 57:1–42, 1992.
11. I. Rahwan and K. Larson. Argumentation and game theory. pages 321–339. Springer, 2009.
12. R. Riveret, H. Prakken, A. Rotolo, and G. Sartor. Heuristics in argumentation: A game theory investigation. In *COMMA*, pages 324–335, 2008.
13. G. Simari and R. Loui. A mathematical treatment of defeasible reasoning and its implementation. *Artificial Intelligence Journal*, 53:125–157, 1992.
14. M. Thimm and A. J. Garcia. Classification and strategical issues of argumentation games on structured argumentation frameworks. In *Proceedings of the Ninth International Joint Conference on Autonomous Agents and Multi-Agent Systems 2010 (AAMAS'10)*, 2010.
15. H. van Ditmarsch and T. French. Becoming aware of propositional variables. *Logic and Its Applications*, pages 204–218, 2011.
16. G. Vreeswijk. Abstract argumentation systems. *Artificial Intelligence Journal*, 90:225–279, 1997.