Merging argumentation systems with weighted argumentation systems: a preliminary study

C. Cayrol, M.-C. Lagasquie-Schiex
IRIT-UPS, Université de Toulouse
ccayrol@irit.fr, lagasq@irit.fr

Rapport IRIT

numéro IRIT/RR- -2011-18- -FR

Juin 2011

Abstract

In this paper, we address the problem of merging argumentation systems in a multi-agent setting. Each agent's system may be built from different sets of arguments and/or different interactions between these arguments. The merging process must lead to solve conflicts between the agents and to identify argumentation systems representing the knowledge of the group of agents.

Previous work [10] has proposed a two-step merging process in which conflicts about an interaction result in a new kind of interaction, called ignorance. However, this merging process is computationally expensive, and does not provide a single resulting argumentation system.

We propose a novel approach to overcome these limitations by introducing a refinement of the ignorance relation under the form of a weighted attack. Our merging process takes only one step and provides a single weighted argumentation system, which is easy to compute.

Contents

1 Introduction

Argumentation has become an influential approach in Artificial Intelligence to model cognitive tasks such as inconsistency handling and defeasible reasoning (e.g. [17, 24, 1, 26]), decision making (e.g. [13, 14, 16, 15]), or negotiation between agents (e.g. [23, 4, 3]). Argumentation is based on the evaluation of interacting arguments, which may support opinions, claims or decisions. Usually, the interaction takes the form of conflicts between arguments, and the fundamental issue is the selection of acceptable sets of arguments, based on the way they interact. Most of the argumentation-based proposals are instantiations of the abstract system proposed by Dung [11], which is reduced to a set of arguments (completely abstract entities) and a binary relation, called attack, which captures the conflicts between arguments. The increasing interest for the argumentation formalism has led to numerous extensions of the basic abstract system which are more appropriate to the applications.

Among them we are interested in a framework proposed for the *merging of argumentation systems*. Indeed in [10], a multi-agent setting has been considered in order to define an argumentation system for a group of agents from their individual argumentation systems. This amounts to make precise the set of arguments of the group and the global attack relation for the group. This can be seen as an idealized, "fully-informed", negotiation protocol: all agents give their arguments and attack relation, and agree to find a result for the whole group. In [10], it has been shown that a direct voting on the (sets of) arguments acceptable for each agent leads to counter-intuitive results because (1) voting make sense only if all agents consider the same set of arguments and (2) attack relations cannot be taken into consideration any more after the computation of the acceptable (sets of) arguments (this last point leads to let aside much significant information for defining the acceptable sets of arguments at the group level). So, [10] has proposed a two-step merging process which first expands each agent's argumentation system for taking into account all the arguments (1), then computes a set of argumentation systems which are as close as possible to these expansions (2). For the first step, [10] has extended Dung's argumentation system with a new kind of interaction between arguments, the *ignorance relation*, so as to reflect that an agent cannot conclude from the other agents that there is or not an attack between two arguments.

However, [10]'s approach suffers from two important drawbacks: computing the outputs of the merging process (a set of argumentation systems) is expensive and in the general case the merging process does not provide a single argumentation system, which complicates the definition of acceptability of (sets of) arguments for the group.

We propose a novel approach to the merging of argumentation systems which overcomes these limitations. Our idea is to refine the notion of ignorance. Indeed in [10], ignorance is viewed as an intermediary between attack and absence of attack. We propose to replace ignorance by a kind of gradual attack, under the form of a *weighted attack*. Recent works have proposed extensions of Dung's argumentation system capable of handling varied-strength attacks ([18, 12, 6]). The idea is that attacks may have different strength and can be compared according to their relative strength. Our purpose is to take advantage of these works, in order to define a merging process such that the resulting system is a unique weighted argumentation system, easy to compute and to use.

In Section 2, we give the relevant background on classical argumentation systems and merging of argumentation systems. Section 3 introduces the basic ideas for a one-step merging process. Then, in Section 4, we propose a new approach for merging argumentation system into a weighted argumentation system. An extension for merging bipolar argumentation systems can be easily defined (see Section 5). Section 6 concludes the paper.

2 Background

2.1 Argumentation systems

We first focus on Dung's theory of argumentation [11] in which only one interaction between arguments is taken into account.

Def 1 *An abstract* argumentation system $\mathbf{AS} = \langle \mathbf{A}, \mathbf{R} \rangle$ *over* $\mathbf{A}$ *is given by a finite*[1] *set* $\mathbf{A}$ *of arguments and a binary relation* $\mathbf{R}$ *on* $\mathbf{A}$ *called an* attack relation. *Consider* a_i *and* $a_j \in \mathbf{A}$. $a_i \mathbf{R} a_j$ *means that* a_i *attacks* a_j *(also denoted by* $(a_i, a_j) \in \mathbf{R}$*).*

$\langle \mathbf{A}, \mathbf{R} \rangle$ clearly defines a directed graph $\mathbf{G}$ called the *interaction graph* whose nodes are arguments and edges correspond to the attack relation.

Whether a set of arguments can be accepted depends on the way arguments interact within the set but also w.r.t. the other arguments of $\mathbf{A}$. Collective acceptability is based on two key notions: conflict-freeness and collective defence.

Def 2 *Let* $\langle \mathbf{A}, \mathbf{R} \rangle$ *be an argumentation system.*

Conflict-free set *A set* $\mathbf{E} \subseteq \mathbf{A}$ *is* conflict-free *if and only if* $\nexists a, b \in \mathbf{E}$ *such that* $a\mathbf{R}b$.

Collective defence *Consider* $\mathbf{E} \subseteq \mathbf{A}$, $a \in \mathbf{A}$. $\mathbf{E}$ (collectively) defends *a if and only if* $\forall b \in \mathbf{A}$, *if* $b\mathbf{R}a, \exists c \in \mathbf{E}$ *such that* $c\mathbf{R}b$. $\mathbf{E}$ defends all its elements *if and only if* $\forall a \in \mathbf{E}$, $\mathbf{E}$ *collectively defends* a.

[1] A usual restriction which corresponds to the real world.

[11] defines several semantics for collective acceptability based on those two notions; among them the *admissible semantics* and the *preferred semantics.*

Def 3 *Let $\langle \mathbf{A}, \mathbf{R} \rangle$ be an argumentation system.*

Admissible semantics *A set $\mathbf{E} \subseteq \mathbf{A}$ is* admissible *iff $\mathbf{E}$ is conflict-free and $\mathbf{E}$ defends all its elements.*

Preferred semantics *A set $\mathbf{E} \subseteq \mathbf{A}$ is a* preferred extension *iff $\mathbf{E}$ is maximal for set inclusion among the admissible sets.*

2.2 Merging of argumentation systems

We consider a multi-agent setting where each agent has its own argumentation system. In the general case, we cannot assume that one agent knows all the arguments which are known by the other agents. Moreover, even if two agents share a set of arguments, they may disagree on the interactions between these arguments. This problem already occurs in argumentation-based dialogues. Another interesting issue is the combination of the different argumentation systems in order to achieve a collective output, that is a collective decision about which arguments should be accepted.

This issue has been discussed in [10], where an approach for merging argumentation systems has been proposed.

