

On the bipolarity in argumentation frameworks

L. Amgoud and C. Cayrol and M.C. Lagasquie-Schiex

IRIT-UPS, 118 route de Narbonne
31062 Toulouse, Cédex, France

Abstract

In this paper, we propose a survey of the use of bipolarity in argumentation frameworks, *i.e.* the presence of two kinds of entities (*a positive* entity and *a negative* entity). An argumentation process follows three steps: building the arguments and the interactions between them, valuating the arguments using or not the interactions and finally defining the acceptability of the arguments. This paper shows on various applications and with some formal definitions that bipolarity appears (in some cases since always) and can be used in each step of this process under different forms.

Introduction

A rational agent can express claims and judgements, aiming at reaching a decision, a conclusion, or informing, convincing, negotiating with other agents. Pertinent information may be insufficient or contrastedly there may be too much relevant but partially incoherent information. And, in case of multi-agent interaction, conflicts of interest are inevitable. So, agents can be assisted by argumentation, a process based on the exchange and the valuation of interacting arguments which support opinions, claims, proposals, decisions, ... The argumentation process has been applied in various domains and applications (see (Fox & Parsons 1997; Parsons 1997; Amgoud, Maudet, & Parsons 2000; Parsons, Sierra, & Jennings 1998; Karacapilidis & Papadias 2001; Gordon & Karacapilidis 1997; Verheij 2002)).

More generally, when abstracting from the internal structure of the arguments, the fundamental characteristics of an argumentation framework is the presence of interactions between arguments. For instance, in (Dung 1995)'s work, there only is one kind of interaction: a defeat relation between arguments. In some other works (see (Karacapilidis & Papadias 2001; Verheij 2002)), there also exists another kind of interaction: a support relation.

This suggests a notion of bipolarity, *i.e.* the existence of two independent kinds of information which have a diametrically opposed nature and which represent repellent forces. This notion of bipolarity appears in many domains and it is essential in order to represent realistic knowledge (see discussions in (Boutilier 1994; Tan & Pearl 1994; Lang, Van der Torre, & Weydert 2002; Benferhat *et al.* 2002)). For instance, in (Benferhat *et al.* 2002), the authors distinguish between two kinds of preferences: positive

preferences (what the agent really wants) and negative preferences (what the agent rejects). This distinction between positive and negative preferences is supported by studies in cognitive psychology which have shown that these two types of preferences are independent and processed separately in the mind. Note that bipolarity is not always related to the notion of preference.

So, the purpose of this paper is to present a *survey* on the use of bipolarity in argumentation frameworks.

The first section is devoted to the background about abstract argumentation frameworks. Then, the outline of this paper follows the structure of the argumentation process which can be viewed under the form of three steps:

1. the *building of the arguments* and, using the structure of the arguments, the *definition of different interactions between arguments*. An argument can take different forms depending on the domain, but it is generally required to be a structured set of linked propositions or claims. This step has a bipolar aspect illustrated with some applications.
2. the *valuation of these arguments* which can be based only on the interactions between arguments, or can also take into account an intrinsic strength for each argument. The resulting valuation can be crisp or gradual. In this step, the bipolarity appears on the form of the interactions between arguments (support and defeat relations between arguments). We present different formal approaches which take into account this bipolar aspect.
3. the *selection of some arguments using the definition of the acceptability*. In this step, different classes of arguments can be distinguished with different levels of acceptability, and the valuation results can be used in order to define these levels. The bipolar aspect of this step is illustrated by some examples.

The abstract framework developed by Dung

Abstracting from the structure of the arguments (so, after the step 1 of the argumentation process), in (Dung 1995) Dung has studied the interactions between arguments. However, his work has focused only on the *defeat relation* between arguments. We briefly recall that abstract framework (it will be extended in the section devoted to the valuation process):

- An argumentation framework is a pair $\langle \mathcal{A}, \mathcal{R} \rangle$ of a set \mathcal{A}

of arguments and a binary relation \mathcal{R} on \mathcal{A} called a *defeat relation*. $A_i \mathcal{R} A_j$ means that A_i defeats A_j .

- The notion of *defence* is defined from the notion of defeat by: an argument A_i defends A_j against B iff $B \mathcal{R} A_j$ and $A_i \mathcal{R} B$.

