Extended Explanatory Argumentation Frameworks

Jérémie Dauphin and Marcos Cramer

University of Luxembourg[**]

Abstract. Multiple extensions of Dung's argumentation frameworks (AFs) have been proposed in order to model features of argumentation that cannot be directly modeled in AFs. One technique that has already previously proven useful to study and combine such extensions is the meta-argumentation methodology involving the notion of a flattening. In order to faithfully model the interaction between explanation argumentation in scientific debates, Šešelja and Straßer have introduced *Explanatory Argumentation Frameworks* (EAFs). In this paper, we first prove that the flattening technique works as expected for recursive (higher-order) attacks. Then we apply this technique in order to combine EAFs with multiple other extensions that have been proposed to AFs, namely with recursive attacks, joint attacks and a support relation between arguments. This gives rise to *Extended Explanatory Argumentation Frameworks* (EEAFs). We illustrate the applicability of EEAFs by using them to model a piece of argumentation from a research-level philosophy book.

1 Introduction

Dung's argumentation frameworks (AFs) [7] are a powerful and flexible formal tool for formally modelling argumentative discourse. However, various researchers have felt the need to extend AFs in order to model features of argumentation that cannot be directly modeled in AFs, e.g. by enriching them with recursive (higher-order) attacks [2], joint attacks [9], a support relation between arguments [5,4], or explanatory features [10].

One technique that has already previously proven useful to study and combine such extensions is the meta-argumentation methodology involving the notion of a *flattening* [3]. A flattening is a function that maps some extended variant of argumentation frameworks into standard AFs. If the definition of the various argumentation semantics for that extended variant of AFs is independent from the definition of that flattening function, one wants the flattening to satisfy the property that it preserves these semantics, in the sense that applying the flattening function, then calculating the extensions according to some argumentation

semantics, and finally unflattening the extensions should yield the same result as directly calculating the extensions according to the corresponding argumentation semantics for the extended variant of argumentation frameworks. However, flattenings can also be used to define argumentation semantics for extended variants of AFs for which there is no definition of the semantics independent of flattenings. This approach has proven particularly useful for combining multiple extensions of AFs [3], because in this case, it is often much clearer what the "right" definition of a flattening is than what the "right" direct definition of the various argumentation semantics is.

Previous work on flattening argumentation frameworks with recursive attacks (AFRAs) was limited to second-order attacks [3,1], even though the original definition of recursive attacks was for arbitrarily deeply nested higher-order attacks [2]. This means that for the purpose of defining the flattening, attacking an attack between two arguments was allowed, but attacking such a second-order attack was already not allowed. In Section 3.1, we show how to define a flattening of arbitrary AFRAs, and prove that it conforms with the direct definition of the semantics of AFRAs.

The rest of the paper is devoted to applying the meta-argumentation methodology of flattening and unflattening in order to incorporate recursive attacks, joint attacks and a support relation between arguments into *Explanatory Argumentation Frameworks* (EAFs), which have been proposed by Šešelja and Straßer [10] in order to faithfully model the interaction between explanation and argumentation in scientific debate. EAFs feature *explananda* and an *explanatory relation* that can hold either between an argument and an explanandum, or between two arguments. We use the terms *Extended Explanatory Argumentation Frameworks* (EEAFs) for this enriched formalism that incorporates recursive attacks, joint attacks and a support relation into EAFs.

The explanatory relation from EAFs cannot be easily flattened. Therefore, for defining the semantics of EEAFs, we apply the meta-argumentation methodology by allowing the output of the flattening function to be an EAF rather than an AF. In other words, we flatten away recursive attacks, joint attacks and the support relation, but we do not flatten away explanations, instead making use of the semantics of EAFs instead of the semantics of standard AFs.

Finally, we illustrate the applicability of EEAFs by using them to model a piece of argumentation from the introduction to Hartry Field's book *Saving Truth from Paradox* [8], an important, relatively recent, monograph about semantic paradoxes, a major research topic within the field of philosophical logic.

The rest of the paper is structured as follows: In Section 2, we describe the various proposed extensions to AFs and outline the meta-argumentation methodology of flattening and unflattening. In Section 3, we first extend the meta-argumentation methodology to arbitrarily deeply nested AFRAs, and then use this methodology to formally define the semantics of EEAFs. In Section 4 we present an example that illustrates the applicability of EEAFs, before concluding the paper in Section 5.

2 Basics of Formal Argumentation

We use the standard notions of abstract argumentation frameworks as defined by P. M. Dung in 1995 [7].

2.1 Explanatory Argumentation Frameworks

In scientific debates, the discussions are usually centered around some phenomenons or evidence and the different parties propose theories to explain them. With this idea in mind, D. Šešelja and C. Straßer have extended abstract argumentation framework with explanatory features [10]. In these frameworks, there are not only arguments but also explananda. These are scientific phenomenons of which, unlike arguments, the acceptability is not being questioned.

Definition 1. An *explanatory argumentation framework* (EAF) is a tuple $\langle \mathcal{A}, \mathcal{X}, \rightarrow, \dashrightarrow, \sim \rangle$, where $\mathcal{A}$ is a set of arguments, $\mathcal{X}$ is a set of explananda, $\rightarrow \subseteq \mathcal{A} \times \mathcal{A}$ is an attack relation, $\dashrightarrow \subseteq \mathcal{A} \times (\mathcal{A} \cup \mathcal{X})$ is an explanation relation from arguments to either explananda or other arguments, and $\sim \subseteq \mathcal{A} \times \mathcal{A}$ is a symmetric incompatibility relation.

Note that the incompatibility relation's purpose is to differentiate between the opposing theories, as scientists usually do not accept multiple explanations of a given phenomenon at the same time.

Definition 2. Let $\langle \mathcal{A}, \mathcal{X}, \rightarrow, \dashrightarrow, \sim \rangle$ be an EAF. A set of arguments $S \subseteq \mathcal{A}$ is said to be *conflict-free* if and only if there are no arguments $a, b \in S$ such that $(a, b) \in \rightarrow \cup \sim$.