Firstly, [10] has shown that the merging process cannot be simply done at the output level of each system (*i.e.* on the sets of accepted arguments). So, the proposal rather consists in combining the interactions, before computing the accepted arguments. It takes two steps: First each agent's argumentation system is expanded into an argumentation system including all the arguments known by other agents. Then a set of argumentation systems which are as close as possible to these expansions is computed.

In the first step, called expansion, each agent adds new arguments and new interactions to her previous knowledge. These new interactions result from a kind of merging of the knowledge held by the other agents. Namely, let us consider a situation where the argumentation system of an agent ag contains neither the argument a nor the argument b, but some other agents know these arguments. Different cases may occur: If all the other agents who know a and b agree on an attack (resp. the absence of attack) from a to b, the agent ag will add a, b and $a\mathbf{R}b$ (resp. a, b and no interaction between them) to her system. Otherwise, ag will add a, b and a new kind of interaction meaning that she lacks knowledge for establishing either that there is an attack or that there is no attack from a to b. In [10], this new kind of interaction is called *ignorance*.

Formally, it leads to consider a new type of argumentation systems, called *partial argumentation systems.*

Def 4 ([10] Partial argumentation system (PAS)) *A* partial argumentation system *over* $\mathbf{A}$ *is a quadruple* $\mathbf{PAS} = \langle \mathbf{A}, \mathbf{R}, \mathbf{I}, \mathbf{N} \rangle$ *where*

- $\mathbf{A}$ *is a finite set of arguments,*
- $\mathbf{R}, \mathbf{I}, \mathbf{N}$ *are binary relations on* $\mathbf{A}$*:*
  - $\mathbf{R}$ *is the* attack relation*,*
  - $\mathbf{I}$ *is called the* ignorance relation *and is such that* $\mathbf{R} \cap \mathbf{I} = \varnothing$*,*
  - *and* $\mathbf{N} = (\mathbf{A} \times \mathbf{A}) \setminus (\mathbf{R} \cup \mathbf{I})$ *is called the* non-attack relation*.*

$\mathbf{N}$ *is deduced from* $\mathbf{A}$*,* $\mathbf{R}$ *and* $\mathbf{I}$*, so a partial argumentation system can be fully specified by* $\langle \mathbf{A}, \mathbf{R}, \mathbf{I} \rangle$*.*

Note that an argumentation system $\mathbf{AS} = \langle \mathbf{A}, \mathbf{R} \rangle$ can be considered as a particular case of a $\mathbf{PAS}$ by taking $\mathbf{I} = \varnothing$ and $\mathbf{N} = (\mathbf{A} \times \mathbf{A}) \setminus \mathbf{R}$.

Now, the notion of expansion recalled above can be formalized. Intuitively, the expansion of an argumentation system $\langle \mathbf{A}, \mathbf{R} \rangle$ given a profile of such systems is obtained by adding a pair of arguments (a, b) (where at least one of a, b is not in $\mathbf{A}$) into the attack (resp. the non-attack relation) *provided that all other agents of the profile who know the two arguments agree on the existence of the attack*[2] (resp. *the non-attack*); otherwise, it is added to the ignorance relation.

Def 5 ([10] Consensual expansion) *Let* $\mathcal{P} = \langle \mathbf{AS}_1, \ldots, \mathbf{AS}_n \rangle$ *be a profile of* n *argumentation systems such that* $\mathbf{AS}_i = \langle \mathbf{A}_i, \mathbf{R}_i \rangle$*. Let* $\mathbf{AS} = \langle \mathbf{A}, \mathbf{R} \rangle$ *be an argumentation system. Let* $conf(\mathcal{P}) = (\bigcup_i \mathbf{R}_i) \cap (\bigcup_i \mathbf{N}_i)$ *be the set of interactions for which a conflict exists within the profile. The* consensual expansion *of* $\mathbf{AS}$ *over* $\mathcal{P}$ *is the* $\mathbf{PAS}$ *specified by the tuple* $\langle \mathbf{A}', \mathbf{R}', \mathbf{I}' \rangle$ *with:*

- $\mathbf{A}' = \mathbf{A} \cup \bigcup_i \mathbf{A}_i$*,*
- $\mathbf{R}' = \mathbf{R} \cup ((\bigcup_i \mathbf{R}_i \setminus conf(\mathcal{P})) \setminus (\mathbf{A} \times \mathbf{A}))$*,*
- $\mathbf{I}' = conf(\mathcal{P}) \setminus (\mathbf{A} \times \mathbf{A})$*.*

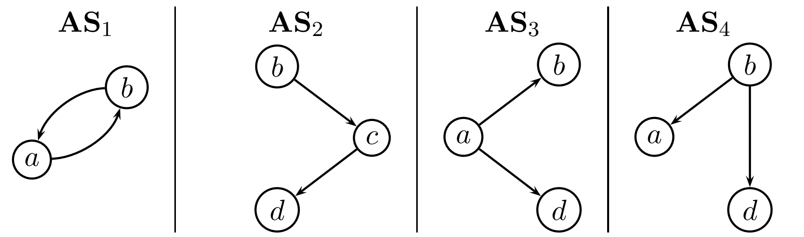
Note that the computation of the consensual expansion does not modify the knowledge previously contained in $\mathbf{AS}$.

The consensual expansion is used at the first step of the merging process as follows. We consider a finite set of agents ag_i, each one with her argumentation system $\mathbf{AS}_i$. Each agent ag_i computes the partial argumentation system $\mathbf{PAS}_i$ over $\mathbf{A} = \bigcup_i \mathbf{A}_i$, which is the consensual expansion of $\mathbf{AS}_i$ over the profile $\langle \mathbf{AS}_1, \ldots, \mathbf{AS}_n \rangle$.

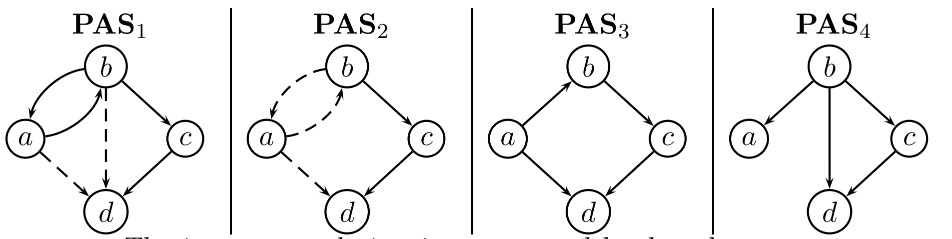
This expansion policy is sensible as soon as each agent has a minimum level of confidence in the other agents: if a piece of information conveyed by one agent is not conflicting with the information stemming from the other agents, every agent of the group is ready to accept it.

[2] *i.e.*, if $a, b \in \mathbf{A}_i$, then $(a, b) \in \mathbf{R}_i$.

Ex 1 *Consider the profile consisting of the following four argumentation systems:*



For each i, the consensual expansion $\mathbf{PAS}_i$ of $\mathbf{AS}_i$ is given by:



The ignorance relation is represented by dotted arrows.

