Epistemic Argumentation Framework

Chiaki Sakama[1] and Tran Cao Son[2]

[1] Wakayama University, Wakayama 640 8510, Japan
sakama@wakayama-u.ac.jp
[2] New Mexico State University, Las Cruces, NM 88003, USA
tson@cs.nmsu.edu

Abstract. The paper introduces the notion of an *epistemic argumentation framework* (EAF) as a means to integrate the beliefs of a reasoner with argumentation. Intuitively, an EAF encodes the beliefs of an agent who reasons about arguments. Formally, an EAF is a pair of an argumentation framework and an *epistemic constraint*. The semantics of the EAF is defined by the notion of an ω-*epistemic labelling set*, where ω is complete, stable, grounded, or preferred, which is a set of ω-labellings that collectively satisfies the epistemic constraint of the EAF. The paper shows how EAF can represent different views of reasoners on the same argumentation framework. It also includes representing preferences in EAF and multi-agent argumentation. Finally, the paper discusses the complexity of the problem of determining whether or not an ω-epistemic labelling set exists.

Keywords: argumentation framework · epistemic information · multiple agents · preferences

1 Introduction

Rational agents often claim that they make their decision based on their knowledge and beliefs when facing alternative and conflicting choices. Consider two examples:

- On January 15, 2019, British Prime Minister's Theresa May suffered a humiliating defeat in the vote on the Brexit deal; 432 Members of Parliament (MPs) voted against the deal while 202 were for it.[3] The MPs who voted against the deal believe that the deal is bad for Britain. Those who voted for the deal believe that the deal is the best that Britain can get.
- In the US presidential election, a voter selects one candidate from a set of candidates (often only two candidates). Everyone claims that he/she has made the "right choice."

In each scenario above, an agent (an MP or a voter) listens to various arguments, which either support or reject a potential decision, and then opts for one among the possibilities, which he/she believes is the right choice. In each situation, the arguments supporting/against a choice, their counter-arguments, etc. can be easily encoded in an *abstract argumentation framework* (AF) introduced in [12]. For instance, $AF = (\{(a)ccept,\ (r)eject\}, \{(a,r),(r,a)\})$, having two arguments mutually attacking each other, represents (in its most condensed form) the AF that the MPs have for

[3] "*Brexit vote*", Jan. 15th, 2019. washingtonpost.com

making their choice about the Brexit's deal. Given arguments made by each agent in each scenario, an argumentation semantics of the corresponding AF provides the result of rational reasoning. The stable semantics of the above AF supports two alternative choices, while the ground semantics of the AF supports "no decision". As such, it would likely result in the unanimous choice by all agents who participate in argumentation and claim that they are rational.

The above discussion raises the question "how to express an agent's opinion for supporting an argument among conflicting arguments in the outcome of an AF?" Arguably, there are two possibilities: the agent modifies the AF so that the new AF supports his/her choice or the agent is simply biased towards his/her conclusion. In the first case, nothing other than the agent's beliefs could influence his/her choice of arguments and/or attacks that lead to the new AF, which ultimately leads to his/her conclusion. In this approach, a modified AF represents objective evidences and subjective beliefs indistinguishably. If one merges objective evidences (normally invariant) and subjective beliefs (possibly variant) in a single AF, however, it must be revised whenever an agent changes its own belief. Moreover, it would become hard to distinguish subjective beliefs from objective evidences in a personally customized AF. In this respect, it is desirable to have a mechanism that can distinguishably represent subjective beliefs (or biases) of agents as well as objective evidences as an AF.

In the second case, biases, reflecting beliefs of agents, could be viewed as agents' preferences. Furthermore, there is a huge amount of literature in AF on dealing with preferences in argumentation. It is therefore instructive to consider whether previously developed approaches to dealing with preferences would be sufficient to capture biases. In most approaches in abstract AF, the key idea is to extend an AF with a syntactic component that records the preferences such as a preference relation among arguments or an attack relation between arguments and attacks, and then define a new semantics for this extended AF (detailed discussion is in Section 4). Approaches to dealing with preferences have thus far only considered biases/preferences between arguments (e.g., prefer an argument over another one) or preferences between arguments and attacks. However, it is difficult to apply those approaches to represent preferences in a complicated situation. Suppose the following scenario: a person, who goes to a restaurant, has a preference on the combination of food and drink: white wine for fish and red wine for meat. However, the person wants no red wine other than French one, so he/she will take white wine for meat if French red wine is unavailable. It is hard to specify such conditional preference using preference relations among individual arguments. Then we represent preferences as a formula over epistemic literals.

In this paper, we propose an approach to incorporate agents' *beliefs* into an argumentation framework (AF). Specifically, we propose an extension of AF, called *epistemic argumentation framework* (EAF). EAF introduces the third component to an AF, *an epistemic constraint*, that represents the belief of an agent given an AF. We study formal properties of EAF and show that it can be used in representing preferences and decision making in multiagent environments. We also investigate computational complexity and discuss related issues. The rest of the paper is organized as follows. Section 2 reviews basic notions of argumentation frameworks used in this paper. Section 3 introduces epistemic argumentation frameworks and addresses its applications.

Section 4 discusses related issues and Section 5 concludes the paper. Due to space limit, we omit proofs of propositions, which will be provided in the full paper.

2 Argumentation Framework

This paper uses *(abstract) argumentation frameworks* introduced by [12].

An *argumentation framework* (AF) is a pair (Ar, att) where Ar is a (finite) set of *arguments* and $att \subseteq Ar \times Ar$. We write $a \rightarrow b$ (say, *a attacks b*) iff $(a, b) \in att$. We say that a *indirectly attacks* b if there is a finite sequence $x_0, ..., x_{2n+1}$ $(n \geq 1)$ such that $a = x_0$ and $b = x_{2n+1}$ and for each $0 \leq i \leq 2n$, $(x_i, x_{i+1}) \in att$.