Note that the definition of admissible sets still stands but with the revised definition of conflict-freeness.

Definition 3. An *explanation* $X[e]$ for $e \in \mathcal{X}$ offered by a set of arguments S is a subset S' of S such that there exists a unique argument $a \in S'$ such that $a \dashrightarrow e$ and for all $a' \in S' \setminus a$, there exists a path in $\dashrightarrow$ from a' to a.
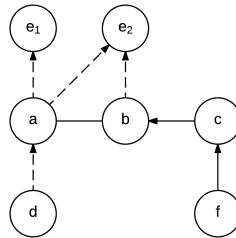

Fig. 1. Example EAF1

Example 1. Consider the EAF on Figure 1. Note that the incompatibility relation has been represented by a straight line with no arrow between a and b.

Here we have two explananda, e_1 and e_2. a explains both e_1 and e_2 while b explains only e_2. Consider the conflict-free set $\{a, d, f\}$. It contains two explanations for e_1, namely $X_1[e_1] = \{a\}$ and $X_2[e_1] = \{a, d\}$. Similarly, it offers two explanations for e_2. The conflict-free set $\{b, f\}$ however offers an explanation only for e_2.

For our goal of selecting the best theory from our model, we need a way to compare how much and how well a given set of arguments is able to explain.

Definition 4. A set of arguments S_1 is *explanatory more powerful* than a set of arguments S_2 ($S_1 >_p S_2$) if and only if the set of explananda for which S_1 offers an explanation is a strict super-set of the set of explananda for which S_2 offers an explanation.

An explanation $X_1[e]$ is *explanatory deeper* than another explanation $X_2[e]$ ($X_1[e] >_d X_2[e]$) if and only if $X_2[e] \subset X_1[e]$.

In our previous example, we have that $\{a, d\} >_p \{b\}$ since $\{a, d\}$ offers an explanation for $\{e_1, e_2\}$ while $\{b\}$ only offers an explanation for $\{e_2\}$. Additionally, we have that $\{a, d\} >_d \{a\}$ and $\{a, d, f\} >_d \{a, f\}$.

Šešelja and Straßer [10] then propose two procedures for the selection of the best sets of arguments with respect to these notions. We have redefined them as extensions, in order to be more in line with abstract argumentation extensions, while preserving their concepts.

Definition 5. Let $\langle \mathcal{A}, \mathcal{X}, \rightarrow, \dashrightarrow, \sim \rangle$ be an EAF and $S \subseteq \mathcal{A}$ a set of arguments.

1. We say that S is *satisfactory* iff S is admissible and there is no $S' \subseteq \mathcal{A}$ such that $S' >_p S$ and S' is admissible.
2. We say that S is *insightful* iff S is satisfactory and there is no $S' \subseteq \mathcal{A}$ such that $S' >_d S$ and S' is satisfactory.
3. We say that S is an *argumentative core extension* (*AC-extension*) of Δ iff S is satisfactory and there is no $S' \supset S$ such that S' is satisfactory.
4. We say that S is an *explanatory core extension* (*EC-extension*) of Δ iff S is insightful and there is no $S' \subset S$ such that S' is insightful.

In our example, the AC-extension is $\{a, d, f\}$, while the EC-extension is $\{a, d\}$.

2.2 Argumentation Frameworks with Recursive Attacks

While EAFs add explanatory features to abstract argumentation frameworks, Baroni et al. [2] have developed an extension which enhances the expressive power of the attack relation. In their frameworks, they allow for attacks to target other attacks. This way, an argument may refute an attack relation between two other arguments without contesting the acceptability of any of them.

Definition 6. An *Argumentation Framework with Recursive Attacks* (AFRA) is a pair $\langle \mathcal{A}, \rightarrow \rangle$ where $\mathcal{A}$ is a set of arguments and $\rightarrow \subseteq \mathcal{A} \times (\mathcal{A} \cup \rightarrow)$ is an attack relation from arguments to either arguments or attacks.

For a given attack $\alpha = (A, X) \in \,\rightarrow$, we say that the source of α is $src(\alpha) = A$ and its target is $trg(\alpha) = X$.

Now that attacks can be targeted, we need to extend our notions of acceptance to also include them.

Definition 7. Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $\varphi \in \,\rightarrow, \psi \in (\mathcal{A} \cup \rightarrow)$ and $S \subseteq (\mathcal{A} \cup \rightarrow)$. We say that φ *defeats* ψ iff either $\psi = trg(\varphi)$ or $src(\psi) = trg(\varphi)$.
Additionally, we say that S is *conflict-free* iff there do not exist $\varphi, \psi \in S$ such that φ defeats ψ.

The notions of defense and admissibility then follows with a similar idea as in standard abstract argumentation frameworks.

Definition 8. Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $\varphi \in (\mathcal{A} \cup \rightarrow)$ and $S \subseteq (\mathcal{A} \cup \rightarrow)$. We say that S *defends* φ iff for every $\psi \in \,\rightarrow$ such that ψ defeats φ, there exists a $\delta \in S$ such that δ defeats ψ. We say that S is *admissible* iff S is conflict-free and defends its elements.

The complete semantics then follows with a similar definition as in classical abstract argumentation but using the adapted notions just defined.

Definition 9. Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA and $S \subseteq (\mathcal{A} \cup \rightarrow)$. We say that S is a *complete extension* of F iff S is admissible and contains every $\varphi \in (\mathcal{A} \cup \rightarrow)$ it defends.

2.3 Support in Abstract Argumentation

While classical abstract argumentation revolves around attacks, there has been research on extending it with a positive relation of support between arguments. We will first examine the formalism introduced by Cayrol and Lagasquie-Schiex called bipolar argumentation framework [5], as summarized by G. Boella et al. in [4].

Definition 10. A *bipolar argumentation framework* (BAF) is a triple $\langle \mathcal{A}, \rightarrow, \Rightarrow \rangle$ where $\mathcal{A}$ is a set of arguments, $\rightarrow \subseteq \mathcal{A} \times \mathcal{A}$ is an attack relation and $\Rightarrow \subseteq \mathcal{A} \times \mathcal{A}$ is a support relation.