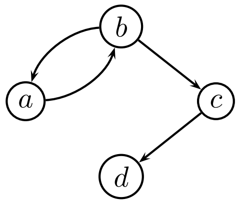
For instance, $\mathbf{PAS}_1$ is obtained by:

- *adding the arguments c and d (unknown by Agent 1 and known by other agents),*
- *keeping the attacks between a and b and the non-attacks from a to a and from b to b (known by Agent 1),*
- *adding the attacks from b to c and c to d (known by Agent 2 and without conflict with the other agents) and also many non-attacks (from c to c, from d to b, $\ldots$),*
- *adding the ignorance from b to d resulting from a conflict between Agents 2 and 3 and Agent 4: for Agents 2 and 3, there is a non-attack from b to d and for Agent 4 there is an attack from b to d,*
- *adding the ignorance from a to d resulting from a conflict between Agent 3 and Agent 4: for Agent 3, there is an attack from a to d and for Agent 4, there is a non-attack from a to d.*

The second step of the merging process aims at characterizing the argumentation systems which are as "close" as possible to the profile $\mathcal{P}'$ of partial argumentation systems, obtained at the first step and taken as a whole. This can be realized using distances between an AS and a PAS and aggregation functions for computing the distance between an AS and a profile of partial argumentation systems.

Let us now illustrate the whole process on Example 1 using the edit distance[3] and the sum aggregation function in order to take into account the number of agents.

Ex 1 (cont'd) *First, let us consider the following argumentation system* $\mathbf{AS}'_1 = \langle\{a,\ b,\ c,\ d\},\ \{(a,b),(b,a),\ (b,c),(c,d)\}\rangle$*, represented by:*

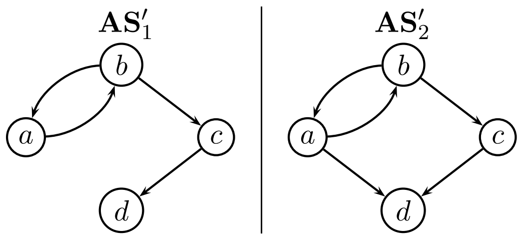


The edit distance between $\mathbf{AS}'_1$ *and each of the PASs obtained by consensual expansion over the profile* $\langle \mathbf{AS}_1, \mathbf{AS}_2, \mathbf{AS}_3, \mathbf{AS}_4\rangle$ *is:*

- $de(\mathbf{AS}'_1, \mathbf{PAS}_1) = 1$,
- $de(\mathbf{AS}'_1, \mathbf{PAS}_2) = 1.5$,
- $de(\mathbf{AS}'_1, \mathbf{PAS}_3) = 2$,
- $de(\mathbf{AS}'_1, \mathbf{PAS}_4) = 2$.

so we obtain $\Sigma_{i=1}^{4} de(\mathbf{AS}'_1, \mathbf{PAS}_i) = 6.5$.

Then, by computing such distances for all candidate ASs (i.e., all ASs over $\{a, b, c, d\}$*), we obtain as the result of the merging process the two following ASs:*



The above example highlights the drawbacks discussed in the introduction of the paper. A distance must be computed between the profile of PASs and each argumentation system which is a candidate for representing the profile. Moreover, several resulting ASs may be obtained.

3 Towards a one-step merging process

A first idea for simplifying the whole process could be to propose a rather different approach which consists in bypassing the first step and directly

[3]The edit distance between an AS and a PAS is computed using the following costs: an attack (resp. a non-attack) in one system which is a non-attack (resp. an attack) in the other system induces a cost of 1, and an ignorance in one PAS which is an attack or a non-attack in the AS induces a cost of 0.5. Then the edit distance is the sum of these costs.

computing a single PAS over the set of all the arguments. That would correspond to a fusion operation defined as follows:

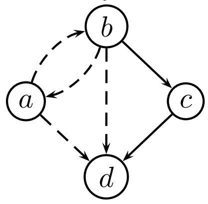
Def 6 (Consensual fusion) *Let $\mathcal{P} = \langle \mathbf{AS}_1, \ldots, \mathbf{AS}_n \rangle$ be a profile of n argumentation systems such that $\mathbf{AS}_i = \langle \mathbf{A}_i, \mathbf{R}_i \rangle$. Let $conf(\mathcal{P}) = (\bigcup_i \mathbf{R}_i) \cap (\bigcup_i \mathbf{N}_i)$ be the set of interactions for which a conflict exists within the profile.*

The consensual fusion *of $\mathcal{P}$ is the* **PAS** *specified by the tuple $\langle \mathbf{A}, \mathbf{R}, \mathbf{I} \rangle$ with:*

- $\mathbf{A} = \bigcup_i \mathbf{A}_i$,
- $\mathbf{R} = (\bigcup_i \mathbf{R}_i \setminus conf(\mathcal{P}))$,
- $\mathbf{I} = conf(\mathcal{P})$.

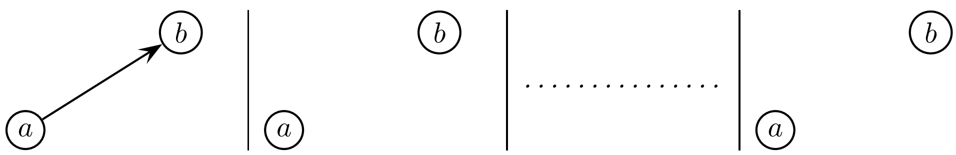
This consensual fusion results in a single PAS and is easy to compute. However, the resulting PAS is little informative, since the ignorance relation may contain many pairs of arguments.

Ex 1 (cont'd) *Using the consensual fusion leads to the following PAS:*

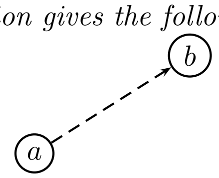


Another drawback is illustrated by the following example:

Ex 2 *Consider n agents with their respective ASs (only one agent considers that there is an attack between a and b, and the other $n-1$ agents consider that there is not):*



Then the consensual fusion gives the following PAS:



For any two arguments a and b, $a\mathbf{I}b$ occurs in the consensual fusion as soon there is a conflict between the agents, regardless the number of agents who know that a attacks b and the number of agents who know that a does not attack b.

Taking into account these numbers will enable to graduate the interaction between a and b. We consider that in case of a conflict between the agents, we obtain a (more or less) weak attack. The strongest attack corresponds to the case when all the agents agree on the existence of the attack. So, our proposal is to replace the ignorance relation in the PAS by a weighted attack relation.

4 Merging of argumentation systems into a weighted argumentation system

Some background about weighted argumentation systems is first presented, then we propose to merge a profile of ASs into a weighted argumentation system.

4.1 Weighted argumentation systems

[12] has proposed weighted argument systems, in which attacks are associated with a numeric weight, indicating how reluctant one would be to disregard the attack. However, these weights are not used for disregarding or comparing defences. Given a threshold, the idea is to combine the weights additively, and to disregard subsets of attacks which sum to no more than the threshold.

Some other works ([18, 20, 19, 6]) explore the common basic idea to use the relative strength of the attacks for disregarding some of the defences: classically, if an argument a is attacked by an argument b, any attacker of b is relevant for inhibiting the attack on a, thus defending a; with attacks of different strengths, it is natural to require that the attack on b is strong enough to reinstate a.