For the semantics of AFs, we use the labelling-based semantics [10]. A *labelling* of (Ar, att) is a (total) function $\mathcal{L} : Ar \rightarrow \{\, \texttt{in}, \texttt{out}, \texttt{und} \,\}$. When $\mathcal{L}(a) = \texttt{in}$ (resp. $\mathcal{L}(a) = \texttt{out}$ or $\mathcal{L}(a) = \texttt{und}$) for an argument $a \in Ar$, it is written as $\texttt{in}(a)$ (resp. $\texttt{out}(a)$ or $\texttt{und}(a)$). In this case, the argument a is said to be *accepted* (resp. *rejected* or *undecided*) in $\mathcal{L}$. Given $AF = (Ar, att)$ and a labelling $\mathcal{L}$, define $\texttt{in}(\mathcal{L}) = \{\, x \mid \mathcal{L}(x) = \texttt{in} \text{ for } x \in Ar \,\}$, $\texttt{out}(\mathcal{L}) = \{\, x \mid \mathcal{L}(x) = \texttt{out} \text{ for } x \in Ar \,\}$, and $\texttt{und}(\mathcal{L}) = \{\, x \mid \mathcal{L}(x) = \texttt{und} \text{ for } x \in Ar \,\}$. A labelling $\mathcal{L}$ of (Ar, att) is also represented as a set $S(\mathcal{L}) = \{\, \lambda(x) \mid \mathcal{L}(x) = \lambda \text{ for } x \in Ar \,\}$. We say that $\lambda(x)$ represents the *justification state* of $x \in Ar$.

A labelling $\mathcal{L}$ of $AF = (Ar, att)$ is a *complete labelling* if for each argument $a \in Ar$, it holds that:

- $\mathcal{L}(a) = \texttt{in}$ iff $\mathcal{L}(b) = \texttt{out}$ for every $b \in Ar$ such that $(b, a) \in att$.
- $\mathcal{L}(a) = \texttt{out}$ iff $\mathcal{L}(b) = \texttt{in}$ for at least one $b \in Ar$ such that $(b, a) \in att$.
- $\mathcal{L}(a) = \texttt{und}$, otherwise.

Let $\mathcal{L}$ be a complete labelling of AF. Then,

- $\mathcal{L}$ is a *stable labelling* iff $\texttt{und}(\mathcal{L}) = \emptyset$.
- $\mathcal{L}$ is a *grounded labelling* iff $\texttt{in}(\mathcal{L}) \subseteq \texttt{in}(\mathcal{L}')$ for any complete labelling $\mathcal{L}'$ of AF.
- $\mathcal{L}$ is a *preferred labelling* iff there is no complete labelling $\mathcal{L}'$ of AF such that $\texttt{in}(\mathcal{L}) \subset \texttt{in}(\mathcal{L}')$.

We often abbreviate complete, stable, grounded, and preferred labelling as *co, st, gr*, and *pr*, respectively.

3 Epistemic Argumentation Framework

3.1 Epistemic Labelling Set

Given $AF = (Ar, att)$, define $\mathcal{A}_{AF} = \{\, \texttt{in}(a), \texttt{out}(a), \texttt{und}(a) \mid a \in Ar \,\}$. An *epistemic atom* over AF is of the form $\mathbf{K}\,\varphi$ or $\mathbf{M}\,\varphi$ where φ is a propositional formula over $\mathcal{A}_{AF}$. An *epistemic literal* is an epistemic atom or its negation. An *epistemic formula* (over $\mathcal{A}_{AF}$) is a propositional formula constructed over epistemic literals together with $\top$ (true) and $\bot$ (false). Intuitively, $\mathbf{K}\,\varphi$ (resp. $\mathbf{M}\,\varphi$) states that the agent believes that φ

is *true* (resp. *possibly true*).[4] We will use epistemic formulas to represent the epistemic side of an agent given an AF.

Let φ be a propositional formula over $\mathcal{A}_{AF}$ and $\mathcal{L}$ be a labelling over AF. Then $S(\mathcal{L})$ is considered an interpretation of φ. We say that φ is true in $\mathcal{L}$, denoted by $\mathcal{L} \models \varphi$, if φ is interpreted to be true under $S(\mathcal{L})$.

Definition 1 (satisfaction) A set SL of labellings *satisfies* an epistemic formula φ, denoted by $SL \models \varphi$, if one of the following conditions holds:

(i) $\varphi = \top$,
(ii) $\varphi = \mathbf{K}\,\psi$ and $\mathcal{L} \models \psi$ for every $\mathcal{L} \in SL$,
(iii) $\varphi = \mathbf{M}\,\psi$ and $\mathcal{L} \models \psi$ for some $\mathcal{L} \in SL$,
(iv) $\varphi = \neg\psi$ and $SL \not\models \psi$,
(v) $\varphi = \varphi_1 \wedge \varphi_2$ and $(SL \models \varphi_1$ and $SL \models \varphi_2)$,
(vi) $\varphi = \varphi_1 \vee \varphi_2$ and $(SL \models \varphi_1$ or $SL \models \varphi_2)$.

An epistemic formula φ is *consistent* if there exists a (non-empty) set SL of labellings such that $SL \models \varphi$; otherwise, φ is *inconsistent*. Some basic properties are addressed.

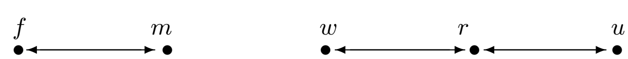
Proposition 1 *Let SL be a set of labellings. For any propositional formula φ and ψ over $\mathcal{A}_{AF}$,*

(i) $SL \models \neg\mathbf{M}\,\varphi$ *iff* $SL \models \mathbf{K}\,\neg\varphi$,
(ii) $SL \models \neg\mathbf{K}\,\varphi$ *iff* $SL \models \mathbf{M}\,\neg\varphi$,
(iii) $SL \models \mathbf{M}\,(\varphi \vee \psi)$ *iff* $SL \models \mathbf{M}\,\varphi \vee \mathbf{M}\,\psi$,
(iv) $SL \models \mathbf{K}\,(\varphi \wedge \psi)$ *iff* $SL \models \mathbf{K}\,\varphi \wedge \mathbf{K}\,\psi$.