Boella et al. [4] treat support in a deductive sense and thus introduce mediated attacks. The intuition behind these attacks is that if from a we can deduce b, then if we do not have b, we also cannot have a.

Definition 11. Let $\langle \mathcal{A}, \rightarrow, \Rightarrow \rangle$ be a bipolar argumentation framework. For $a, b \in \mathcal{A}$, there is a *mediated attack* from a to b if and only if there is a sequence $a_1 \Rightarrow a_2, ..., a_{n-1} \Rightarrow a_n$ such that $n \geq 2$, $a = a_1$ and $b \rightarrow a_n$.

They then define the semantics of bipolar argumentation frameworks with respect to their flattening. The flattened framework will consist of meta-arguments and an attack relation only, with the support relation from the BAFs being represented as a combination of auxiliary meta-arguments and attack relations.

Definition 12. Given a bipolar argumentation framework $\langle \mathcal{A}, \rightarrow, \Rightarrow \rangle$, the set of corresponding meta-arguments MA is $\{acc(a) \mid a \in \mathcal{A}\} \cup \{X_{a,b}, Y_{a,b} \mid a, b \in \mathcal{A}\} \cup \{Z_{a,b} \mid a, b \in \mathcal{A}\}$ and $\rightarrow_2 \subseteq MA \times MA$ is a binary relation on MA such that:

  – For all $a, b \in \mathcal{A}$ such that $a \rightarrow b$, we have $acc(a) \rightarrow_2 X_{a,b}$, $X_{a,b} \rightarrow_2 Y_{a,b}$ and $Y_{a,b} \rightarrow_2 acc(b)$
  – For all $a, b \in \mathcal{A}$ such that $a \Rightarrow b$, we have $acc(b) \rightarrow_2 Z_{a,b}$ and $Z_{a,b} \rightarrow_2 acc(a)$

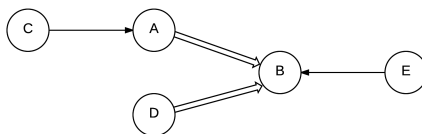**Example 2.** The example represented in Figure 2 is flattened in Figure 3:

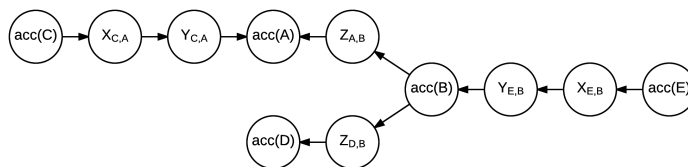**Fig. 2.** Example bipolar argumentation framework

Fig. 3. Flattened BAF from Figure 2

In the flattening, the mediated attacks are made apparent. By applying the semantics of classical abstract argumentation frameworks we can then retrieve the corresponding extensions of the BAF.

Note that Baroni et al [6] have combined higher-order attacks and supports, with semantics defined directly on the higher-level frameworks. However, unlike with the flattening approach, it is unclear how to take further features into account in those direct semantics.

2.4 Joint attacks

Another extension of AFs allows for joint attacks, where multiple arguments join forces to attack another argument.

D. Gabbay [9] calls this kind of relation a *joint attack*. He defines it as follows:

Definition 13. A *higher level argumentation framework* is a triple $(S, S^0, \rightarrow)$, where $S \neq \emptyset$ is a set of arguments, S^0 is the family of all finite non-empty subsets of S and $\rightarrow \subseteq S^0 \times S$ is an attack relation.
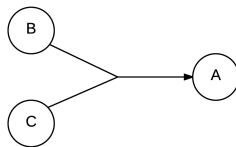
Fig. 4. Higher level argumentation framework

For simplicity of notation we will identify the singleton set $\{x\}$ with x.

Similarly as before, the semantics of higher level networks will be defined in terms of their flattening. We define the flattening as follows:

Definition 14. Given a higher level argumentation framework $(S, S^0, \rightarrow)$, the set of corresponding meta-arguments MA is $\{acc(a), rej(a) \mid a \in \mathcal{A}\} \cup \{e(X) \mid X \in S^0\}$ and $\rightarrow_2 \subseteq MA \times MA$ is a binary relation on MA such that:

– For all $a \in \mathcal{A}$, we have $acc(a) \rightarrow_2 rej(a)$
– For all $X \in S^0$, and every $b \in \mathcal{A}$ such that $X \rightarrow b$, we have that $e(X) \rightarrow_2 acc(b)$ and $rej(a) \rightarrow_2 e(X)$ for every $a \in X$.

In the flattening, the success of a joint attack depends solely on the acceptance of the meta-argument $e(X)$, which itself depends on the acceptance of every argument in the coalition.

The flattening of the framework from Figure 4 is depicted in Figure 5.
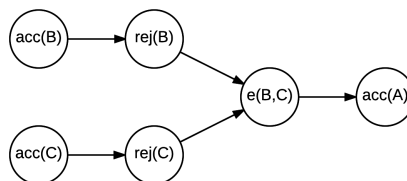

Fig. 5. Flattened version of the framework from Figure 4

3 Aggregating multiple extensions of abstract argumentation frameworks: EEAFs

In this section, we will introduce Extended Explanatory Argumentation Frameworks (EEAFs), an extension of EAFs from Section 2.1 with meta-argumentation features such as higher order attacks, support and joint attacks.

In order to motivate the semantics of EEAFs based on a flattening function, we will start by suggesting a flattening for AFRAs of any order. We will prove that this flattening leads to the same extensions as the AFRA semantics defined by Baroni et al. [2].

3.1 Flattening AFRAs

Boella et al. [3] define a flattening function for second-order AFRAs, which allows one to obtain for a given AFRA an equivalent abstract argumentation framework. We will now propose a flattening function for AFRAs of any order.

We will first define a function m which will associate each argument and each attack relation to the corresponding meta-argument. For an argument a, it will be the meta-argument $acc(a)$, while for an attack, it will be the Y auxiliary argument, since its acceptability is synonym of success for the attack.