In Dung's framework, only the selection step of an argumentation process is taken into account¹. In this work, the *acceptability of an argument* depends on its membership of some sets (acceptable sets or extensions) characterised by particular properties. It is a collective acceptability. The main characteristic properties are:

- *Conflict-free*: a subset S of \mathcal{A} is conflict-free iff there exist no A_i, A_j in S such that $A_i \mathcal{R} A_j$.
- *Defends collectively*: a subset S of \mathcal{A} defends collectively an argument A_i iff for each argument B , if $B \mathcal{R} A_i$ there exists C in S such that $C \mathcal{R} B$.

Then several *semantics for acceptability* have been defined by (Dung 1995):

- *Admissible*: a subset S of \mathcal{A} is an admissible set iff S is conflict-free and S defends collectively all its elements.
- *Preferred*: a subset S of \mathcal{A} is a preferred extension iff S is maximal for the set inclusion among the admissible sets of \mathcal{A} .
- *Grounded*: a subset S of \mathcal{A} is the grounded extension iff S is the least fixed point of the characteristic function F of $\langle \mathcal{A}, \mathcal{R} \rangle$ ($F: 2^{\langle \mathcal{A}, \mathcal{R} \rangle} \rightarrow 2^{\langle \mathcal{A}, \mathcal{R} \rangle}$ with $F(S) = \{A \text{ such that } A \text{ is defended collectively by } S\}$).

These semantics are used in the selection step of the argumentation process and, sometimes, they are associated with the results of the valuation step (see for example (Cayrol & Lagasque-Schiex 2003a)). They will be used in the section devoted to the selection process.

Bipolarity at the argument level

In an argumentation process some arguments are constructed in *favour* of a conclusion and others are constructed *against* that conclusion (called also defeaters). One might say that arguments are presented in a bipolar way since arguments in favour of a conclusion can be considered as *positive* and arguments against the conclusion as *negative* ones. In some sense this is true since the two kinds of arguments have different and opposite roles. However, generally these arguments are defined and handled in the same way.

Within the negotiation framework, various natures of arguments have been distinguished in (Amgoud & Prade 2004a): *instrumental arguments, explanatory arguments, treats and rewards*. For each nature of argument, only one definition is proposed (for supporting or attacking) a conclusion. Let's consider the case of explanatory arguments. Explanations constitute the most common category of arguments. In classical argumentation-based frameworks which have been developed for handling inconsistency in knowledge bases, each conclusion is justified by arguments. They represent the reasons to believe in the conclusion. Such arguments

¹However, some notions proposed by Dung can be used in the valuation of arguments (see (Cayrol & Lagasque-Schiex 2003b)).

have a deductive form. Indeed, from premises, a conclusion is entailed.

Let \mathcal{L} be a propositional language. \vdash denotes classical inference and \equiv denotes logical equivalence. Let $\mathcal{K} = \{k_j; j = 1, \dots, l\}$ be a base representing the available knowledge of an agent with k_j are formulas of \mathcal{L} .

Definition 1 (Explanatory argument) An explanatory argument is a pair $\langle H, h \rangle$ such that:

- $H \subseteq \mathcal{K}$.
- $H \vdash h$.
- H is consistent and minimal (for set inclusion) among the consistent sets H satisfying the two previous items.

H is the support of the argument and h its conclusion.

Example 1 Let $\mathcal{K} = \{p, p \rightarrow b, p \rightarrow \neg f, b \rightarrow f\}$ where p means penguin, b means bird, f means fly. Let $H = \{p, p \rightarrow b, b \rightarrow f\}$, $H' = \{p, p \rightarrow \neg f\}$ be two subsets of \mathcal{K} . The fact that p flies is justified by the explanatory argument $\langle H, f \rangle$. However, the conclusion f has a counter-argument which is $\langle H', \neg f \rangle$.

In the case of argumentation-based decision, things seem different. In (Amgoud & Prade 2004b) the decisions are based on the comparison of arguments and counter-arguments which are defined in different ways. The idea is that a decision is justified if it leads to the satisfaction of some desires. Let $\mathcal{D} = \{d_i; i = 1, \dots, m\}$ represent the desires of the decision-maker and \mathcal{D}_e be a set of decisions. Elements of \mathcal{D} and \mathcal{D}_e are formulas of \mathcal{L} .

Definition 2 (Argument for a decision) An argument in favour of a decision is a triple $A = \langle S, C, d \rangle$ such that:

- $d \in \mathcal{D}_e$
- $S \subseteq \mathcal{K}$ and $C \subseteq \mathcal{D}$
- $S \cup \{d\}$ is consistent
- $S \cup \{d\} \vdash C$
- S is minimal and C is maximal (for set inclusion) among the sets S and C satisfying the four previous items.

