Merging Argumentation Systems

Sylvie Coste-Marquis[1] **Caroline Devred**[1] **Sébastien Konieczny**[1]
Marie-Christine Lagasquie-Schiex[2] **Pierre Marquis**[1]
[1] CRIL-CNRS, Université d'Artois, 62307 Lens, France
{coste,devred,konieczny,marquis}@cril.univ-artois.fr
[2] IRIT-CNRS, UPS, 31062 Toulouse, France
lagasquie@irit.fr

Abstract

In this paper, we address the problem of deriving sensible information from a collection of argumentation systems coming from different agents. A general framework for merging argumentation systems from Dung's theory of argumentation is presented. Each argumentation system gives both a set of arguments and the way they interact (i.e. attack or non-attack) according to the corresponding agent. The aim is to define the argument system (or the set of argument systems) that best represents the group. Our framework is general enough to handle the case when agents do not share the same set of arguments. Merging argumentation systems is shown as a valuable approach for defining (sets of) arguments acceptable by the group.

Introduction

Argumentation is based on the exchange and valuation of interacting arguments which may represent information of various kinds, especially beliefs or goals. Argumentation has been applied, among others, in the legal domain, for collective decision support systems or for negotiation support.

Several theories of argumentation exist; each of them makes explicit the nature of arguments, the way arguments are generated, how they interact and how to evaluate them, and finally what are the most acceptable arguments. A key issue is the interaction between arguments which is typically based on a notion of attack; for example, when an argument takes the form of a logical proof, arguments for a statement and arguments against it can be put forward. In that case, the attack relation relies on logical inconsistency.

The theory of argumentation frameworks as introduced by (Dung 1995) is abstract enough to manage without any assumptions on the nature of arguments or the attack relation. As such, it includes several formal systems developed so far for commonsense reasoning or logic programming (Dung 1995). For instance, when an agent has conflicting pieces of belief (viewed as arguments), a (nontrivial) set of plausible consequences can be derived from the most acceptable arguments for the agent (additional information like a plausibility ordering are often taken into account in the evaluation phase). Much work has been devoted to this issue (see for example (Dung 1995; Krause *et al.* 1995; Prakken & Sartor 1997; Pollock 2001; Amgoud & Cayrol 2002; Prakken & Vreeswijk 2002)).

In a multi-agent setting, argumentation can also be used to represent (part of) some information exchange processes, like negotiation, or persuasion (see for example (MacKenzie 1979; Walton & Krabbe 1995; Gordon 1995; Parsons & Jennings 1996; Amgoud, Maudet, & Parsons 2000; Amgoud & Parsons 2002; Amgoud & Prade 2004)). For instance, a negotiation process between two agents about whether some piece of belief must be considered as true given some evidence can be modelled as a two-player game where each move consists in reporting an argument which attacks arguments given by the opponent.

In this paper, we also consider argumentation in a multi-agent setting, but from a very different perspective. The purpose is to define argumentation systems 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 the agents give their arguments and attack relation, and agree to find a result for the whole group. Compared with negotiation protocols, where agents only exchange *parts* of their arguments/attack relation, the result obtained by merging argumentation systems can be viewed as an "ideal" result one can achieve if the interaction between agents is not polluted by some fixed protocol (the distinction is similar to the one between games under complete information vs games under incomplete information). Merging argumentation systems can also be used to define a baseline for comparing different negotiation protocols: the closer the result of the negotiation protocol to the result of the merging process, the best.

Our framework is general enough to handle the case when agents do not share the same set of arguments and disagree on the attack relation: each agent may have her own view on what an attack is and as a consequence, an agent may believe that one argument attacks another argument, while another agent may believe that this is not the case. Merging argumentation systems can be used to define (sets of) arguments acceptable by the group. By means of example, we show that it leads to results which are much more expected than those furnished by a direct voting on the (sets of) arguments acceptable by each agent. Finally, we also briefly explain how the framework can be further refined to take into account attacks strengths for each agent and each pair of arguments, and even computational resources at the group level.

Formal preliminaries

We focus on Dung's theory of argumentation (Dung 1995).

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

$\langle A, R\rangle$ clearly defines a directed graph $\mathcal{G}$ called the *attack graph*.

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 A. Collective acceptability is based on two key notions: lack of conflict and collective defence.