Definition 15. Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA. The set of corresponding meta-arguments is $MA = \{acc(a) \mid a \in \mathcal{A}\} \cup \{X_{a,\psi}, Y_{a,\psi}\} \mid a \in \mathcal{A}, \ \psi \in (\mathcal{A} \cup \rightarrow)\}$. We define a partial function m: $(\mathcal{A} \cup \rightarrow) \mapsto MA$, such that:

- if $\varphi \in \mathcal{A}$, then $m(\varphi) = acc(\varphi)$.
- if $\varphi \in \rightarrow$ such that for some $\psi \in \mathcal{A}$ and some $\delta \in (\mathcal{A} \cup \rightarrow)$, $\varphi = (\psi, \delta)$, then $m(\varphi) = Y_{\psi,\delta}$.

We define the *flattening function* f to be $f(F) = \langle MA, \rightarrow_2 \rangle$, where $\rightarrow_2 \subseteq MA \times MA$ is a binary relations on MA such that

$$acc(a) \rightarrow_2 X_{a,\psi}, X_{a,\psi} \rightarrow_2 Y_{a,\psi} \text{ and } Y_{a,\psi} \rightarrow_2 m(\psi) \text{ for all } a \in \mathcal{A}, \psi \in (\mathcal{A} \cup \rightarrow)$$

One can then apply the classical abstract argumentation semantics such as complete, stable, preferred and grounded. We then need to define a function which can transform a meta-extension from the flattened AFRA to an extension for the original AFRA. A similar unflattening function has been introduced in [3], and has been slightly modified here to also unflatten attacks.

Definition 16. Given a set of meta-arguments $B \subseteq MA$, we define the *unflattening function g* as:

$$g(B) = \{a \mid acc(a) \in B\} \cup \{(a, \psi) \mid Y_{a,\psi} \in B\}$$

We also define a function $\bar{f}$ which provides a correspondence between a set of arguments and attacks from an AFRA and a set of meta-arguments from its flattened version.

Definition 17. Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $f(F) = \langle MA, \rightarrow_2 \rangle$ its flattening and $S \subseteq (\mathcal{A} \cup \rightarrow)$. We define the *correspondence function $\bar{f}$* : $\mathbb{P}(\mathcal{A} \cup \rightarrow) \mapsto \mathbb{P}(MA)$ as follows:

$$\bar{f}(S) = \{acc(a) \mid a \in S \cap \mathcal{A}\} \cup \{Y_{a,\psi} \mid (a, \psi) \in S \cap \rightarrow\} \cup \\ \{X_{b,\psi} \mid (a, b) \in S \cap \rightarrow, \ \psi \in \rightarrow\}$$

Notice that $g(\bar{f}(S)) = S$. We add the extra $X_{i,j}$ meta-arguments in order to represent the indirect attacks which the arguments in S might carry out, i.e. the attacks which are indirectly attacked by arguments in S due to them attacking the source of these attacks.

In [2], Baroni et al. define the semantics of AFRAs without having recourse to flattening. We will show that the process of flattening, applying complete semantics on the flattened frameworks and then unflattening it is equivalent to the directly applying the semantics they define for the complete semantics. We will show this gradually by first stating and proving three lemmas:

Lemma 1. Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $f(F) = \langle MA, \rightarrow_2 \rangle$ and $S \subseteq (\mathcal{A} \cup \rightarrow)$. S is conflict-free in F if and only if $\bar{f}(S)$ is conflict-free in $f(F)$.

> **Proof:**
> Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $f(F) = \langle MA, \rightarrow_2 \rangle$ and $S \subseteq (\mathcal{A} \cup \rightarrow)$.
>
> 1. $\Rightarrow$: Assume that S is conflict-free in F. Then, there is no $\varphi, \psi \in S$ such that $trg(\varphi) = \psi$ or $trg(\varphi) = src(\psi)$. Suppose for a contradiction that $\bar{f}(S)$ is not conflict-free in $f(F)$. This means that there exists two arguments $p, q \in \bar{f}(S)$ such that $p \rightarrow_2 q$. By the construction of $\rightarrow_2$ defined by the flattening function, there are only four possible cases, which all lead to the contradiction that S is not conflict-free. Therefore $\bar{f}(S)$ is conflict-free.
> 2. $\Leftarrow$: Suppose $\bar{f}(S)$ is conflict-free. Suppose for a contradiction that S is not conflict-free. Then, there exists $(a, \varphi), (b, \psi) \in S$ such that $\varphi = (b, \psi)$ or $\varphi = b$.
>   In both cases we can reach the contradiction that $\bar{f}(S)$ is not conflict-free, therefore S is conflict-free. $\qquad\square$

Lemma 2. Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $f(F) = \langle MA, \rightarrow_2 \rangle$, $\varphi \in (\mathcal{A} \cup \rightarrow)$ and $S \subseteq (\mathcal{A} \cup \rightarrow)$. We have that:

$$\varphi \text{ is defended by } S \text{ in } F \text{ and if } \varphi = (a, \psi) \in \rightarrow, \text{ we have } a \in S, \text{ iff}$$
$$m(\varphi) \text{ is defended by } \bar{f}(S) \text{ in } f(F) \text{ and if } \varphi \in \rightarrow, \text{ then } acc(src(\varphi)) \text{ is also}$$
$$\text{defended by } \bar{f}(S).$$