Definition 2 (epistemic argumentation framework) An *epistemic argumentation framework* (EAF) is a triple (Ar, att, φ) where $AF = (Ar, att)$ is an argumentation framework and φ is an epistemic formula (called an *epistemic constraint*).

Intuitively, an EAF (Ar, att, φ) represents the view of an agent who, given $AF = (Ar, att)$, believes that φ is true. So, an EAF consists of two different types of information: an objective evidence AF and a subjective belief φ of an agent. We also refer to an EAF by (AF, φ) whenever it is clear from the context what AF refers to.

Example 1 In the introductory example, consider an AF with the set of arguments $\{(f)ish,\ (m)eat,\ (w)hite,\ (r)ed,\ (u)navailable\}$ and the set of attacks $\{(f,m),\ (m,f),\ (w,r),\ (r,w),\ (r,u),\ (u,r)\}$.



Then, some EAFs are defined as follows:

- $EAF_1 = (AF, \mathbf{M}\,\mathtt{in}(r))$ represents the view of an agent who believes that r is possibly accepted.

[4] By the meaning, it might be better to write $\mathbf{B}\varphi$ rather than $\mathbf{K}\,\varphi$, but we use $\mathbf{K}$ because we implement it using epistemic logic programs in which $\mathbf{K}$ and $\mathbf{M}$ are used (see Section 5).

- $EAF_2 = (AF, \mathbf{K}\,\mathtt{in}(w) \vee \mathbf{K}\,\mathtt{in}(r))$ represents the view of an agent who believes that either w or r should be accepted.
- EAF_3=$(AF, \mathbf{K}(\mathtt{in}(m) \wedge \neg\mathtt{in}(u) \rightarrow \mathtt{in}(r)) \wedge \mathbf{K}(\mathtt{in}(f) \rightarrow \mathtt{in}(w)))$ represents the view of an agent whose belief is given by the statement: "*if m is accepted and u is unaccepted, then r should be accepted; and if f is accepted then w should be accepted*."

Next we define the semantics of an EAF.

Definition 3 (epistemic labelling set) Let $EAF = (AF, \varphi)$ and $\omega \in \{$*co, st, gr, pr*$\}$. A set SL of labellings is an ω-*epistemic labelling set* of (AF, φ) if (i) each $\mathcal{L} \in SL$ is an ω-labelling of AF, and (ii) SL is a $\subseteq$-maximal set of ω-labellings of AF that satisfies φ. An EAF possibly has multiple ω-epistemic labelling sets.

Intuitively, an ω-epistemic labelling set is a collection of ω-labellings that reflects the belief of an agent. In particular, $EAF = (AF, \top)$ has the unique ω-epistemic labelling set that coincides with the set of ω-labellings of AF. In what follows, we assume $\omega \in \{$*co, st, gr, pr*$\}$ unless stated otherwise. By definition, EAF always has an ω-epistemic labelling set (possibly as an empty set).

Proposition 2 *$EAF = (AF, \bot)$ has the ω-epistemic labelling set $\emptyset$.*

Our primary interest is an EAF that has non-empty ω-epistemic labelling sets.

Example 2 Consider the EAFs of Example 1 under the stable semantics. First, AF in the EAFs has four stable labellings:

$$L_1 = \{\, \mathtt{in}(f),\, \mathtt{out}(m),\, \mathtt{out}(w),\, \mathtt{in}(r),\, \mathtt{out}(u)\, \},$$
$$L_2 = \{\, \mathtt{out}(f),\, \mathtt{in}(m),\, \mathtt{out}(w),\, \mathtt{in}(r),\, \mathtt{out}(u)\, \},$$
$$L_3 = \{\, \mathtt{in}(f),\, \mathtt{out}(m),\, \mathtt{in}(w),\, \mathtt{out}(r),\, \mathtt{in}(u)\, \},$$
$$L_4 = \{\, \mathtt{out}(f),\, \mathtt{in}(m),\, \mathtt{in}(w),\, \mathtt{out}(r),\, \mathtt{in}(u)\, \}.$$

This implies EAF_1 has a unique stable epistemic labelling set $\{L_1, L_2, L_3, L_4\}$; EAF_2 has two stable epistemic labelling sets $\{L_1, L_2\}$ and $\{L_3, L_4\}$; and EAF_3 has a unique stable epistemic labelling set $\{L_2, L_3, L_4\}$. Suppose that it turns that French red wine is unavailable. The situation is represented by

$$EAF_4 = (AF,\, \mathbf{K}(\mathtt{in}(m) \wedge \neg\mathtt{in}(u) \rightarrow \mathtt{in}(r)) \,\wedge\, \mathbf{K}(\mathtt{in}(f) \rightarrow \mathtt{in}(w)) \,\wedge\, \mathbf{K}\,\mathtt{in}(u)\,).$$

Then EAF_4 has a unique stable epistemic labelling set $\{L_3, L_4\}$.

As shown in the above example, EAF can represent belief change of an agent by revising an epistemic constraint without modifying AF. The revised EAF then produces new epistemic labelling sets that reflect new belief states of an agent. In Example 2, EAF_4 introduces an additional constraint $\mathbf{K}\,\mathtt{in}(u)$ to EAF_3, which results in eliminating L_2 from the stable epistemic labelling set of EAF_3. For two epistemic formulas φ_1 and φ_2, we say that φ_1 is *stronger* than φ_2 if $\varphi_1 \models \varphi_2$ (in the sense of classical logic). Introducing a stronger constraint to EAF eliminates elements of SL in general.

Proposition 3 *Let $EAF_1 = (AF, \varphi_1)$ and $EAF_2 = (AF, \varphi_2)$ be two EAFs such that φ_1 is stronger than φ_2. Then, for each ω-epistemic labelling set SL_1 of EAF_1 there exists some ω-epistemic labelling set SL_2 of EAF_2 such that $SL_1 \subseteq SL_2$.*

In argumentation frameworks, stable, grounded, or preferred labellings are complete labellings. In epistemic argumentation frameworks, a similar result holds.