In this paper, we are interested by using an argumentation system with varied-strength attacks considering that this strength is given by numbers as in [12] and is used for defining a new notion of defence.

Def 7 ([12] Weighted argumentation system) *A* weighted argumentation system *over* $\mathbf{A}$ *is a triple* $\langle \mathbf{A}, \mathbf{R}, w \rangle$*, denoted by* **WAS***, where*

- $\mathbf{A}$ *is a finite set of arguments,*
- $\mathbf{R}$ *is an attack relation on* $\mathbf{A}$ *and*
- w *is a function* $\mathbf{R} \to \mathbb{R}$ *assigning real valued weights to attacks.*

A weighted argumentation system can be considered as a particular case of the argumentation system with varied-strength attacks of [19, 6], using the natural ordering $\geq$ on $\mathbb{R}$. Then, considering that greater is the weight $w(a, b)$, stronger is the attack from a to b, the notion of vs-defence can be defined as:

Def 8 ([6] vs-defence – vs means "various-strength") *Let* $\mathbf{WAS} = \langle \mathbf{A}, \mathbf{R}, w \rangle$. *Let* a, b, $c \in \mathbf{A}$ *such that* $c\mathbf{R}b$ *and* $b\mathbf{R}a$. c vs-defends *a against b (or c is a* vs-defender *of a against b) iff* $w(c, b) \geq w(b, a)$ *(i.e. the attack from b to a is not strictly stronger than the one from c to b).*

Then using the vs-defence, vs-admissible semantics can be defined:

Def 9 ([6] vs-admissibility) *Let* $\mathbf{WAS} = \langle \mathbf{A}, \mathbf{R}, w \rangle$. *Let* $\mathbf{E} \subseteq \mathbf{A}$.

- $\mathbf{E}$ *is* conflict-free *in* $\mathbf{WAS}$ *iff* $\nexists a, b \in \mathbf{E}$ *such that* $b\mathbf{R}a$.
- $\mathbf{E}$ vs-defends *a iff* $\forall b \in \mathbf{A}$, *if* $b\mathbf{R}a$ *then* $\exists c \in \mathbf{E}$ *such that c vs-defends a against b.*
- $\mathbf{E}$ *is* vs-admissible *iff* $\mathbf{E}$ *is conflict-free and* $\forall a \in \mathbf{E}$, $\mathbf{E}$ vs-defends a.

In the following, we propose to merge a profile of ASs into a WAS such that the weight on an attack in the WAS represents the strength of the opinions of the group of agents concerning this attack.

4.2 Definition

We use a global strategy for computing the argumentation system issued from the group of agents. A natural idea is to consider each agent as a voting person and to add the votes for and against each interaction.

Def 10 (Merging of n ASs) *Let* $\mathcal{P} = \langle \mathbf{AS}_1, \ldots, \mathbf{AS}_n \rangle$ *be a profile of n argumentation systems with* $\mathbf{AS}_i = \langle \mathbf{A}_i, \mathbf{R}_i \rangle$. *The merging of $\mathcal{P}$ is the WAS defined by* $\mathbf{WAS} = \langle \mathbf{A}, \mathbf{R}, w \rangle$ *with:*

- $\mathbf{A} = \bigcup_{i=1}^{n} \mathbf{A}_i$
- $\mathbf{R} = \bigcup_{i=1}^{n} \mathbf{R}_i$
- $w : \mathbf{R} \rightarrow]0, +1]$ *defined by* $\forall (a, b) \in \mathbf{R}$, $w(a, b) = \frac{\Sigma_{i=1}^{n} w_i(a,b)}{n(a,b)}$ *with*
  - $w_i(a, b) = 1$ *if* $(a, b) \in \mathbf{R}_i$,
  - $w_i(a, b) = 0$ *if* $(a, b) \notin \mathbf{R}_i$
  - *and* $n(a, b)$ *is the number of* $\mathbf{AS}_i$ *such that* $a \in \mathbf{A}_i$ *and* $b \in \mathbf{A}_i$.

The following observation follows trivially from Definition 10.

Observation: Let $a, b \in \mathbf{A}$.

- If $(a, b) \in \mathbf{R}$, there exists at least one $\mathbf{AS}_i$ such that $(a, b) \in \mathbf{R}_i$, and so $n(a, b) > 0$.
- $(a, b) \notin \mathbf{R}$ means that for each $\mathbf{AS}_i$, either $a, b \in \mathbf{A}_i$ and a does not attack b in $\mathbf{AS}_i$, or $(a, b) \notin (\mathbf{A}_i \times \mathbf{A}_i)$.

Note that there is only one WAS representing the merging of a profile of ASs. And this process can be done in linear time which is not the case of the merging process proposed in [10].

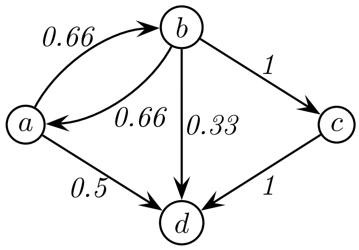
As for Dung's system, a WAS can be represented by a directed graph whose nodes are arguments and whose edges represent attacks and are weighted. As explained above, a missing edge has two possible interpretations: either all the agents who know both arguments agree on the non-attack, or none agent knows both arguments.

Finally, note that, in the above definition, the agents are treated equally. Other definitions could be given for coping for instance with a reliability level.

4.3 Examples

Let us consider the profile of ASs given in Example 1, merging this profile produces the following WAS:

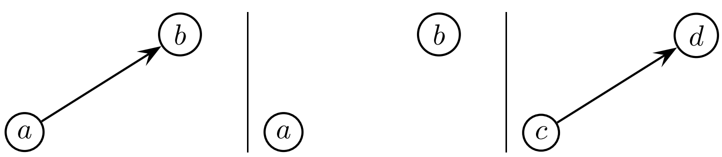
Ex 1 (cont'd)



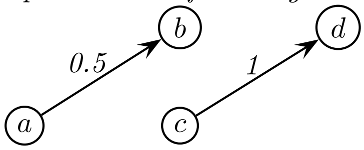
Using the notion of vs-defence, it follows that a does not vs-defend c. This conclusion is not really surprising because among the three agents who know a and b only two agree on the attack from a to b. Then, we obtain two maximal (for set-inclusion) maximal vs-admissible sets $\{a\}$ *and* $\{b\}$*.*

In the following, we give some small examples which illustrate some basic cases of merging.