> **Proof**:
> Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $f(F) = \langle MA, \rightarrow_2 \rangle$, $\varphi \in (\mathcal{A} \cup \rightarrow)$ and $S \subseteq (\mathcal{A} \cup \rightarrow)$.
>
> 1. $\Rightarrow$: Suppose that φ is defended by S in F and if $\varphi = (a, \psi) \in \rightarrow$, we have $a \in S$. Consider $m(\varphi)$ in $f(F)$. Suppose for some $p \in MA$, $p \rightarrow_2 m(\varphi)$. By the construction of $\rightarrow_2$ defined by the flattening function, either $p = Y_{a,\varphi}$ for some $a \in \mathcal{A}$, or $p = X_{src(\varphi),trg(\varphi)}$.
>   In both cases, $m(\varphi)$ is defended by $\bar{f}(S)$. Hence, if φ is defended by S in F, then $m(\varphi)$ is defended by $\bar{f}(S)$ in $f(F)$. We now have to show that if $\varphi \in \rightarrow$, then $acc(src(\varphi))$ is also defended by $\bar{f}(S)$.
>   Suppose $\varphi \in \rightarrow$ and $p \in MA$ such that $p \rightarrow_2 acc(src(\varphi))$. Then, p must be of the form $Y_{a,src(\varphi)}$ for some $a \in \mathcal{A}$, and hence there exists $(a, src(\varphi)) \in \rightarrow$. Since $(a, src(\varphi))$ defeats φ, there exists some $\delta \in S$ such that δ defeats $(a, src(\varphi))$. We distinguish two cases:

Either $\delta = (b,a)$ or $\delta = (b,(a,src(\varphi)))$ for some $b \in \mathcal{A}$. In both cases, $acc(src(\varphi))$ is also defended by $\bar{f}(S)$.

Therefore, if φ is defended by S in F and if $\varphi = (a,\psi) \in \rightarrow$, we have $a \in S$, then $m(\varphi)$ is defended by $\bar{f}(S)$ in $f(F)$ and if $\varphi \in \rightarrow$, then $acc(src(\varphi))$ is also defended by $\bar{f}(S)$.

2. $\Leftarrow$: Suppose $m(\varphi)$ is defended by $\bar{f}(S)$ in $f(F)$ and if $\varphi \in \rightarrow$, then $acc(src(\varphi))$ is also defended by $\bar{f}(S)$. Consider φ in F. Suppose that for some $\psi \in \rightarrow$, ψ defeats φ. This means that either $\psi = (a,\varphi)$ or $\psi = (a,src(\varphi))$ for some $a \in \mathcal{A}$. In both cases, we can conclude that there exists a $\delta \in S$ such that δ defeats ψ by contradiction. Therefore, φ is defended by S.

   We now have to show that if $\varphi = (a,\psi) \in \rightarrow$, we have $a \in S$, still under the assumption that $m(\varphi)$ is defended by $\bar{f}(S)$ in $f(F)$ and if $\varphi \in \rightarrow$, then $acc(src(\varphi))$ is also defended by $\bar{f}(S)$.

   Suppose that $\varphi = (a,\psi) \in \rightarrow$. Then, by the construction of $\rightarrow_2$ defined by the flattening function, we have $X_{a,\psi} \rightarrow_2 Y_{a,\psi}$. Since $m(\varphi) = Y_{a,\psi}$ is defended by $\bar{f}(S)$, there exists $p \in \bar{f}(S)$ such that $p \rightarrow_2 X_{a,\psi}$. By the construction of $\rightarrow_2$, the only possibility is $p = acc(a)$. Hence, $acc(a) \in \bar{f}(S)$. Therefore, we have $a \in S$.

Thus, we can conclude that φ is defended by S in F and if $\varphi = (a,\psi) \in \rightarrow$, we have $a \in S$, if and only if $m(\varphi)$ is defended by $\bar{f}(S)$ in $f(F)$ and if $\varphi \in \rightarrow$, then $acc(src(\varphi))$ is also defended by $\bar{f}(S)$. $\qquad\square$

Lemma 3. Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $f(F) = \langle MA, \rightarrow_2 \rangle$ and $S \subseteq (\mathcal{A} \cup \rightarrow)$. We have that:

S is admissible in F and for every $(a,\psi) \in (S \cap \rightarrow)$, we have that $a \in S$
$$\text{if and only if}$$
$\bar{f}(S)$ is admissible in $f(F)$.

Proof: Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $f(F) = \langle MA, \rightarrow_2 \rangle$ and $S \subseteq (\mathcal{A} \cup \rightarrow)$.

1. $\Rightarrow$:Suppose $\bar{f}(S)$ is admissible in $f(F)$. Then, $\bar{f}(S)$ is conflict-free. Hence, according to Lemma 1, S is also conflict-free.

   Let $\varphi \in S$. We need to show that φ is defended by S. We do this by applying Lemma 2, i.e. by establishing that $m(\varphi)$ is defended by $\bar{f}(S)$ in $f(F)$ and if $\varphi \in \rightarrow$, then $acc(src(\varphi))$ is also defended by $\bar{f}(S)$. We have $m(\varphi) \in \bar{f}(S)$ and $m(\varphi)$ is defended by $\bar{f}(S)$ since $\bar{f}(S)$ is admissible. By the definition of $\bar{f}$, for every $(a,\psi) \in (S \cap \rightarrow)$, we have $Y_{a,\psi} \in \bar{f}(S)$. Therefore, $acc(a) \in \bar{f}(S)$, since it is the only argument which can defend $Y_{a,\psi}$ from $X_{a,\psi}$'s attack and $\bar{f}(S)$ is admissible. This means that $acc(a)$ is defended by $\bar{f}(S)$. Thus, according to Lemma 2, every $\varphi \in S$ is defended by S, which means that S is admissible, and for every $(a,\psi) \in (S \cap \rightarrow)$, we have that $a \in S$.

2. $\Leftarrow$:Suppose S is admissible in F and for every $(a,\psi) \in (S \cap \rightarrow)$, we have that $a \in S$. Then, S is conflict-free and so, according to Lemma 1, $\bar{f}(S)$ is also conflict-free.

Let $p \in \bar{f}(S)$. p is either of the form $m(\varphi)$ for some $\varphi \in S$, or of the form $X_{a,b}$ for some $a, b \in MA$ and $(\psi, a) \in S$.

In both cases, p is defended by $\bar{f}(S)$. Hence, $\bar{f}(S)$ is admissible in $f(F)$.