Proposition 4 *Let (AF, φ) be an EAF. If a non-empty set SL of labellings is a stable, grounded, or preferred epistemic labelling set of (AF, φ), then $\mathcal{L} \in SL$ is an element of a complete epistemic labelling set of (AF, φ).*

We next consider a sufficient condition for the uniqueness of ω-epistemic labelling sets.

Lemma 5 *Let φ be a conjunction of epistemic literals over $\mathcal{A}_{AF}$. If two sets of labellings SL_1 and SL_2 satisfy φ (i.e., $SL_1 \models \varphi$ and $SL_2 \models \varphi$), then $SL_1 \cup SL_2 \models \varphi$.*

Using the lemma, we can prove the next result.

Proposition 6 *Let (AF, φ) be an EAF such that φ is a conjunction of epistemic literals. Then (AF, φ) has a unique ω-epistemic labelling set.*

Assume that φ is a DNF in which each disjunct is a conjunction of epistemic literals. Due to Proposition 1, we can assume that each disjunct in φ is of the form $\mathbf{K}\psi_0 \wedge \mathbf{M}\psi_1 \wedge \cdots \wedge \mathbf{M}\,\psi_n$[5] where ψ_i $(0 \leq i \leq n)$ is a propositional formula over $\mathcal{A}_{AF}$, which will be denoted by $EC(\psi_0; \psi_1, \ldots, \psi_n)$. We can prove:

Lemma 7 *Let SL be a set of labellings such that $SL \models EC(\psi_0; \psi_1, \ldots, \psi_n)$. Then, for each $i = 1, \ldots, n$, there exists some $\mathcal{L} \in SL$ such that $\mathcal{L} \models \psi_0 \wedge \psi_i$.*

Proposition 8 *Let $\varphi = \bigvee_{j=1}^{k} EC(\psi_j; \psi_1^j, \ldots, \psi_{n_j}^j)$ $(k \geq 1)$ be an epistemic formula. Then, EAF=(AF, φ) has a non-empty ω-epistemic labelling set if there exists an integer j $(1 \leq j \leq k)$ such that for each $1 \leq i \leq n_j$, AF has an ω-labelling $\mathcal{L}$ and $\mathcal{L} \models \psi_j \wedge \psi_i^j$.*

Each AF semantics imposes some specific condition on every argument, e.g., the stable semantics allows no argument to be undecided, while the grounded semantics keeps controversial arguments undecided. EAF is useful for selecting intended labellings from the set of all possible labellings.

Example 3 Consider the AF in Example 1. Since the availability of French red wine is unknown before visiting a restaurant, an agent wants to keep the argument u undecided. The situation is specified as the epistemic constraint $\varphi = \mathbf{K}\mathtt{und}(u)$. Then (AF, φ) has the single preferred epistemic labelling set $\{\{\mathtt{in}(f), \mathtt{out}(m), \mathtt{und}(w), \mathtt{und}(r), \mathtt{und}(u)\}, \{\mathtt{out}(f), \mathtt{in}(m), \mathtt{und}(w), \mathtt{und}(r), \mathtt{und}(u)\}\}$.

[5] $\neg\mathbf{M}\,\psi$ (resp. $\neg\mathbf{K}\,\psi$) is converted to $\mathbf{K}\,\neg\psi$ (resp. $\mathbf{M}\,\neg\psi$), and $\mathbf{K}\,\psi_1 \wedge \mathbf{K}\,\psi_2$ is converted to $\mathbf{K}\,(\psi_1 \wedge \psi_2)$.

3.2 Representing Preference

Preference among arguments can be specified in EAF as follows. Let $\succeq$ be a pre-order (i.e., reflexive and transitive) relation over $Ar \times Ar$ such that $(x, y) \in \succeq$ implies that x indirectly attacks y or vice versa. $x \succeq y$ means that an argument x is at least as preferred as y. We write $x \succ y$ if $x \succeq y$ and $y \not\succeq x$.

Definition 4 (preference over arguments) Given $AF = (Ar, att)$ and a preorder relation $\succeq \subseteq Ar \times Ar$, define $EAF = (AF, \varphi_A)$ where

$$\varphi_A = \bigwedge_{x \succ y} \mathbf{K}(\mathtt{in}(y) \supset \mathtt{in}(x)).$$

Intuitively speaking, φ_A represents that an argument x should be accepted whenever another argument y of lower preference is accepted. Note that the preference is specified as $x \succ y$ but not as $x \succeq y$ in φ_A. When both $x \succeq y$ and $y \succeq x$ exist, there is no reason to prefer one of them. In this case, the conjunct involved x and y in φ_A is $\top$.

Proposition 9 *Let $EAF = (AF, \varphi_A)$ be an EAF defined as above. Then, for any ω-epistemic labelling set SL of EAF, there is no $\mathcal{L} \in SL$ such that $\mathtt{in}(x) \notin \mathcal{L}$ and $\mathtt{in}(y) \in \mathcal{L}$ for any $x \succ y$.*

Example 4 Consider $AF = (\{a, r\}, \{(a, r), (r, a)\})$ with $r \succ a$. Then $EAF = (AF, \varphi_A)$ with $\varphi_A = \mathbf{K}(\mathtt{in}(a) \supset \mathtt{in}(r))$ has the unique stable epistemic labelling set $\{\{\mathtt{in}(r), \mathtt{out}(a)\}\}$, and the unique complete epistemic labelling set $\{\{\mathtt{in}(r), \mathtt{out}(a)\}, \{\mathtt{und}(r), \mathtt{und}(a)\}\}$.

In Example 4, the complete epistemic labelling set contains $\{\mathtt{und}(r), \mathtt{und}(a)\}$. This can be eliminated by introducing the constraint $\varphi_A = \mathbf{K}(\mathtt{in}(a) \vee \mathtt{und}(a) \supset \mathtt{in}(r))$.