Ex 3 *Consider 3 agents with their respective ASs:*



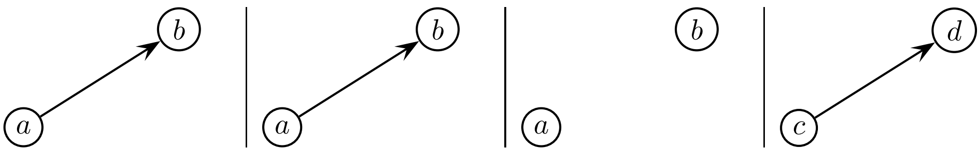
The merging process produces the following WAS:



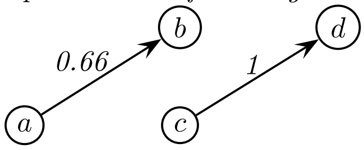
This result confirms two points:

- *only one agent knows c and d, so the interaction between c and d for the group is the interaction known by this agent;*
- *the two other agents do not agree concerning the interaction between a and b, so the resulting interaction is a weak attack.*

Ex 4 *Consider 4 agents with their respective ASs:*

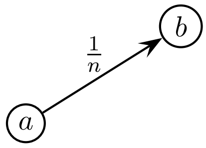


The merging process produces the following WAS:



This result illustrates the computation of the weights and the importance of the non-interaction that is taken into account across the number of the agents knowing the couple of arguments (a, b)*.*

Ex 2 (cont'd) *The merging process produces the following WAS:*



This result shows an important difference with regard to [10]. With the new merging process, we are able to take into account the difference between the number of agents "in favor of the presence of an interaction" and the number of agents "against this presence". It was not the case in [10]: each conflict about the interpretation of an interaction between two arguments was always represented by an ignorance relation between these arguments.

So [10], with the edit distance and the sum aggregation function, the merging process produces one argumentation system without attack from a to b.

4.4 Properties

Let $\mathcal{P} = \langle \mathbf{AS}_1, \ldots, \mathbf{AS}_n \rangle$ be a profile of n ASs with $\mathbf{AS}_i = \langle \mathbf{A}_i, \mathbf{R}_i \rangle$. Let $\mathbf{WAS} = \langle \mathbf{A}, \mathbf{R}, w \rangle$ denote the merging of $\mathcal{P}$ according to Definition 10.

It is easy to prove that Definition 10 refines Definition 6:

Prop 1 *Let* $\langle \mathbf{A}, \mathbf{R}', \mathbf{I} \rangle$ *be the PAS obtained by the* consensual fusion of $\mathcal{P}$ (Definition 6). We have:

- $(a\mathbf{R}'b)$ in **PAS** iff $(a\mathbf{R}b)$ and $w(a,b) = 1$ in **WAS**,
- $(a\mathbf{I}b)$ in **PAS** iff $(a\mathbf{R}b)$ and $w(a,b) < 1$ in **WAS**.

Proof: It follows directly from Definitions 6 and 10. □

As a direct consequence, in the case when there is no conflict between the agents, the WAS corresponding to the merging of the ASs is an AS. That is to say that all the weights are equal to 1.

Prop 2 *The merging of* $\mathcal{P}$ *according to Definition 10 results in an AS iff* $conf(\mathcal{P})$[4] *is empty. And, in that case, the resulting WAS is the union of the ASs, namely* $\langle \bigcup_{i=1}^{n} \mathbf{A}_i, \bigcup_{i=1}^{n} \mathbf{R}_i \rangle$.

Proof: An AS being a particular WAS with $w : \mathbf{R} \rightarrow \{1\}$, the property follows directly from Definition 10. □

A similar result has been given in [10] (see Proposition 35 of [10]) where the case when $conf(\mathcal{P})$ is empty corresponds to a *concordant* profile (see Definition 26 and Proposition 27 in [10]).

The following properties enable to compare the pieces of information contained in **WAS** and in the $\mathbf{AS}_i$. The first one concerns the non-attacked arguments:

Prop 3 *Let* $a \in \mathbf{A}$. *a is not attacked in* **WAS** *iff* $\forall i$ *such that* $a \in \mathbf{A}_i$, *a is not attacked in* $\mathbf{AS}_i$.

Proof: Following Definition 10, we have: there exists $b \in \mathbf{A}$ such that $b\mathbf{R}a$ iff there exists i such that $b\mathbf{R}_i a$. The proof follows by taking the contrapositive. □

The second property shows how WAS takes into account unanimity between the agents. It follows directly from Definition 10:

[4] See Definition 5.

Prop 4 *Let $a, b \in \mathbf{A}$.*

- *Let $(a, b) \in \mathbf{R}$: $w(a, b) = 1$ iff $\forall \mathbf{AS}_i$ such that $a, b \in \mathbf{A}_i$, $(a, b) \in \mathbf{R}_i$.*
- *$(a, b) \notin \mathbf{R}$ iff $\forall \mathbf{AS}_i$ such that $a, b \in \mathbf{A}_i$, $(a, b) \notin \mathbf{R}_i$.*

As a consequence of Property 4, situations when agents disagree can also be characterized:

Conseq 1 *Let $(a, b) \in \mathbf{R}$. $w(a, b) < 1$ iff there exists $\mathbf{AS}_i$ such that $(a, b) \in \mathbf{R}_i$ and there exists $\mathbf{AS}_j$ such that $(a, b) \notin \mathbf{R}_j$.*

Moreover, the pieces of information contained in **WAS** can be also compared with the result of the merging process as defined in [10].

Prop 5 *Let $\mathbf{PAS}_i$ be the consensual expansion of $\mathbf{AS}_i$ over $\mathcal{P}$. Let $a, b \in \mathbf{A}$.*

- *If $(a, b) \in \mathbf{R}$ and $w(a, b) = 1$, then a attacks b in each $\mathbf{PAS}_i$.*
- *If $(a, b) \notin \mathbf{R}$ then in each $\mathbf{PAS}_i$ there is no interaction (neither attack, nor ignorance) between a and b.*

Proof: It follows directly from Definitions 5 and 10. □

Prop 6 *Let $\langle \mathbf{AS}'_1, \ldots, \mathbf{AS}'_k \rangle$ be the resulting ASs given by the merging based on consensual expansion and the edit distance. Let $a, b \in \mathbf{A}$.*

- *If $(a, b) \in \mathbf{R}$ and $w(a, b) = 1$, then a attacks b in each $\mathbf{AS}'_i$ (the converse does not hold).*
- *If $(a, b) \notin \mathbf{R}$, then in each $\mathbf{AS}'_i$ there is no attack between a and b (the converse does not hold).*

Proof:

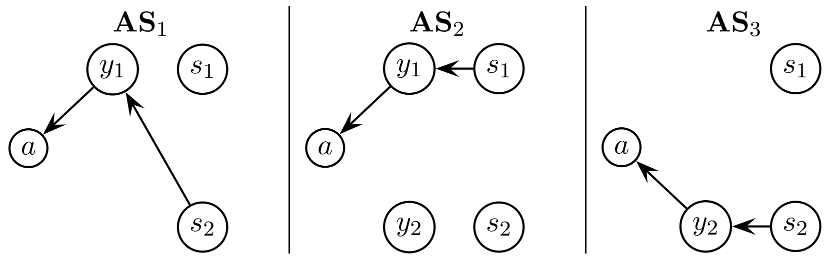
- Following Property 4, if $w(a, b) = 1$, then (a, b) belongs to the set of attacks that are not questioned by other agents. Then, using Proposition 36 given in [10], we can conclude that (a, b) belongs to the attack relation of each resulting $\mathbf{AS}'_i$.
- In a similar way, if $(a, b) \notin \mathbf{R}$, then (a, b) belongs to the set of non-attacks that are not questioned by other agents. Proposition 36 given in [10] enables to conclude that (a, b) belongs to the non-attack relation of each resulting $\mathbf{AS}'_i$.
- The converses do not hold (see Example 1 considering (a, b) for the first point and (b, d) for the second one).