Therefore, S is admissible in F and for every $(a, \psi) \in (S \cap \rightarrow)$, we have that $a \in S$, if and only if $\bar{f}(S)$ is admissible in $f(F)$. □

Theorem 1. Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $f(F) = \langle MA, \rightarrow_2 \rangle$ and $S \subseteq (\mathcal{A} \cup \rightarrow)$. S is a complete extension of F if and only if $\bar{f}(S)$ is a complete extension of $f(F)$.

Proof:

Let $F = \langle \mathcal{A}, \rightarrow \rangle$ be an AFRA, $f(F) = \langle MA, \rightarrow_2 \rangle$ and $S \subseteq (\mathcal{A} \cup \rightarrow)$.

1. $\Rightarrow$: Suppose S is a complete extension of F. For every $(a, \psi) \in (S \cap \rightarrow)$, by the definition of defeat, a is defended by S, and thus $a \in S$. Therefore, by Lemma 3, $\bar{f}(S)$ is admissible.

   Take some arbitrary $p \in MA$ and suppose that p is defended by $\bar{f}(S)$. Then, either $p = m(\varphi)$ for some $\varphi \in (\mathcal{A} \cup \rightarrow)$, or $p = X_{a,b}$ for some $a, b \in \mathcal{A}$.

   (a) Suppose that $p = m(\varphi)$ for some $\varphi \in (\mathcal{A} \cup \rightarrow)$. Now assume that $\varphi \in \rightarrow$. Then, $m(\varphi) = Y_{src(\varphi),trg(\varphi)}$. By construction of $\rightarrow_2$, we have that $X_{src(\varphi),trg(\varphi)} \rightarrow_2 Y_{src(\varphi),trg(\varphi)}$. The only argument which can defend $Y_{src(\varphi),trg(\varphi)}$ from $X_{src(\varphi),trg(\varphi)}$ is $acc(src(\varphi))$. Since $\bar{f}(S)$ defends $Y_{src(\varphi),trg(\varphi)}$, we have that $acc(src(\varphi)) \in \bar{f}(S)$. As $\bar{f}(S)$ is admissible, $acc(src(\varphi))$ is defended by $\bar{f}(S)$. Hence, if $\varphi \in \rightarrow$, then $acc(src(\varphi))$ is defended by $\bar{f}(S)$.

   Therefore, by Lemma 2, φ is defended by S. Since S is a complete extension, this means that $\varphi \in S$. Therefore, $p = m(\varphi) \in \bar{f}(S)$.

   (b) Now suppose that $p = X_{a,b}$ for some $a, b \in \mathcal{A}$. According to our assumptions, $\bar{f}(S)$ defends $X_{a,b}$. By construction of $\Rightarrow_2$, the only argument which attacks $X_{a,b}$ is $acc(a)$. Hence, there exists $Y_{c,a} \in \bar{f}(S)$ for some $c \in \mathcal{A}$. So, by definition of $\bar{f}$, we have that $p = X_{a,b} \in \bar{f}(S)$.

   In either case, we have that $p \in \bar{f}(S)$. Hence, $\bar{f}(S)$ contains all arguments it defends. Since it is also admissible, $\bar{f}(S)$ is a complete extension of $f(F)$.

2. $\Leftarrow$: Suppose that $\bar{f}(S)$ is a complete extension of $f(F)$. Then, $\bar{f}(S)$ is admissible and contains all arguments it defends. According to Lemma 3, we have that S is admissible and for every $(a, \psi) \in (S \cap \rightarrow)$, we have that $a \in S$. Suppose that for some $\varphi \in (\mathcal{A} \cup \rightarrow)$, φ is defended by S. Hence, by Lemma 2, $m(\varphi)$ is defended by $\bar{f}(S)$. Since $\bar{f}(S)$ is a complete extension of $f(F)$, $m(\varphi) \in \bar{f}(S)$. Hence, by construction of $\bar{f}(S)$, we have that $\varphi \in S$. Therefore, for any $\varphi \in (\mathcal{A} \cup \rightarrow)$ such that φ is defended by S, we have $\varphi \in S$. Since S is also admissible, S is a complete extension of F.

Hence, S is a complete extension of F if and only if $\bar{f}(S)$ is a complete extension of $f(F)$. □

3.2 Extended Explanatory Argumentation Frameworks

We will now extend EAFs, as seen in Section 2.1, by integrating them with the meta-argumentation techniques we have discussed so far.

Definition 18. An *extended explanatory argumentation framework* (EEAF) is a tuple $\langle \mathcal{A}, \mathcal{X}, \rightarrow, \dashrightarrow, \sim, \Rightarrow \rangle$, where $\mathcal{A}$ is a set of arguments, $\mathcal{X}$ is a set of explananda, $\dashrightarrow \subseteq (\mathcal{A} \times \mathcal{A}) \cup (\mathcal{A} \times \mathcal{X})$ is an explanatory relation, $\rightarrow \subseteq (\mathbb{P}(\mathcal{A}) \cup \dashrightarrow \cup \rightarrow) \times (\mathcal{A} \cup \dashrightarrow \cup \rightarrow \cup \Rightarrow)$ is a higher-order attack relation, $\sim \subseteq \mathcal{A} \times \mathcal{A}$ is an incompatibility relation and $\Rightarrow \subseteq \mathcal{A} \times \mathcal{A}$ is a support relation.

We then define the semantics of EEAFs in terms of their flattening.