Preference over arguments is generalized to preference over justification states of arguments as follows. A pre-order relation $\sqsupseteq$ over justification states of arguments is a collection of elements of the form $\lambda(x) \sqsupseteq \mu(y)$ where $\lambda,\ \mu \in \{\mathtt{in}, \mathtt{out}, \mathtt{und}\}$, meaning that $\lambda(x)$ is at least as preferred as $\mu(y)$ for arguments x and y. We write $\lambda(x) \sqsupset \mu(y)$ if $\lambda(x) \sqsupseteq \mu(y)$ and $\mu(y) \not\sqsupseteq \lambda(x)$.

Definition 5 (preference over justification states) Given $AF = (Ar, att)$ and a preorder relation $\sqsupseteq \subseteq \mathcal{A}_{AF} \times \mathcal{A}_{AF}$, define $EAF = (AF, \varphi_J)$ where

$$\varphi_J = \bigwedge_{\lambda(x) \sqsupset \mu(y)} \mathbf{K}\,(\mu(y) \supset \lambda(x)).$$

φ_J states that if the justification state $\lambda(x)$ is preferred to $\mu(y)$ for $x, y \in Ar$, then $\mathcal{L} \models \mu(x)$ implies $\mathcal{L} \models \lambda(x)$ for any $\mathcal{L} \in SL$ where SL is any ω-epistemic labelling set of EAF.

By definition, Def. 4 is considered a special case of Def. 5 with $\mu = \lambda = \mathtt{in}$.

Proposition 10 *Let $EAF = (AF, \varphi_J)$ be an EAF defined as above. Then, for any ω-epistemic labelling set SL of EAF, there is no $\mathcal{L} \in SL$ such that $\lambda(x) \notin \mathcal{L}$ and $\mu(y) \in \mathcal{L}$ for any $\lambda(x) \sqsupset \mu(y)$. In particular, $\mu(y) \notin \mathcal{L}$ for any $\mathcal{L} \in SL$ if $x = y$.*

Example 5 Suppose that in Example 4, an MP prefers keeping the decision undecided if possible. This is represented by $\sqsupset = \{(\texttt{und}(x), \texttt{in}(x)), (\texttt{und}(x), \texttt{out}(x)) \mid x \in \{a, r\}\}$ which is translated to $\varphi_J = \bigwedge_{x \in \{a,r\}} \mathbf{K}\, (\texttt{in}(x) \supset \texttt{und}(x)) \wedge \mathbf{K}\, (\texttt{out}(x) \supset \texttt{und}(x))$. Then $EAF = (AF, \varphi_J)$ has the unique complete epistemic labelling set $\{\{\texttt{und}(r), \texttt{und}(a)\}\}$. Furthermore, $\emptyset$ is the stable epistemic labelling set, since there is no choice to make a and r undecided.

In this way, EAF enables us to specify preference over not only arguments but also justification states of arguments. Furthermore, it could also be useful to introduce preferences among epistemic formulas. For instance, we could write $\mathbf{K}\, \lambda(x) > \mathbf{K}\, \mu(x)$ for some argument x to indicate that we prefer SL_1 over SL_2 whenever $SL_1 \models \mathbf{K}\, \lambda(x)$ and $SL_2 \models \mathbf{K}\mu(x)$ for two arbitrary ω-epistemic labelling sets SL_1 and SL_2. We leave such extensions for future work.

3.3 Multiple Agents

Suppose that two agents share $AF = (\{a, r\}, \{(a, r), (r, a)\})$. If they have the same belief represented by the epistemic constraint $\varphi = \mathbf{K}\texttt{in}(a)$, the EAF (AF, φ) has the single epistemic complete labelling set $\{\{\texttt{in}(a), \texttt{out}(r)\}\}$ and the agents agree on accepting a. On the other hand, if two agents have conflicting beliefs $\varphi_1 = \mathbf{K}\texttt{in}(a)$ and $\varphi_2 = \neg\mathbf{K}\texttt{in}(a)$ respectively, then they do not agree on accepting a or r. In this section, we assume multiple agents who share the same AF while having different beliefs in general. The situation is represented by the collection of EAFs (AF, φ_i) $(1 \leq i \leq n)$. First, we define two different types of agreements.

Definition 6 (agreement) Let $AF = (Ar, att)$ and $EAF_1 = (AF, \varphi_1), \ldots, EAF_n = (AF, \varphi_n)$ $(n \geq 1)$. Then $EAF_1, \ldots, EAF_n$ *credulously agree on* $\lambda(a)$ for $a \in Ar$ where $\lambda \in \{\texttt{in}, \texttt{out}, \texttt{und}\}$ *under ω-epistemic labelling* if each EAF_i $(i = 1, \ldots, n)$ has an ω-epistemic labelling set SL_i such that $SL_i \models \mathbf{M}\lambda(a)$. In contrast, $EAF_1, \ldots, EAF_n$ *skeptically agree on* $\lambda(a)$ *under ω-epistemic labelling* if for any ω-epistemic labelling set SL_i of EAF_i $(i = 1, \ldots, n)$ $SL_i \models \mathbf{K}\lambda(a)$.

The above definition characterizes two different situations (credulous or skeptical) in which agents reach an agreement on $\lambda(a)$. For simplicity reasons, Def.6 assumes that different agents employ the same ω-epistemic labelling, but the definition is easily extended to a case in which agents employ different ω-labellings.

Proposition 11 *Let $AF = (Ar, att)$ and $EAF_1 = (AF, \varphi_1)$, …, $EAF_n = (AF, \varphi_n)$ $(n \geq 1)$. Then, $EAF_1, \ldots, EAF_n$ skeptically agree on $\lambda(a)$ for $a \in Ar$ under ω-epistemic labelling iff EAF_i and $EAF'_i = (AF, \varphi_i \wedge \mathbf{K}\, \lambda(a))$ $(i = 1, \ldots, n)$ have the same ω-epistemic labelling sets.*