Definition 2 *Let $\langle A, R\rangle$ be an argumentation system.*

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

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

(Dung 1995) defines several semantics for collective acceptability based on those two notions; among them the *admissible semantics* and the *preferred semantics*.

Definition 3 *Let $\langle A, R\rangle$ be an argumentation system.*

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

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

Definition 4 *An* acceptability relation*, denoted by Acc_{AF}, for a given argumentation system $\mathsf{AF} = \langle A, R\rangle$, is a total function from 2^A to $\{true, false\}$ which associates each subset E of A with true if E is an acceptable set for AF and with false otherwise.*

Usually, an acceptability relation is based on a specific semantics (plus possibly a selection principle). For instance, a set of arguments can be considered acceptable iff it coincides with one of the extensions (for the chosen semantics). Alternatively, a set of arguments can be considered acceptable iff it is included in one (credulous selection) or all (skeptical selection) the extensions. Whatever the way it is defined, an acceptability relation can be viewed as a choice function among the elements of 2^A.

Motivation

When merging a profile of argumentation systems (i.e., a vector $\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n\rangle$ of such systems $\mathsf{AF}_i = \langle A_i, R_i\rangle$ where each index i corresponds to a specific agent), one basically wants to define an argumentation system (or a set of such systems) which reflects at best how arguments interact for the whole group of agents.Among other things, this can be used to define the sets of arguments considered acceptable for the group.

Definition 5 *A* joint acceptability relation *for a multiset $\{\mathsf{AF}_1, \ldots, \mathsf{AF}_n\}$ of AFs, denoted by $Acc_{\{\mathsf{AF}_1,\ldots,\mathsf{AF}_n\}}$, is a total function from $2^{\bigcup_i A_i}$ to $\{true, false\}$ which associates each subset E of $\bigcup_i A_i$ with true if E is a jointly acceptable set for $\{\mathsf{AF}_1, \ldots, \mathsf{AF}_n\}$ and with false otherwise.*

For instance, a joint acceptability relation for a multiset $\{\mathsf{AF}_1, \ldots, \mathsf{AF}_n\}$ can be defined by the acceptability relations Acc_{AF_i} (based themselves on some semantics and some selection principles), which can coincide for every AF_i (but this is not mandatory) and a voting method $V : \{true, false\}^n \mapsto \{true, false\}$:

$$Acc_{\{\mathsf{AF}_1,\ldots,\mathsf{AF}_n\}}(E) = V(Acc_{\mathsf{AF}_1}(E), \ldots, Acc_{\mathsf{AF}_n}(E)).$$

So, a set of arguments can be considered acceptable for the group iff it is acceptable for "sufficiently many" agents from the group. The voting method under consideration makes precise what "sufficiently many" means: for instance, one agent, every agent, k agents (where k is fixed a priori), $k\%$ of the total number of agents (including the (weak) majority rule when $k = 50$ and the strict majority rule when $k = 50 + \epsilon$). However, this approach may easily lead to counterintuitive results:

Example 1 *Let $\mathsf{AF}_1 = \langle \{a, b, e, f\}, \{(a, b), (b, a), (e, f)\}\rangle$, $\mathsf{AF}_2 = \langle \{b, c, d, e, f\}, \{(b, c), (c, d), (f, e)\}\rangle$ and $\mathsf{AF}_3 = \langle \{e, f\}, \{(e, f)\}\rangle$ be three argumentation systems. Whatever the semantics (among Dung's ones), c belongs to no extension of AF_2, hence it cannot be elected as a member of an acceptable set whatever the voting method (under the reasonable assumption that it is a choice function based on extensions). However since c (resp. a) is not among the arguments reported by the first agent (resp. the second and the third ones), it can be sensible to assume that both agents (having no other information than the ones given by the other agents) agree on the fact that a attacks b and b attacks c. Indeed, this assumption does not contradict what the three agents report as to the way arguments interact. Under this assumption, $\{c\}$ should be considered as credulously acceptable by the group.*

Such a naive voting approach suffers from two major drawbacks:

Problem 1 Voting makes sense only if all agents consider the same set of arguments A at start (otherwise, the set 2^A of alternatives is not common to all agents). However, it can be the case that the sets of arguments reported by the agents differ one another.

Problem 2 Voting relies only on the selected extensions: the attack relations (from which extensions are characterized) are not taken into consideration any more once extensions have been computed. This leads to let aside much significant information which could be exploited to define the sets of acceptable arguments at the group level.