Definition 19. Let $F = \langle \mathcal{A}, \mathcal{X}, \rightarrow, \dashrightarrow, \sim, \Rightarrow \rangle$ be an EEAF. The set of meta-arguments corresponding to F is $MA = \{acc(a), rej(a) \mid a \in \mathcal{A}\} \cup \{X_{m(\varphi), m(\psi)}, Y_{m(\varphi), m(\psi)} \mid \varphi \in (\mathcal{A} \cup \rightarrow \cup \dashrightarrow), \psi \in (\mathcal{A} \cup \rightarrow \cup \dashrightarrow \cup \Rightarrow)\} \cup \{e(S) \mid S \subseteq \mathcal{A}$ with at least two elements$\} \cup \{P_{a,\psi}, Q_{a,\psi} \mid a \in \mathcal{A}, \ \psi \in (\mathcal{A} \cup \mathcal{X})\} \cup \{Z_{a,b} \mid a, b \in \mathcal{A}\}$ and the set of meta-explananda is $MX = \mathcal{X}$. We define a partial function m which assigns for each element of the framework a corresponding meta-argument.

$$m: (\mathcal{A} \cup \rightarrow \cup \dashrightarrow \cup \Rightarrow) \mapsto MA.$$

such that:

– if $\varphi \in \mathcal{A}$, then $m(\varphi) = acc(\varphi)$;
– if $\varphi \in \mathcal{X}$, then $m(\varphi) = \varphi$;
– if $\varphi \in \Rightarrow$ such that for some $a, b \in \mathcal{A}$, $\varphi = (a \Rightarrow b)$, then $m(\varphi) = Z_{a,b}$;
– if $\varphi \in \rightarrow$ such that for some $S \subseteq \mathcal{A}$ with at least two elements and some $\psi \in (\mathcal{A} \cup \dashrightarrow \cup \rightarrow \cup \Rightarrow)$, $\varphi = (S \rightarrow \psi)$, then $m(\varphi) = e(S)$;
– if $\varphi \in \rightarrow$ such that for some $\psi \in (\mathcal{A} \cup \dashrightarrow \cup \rightarrow)$ and some $\delta \in (\mathcal{A} \cup \dashrightarrow \cup \rightarrow \cup \Rightarrow)$, $\varphi = (\psi \rightarrow \delta)$, then $m(\varphi) = Y_{\psi, \delta}$;
– if $\varphi \in \dashrightarrow$ such that for some $a \in \mathcal{A}$ and $\psi \in (\mathcal{A} \cup \mathcal{X})$, $\varphi = (a \dashrightarrow \psi)$, then $m(\varphi) = P_{a,\psi}$.

We define the *flattening function* f to be $f(F) = \langle MA, \mathcal{X}, \rightarrow_2, \dashrightarrow_2, \sim_2 \rangle$, where $\rightarrow_2, \sim_2 \subseteq MA \times MA$ and $\dashrightarrow_2 \subseteq MA \times (MA \cup \mathcal{X})$ are such that:

– $X_{m(\varphi), m(\psi)} \rightarrow_2 Y_{m(\varphi), m(\psi)}, Y_{m(\varphi), m(\psi)} \rightarrow_2 m(\psi)$ for all $\varphi, \psi \in (\mathcal{A} \cup \rightarrow \cup \dashrightarrow \cup \Rightarrow)$;
– $m(\varphi) \rightarrow_2 X_{m(\varphi), m(\psi)}$ if and only if $\varphi \rightarrow \psi$ and φ is not a set of arguments with at least two elements;
– $acc(a) \rightarrow_2 rej(a)$ for all $a \in \mathcal{A}$;
– $e(S) \rightarrow_2 m(\varphi)$ if and only if $S \rightarrow \varphi$ for $S \subseteq \mathcal{A}$ with at least 2 elements;
– $rej(a) \rightarrow_2 e(S)$ if and only if $a \in S$;
– $Z_{a,b} \rightarrow_2 acc(a)$ for all $a, b \in \mathcal{A}$;
– $acc(b) \rightarrow_2 Z_{a,b}$ if and only if $a \Rightarrow b$;
– $acc(a) \dashrightarrow_2 P_{a,\varphi}, P_{a,\varphi} \dashrightarrow_2 m(\varphi), acc(a) \rightarrow_2 Q_{a,\varphi}$ and $Q_{a,\varphi} \rightarrow_2 P_{a,\varphi}$ if and only if $a \dashrightarrow \varphi$;
– $acc(a) \sim_2 acc(b)$ if and only if $a \sim b$.

Notice that the set of meta-arguments MA and the correspondence function m are defined through a simultaneous inductive definition, which is well-founded, because $\rightarrow$ is a well-founded relation (assuming that the set theory presupposed in Definition 6 is a standard set theory like ZFC that satisfies the Axiom of Foundation).

Note that we do not fully flatten the explanatory relation and flatten EEAFs into EAFs instead of AFs. This is due to the fact that the explanatory relation is not easily flattened, and extensions can still be extracted from explanatory argumentation frameworks via the two EAF extensions which are well-suited for our task. In order to do this, we need to define an unflattening function which will map a set of meta-arguments from a flattened EEAF to the corresponding set of arguments from the original EEAF.

Definition 20. Given an EEAF F and a set of meta-arguments $B \subseteq MA$ such that MA corresponds to F, we define the *unflattening function g* to be:

$$g(B) = \{a \mid acc(a) \in B\}$$

Notice that in the unflattening, we only care about the arguments and do not unflatten the meta-arguments which represent the other elements of EEAFs. This is due to the fact that we are only interested in selecting the arguments of the EEAF, which make up the argumentative and explanatory cores.

Definition 21. Let F be an EEAF and $G = f(F)$ its flattening. We say that $S \subseteq \mathcal{A}$ is an *AC-extension* of F iff $S = g(S')$, where S' is an AC-extension of G. Similarly, we say that $S \subseteq \mathcal{A}$ is an *EC-extension* of F iff $S = g(S')$, where S' is an EC-extension of G.

4 Applying EEAFs to the liar paradox

Let us now move on to an example, which focuses on two groups of solutions for the liar paradox. The arguments are extracted from *Saving Truth from Paradox* [8]. The first group is the solutions which weaken classical logic, namely the paracomplete, paraconsistent and semi-classical solutions. The second group is comprised of the underspill and overspill solutions.

We have the following arguments:

- E_p: This explanandum represents the paradox.
- A: The paracomplete, paraconsistent and semi-classical solutions which provide explanations for the paradox by weakening classical logic.
- B: The underspill and overspill solutions which provide their own explanation of the paradox by suggesting that for some predicates F, F is true of some objects that aren't F or vice-versa.
- C: We did not change logic to hide the defects in other flawed theories such as Ptolemaic astronomy, so why should we change the logic simply to hide these paradoxes?