Proposition 12 *Let $AF = (Ar, att)$ and $EAF_1 = (AF, \varphi_1)$, …, $EAF_n = (AF, \varphi_n)$ $(n \geq 1)$. If $EAF_1, \ldots, EAF_n$ credulously agree on $\lambda(a)$ for $a \in Ar$ under ω-epistemic labelling, then $(AF, \varphi_1 \vee \cdots \vee \varphi_n)$ has an ω-epistemic labelling set SL such that $SL \models \mathbf{M}\lambda(a)$. Conversely, if $(AF, \varphi_1 \wedge \cdots \wedge \varphi_n)$ has an ω-epistemic labelling set SL such that $SL \models \mathbf{M}\lambda(a)$, then $EAF_1, \ldots, EAF_n$ credulously agree on $\lambda(a)$ under ω-epistemic labelling.*

Algorithm 1: $\texttt{Existence}(EAF, \omega)$

Input: ω, $EAF = (AF, \varphi)$.
Output: **true** if EAF has a (non-empty) ω-epistemic labelling set; **false** otherwise.
1 Convert to DNF: $\varphi = \vee_{j=1}^{k} EC(\psi_j; \psi_1^j, \ldots, \psi_{n_j}^j)$
2 where $EC(\psi; \psi_1, \ldots, \psi_k) = \mathbf{K}\psi \wedge \bigwedge_{i=1}^{k} \mathbf{M}\psi_i$
3 **for** $j = 1$ *to* k **do**
4 $num\_labelling := 0$
5 **for** $i = 1$ *to* n_j **do**
6 **if** $\texttt{D}(\omega, AF, \psi_j \wedge \psi_i^j) = \textbf{true}$ **then**
7 $num\_labelling := num\_labelling + 1$
8 **if** $num\_labelling = n_j$ **then return true**
9 **return false**

We next show that EAF can be used for formalizing majority voting. In the presence of $EAF_i = (AF, \varphi_i)$ $(1 \leq i \leq n)$, define:

$$M_\psi^\omega = \{\, i \mid EAF_i \text{ has an } \omega\text{-epistemic labelling set } SL \text{ s.t. } SL \models \mathbf{M}\,\psi \,\},$$
$$N_\psi^\omega = \{\, i \mid \text{for each } \omega\text{-epistemic labelling set } SL \text{ of } EAF_i,\ SL \models \mathbf{K}\,\psi \,\}.$$

Definition 7 (majority voting) Let $AF{=}(Ar, att)$ and $EAF_i = (AF, \varphi_i)$ for $(1 \leq i \leq n)$. For $a \in Ar$, $\lambda(a)$ is *credulously* (resp. *skeptically*) adopted by *majority voting under ω-epistemic labelling* iff the cardinality of the set $M^\omega_{\lambda(a)}$ (resp. $N^\omega_{\lambda(a)}$) is greater than the cardinality of the set $M^\omega_{\mu(a)}$ (resp. $N^\omega_{\mu(a)}$) where $\lambda, \mu \in \{\texttt{in}, \texttt{out}, \texttt{und}\}$ and $\lambda \neq \mu$.

When $|M^\omega_{\lambda(a)}|{=}n$ (resp. $|N^\omega_{\lambda(a)}|{=}n$) in Def. 7, $EAF_1, \ldots, EAF_n$ credulously (resp. skeptically) agree on $\lambda(a)$.

Example 6 Consider $AF = (\{a, r\}, \{(a, r), (r, a)\})$ and three EAFs: $EAF_1 = (AF, \mathbf{K}\ \texttt{in}(a))$, $EAF_2 = (AF, \neg\mathbf{M}\ \texttt{und}(a))$, and $EAF_3 = (AF, \mathbf{K}\ \texttt{und}(a))$. Then $\texttt{in}(a)$ is credulously adopted by majority voting under the complete epistemic labelling, while it is not skeptically adopted.

3.4 Complexity

We assume that the readers are familiar with the well-known notations in computational complexity (e.g., P-c, NP-c, coNP-c, etc.). Let $\omega \in \{gr, st, co, pr\}$ and $EAF = (AF, \varphi)$. Due to Proposition 8, we can check for the existence of a non-empty ω-epistemic labelling set using Algorithm 1, assuming the existence of a procedure $\texttt{D}(\omega, AF, \psi)$ that determines the existence of an ω-labelling $\mathcal{L}$ of AF such that $\mathcal{L} \models \psi$.

In essence, Algorithm 1 shows that checking whether EAF has a non-empty ω-epistemic labelling set can be reduced to checking whether a labelling $\mathcal{L}$ of AF satisfies a formula over $\mathcal{A}_{AF}$. In line 1 we assume that φ has at most k disjuncts, and each contains at most p conjuncts, where p and k are polynomial in the size of the AF and refer to φ as a (k, p)-DNF.[6] Under this assumption, Algorithm 1 will call $\texttt{D}(\omega, AF, \psi)$ at most $k \times p$ times. Consider the following decision problem:

> $\texttt{Exists}_{\omega}^{(k,p)}$: Given an $AF = (Ar, att)$ and a (k, p)-DNF epistemic formula φ over $\mathcal{A}_{AF}$, does (AF, φ) have a non-empty ω-epistemic labelling set?

The above discussion gives us the next result.

Proposition 13 $\texttt{Exists}_{\omega}^{(k,p)}$ *is P-c for* $\omega = gr$ *and NP-c for* $\omega \in \{co, st, pr\}$.

The proof of the above results relies on the following facts: (*i*) the grounded labelling of AF can be computed in polynomial time and is unique; and given a labelling $\mathcal{L}$ and a propositional formula ψ over $\mathcal{A}_{AF}$, (*ii*) checking whether there exists an ω-labelling satisfying a formula is NP-c for $\omega \in \{st, co, pr\}$ (by the result $\texttt{Cred}_{\sigma}$ in [13, Table 1] or in [14]); (*iii*) checking whether a given labelling $\mathcal{L}$ satisfies a propositional formula over $\mathcal{A}_{AF}$ is polynomial.

4 Related Work

EAF could be viewed as an approach to limiting the set of extensions (or labellings) of an argumentation framework for semantical consideration and this is similar, at least in the spirit, to argumentation with preferences and probabilistic argumentation. By introducing epistemic constraints, it is similar to works focusing on a reasoner's belief. The key difference between EAF and the other approaches can be summarized as follows.