Partial argumentation systems

In order to define the merging of argumentation systems in a satisfying manner, one must find a way to solve Problem 1 above. Taking the union of the argumentation systems AF_1, ..., AF_n, i.e. considering the system $\mathsf{AF} = \bigcup_{i=1}^{n} \langle A_i, R_i\rangle$ defined by $\mathsf{AF} = \langle \bigcup_{i=1}^{n} A_i, \bigcup_{i=1}^{n} R_i\rangle$ would be over simplistic in many cases. Indeed, it would lead to assimilate pieces of information of very different nature: ignorance about attack and absence of attack.

Example 1 (continued) *Agent 1 does not report c at start, hence one may assume that she ignores whether some interactions between c and the arguments she reports exist or not. For instance, she ignores whether a attacks c or not. Now, what's about the interaction between f and e from the point of view of agent 1? Since 1 reports both e and f but no attack (f, e), the conclusion is that 1 believes that f does not attack e. Accordingly, there is no conflict between the beliefs of agents 1 and 2 concerning interactions with c while there is a conflict between their beliefs as to the way e and f interact.*

Handling the two kinds of information within a uniform setting calls for an extension of the notion of argumentation systems, that we call *partial argumentation systems*.

Definition 6 *A* partial argumentation system *over A is a quadruple $\mathsf{PAF} = \langle A, R, I, N\rangle$ where A is a set of arguments, R, I, N are binary relations on A. R is the attack relation, I is called the ignorance relation and is such that $R \cap I = \varnothing$. $N = (A \times A) \setminus (R \cup I)$ is called the non-attack relation.*
N is deduced from A, R and I, so a partial argumentation system can be fully specified by $\langle A, R, I\rangle$. We will use both notations in the following.

Each AF is a particular PAF for which the set I is empty. In an AF, the N relation also exists even if it is not given explicitly ($N = A \times A \setminus R$, $I = \emptyset$).

Each PAF over A can be viewed as a compact representation of a set of AFs over A, called its *completions*:

Definition 7 *Let $\mathsf{PAF} = \langle A, R, I\rangle$. Let $\mathsf{AF} = \langle A, S\rangle$. AF is a* completion *of PAF iff $R \subseteq S \subseteq R \cup I$.*
The set of the completions of PAF is denoted $\mathcal{C}(\mathsf{PAF})$.

Now, Problem 1 can be addressed by first associating each argumentation system AF_i with a corresponding PAF_i so that all PAF_i are about the same set of arguments $\bigcup_{i=1}^{n} A_i$. We define the notion of *expansion* of an AF:

Definition 8 *Let $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n\rangle$ be a profile of n AFs with each $\mathsf{AF}_i = \langle A_i, R_i, N_i\rangle$. An* expansion *of an $\mathsf{AF} = \langle A, R\rangle$ given $\mathcal{P}$ is any PAF $\mathtt{exp}(\mathsf{AF}, \mathcal{P})$ defined by $\langle A \cup \bigcup_i A_i, R', I', N'\rangle$ s.t. $R \subseteq R'$ and $(A \times A) \setminus R \subseteq N'$.*
By convention, if $\mathcal{P} = \varnothing$ then $\mathtt{exp}(\mathsf{AF}, \mathcal{P}) = \mathsf{AF}$. $\mathtt{exp}$ is referred to as an expansion function.

In order to be general enough, this definition does not impose many constraints on the resulting PAF: what is important is to preserve the attack and non-attack relations from the initial AF while extending its set of arguments. Many policies can be used to give rise to expansions of different kinds, reflecting the various attitudes of agents in light of "new" arguments; for instance, if a is any argument considered by agent i at start and a "new" argument b has to be incorporated, agent i can (among other things):

- always reject b (e.g. adding (b, b) to its relation R'_i),
- always accept b (adding (a, b), (b, a) and (b, b) to its non-attack relation N'_i),
- just express its ignorance about b (adding (a, b), (b, a) and (b, b) to its ignorance relation I'_i).

Each agent may also compute the exact interaction between a and b when the attack relation is not primitive but defined from more basic notions (as in the approach by Elvang-Gøransson et al., see e.g. (Elvang-Gøransson, Fox, & Krause 1993a; 1993b; Elvang-Gøransson & Hunter 1995)). Note that if she has limited computational resources, agent i can compute exact interactions as far as she can, then express ignorance for the remaining ones.