    – *D*: There is no known way of saving these flawed theories such as Ptolemaic astronomy and even if there was, there is little benefit to doing so.
    – *F*: We have worked out the details of the new logics and they allow us to conserve the theory of truth.
    – *G*: Changing the logic implies changing the meaning.
    – *H*: Change of meaning is bad.
    – *I*: The change is mere.
    – *J*: This is no 'mere' relabelling.
    – *K*: Change of truth schema is a change of the meaning of 'true'.
    – *L*: The paradox forces a change of meaning.

The framework is represented in Figure 6 and its flattening in Figure 7. We have omitted less-relevant auxiliary arguments for the sake of visibility.
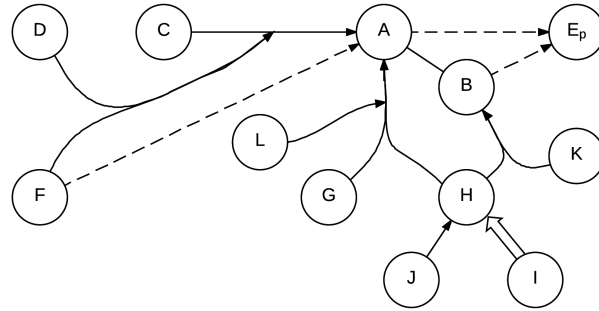

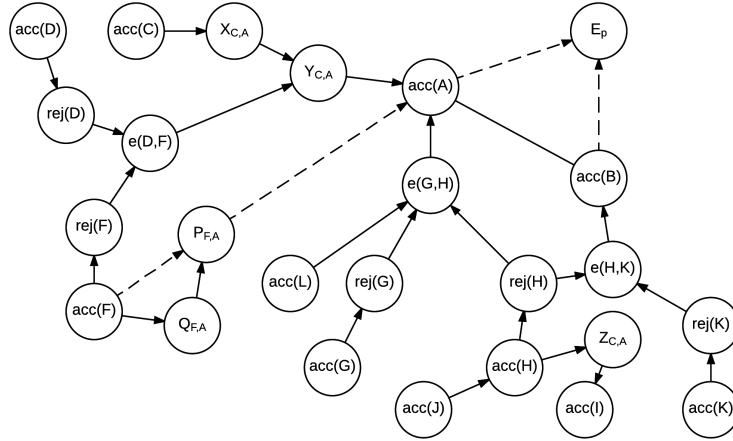**Fig. 6.** EEAF representing the reasoning behind the excerpt

Fig. 7. Flattened EEAF representing the reasoning behind the excerpt

We get that the AC-extensions are $\{A, C, D, F, L, G, J, K\}$ and $\{B, C, D, F, L, G, J, K\}$. We can distinguish here the two rivaling solutions which are both

selected. This is due to the fact that even though the author might have a preference for one or another, in the excerpt we have analyzed, he is merely defending the solutions represented in A from attacks and making no argument which attacks the solutions represented in B.

The EC-extensions are $\{A, D, F, L\}$, $\{A, D, F, J\}$ and $\{B, J\}$. Notice that there are two different EC-extensions which contain A, as there are two arguments which individually defend A from the coalition attack of $\{G, H\}$.

5 Conclusion and future work

We have examined several extensions of abstract argumentation frameworks that add explanatory features, recursive attacks, support and joint attacks. In the cases of recursive attacks, support and joint attacks, we have presented a flattening function, which allows us to instantiate these extended framework as standard AFs. We have shown that in the case of AFRAs, the complete semantics defined in terms of the flattening is equivalent to the complete semantics which has been defined directly on AFRAs. We have then aggregated these extensions into one framework, EEAFs, and defined the semantics in terms of its flattening to EAFs. Finally, we have explored an application of EEAFs to argumentation from a research-level philosophy book.

Concerning future work in the line of research of this paper, we plan to extend the result about the flattening of AFRAs to other argumentation semantics than the complete semantics. Furthermore, it might be interesting to investigate flattening the explanatory relation and explananda. Due to their intricate nature, it is not obvious how to flatten them and obtain semantics equivalent to the ones defined on EAFs. Another point of interest would be to apply EEAFs to other areas of scientific debates and examine whether the current features provide enough expressive power.

References

1. Pietro Baroni, Guido Boella, Federico Cerutti, Massimiliano Giacomin, Leendert van der Torre, and Serena Villata. On the input/output behavior of argumentation frameworks. *Artif. Intell.*, 217:144–197, 2014.
2. Pietro Baroni, Federico Cerutti, Massimiliano Giacomin, and Giovanni Guida. Encompassing attacks to attacks in abstract argumentation frameworks. In *EC-SQARU*, pages 83–94. Springer, 2009.
3. Guido Boella, Dov M Gabbay, Leendert van der Torre, and Serena Villata. Meta-argumentation modelling i: Methodology and techniques. *Studia Logica*, 93(2-3):297–355, 2009.
4. Guido Boella, Dov M Gabbay, Leendert van der Torre, and Serena Villata. Support in abstract argumentation. *COMMA*, 216:111–122, 2010.
5. C. Cayrol and M.C. Lagasquie-Schiex. Bipolarity in argumentation graphs: Towards a better understanding. *IJAR*, 54(7):876–899, 2013.

6. Andrea Cohen, Sebastian Gottifredi, Alejandro Javier García, and Guillermo Ricardo Simari. On the acceptability semantics of argumentation frameworks with recursive attack and support. In *Computational Models of Argument - Proceedings of COMMA 2016*, pages 231–242, 2016.

7. Phan Minh Dung. 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.

8. Hartry Field. *Saving truth from paradox*. Oxford University Press, 2008.

9. Dov M Gabbay. Fibring argumentation frames. *Studia Logica*, 93(2-3):231–295, 2009.

10. Dunja Šešelja and Christian Straßer. Abstract argumentation and explanation applied to scientific debates. *Synthese*, 190(12):2195–2217, 2013.