Constrained argumentation frameworks (CAF) proposed in [11] are syntactically similar to EAF. Both are of the form $\langle A, R, C \rangle$ where (A, R) is an AF and C is a propositional formula (over A) in a CAF whilst it is an epistemic formula (over $\mathcal{A}_{AF}$) in an EAF. The key distinction between CAF and EAF lies in the use of the constraint. In CAF, C is imposed on extensions of the AF leading to a new set of extensions of the original AF. In contrast, φ does not change the labellings of the original AF in an EAF (AF, φ). Another extension of Dung's AF is *abstract dialectical framework* (ADF) [9] where each argument has an associated acceptance condition expressed by a propositional formula over the existing arguments. In EAF individual arguments do not have acceptance conditions, while epistemic constraints specify beliefs concerning which arguments are to be (un)accepted in the final outcome.

Probabilistic argumentation as proposed in [16,17] focuses on the uncertainty of arguments rather than reasoners' beliefs. This approach represents the beliefs of agents by a probability assignment to arguments [16] or an epistemic labelling [17]. It provides methods for computing epistemic extensions of an AF which contain arguments with probability greater than a certain threshold or assigning labels to arguments in accordance to the probability of the labelling, i.e., it merges an objective evidence and subjective beliefs in a single framework, which is in contrast to our approach. Moreover, it differs from EAFs significantly as its extensions might not correspond to any

[6] The DNF of a formula φ might have exponential number of disjuncts in general, however, it would be a rare case that belief of an agent is expressed by an exponential formula.

type of extensions of the original AF. On the other hand, there would be a connection between probabilistic argumentation and EAF. For instance, we consider that for each $EAF = (AF, \varphi)$ and ω, there would exist a probabilistic distribution P with respect to AF with the property that x is believed wrt P ($P(x) > 0.5$)) then $in(x)$ is skeptically entailed by every ω-labelling set of EAF. We believe that the inverse could be true as well. We leave the precise formulation and proof of this interesting problem for future work. Recent work in this direction has introduced *epistemic attack semantics* that considers extended probability distribution, which assigns degrees of belief to arguments and attacks [22] which is then further investigated in dynamic setting [18]. Whether formulas in EAF could sufficiently model this type of extension is an open question that we intend to pursuit as well.

Argumentation with preferences or priorities has been studied extensively in recent years. Preference over arguments is introduced as a preorder relation over arguments in [2,3,4,19], while a new attack relation that ranges from arguments to attacks is used in [20]. Our representation of preferences is close to the approach employing a preorder but there are differences from them. For instance, given $AF = (\{a, b\}, \{(a, b)\})$ with the preference $a \preceq b$, Kaci and van der Torre [19] provide its semantics by extensions of $AF_1 = (\{a, b\}, \{\})$, and Amgoud and Vesic [3] convert AF to $AF_2 = (\{a, b\}, \{(b, a)\})$. As such, the structure of the original argumentation graph is changed, and as a result, extensions of the preference-based AF are not extensions selected from those of AF. Wakaki [23] introduces *preference-based AF* (PAF) which, as we do, selects extensions based on preference relation over arguments. Our representation of preference in EAF is different from PAF in the sense that EAF can represent preference over not only arguments but justification states. *Value-based argumentation framework* (VAF) [6] represents preference in AF by assigning values to arguments. In VAF acceptable arguments may change depending on the order of values. Arguments acceptable irrespective of any value order are called *objectively acceptable* and those acceptable for some order are called *subjectively acceptable*. In EAF justification states of arguments change depending on epistemic constraints, so the effect of epistemic constraints in EAF is similar to the effect of value in VAF. On the other hand, VAF may produce extensions that are not those of the original AF, while EAF produces labellings that are also labellings of the original AF. Airiau $et\,al.$ [1] consider the problem such that given a profile of argumentation frameworks $(AF_1, \ldots, AF_n)$, one for each agent, can this profile be explained in terms of a single master argumentation framework, an association of arguments with values, and a profile of preference orders over values $(\succeq_1, \ldots, \succeq_n)$, one for each agent? Their approach represents individual views of a common AF by preference orders over values, which is in contrast with our approach in which individual views are encoded by epistemic formulas over arguments. Visser $et\,al.$ [24] introduce an *epistemic argumentation framework* for reasoning about preferences with uncertain information. They provide languages and inference schemes for instantiated AFs, which is in contrast with our framework for abstract argumentation.

Schwarzentruber $et\,al.$ [21] introduce a logical framework for reasoning about arguments owned by agents and their knowledge about other agents' arguments. They introduce epistemic logics to represent belief state of agents in dialogues and define Kripke semantics. For instance, they represent that "an agent 1 believes that there exists an argument about global warming (*gw*) owned by an agent 2" by the formula: $B_1(\langle U \rangle(gw \wedge ownedby(2)))$. Our approach is different from theirs in two ways: first EAF is an extension of AF and we do not use modal logic based on Kripke structures. Second, our primary interest in this paper is to represent an agent's own beliefs, and we do not consider reasoning about beliefs of other agents. Finally, we note that an EAF realizes meta-level reasoning about arguments in abstract argumentation frameworks. In this sense, it could be viewed as a kind of *meta-level arguments* discussed in [7].

5 Conclusion and Future Work

An epistemic argumentation framework introduces belief of agents to argumentation frameworks. A unique feature of EAF is that it can represent arguments and attacks as objective evidence in AF, while at the same time, it can encode subjective beliefs of individual agents by epistemic constraints over the outcome. By separating objective knowledge and subjective beliefs, individual agents could produce different conclusions based on their biases toward a common AF. Such a situation happens, for instance, in a court case where jurors share the same open AF while could reach different conclusions based on their biases. Moreover, the separation has an advantage that an individual agent can easily revise his/her belief without changing the structure of an AF.