In the following, we specifically focus on *consensual expansions*. Intuitively, the consensual expansion of an argumentation system $\mathsf{AF} = \langle A, R\rangle$ 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 A) into the attack (resp. the non-attack relation) *provided that all other agents of the profile that know the two arguments agree on the attack* [1] (resp. the non-attack); otherwise, it is added to the ignorance relation:

Proposition 1 *Let $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n\rangle$ be a profile of n AFs with each $\mathsf{AF}_i = \langle A_i, R_i\rangle$. Let $\mathsf{AF} = \langle A, R\rangle$, and let N be the corresponding non-attack relation. Let $conf(\mathcal{P}) = (\bigcup_i R_i) \cap (\bigcup_i N_i)$ be the set of interactions for which a conflict exists within the profile. The* consensual expansion *of AF over $\mathcal{P}$ noted $\mathtt{exp}_C = \langle A', R', I', N'\rangle$ with:*

- $A' = A \cup \bigcup_i A_i$,
- $R' = R \cup ((\bigcup_i R_i \setminus conf(\mathcal{P})) \setminus N)$,
- $I' = conf(\mathcal{P}) \setminus (R \cup N)$,
- $N' = (A' \times A') \setminus (R' \cup I')$.

is an expansion of AF over $\mathcal{P}$ in the sense of Definition 8.

Merging operators

In order to deal with Problem 2, we suggest to merge interactions instead of sets of acceptable arguments. The goal is to characterize argumentation systems which are as close as possible to the given profile of argumentation systems, taken as a whole. A way to achieve it consists in defining a notion of "distance" between an AF and a profile of AFs, or more generally between a PAF and a profile of PAFs. This calls for a notion of pseudo-distance between two PAFs, and a way to combine such pseudo-distances:

Definition 9 *A* pseudo-distance *d between PAFs over A is a mapping which associates a real number to each pair of PAFs over A and satisfies the properties of symmetry ($d(x, y) = d(y, x)$) and minimality ($d(x, y) = 0$ iff $x = y$). d is a distance if it satisfies also the triangular inequality ($d(x, y) \leq d(x, y) + d(y, z)$).*

[1] i.e. if $a, b \in A_i$, then $(a, b) \in R_i$

Definition 10 *An* aggregation function *is a mapping $\otimes$ from $(\mathbb{R}+)^n$ to $(\mathbb{R}+)$ (strictly speaking, it is a family of relations, one for each n), that satisfies non-decreasingness (if $x_i \geq x'_i$, then $\otimes(x_1, \ldots, x_i, \ldots, x_n) \geq \otimes(x_1, \ldots, x'_i, \ldots, x_n)$), minimality ($\otimes(x_1, \ldots, x_n) = 0$ if $x_i = 0, \forall i$), and identity ($\otimes(x) = x$).*

The merging of a profile of AFs is defined as a set of AFs:

Definition 11 *Let $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ be a profile of n AFs. Let d be any pseudo-distance between PAFs, let $\otimes$ be an aggregation function, and let $\mathtt{exp}_1, \ldots, \mathtt{exp}_n$ be n expansion functions. The* merging *of $\mathcal{P}$ is the set*

$$\Delta_d^{\otimes}(\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle, \langle \mathtt{exp}_1, \ldots, \mathtt{exp}_n \rangle) =$$

$$\{\mathsf{AF} \textit{ over } \bigcup_i A_i \mid \mathsf{AF} \textit{ minimizes } \otimes_{i=1}^n d(\mathsf{AF}, \mathtt{exp}_i(\mathsf{AF}_i, \mathcal{P}))$$

Thus, merging a profile of AFs $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ is a two-step process:

expansion : An expansion of each AF_i over $\mathcal{P}$ is first computed. Note that nothing prevents from considering expansion functions that are specific to each agent. What is important is that $\mathtt{exp}_i(\mathsf{AF}_i, \mathcal{P})$ is a PAF over $A = \bigcup_i A_i$.

fusion : The AFs over A that are selected as result of the merging are the ones that best represent $\mathcal{P}$ (i.e. that are the "closest" to $\mathcal{P}$).

In the following, we will focus on the case when every agent uses the consensual expansion. So to avoid too heavy notations, we remove $\langle \mathtt{exp}_1, \ldots, \mathtt{exp}_n \rangle$ from the list of parameters of merging operators.