$S = \text{Support}(A)$ is the support of the argument and $C = \text{Consequences}(A)$ its consequences (the desires which are reached by the decision d).

Arguments against a decision, however, show clearly that desires will not be satisfied by that decision. Formally:

Definition 3 (Argument against a decision) An argument against a decision is a triple $A = \langle S, C, d \rangle$ such that:

- $d \in \mathcal{D}_e$
- $S \subseteq \mathcal{K}$ and $C \subseteq \mathcal{D}$
- $S \cup \{d\}$ is consistent
- $\forall g_i \in C, S \cup \{d\} \vdash \neg g_i$
- S is minimal and C is maximal (for set inclusion) among the sets S and C satisfying the four previous items.

$S = \text{Support}(A)$ is the support of the argument and $C = \text{Consequences}(A)$ its consequences (the desires which are not satisfied by the decision d).

Example 2 The example is about taking an umbrella or not, knowing that the sky is cloudy. The knowledge base is $\mathcal{K} = \{u \rightarrow \neg w, r \wedge \neg u \rightarrow w, c, \neg r \rightarrow \neg w, c \rightarrow r\}$ with: r : it rains, w : being wet, u : taking an umbrella, c : the sky is cloudy.

The base of desires is $\mathcal{D} = \{\neg w, \neg u\}$. There is one argument in favour of the decision “ u ”: $\langle \{u \rightarrow \neg w\}, \{\neg w\}, u \rangle$ and one argument against the decision “ u ”: $\langle \emptyset, \{\neg u\}, u \rangle$.

Note also that the above two arguments are handled in different ways. In (Amgoud & Prade 2004a), it has been showed that the use of arguments in favour of a decision is sufficient to capture the results of the pessimistic criteria defined in (Dubois & Prade 1995) in qualitative decision making. Whereas, the use of arguments against a decision allows us to capture the results of the optimistic criteria.

Valuation in a bipolar argumentation framework

In the basic argumentation context recalled in the first section, only one kind of interaction is explicitly represented by the *defeat* relation. Within that framework, several approaches have been proposed for valuating the arguments, using or not the defeat relation², see for example (Krause *et al.* 1995; Parsons 1997; Prakken & Sartor 1997; Jakobovits & Vermeir 1999; Besnard & Hunter 2001; Cayrol & Lagasquie-Schiex 2003b; Amgoud 1999).

In this context, the support of an argument A by another argument B can be represented only if B defends A in the sense of (Dung 1995). So, this support is a *dependent* notion of the defeat relation. It is not necessary a bad thing (many systems use this approach), but we are interested in modelling situations where two *independent* kinds of interactions are available: a positive and a negative one (see for example in the medical domain the work (Karacapilidis & Papadias 2001)). So, following (Karacapilidis & Papadias 2001; Verheij 2002), we propose a new argumentation framework: an abstract bipolar argumentation framework and an associated valuation procedure.

An abstract bipolar argumentation framework

An abstract bipolar argumentation framework is an extension of the basic argumentation framework introduced by (Dung 1995) in which we use a new kind of interactions between arguments: the *support* relation which represents the support, the help brought by some arguments to other arguments³. This new relation is assumed to be totally independent of the defeat relation. So, we have a bipolar representation of the interactions between arguments.

Formal definition of an abstract bipolar argumentation framework An abstract bipolar argumentation framework $\langle \mathcal{A}, \mathcal{R}_{def}, \mathcal{R}_{sup} \rangle$ is a set \mathcal{A} of arguments, a binary

²So, the value of an argument may depend on its interactions with the other arguments, or may depend on an intrinsic strength of the argument which can be defined for example by an explicit preference.

³If the support relation is removed, we retrieve Dung’s framework.

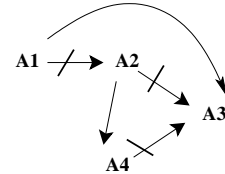
relation \mathcal{R}_{def} on \mathcal{A} called a *defeat relation* and another binary relation \mathcal{R}_{sup} on \mathcal{A} called a *support relation*: consider A_i and $A_j \in \mathcal{A}$, $A_i \mathcal{R}_{def} A_j$ (resp. $A_i \mathcal{R}_{sup} A_j$) means that A_i defeats A_j (resp. A_i supports A_j).