This paper addresses declarative aspects of EAFs. From the procedural viewpoint, a system for computing epistemic labelling sets is built on top of *answer set solvers* [8]. More precisely, suppose an EAF (AF, φ) where φ is a CNF $\varphi = \psi_1 \wedge \cdots \wedge \psi_n$ in which $\psi_i \, (1 \leq i \leq n)$ is a disjunction of simple epistemic literals of the form $\mathbf{E}\lambda(x)$ or $\mathbf{E}\neg\lambda(x)$ where $\mathbf{E} \in \{\mathbf{M}, \mathbf{K}\}$ and $\lambda \in \{\mathtt{in}, \mathtt{out}, \mathtt{und}\}$. In this case, the EAF is transformed to an *epistemic logic program* [15] Π and ω-epistemic labelling sets are computed by *world views* of Π. We will address the issue in the full paper.

In this paper, we focus on representing an agent's own belief in EAFs. On the other hand, EAF could be extended to reasoning about beliefs of other agents and representing an agent's own belief based on beliefs of other agents. This type of belief contains a constraint such that "$\mathbf{K}_1\mathtt{in}(a) \supset \mathbf{M}_2\mathtt{in}(a)$" (if an agent 1 supports the acceptance of an argument a then an agent 2 would not argue against it). EAF is used for characterizing several problems in argumentation. For instance, the *enforcement* [5] of an argument a in AF is captured as finding an EAF $(AF', \mathbf{M}\,\mathtt{in}(a))$ having a non-empty ω-epistemic labelling set where AF' is an expansion of AF. We introduce EAF for complete, stable, grounded, or preferred semantics, but the framework is extended to other semantics such as semi-stable, stage, ideal, etc. Those issues are left for future work.

Acknowledgment
This work is supported by JSPS KAKENHI Grant Numbers JP17H00763 and JP18H03288. The authors also acknowledge the partial support from the NSF grants HRD-1345232, IIS-1812628, and OIA-1757207.

References

1. Airiau, S., Bonzon, E., Endriss, U., Maudet, N., Rossit, J.: Rationalisation of profiles of abstract argumentation frameworks: characterisation and complexity. J. Artif. Intell. Res. **60**,

149–177 (2017)
2. Amgoud, L., Cayrol, C.: On the acceptability of arguments in preference-based argumentation. In: Proc. 14th Conf. Uncertainty in Artificial Intelligence, pp. 1–7 (1998)
3. Amgoud, L., Vesic, S.: On the role of preference in argumentation frameworks. In: Proc. 22nd IEEE Int. Conf. Tools with Artificial Intelligence. pp. 219–222 (2010)
4. Amgoud, L., Vesic, S.: A new approach for preference-based argumentation frameworks. Ann. Math. Artif. Intell. **63**(2), 149–183 (2011)
5. Baumann, R., Brewka, G.: Expanding argumentation frameworks: enforcing and monotonicity results. In: Proc. 3rd Int. Conf. Computational Models of Argument, Frontiers in AI and Applications. vol. 216, pp. 75–86. IOS Press (2010)
6. Bench-Capon, T. J. M.: Value-based argumentation frameworks. In: Proc. 9th Int. Workshop on Non-Monotonic Reasoning. pp. 443–454 (2002)
7. Boella, G., Gabbay, D.M., van der Torre, L., Villata, S.: Meta-argumentation modelling I: methodology and techniques. Studia Logica **93**(1), 297–355 (2009)
8. Brewka, G., Eiter, T., Truszczynski, M.: Answer set programming at a glance. Commun. ACM **54**(12), 92–103 (2011)
9. Brewka, G., Woltran, S.: Abstract dialectical frameworks. In: Proc. 12th Int. Conf. Principles of Knowledge Representation and Reasoning. pp. 102–111 (2010)
10. Caminada, M. W. A., Gabbay, D. M.: A logical account of formal argumentation. Studia Logica **93**(2-3), 109–145 (2009)
11. Coste-Marquis, S., Devred, C., Marquis, P.: Constrained argumentation frameworks. In: Proc. 10th Int. Conf. Principles of Knowledge Rep. and Reasoning. pp. 112–122 (2006)
12. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artif. Intell. **77**, 321 – 357 (1995)
13. Dvorák, W.: On the complexity of computing the justification status of an argument. In: Proc. 1st Int. Workshop on Theory and Applications of Formal Argumentation, Revised Selected Papers, LNCS 7132, pp. 32–49. Springer (2011)
14. Dvorák, W., Dunne, P. E.: Computational problems in formal argumentation and their complexity. In Handbook of Formal Argumentation. pp. 631–688. College Publications (2018)
15. Gelfond, M.: Strong introspection. In: Proc. 9th National Conf. Artificial Intelligence (AAAI), pp. 386–391 (1991)
16. Hunter, A.: A probabilistic approach to modelling uncertain logical arguments. J. Approximate Reasoning **54**, 47–81 (2013)
17. Hunter, A., Thimm, M.: Probabilistic reasoning with abstract argumentation frameworks. J. Artificial Intelligence Research **59**, 565–611 (2017)
18. Hunter, A., Polberg, S., Potyka, N.: Updating belief in arguments in epistemic graphs. In: Proc. 16th Int. Conf. Principles of Knowledge Rep. and Reasoning, pp. 138–147 (2018)
19. Kaci, S., van der Torre, L.: Preference-based argumentation: arguments supporting multiple values. J. Approximate Reasoning **48**, 730–751 (2008)
20. Modgil, S.: Reasoning about preferences in argumentation frameworks. Artif. Intell. **173**(9-10), 901–934 (2009)
21. Schwarzentruber, F., Vesic, S., Rienstra, T.: Building an epistemic logic for argumentation. In: Proc. 13th European Conf. Logics in AI, pp. 359–371. Springer, LNAI 7519 (2012)
22. Thimm, M., Polberg, S., Hunter, A.: Epistemic attack semantics. In: Proc. 7th Int. Conf. Computational Models of Argument, Frontiers in AI and Applications. vol. 305, pp. 37–48. IOS Press (2018)
23. Wakaki, T.: Preference-based argumentation built from prioritized logic programming. J. Logic and Computation **25**(2), 251–301 (2015)
24. Wietske, V., Hindriks, K. V., Jonker, C. M.: Argumentation-based qualitative preference modelling with incomplete and uncertain information. Group Decision and Negotiation **21**(1), 99–127 (2012)