Note that it is possible to refine Definition 11 so as to include integrity constraints into the picture. This can be useful if one has some (unquestionable) knowledge about the expected result (some attacks between arguments which have to hold for the group). It is then enough to look only to the AFs that satisfy the constraints, similarly to what is done in propositional belief base merging (see e.g. (Konieczny & Pino Pérez 2002)). In contrast to the belief base merging scenario, one can also impose constraints on the *structure* of the candidate AFs. In particular, considering only acyclic AFs can prove valuable since (1) such AFs are well-founded, (which means that only one extension has to be considered whatever the underlying semantics – among Dung's ones), and (2) this extension (referred to as the grounded one, see (Dung 1995)) can be computed in time polynomial in the size of the AF (while computing a single extension is intractable for the other semantics in the general case – under the standard assumptions of complexity theory – see (Dunne & Bench-Capon 2002)).

Now, many pseudo-distances between PAFs and many aggregation functions can be used, giving rise to many merging operators. Usual aggregation functions include the sum, the max and the leximax, but nothing prevents from taking advantage of nonsymmetric functions (this is particularly useful to model situations when some agents are more important than others). Some aggregation functions (like sum) enable to take into account the number of agents believing that an argument attacks or not another argument:

Example 1 (continued) *Two agents over three agree on the fact that e attacks f and f does not attack e. It may prove sensible that the group agrees with the majority. Taking simply the union of the three AFs as their merging would lead to consider as well that f attacks e, while more than half of the agents believes that this is not the case.*

The aggregation function is very important for tuning the operator behaviour with the expected one. For example, if one wants to solve conflicts with use of majority, sum is a possible choice. If one wants to be more consensual, by trying to define a result close to every agent, the leximax function can be more interesting. The distinction between majority and arbitration operators as considered in propositional belief base merging (Konieczny & Pino Pérez 2002) also applies here.

In the following, we focus on the *edition distance* between PAFs:

Definition 12 *Let $\mathsf{PAF}_1 = \langle A, R_1, I_1, N_1 \rangle$ and $\mathsf{PAF}_2 = \langle A, R_2, I_2, N_2 \rangle$ be two PAFs over A.*

- *Let a, b be two arguments $\in A$. The* edition distance between PAF_1 and PAF_2 over a, b *is the relation $de_{a,b}$ such that:*
  - $de_{a,b}(\mathsf{PAF1}, \mathsf{PAF2}) = 0$ *iff* $(a, b) \in R_1 \cap R_2$ *or* $I_1 \cap I_2$ *or* $N_1 \cap N_2$,
  - $de_{a,b}(\mathsf{PAF1}, \mathsf{PAF2}) = 1$ *iff* $(a, b) \in R_1 \cap N_2$ *or* $N_1 \cap R_2$,
  - $de_{a,b}(\mathsf{PAF1}, \mathsf{PAF2}) = 0.5$ *otherwise.*
- *The* edition distance between PAF_1 and PAF_2 *is given by*

$$de(\mathsf{PAF1}, \mathsf{PAF2}) = \Sigma_{(a,b) \in A \times A} de_{a,b}(\mathsf{PAF1}, \mathsf{PAF2}).$$

Proposition 2 *The edition distance de between PAFs is a distance.*

Example 2 *Let $\mathsf{AF}_1 = \langle \{a, b\}, \{(a, b), (b, a)\} \rangle$, $\mathsf{AF}_2 = \langle \{b, c, d\}, \{(b, c), (c, d)\} \rangle$, $\mathsf{AF}_3 = \langle \{a, b, d\}, \{(a, b), (a, d)\} \rangle$ and $\mathsf{AF}_4 = \langle \{a, b, d\}, \{(b, d), (b, a)\} \rangle$ be four argumentation systems.*

First, the consensual expansions of the four AFs are:

- $\mathsf{PAF}_1 = \langle \{a, b, c, d\}, \{(a, b), (b, a), (b, c), (c, d)\}, \{(a, d), (b, d)\} \rangle$,
- $\mathsf{PAF}_2 = \langle \{a, b, c, d\}, \{(b, c), (c, d)\}, \{(a, b), (b, a), (a, d)\} \rangle$,
- $\mathsf{PAF}_3 = \langle \{a, b, c, d\}, \{(a, b), (a, d), (b, c), (c, d)\}, \{\} \rangle$,
- $\mathsf{PAF}_4 = \langle \{a, b, c, d\}, \{(b, d), (b, a), (b, c), (c, d)\}, \{\} \rangle$

We obtain $\Delta_{de}^{\Sigma}(\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_4 \rangle)$ as the set containing the two AFs: $\langle \{a, b, c, d\}, \{(a, b), (b, a), (b, c), (c, d)\} \rangle$ and $\langle \{a, b, c, d\}, \{(a, b), (b, a), (b, c), (a, d), (c, d)\} \rangle$.