A bipolar argumentation framework is called *well-founded* if and only if there is no infinite sequence $A_0, A_1, \dots, A_n, \dots$ such that $\forall i, A_i \in \mathcal{A}$ and $A_{i+1} \mathcal{R} A_i$ with $\mathcal{R} = \mathcal{R}_{def}$ or \mathcal{R}_{sup} .

Here, we are not interested in the structure of the arguments and we consider arbitrary defeat and support relations. The only assumption is that \mathcal{R}_{def} and \mathcal{R}_{sup} are independent of each other.

Notations Consider $A \in \mathcal{A}$, $A \mathcal{R}_{def} B$ is represented by $A \not\rightarrow B$ and $A \mathcal{R}_{sup} B$ is represented by $A \rightarrow B$. The set $\{A_i \in \mathcal{A} | A_i \mathcal{R}_{def} A\}$ is denoted by $\mathcal{R}_{def}^-(A)$ and the set $\{A_i \in \mathcal{A} | A \mathcal{R}_{def} A_i\}$ is denoted by $\mathcal{R}_{def}^+(A)$. In the same way, we define $\mathcal{R}_{sup}^-(A)$ and $\mathcal{R}_{sup}^+(A)$. $\langle \mathcal{A}, \mathcal{R}_{def}, \mathcal{R}_{sup} \rangle$ defines a directed graph \mathcal{G}_b called the *bipolar graph*.

Example 3 The framework $\langle \mathcal{A} = \{A_1, A_2, A_3, A_4\}, \mathcal{R}_{def} = \{(A_2, A_3), (A_4, A_3), (A_1, A_2)\}, \mathcal{R}_{sup} = \{(A_2, A_4), (A_1, A_3)\} \rangle$ defines the following graph \mathcal{G}_b with the root A_3 :



Definition 4 (Graphical representation of a bipolar argumentation framework) Let \mathcal{G}_b be the bipolar graph associated with the abstract bipolar argumentation framework $\langle \mathcal{A}, \mathcal{R}_{def}, \mathcal{R}_{sup} \rangle$, we define:

- **Leaf of the bipolar graph** A leaf of \mathcal{G}_b is an argument $A \in \mathcal{A}$ such that $\mathcal{R}_{def}^-(A) = \emptyset$ and $\mathcal{R}_{sup}^-(A) = \emptyset$.
- **Path in the bipolar graph** A path from A to B is a sequence of arguments $\mathcal{P} = A_1 - \dots - A_n$ such that $A = A_1, A_1 \mathcal{R}_1 A_2, \dots, A_{n-1} \mathcal{R}_{n-1} A_n, A_n = B$, and $\forall i = 1, \dots, n-1, \mathcal{R}_i = \mathcal{R}_{def}$ or \mathcal{R}_{sup} . The length of the path is $n-1$ (the number of edges that are used in the path) and will be denoted by $l_{\mathcal{P}}$. The defeat number of the path (resp. support number of the path) is the number of $\mathcal{R}_i = \mathcal{R}_{def}$ (resp. $\mathcal{R}_i = \mathcal{R}_{sup}$) and will be denoted by $n_{def}(\mathcal{P})$ (resp. $n_{sup}(\mathcal{P})$). A homogeneous path from A to B is a path in which all the \mathcal{R}_i are the same. So, we can have homogeneous defeat paths or homogeneous support paths. The set of the paths from A to B will be denoted by $\mathcal{P}(A, B)$.
- **Branch for an argument** A path from A to B is a branch for B iff A is a leaf of \mathcal{G}_b .

We assume that the bipolar graph is acyclic. We propose the notions of direct and indirect defeaters and defenders⁴,

⁴The notions introduced here are inspired by related definitions first introduced in (Dung 1995) but are not strictly equivalent.

completed with the notion of direct and indirect supporters. Note that negative information (defeat edges) is considered as having priority over positive information (support edges). So, we do not have symmetrical definitions for indirect defeaters/defenders and indirect supporters⁵:

Definition 5 (Direct/Indirect Defeaters/Defenders of an argument) Consider $A \in \mathcal{A}$:

- The direct defeaters of A are the elements of $\mathcal{R}_{def}^-(A)$.
- The direct defenders of A are the direct defeaters of the elements of $\mathcal{R}_{def}^-(A)$.
- The indirect defeaters of A are the elements A_i defined by:
 - $\exists \mathcal{P} \in \mathcal{P}(A_i, A)$ such that $n_{def}(\mathcal{P}) = 2k + 1$, with $k \geq 0$ and A_i is not a direct defeater.
- The indirect defenders of A are the elements A_i defined by:
 - $\exists \mathcal{P} \in \mathcal{P}(A_i, A)$ such that $n_{def}(\mathcal{P}) = 2k$, with $k \geq 1$ and A_i is not a direct defender.