□

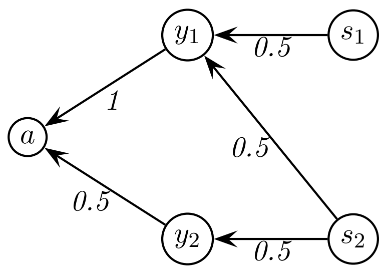
Property 4 focus on the information conveyed by the interactions between arguments and shows that the merging preserves the information on which the agents do not disagree. This result is not surprising since the characteristic feature of our approach is that the merging occurs quite early, at the interaction level.

However, we have no such result when we look at the output of the argumentation systems. As shown by the following example, it may happen that an argument a is accepted[5] by all the agents who know it but a is not accepted in the WAS.

Ex 5 *Consider three agents with their respective ASs:*



The set $\{a, s_1, s_2\}$ *is admissible in each* $\mathbf{AS}_i$*. The merging gives the following WAS:*



After the merging, no vs-defence exists for a *against* y_1*. So* a *cannot belong to a vs-admissible set. This result is not surprising: the attack from* y_1 *to* a *is known in* $\mathbf{AS}_1$ *and* $\mathbf{AS}_2$ *in which* s_1 *and* s_2 *are also known; but* $\mathbf{AS}_1$ *and* $\mathbf{AS}_2$ *do not agree on the attack from* s_1 *to* y_1 *and on the attack from* s_2 *to* y_1*. So the group cannot consider that* s_1 *or* s_2 *are good vs-defenders for* a *against* y_1*.*

Note that this behavior also appears with the merging process proposed in [10].

[5] For instance because it belongs to all preferred extensions.

5 Extension for merging bipolar argumentation system

The bipolar argumentation systems extend classical argumentation systems by considering the existence of another kind of interaction between arguments, the support relation.

5.1 Bipolar argumentation systems

As already said, due for instance to the presence of inconsistency in knowledge bases, arguments may be conflicting. These conflicts are captured by the attack relation in an argumentation system, and may be considered as negative interactions. Then, the concept of defence has been introduced in order to reinstate some of the attacked arguments, namely those whose attackers are in turn attacked. So, most logical theories of argumentation assume that if an argument a_3 defends an argument a_1 against an argument a_2, then a_3 is a kind of support for a_1. The fact that a_3 defends a_1 may be considered as a positive interaction. In the abstract unipolar argumentation system recalled in Section 2.1, only negative interaction is explicitly represented by the *attack* relation, and positive interaction is represented through the notion of defence. So, support and attack are *dependent* notions. It is a parsimonious strategy, but it is not a correct description of the process of argumentation in realistic examples.

So, in many papers (see [16, 25, 26, 2, 7, 8, 9, 22, 5, 21]), an abstract bipolar argumentation system has been proposed, in order to formalize situations where two *independent* kinds of interactions are available: a negative (which modelizes the conflicts) and a positive one (which is not a simple defence). This is an extension of the unipolar abstract argumentation system introduced by [11]:

Def 11 ([7, 8] Abstract bipolar argumentation system) *An* abstract bipolar argumentation system $\mathbf{BAS} = \langle \mathbf{A}, \mathbf{R}_{\mathtt{att}}, \mathbf{R}_{\mathtt{sup}} \rangle$ *consists of a set* $\mathbf{A}$ *of arguments, a binary relation* $\mathbf{R}_{\mathtt{att}}$ *on* $\mathbf{A}$ *called a* attack relation *and another binary relation* $\mathbf{R}_{\mathtt{sup}}$ *on* $\mathbf{A}$ *called a* support relation*.*

Consider a_i and $a_j \in \mathbf{A}$, $a_i \mathbf{R}_{\mathtt{att}} a_j$ (resp. $a_i \mathbf{R}_{\mathtt{sup}} a_j$) means that a_i attacks a_j (resp. a_i supports a_j).

A bipolar argumentation system can still be represented by a directed graph, with two kinds of edges, one for the attack relation and another one for the support relation.

Notations: Consider $a, b \in \mathbf{A}$, $a\mathbf{R}_{\mathtt{att}}b$ is represented by $a \rightarrow b$ and $a\mathbf{R}_{\mathtt{sup}}b$ is represented by $a \dashrightarrow b$. $\langle \mathbf{A}, \mathbf{R}_{\mathtt{att}}, \mathbf{R}_{\mathtt{sup}} \rangle$ defines a directed graph $\mathbf{G_b}$ called the *bipolar interaction graph*.

In order to define acceptability in abstract bipolar argumentation systems, Dung's methodology can be followed using the classical notion of defence which enables to capture reinstatement associated with a notion of conflict-freeness. This last notion generalizes the key concept of attack between two arguments, by combining a sequence of supports with a direct attack. For instance, in [7, 8], one can found the following new types of attacks:

Def 12 ([7, 8] Supported and secondary attack)
A supported attack *for an argument b is a sequence $a_1\mathbf{R}_1 \ldots \mathbf{R}_{n-1}a_n$, $n \geq 3$, with $a_n = b$, such that $\forall i = 1 \ldots n-2$, $\mathbf{R}_i = \mathbf{R}_{\mathtt{sup}}$ and $\mathbf{R}_{n-1} = \mathbf{R}_{\mathtt{att}}$.*

A secondary attack *for an argument b is a sequence $a_1\mathbf{R}_1 \ldots \mathbf{R}_{n-1}a_n$, $n \geq 3$, with $a_n = b$, such that $\forall i = 2 \ldots n-1$, $\mathbf{R}_i = \mathbf{R}_{\mathtt{sup}}$ and $\mathbf{R}_1 = \mathbf{R}_{\mathtt{att}}$.*

Many different approaches proposed distinct semantics using distinct notions of conflict-freeness and acceptability and sometimes it may be useful to attribute a semantics to the support relation (see [22, 5, 21]). For instance, [5] proposes to introduce a special kind of attack reflecting a particular meaning of the support:

Def 13 ([5] Mediated attack) *A* mediated attack *from a to b appears iff there is a sequence $a_1\mathbf{R}^s \ldots \mathbf{R}^s a_{n-1}$, and $a_n\mathbf{R}^a a_{n-1}$, $n \geq 3$, with $a_1 = b$, $a_n = a$.*

5.2 Merging of bipolar argumentation systems

In this section, for merging bipolar argumentation systems, we propose to use a variant of weighted argumentation systems in which the binary relation $\mathbf{R}$ between arguments represents an "interaction" between arguments which is not reduced to an attack.

5.2.1 Definitions

The strategy used for computing the weighted argumentation system issued from the group of agents is exactly the same as the one used for merging ASs: each agent is a voting person and we add the votes for and against each interaction. However, two types of interaction must be considered. Our approach allows for a compensation between support and attack (this idea was already used for valuating BASs – see [7]).