Some properties

Let us now present some properties of consensual expansions and merging operators based on the edition distance, showing them as interesting choices. We first need the notion of conflict-free part of a profile of PAFs:

Definition 13 *Let $\mathcal{P} = \langle \mathsf{PAF}_1, \ldots, \mathsf{PAF}_n \rangle$ be a profile of PAFs. The* conflict-free part *of $\mathcal{P}$ is denoted by $CFP(\mathcal{P})$ and is defined by:*
$CFP(\mathcal{P}) = \langle \bigcup_i A_i, \bigcup_i R_i \setminus \bigcup_i N_i, I_{CFP}, \bigcup_i N_i \setminus \bigcup_i R_i \rangle$, where $I_{CFP} = (\bigcup_i A_i \times \bigcup_i A_i) \setminus ((\bigcup_i R_i \setminus \bigcup_i N_i) \cup (\bigcup_i N_i \setminus \bigcup_i R_i))$.

Proposition 3 *Let $\mathcal{P} = \langle \mathsf{PAF}_1, \ldots, \mathsf{PAF}_n \rangle$ be a profile of PAFs. The* common part *of $\mathcal{P}$ given by $CP(\mathcal{P}) = \langle \bigcap_i A_i, \bigcap_i R_i, \bigcap_i I_i, \bigcap_i N_i \rangle$, is pointwise included into $CFP(\mathcal{P})$, noted $CP(\mathcal{P}) \sqsubseteq CFP(\mathcal{P})$, i.e.:*

- $\bigcap_i R_i \subseteq \bigcup_i R_i \setminus \bigcup_i N_i$;
- $\bigcap_i I_i \subseteq I_{CFP}$;
- $\bigcap_i N_i \subseteq \bigcup_i N_i \setminus \bigcup_i R_i$.

The common part of a profile of n PAFs (resp. AFs) is not always a PAF (resp. an AF). Contrastingly, the conflict-free part of a profile of n PAFs is a PAF (still, the conflict-free part of a profile of n AFs is not always an AF).

A valuable property of any consensual expansion over a profile of AFs is that the conflict-free part is preserved:

Proposition 4 *Let $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ be a profile of AFs. For each $i \in 1 \ldots n$, we have $CFP(\mathcal{P}) \sqsubseteq \mathtt{exp}_C(\mathsf{AF}_i, \mathcal{P})$.*

One can also define a notion of coherence between AFs:

Definition 14 *Let $\mathsf{AF}_1 = \langle A_1, R_1 \rangle$, $\mathsf{AF}_2 = \langle A_2, R_2 \rangle$ be two AFs. AF_1, AF_2 are* coherent *iff $\nexists a, b \in A_1 \cap A_2$ such that $(a, b) \in (R_1 \setminus R_2) \cup (R_2 \setminus R_1)$. Otherwise they are* incoherent.
Let $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ be a profile of AFs. $\mathcal{P}$ is coherent *iff all its AFs are pairwise coherent. Otherwise it is* incoherent.

When a profile of AFs is coherent, its conflict-free part is the union of its elements, and the converse also holds:

Proposition 5 *Let $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ be a profile of AFs. $\mathcal{P}$ is coherent iff $CFP(\mathcal{P}) = \bigcup_i \mathsf{AF}_i$.*

The notion of compatibility of a profile of PAFs over the same set of arguments can also be introduced:

Definition 15 *Let $\mathcal{P} = \langle \mathsf{PAF}_1, \ldots, \mathsf{PAF}_n \rangle$ be a profile of PAFs over a set of arguments A. $\mathsf{PAF}_1, \ldots, \mathsf{PAF}_n$ are said to be* compatible *iff they have at least one common completion. Otherwise they are* incompatible.

Proposition 6 *Let $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ be a profile of AFs. $\mathcal{P}$ is coherent iff $\mathtt{exp}_C(\mathsf{AF}_1, \mathcal{P})$, ..., $\mathtt{exp}_C(\mathsf{AF}_n, \mathcal{P})$ are compatible.*

Let us now give some properties achieved by merging operators based on the edition distance:

Proposition 7 *Let $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ be a profile of AFs. Assume that the expansion function under use for each agent is the consensual one. For any $\mathsf{AF} \in \Delta^{\otimes}_{de}(\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle)$, we have $CFP(\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle) \sqsubseteq \mathsf{AF}$.*

Proposition 8 *Let $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ be a profile of AFs. Assume that the expansion function under use for each agent is the consensual one. If $\mathcal{P}$ is coherent, then $\Delta^{\otimes}_{de}(\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle)$ is reduced to the singleton $\{\bigcup_i \mathsf{AF}_i\}$.*

When sum is used as an aggregation function and all AFs are about the same set of arguments, the merging of a profile can be characterized in a concise way, thanks to the notion of majority graph. Intuitively the majority graph of a profile of AFs over the same set of arguments is obtained by applying the strict majority rule to decide whether a attacks b or not, for every pair (a, b) of arguments. Whenever there is no strict majority, an ignorance edge is generated.

Definition 16 *Let $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ be a profile of AFs over the same set A of arguments. The* majority PAF *$MP(\mathcal{P})$ of $\mathcal{P}$ is the PAF over A such that $\forall a, b \in A$:*[2]

- *$(a, b) \in R$ iff $\#(\{i \in 1 \ldots n \mid (a, b) \in R_i\}) > \#(\{i \in 1 \ldots n \mid (a, b) \in N_i\})$;*
- *$(a, b) \in N$ iff $\#(\{i \in 1 \ldots n \mid (a, b) \in N_i\}) > \#(\{i \in 1 \ldots n \mid (a, b) \in R_i\})$;*
- *$(a, b) \in I$ otherwise.*

Proposition 9 *Let $\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ be a profile of AFs over the same set A of arguments. $\Delta^{\Sigma}_{de}(\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle) = \mathcal{C}(MP(\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle))$.*

Acceptability for merged AFs

Starting from a set of AFs (over possibly different sets of arguments), a merging operator enables to compute another set of AFs (this time, over the same set of arguments) which are the best candidates to represent the AFs of the group. There is an important epistemic difference between those two sets of AFs, the first one reflects different points of view (given by different agents), the second one denotes the uncertainty on the result.

It is interesting to compare the joint acceptability relation for the input profile $\mathcal{P} = \langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ with the joint acceptability relation for $\Delta^{\otimes}_{d}(\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle)$. Unsurprisingly, both predicates are not connected, even in the case when the two joint acceptability relations are based on the same notion of individual acceptability and the same voting method.

Thus, it can be the case that new jointly acceptable sets are obtained after merging while they were not jointly acceptable at start:

Example 2 (continued) *Assume that the individual acceptability relation for all agents is the one which considers a set of arguments as acceptable when it is one of the preferred extensions; assume also that the voting method is (weak) majority. $\{a\}, \{b\}$ are acceptable for agent 1, $\{b, d\}$ for agent 2, $\{a\}$ for agent 3 and $\{b\}$ for agent 4. Hence, the weak majority rule leads to consider only $\{a\}$ or $\{b\}$ acceptable for the group.*

Once the AFs have been merged using Δ^{Σ}_{de}, one obtains two AFs representing the best compromises for the group. Individual acceptability leads to consider $\{a, c\}$ and $\{b, d\}$ as acceptable in both cases. Hence each of those sets is considered jointly acceptable. Thus $\{a, c\}$ is acceptable for the group, while c does not belong to any preferred extensions of the four initial AFs.

Furthermore, one can show that if a set of arguments is included into one of the preferred extensions for an agent, it is not necessarily included into one of the preferred extensions of any AF from the result of the merging. (This remains true for singletons). The converse is also true (see for instance $\{a, c\}$ in the above example).

More surprisingly, even if a set of arguments is included into each preferred extension for an agent, it is not guaranteed to be included into a preferred extension of an AF from

[2] For any set S, $\#(S)$ denotes the cardinal of S.

the result of the merging. Conversely, if a set of arguments is included into every preferred extension of the AFs from the result of the merging, it is not guaranteed to be included into a preferred extension for one of the agents. Intuitively, this can be explained by the fact that if an argument is accepted by all agents *for bad reasons* (for instance, because they lack information about attacks on it), it can be rejected by the group after the merging. More formally, this is due to the fact that nothing ensures that one of the initial AFs will belong to the result of the merging and to the nonmonotonicity of acceptability (in the sense that adding or removing a single attack (a, b) in an AF may heavily change its extensions).