Def 14 (Merging of n BASs) *Let $\mathcal{P} = \langle \mathbf{BAS}_1, \ldots, \mathbf{BAS}_n \rangle$ be a profile of n bipolar argumentation systems with $\mathbf{BAS}_i = \langle \mathbf{A}_i, \mathbf{R}_{\mathtt{att}\,i}, \mathbf{R}_{\mathtt{sup}\,i} \rangle$. The merging of $\mathcal{P}$ is the WAS defined by $\mathbf{WAS} = \langle \mathbf{A}, \mathbf{R}, w \rangle$ with:*

- $\mathbf{A} = \bigcup_{i=1}^{n} \mathbf{A}_i$

- $\mathbf{R} = (\cup_{i=1}^n \mathbf{R}_{\mathtt{att}i}) \bigcup (\cup_{i=1}^n \mathbf{R}_{\mathtt{sup}_i})$
- $w : \mathbf{R} \to [-1, +1]$ *defined by* $\forall (a, b) \in \mathbf{R}$, $w(a, b) = \frac{\Sigma_{i=1}^n w_i(a,b)}{n(a,b)}$ *with*
    - $w_i(a, b) = -1$ *if* $(a, b) \in \mathbf{R}_{\mathtt{att}i}$,
    - $w_i(a, b) = 1$ *if* $(a, b) \in \mathbf{R}_{\mathtt{sup}_i}$,
    - $w_i(a, b) = 0$ *if* $(a, b) \notin (\mathbf{R}_{\mathtt{att}i} \cup \mathbf{R}_{\mathtt{sup}_i})$
    - *and* $n(a, b)$ *is the number of* $\mathbf{BAS}_i$ *such that* $a \in \mathbf{A}_i$ *and* $b \in \mathbf{A}_i$.

As a direct consequence, we have:

Observation: Let $a, b \in \mathbf{A}$.

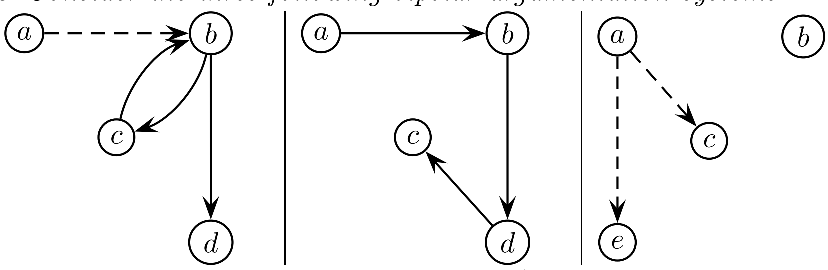
- If $(a, b) \in \mathbf{R}$, there exists at least one $\mathbf{BAS}_i$ such that $(a, b) \in (\mathbf{R}_{\mathtt{att}i} \cup \mathbf{R}_{\mathtt{sup}_i})$, so $n(a, b) \neq 0$.
- $(a, b) \notin \mathbf{R}$ means that for each $\mathbf{BAS}_i$, either $a, b \in \mathbf{A}_i$ and a and b are not in interaction in $\mathbf{BAS}_i$ (*i.e.* $(a, b) \notin (\mathbf{R}_{\mathtt{att}i} \cup \mathbf{R}_{\mathtt{sup}_i})$), or $(a, b) \notin (\mathbf{A}_i \times \mathbf{A}_i)$.

As in the case of ASs, there is always one WAS representing the merging of n BASs. And this WAS can be computed in linear time as for merging of n ASs.

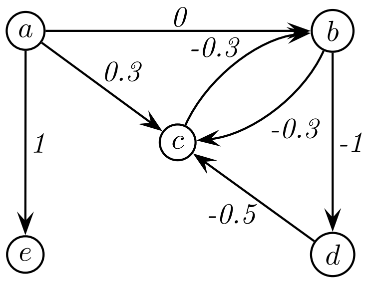
Note again that, in the above definition, the agents are treated equally. Other definitions could be given for coping for instance with a reliability level (for instance, if one considers that the reliability of the support relation is not exactly the same as the attack relation).

As in Section 4.2, the WAS obtained by Definition 14 can be represented by a directed graph whose nodes are arguments and edges are binary interactions between arguments.

Ex 6 *Consider the three following bipolar argumentation systems:*



Following Definition 14, the following WAS corresponds to the merging of the previous BASs:



Some important remarks must be made:

- First, as in the case of merging ASs, a missing edge has two possible interpretations: either all the agents who know both arguments agree on the non-interaction (non-attack and non-support), or none agent knows both arguments.

- Secondly, an important point concerns the interpretation of the edges; each edge represents an *interaction* between arguments and it is natural to associate the meaning of this interaction with its weight, considering that a positive weight corresponds to a support and a negative one to an attack.

- Moreover, by the mechanism of compensation between supports and attacks, it may be the case that a weight equals to 0 which is impossible when merging ASs (see for instance the interaction between a and b in Example 6). This kind of edge reflects a "complete" ignorance of the group (it is impossible to decide which is the interaction between a and b since there are as many agents who know the attack from a to b that agents who know the support from a to b).

- The last point concerns the use of these weights for comparing interactions: let $(a, b) \in \mathbf{R}$, we consider that stronger is $|w(a, b)|$[6], stronger is the interaction from a to b. This is in agreement with the use of WAS for merging several BASs.

Notions at work in ASs, BASs, and WAS can be restated in this variant of WAS (even if some adjustments must be made in order to take into account the negative weights vs the positive ones). In this framework, a comparison of the weights could be introduced in the definitions of supported, secondary or mediated attacks as in the definition of the vs-defence.

Another possibility would be to give a threshold that could be used as in [12] for computing an acceptable acceptability level. This threshold should be also used for removing non-significant interactions. The possibilities are very numerous and it is not the aim of this paper to give them exhaustively.

[6]The notation $|x|$ for $x \in \mathbb{R}$ gives the absolute value of x.

5.2.2 Properties

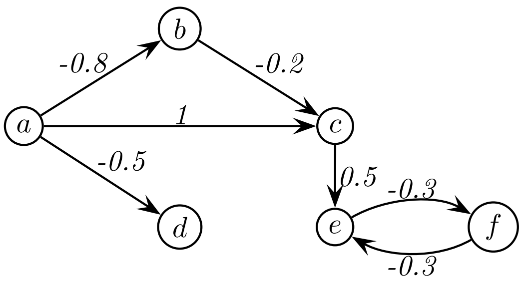
Let **WAS** be the system resulting of the merging of several BASs, the main trivial but interesting property is the correspondence with the original frameworks:

- if all weights are taken in the singleton $\{x\}$, with $x < 0$, **WAS** is a Dung's AS;
- if all weights are strictly negative, **WAS** can be easily translated into a WAS as in [12] replacing each negative weight by its corresponding positive value;
- if all weights are taken in the set $\{x_1, x_2\}$, with $x_1 < 0 < x_2$ and $|x_1| = x_2$, **WAS** is a BAS;
- if all weights are taken in the set $\{x_1, 0\}$, with $x_1 < 0$, **WAS** is a PAS (each edge with a negative weight is an attack and each edge with a weight equals to zero is an ignorance).

5.2.3 Examples

The following examples illustrate different WASs giving a possible reading of each case.

Ex 7



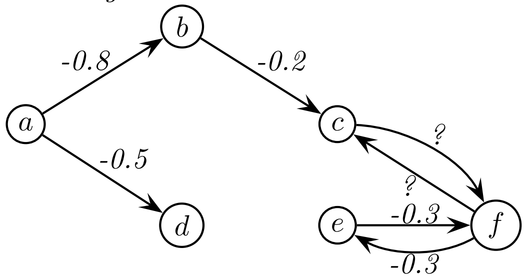
In this example a vs-defends and directly supports c.