Conclusion and perspectives

We have presented a new framework for merging argumentation systems. Our framework is general enough to allow for the representation of many different scenarios. No assumption is made concerning the meaning of the attack relations, so that such relations may differ not only because agents have different points of view on the way arguments interact but more generally may disagree on what an interaction is. Each agent may be associated to a specific expansion function, which enables for encoding many attitudes when facing a new argument. Many different distances between PAFs and many different aggregation functions can be used.

We plan to refine this framework in several directions. Let us briefly sketch two of them:

Merging PAFs. Our framework can be extended to PAFs merging (instead of AFs). This enables us to take into account agents with incomplete belief states regarding the attack relation between arguments. Expansions of PAFs can be defined in a very similar way to expansions of AFs (what mainly changes is the way ignorance is handled). As PAFs are more expressive than AFs, an interesting issue is to define acceptability for PAFs.

Attacks strengths. Assume that each attack believed by agent i is associated to a numerical value reflecting the strength of the attack according to the agent, i.e. the degree to which agent i believes that a attacks b. It is easy to take into account those values by modifying slightly the definition of the edition distance over a pair or arguments (for instance, viewing them as weights once normalized within $[0, 1]$). Another possibility regarding attacks strengths is, from unweighted attack relations, to generate a weighted one, representing different degree of accordance in the group. For instance, each attack (a, b) in the majority PAF of a profile $\langle \mathsf{AF}_1, \ldots, \mathsf{AF}_n \rangle$ can be labelled by the ratio $\frac{\#(\{i \in 1 \ldots n \mid (a,b) \in R_i\})}{n}$ and similarly for the non-attack relation (this leads to consider both the attack and the non-attack relations of the majority PAF as fuzzy relations). Corresponding acceptability relations remain to be defined. This is another perspective of this work.

Acknowledgements

The authors have been partly supported by the Région Nord/Pas-de-Calais, the IRCICA Consortium, the European Community FEDER Program and the IUT de Lens.

References

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

Amgoud, L., and Parsons, S. 2002. Agent dialogues with conflicting preferences. In *Intelligent Agents VIII: Agent theories, architectures and languages*, 190–205.

Amgoud, L., and Prade, H. 2004. Reaching agreement through argumentation: A possibilistic approach. In *KR'04*, 175–182.

Amgoud, L.; Maudet, N.; and Parsons, S. 2000. Arguments, Dialogue and Negotiation. In *ECAI'00*, 338–342.

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

Dunne, P., and Bench-Capon, T. 2002. Coherence in finite argument system. *Artificial Intelligence* 141:187–203.

Elvang-Gøransson, M., and Hunter, A. 1995. Argumentative logics: Reasoning with classically inconsistent information. *Data and Knowledge Engineering* 16:125–145.

Elvang-Gøransson, M.; Fox, J.; and Krause, P. 1993a. Acceptability of arguments as logical uncertainty. In *ECSQARU'93*, 85–90.

Elvang-Gøransson, M.; Fox, J.; and Krause, P. 1993b. Dialectic reasoning with inconsistent information. In *UAI'93*, 114–121.

Gordon, T. F. 1995. *The Pleadings Game. An Artificial Intelligence Model of Procedural Justice*. Kluwer Academic.

Konieczny, S., and Pino Pérez, R. 2002. Merging information under constraints: a qualitative framework. *Journal of Logic and Computation* 12(5):773–808.

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

MacKenzie, J. 1979. Question-begging in non-cumulative systems. *Journal of Philosophical logic* 8:117–133.

Parsons, S., and Jennings, N. R. 1996. Negotiation through argumentation—a preliminary report. In *ICMAS'96*, 267–274.

Pollock, J. L. 2001. Defeasible reasoning with variable degrees of justification. *Artificial Intelligence* 133:233–282.

Prakken, H., and Sartor, G. 1997. Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics* 7:25–75.

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

Walton, D., and Krabbe, E. 1995. *Commitment in dialogue. Basic concepts of interpersonal reasoning*. State University of New York Press.