If we do not take into account the weights in the definition of supported and secondary attacks, there is a supported attack from c to f and a secondary attack from b to e. Otherwise, it might be the case that this supported attack remains whereas the secondary one disappears (the attack from b to c could be considered too weak to have an influence on the argument e supported by c).

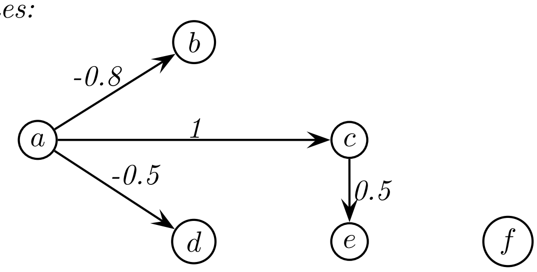
In the context proposed by [5], there could be also a mediated attack from f to c.

Note that, if we choose to introduce all these new attacks, an important problem appears: how to compute the weights of these added attacks (this will be the subject of a future work).

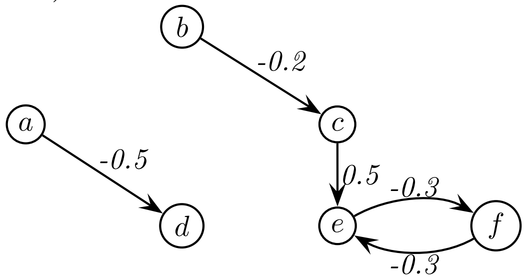
After the addition of supported and mediated attacks and the removal of the supports, the resulting WAS should be:



In another context using a threshold, one can propose a simplification of the WAS considering that all interactions which are lower than the threshold are non significant. For instance, in this case, with a threshold equals to 0.5*, the WAS becomes:*



Always with a threshold equals to 0.5*, but considering than the weights represent costs, the simplification gives the following WAS (we keep the cheapest interactions):*



6 Conclusion

In this paper, we have proposed a new method for merging different argumentation systems in a multi-agent setting. This approach is particularly interesting from a computational point of view because it is a linear process. It is easy to apply and exactly reflects the impact of each agent on the group. It also corrects some disadvantages of the method proposed in [10] (this is, to our knowledge, the only existing method for merging argumentation systems) since it provides only a single output weighted argumentation system.

There are several directions for future works. The first one will be to consider a reliability level for each agent and to take it into account for computing the weights of the interactions for the group. And a second one is to pursue the study of the WAS (in particular when we want to express some bipolar interactions).

References

[1] L. Amgoud and C. Cayrol. Inferring from inconsistency in preference-based argumentation frameworks. *Journal of Automated Reasoning*, 29:125–169, 2002.

[2] L. Amgoud, C. Cayrol, and MC. Lagasquie-Schiex. On the bipolarity in argumentation frameworks. In *Proc. of the 10^{th} NMR-UF workshop*, pages 1–9, 2004.

[3] L. Amgoud, Y. Dimopoulos, and P. Moraïtis. A unified and general framework dor argumentation-based negotiation. In *Proc. of AAMAS*, pages 963–970, 2007.

[4] L. Amgoud, N. Maudet, and S. Parsons. Arguments, Dialogue and Negotiation. In *Proc of* 14^{th} *ECAI*, pages 338–342, 2000.

[5] G. Boella, D. M. Gabbay, L. van der Torre, and S. Villata. Support in abstract argumentation. In *Proc. of COMMA*, pages 111–122. IOS Press, 2010.

[6] C. Cayrol, C. Devred, and MC. Lagasquie-Schiex. Acceptability semantics accounting for strength of attacks in argumentation. In *Proc. of ECAI*, pages 995–996, 2010.

[7] C. Cayrol and M.-C. Lagasquie-Schiex. Gradual valuation for bipolar argumentation frameworks. In *Proc. of ECSQARU*, pages 366–377. Springer-Verlag, 2005.

[8] C. Cayrol and M.-C. Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In *Proc. of ECSQARU*, pages 378–389. Springer-Verlag, 2005.

[9] C. Cayrol and M.-C. Lagasquie-Schiex. Coalitions of arguments: a tool for handling bipolar argumentation frameworks. *International Journal of Intelligent Systems*, 25:83–109, 2010.

[10] S. Coste-Marquis, C. Devred, S. Konieczny, MC. Lagasquie-Schiex, and P. Marquis. On the merging of Dung's argumentation systems. *Artificial Intelligence, Argumentation in Artificial Intelligence*, 171(10-15):730–753, 2007.

[11] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[12] P. E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge. Inconsistency tolerance in weighted argument systems. In *Proc of AAMAS*, 2009.

[13] J. Fox and S. Parsons. On using arguments for reasoning about values and actions. In *Proc. of AAAI-Symposium on qualitative preferences in deliberation and practical reasoning*, pages 55–63, 1997.

[14] T. Gordon and N. Karacapilidis. The zeno argumentation framework. In *Proc. of ICAIL*, pages 10–18. ACM Press, 1997.

[15] A. C. Kakas and P. Moraïtis. Argumentation based decision making for autonomous agents. In *Proc. of AAMAS*, pages 883–890, 2003.

[16] N. Karacapilidis and D. Papadias. Computer supported argumentation and collaborative decision making: the HERMES system. *Information systems*, 26(4):259–277, 2001.

[17] P. Krause, S. Ambler, M. Elvang, and J. Fox. A logic of argumentation for reasoning under uncertainty. *Computational Intelligence*, 11 (1):113–131, 1995.

[18] D. C. Martinez, A. J. Garcia, and G. R. Simari. On defense strength of blocking defeaters in admissible sets. In *Proc. of KSEM (LNAI 4798)*, pages 140–152, 2007.

[19] D. C. Martinez, A. J. Garcia, and G. R. Simari. An abstract argumentation framework with varied-strength attacks. In *Proc of KR*, pages 135–143, 2008.

[20] D. C. Martinez, A. J. Garcia, and G. R. Simari. Strong and weak forms of abstract argument defense. In *Proc of COMMA*, pages 216–227, 2008.

[21] F. Nouioua and V. Risch. Bipolar argumentation frameworks with specialized supports. In *Proc. of ICTAI*, pages 215–218. IEEE Computer Society, 2010.

[22] N. Oren, C. Reed, and M. Luck. Moving between argumentation frameworks. In *Proc. of COMMA*, pages 379–390. IOS Press, 2010.

[23] S. Parsons, C. Sierra, and N. R. Jennings. Agents that reason and negociate by arguing. *Journal of Logic and Computation*, 8(3):261–292, 1998.

[24] H. Prakken and G. Vreeswijk. Logics for defeasible argumentation. In *Handbook of Philosophical Logic*, volume 4, pages 218–319. Kluwer Academic, 2002.

[25] B. Verheij. Automated argument assistance for lawyers. In *Proc. of ICAIL*, pages 43–52. ACM Press, 1999.

[26] B. Verheij. Deflog: on the logical interpretation of prima facie justified assumptions. *Journal of Logic in Computation*, 13:319–346, 2003.