Definition 6 (Direct/Indirect supporters of an argument) Consider $A \in \mathcal{A}$:

- The direct supporters of A are the elements of $\mathcal{R}_{sup}^-(A)$.
- The indirect supporters of A are the elements A_i defined by:
 - $\exists \mathcal{P} \in \mathcal{P}(A_i, A)$ such that $n_{sup}(\mathcal{P}) = l_{\mathcal{P}} \geq 2$.

Remarks and examples In a bipolar argumentation framework, the support relation carries positive information while the defeat relation carries negative information, and positive and negative information are represented in the same structure (the bipolar graph). It is a particularity of the argumentation context (in many other domains, positive and negative information are represented in two distinct frameworks, and sometimes they do not have the same nature).

We present below two illustrative examples.

Example 4 During a discussion between reporters about the publication of an information I concerning the person X , the following arguments are presented:

- A : I is an important information, we must publish it.
- B : I concerns the person X , X is a private person and we can not publish an information about a private person without her agreement, and X does not agree with the publication.
- C : X is a minister, so X is a public person, not a private person.
- D : X has resigned, so X is no more a minister.
- E : her resignation has been refused by the chief of the government.

lent: in (Dung 1995)'s work, direct defeaters (resp. defenders) are also indirect defeaters (resp. defenders) which is not true in our definitions.

⁵As soon as the path $A_i - A$ contains at least one defeat edge, it defines A_i as an indirect defeater or defender for A ! Contrastedly, an indirect supporter A_i for A excludes defeat edges in the path $A_i - A$. This illustrates that bipolarity is well represented by the relations defeat and support while no bipolarity exists with the relations defeat and defend.

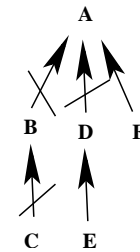
- F : I concerns a problem of public health, so I is an important information.



In this example, B is a direct defeater of A , F is a direct supporter of A , C is a direct defender of A , D is an indirect defeater of A and E is an indirect defender of A .

Example 5 During a discussion between doctors about the installation of a prosthesis on the patient X , the following arguments are presented:

- A : X has difficulties for walking, we must install a prosthesis.
- B : the installation of a prosthesis needs a surgical operation with an anaesthesia which is very risked for the patient.
- C : we can use a local anaesthesia, so there is no more risk.
- D : a surgical operation presents also important risks of post-infections.
- E : there are more and more kinds of nosocomial infections in the hospital and it is very difficult to cure them.
- F : the classical treatments (injections) are unable to cure X 's knee problem.



In this example, B and D are direct defeaters of A , F is a direct supporter of A , C is a direct defender of A and E is an indirect defeater of A .

Gradual valuation in a bipolar argumentation framework

In the context of a bipolar argumentation framework, the valuation follows the same principles that have already been described in (Cayrol & Lagasque-Schiex 2003b) completed with new principles corresponding to the "support" information. Here, we propose two kinds of valuations:

- a *local valuation* in which the value of an argument only depends on the values of the direct defeaters or supporters of this argument.
- a *global valuation* in which the value of an argument represents the set of all the branches leading to this argument.

Local valuation in a bipolar argumentation framework

Let v be a local gradual valuation in a bipolar argumentation framework, v must respect the following principles:

P11 The valuation of an argument is a function of its direct defeaters and of its direct supporters.

P12 If the quality of the support (resp. defeat) increases then the value of the argument increases (resp. decreases).

P13 If the quantity of the supports (resp. defeats) increases then the quality of the support (resp. defeat) increases.

In the respect of these principles, we propose the following formal definition for a local gradual valuation.

Definition 7 Let $\langle \mathcal{A}, \mathcal{R}_{def}, \mathcal{R}_{sup} \rangle$ be a bipolar argumentation framework. We assume that there exists a completely ordered set V with a minimum element (V_{Min}) and a maximum element (V_{Max}).

Consider $A \in \mathcal{A}$ with $\mathcal{R}_{def}^-(A) = \{B_1, \dots, B_n\}$ and $\mathcal{R}_{sup}^-(A) = \{C_1, \dots, C_p\}$. A local gradual valuation on $\langle \mathcal{A}, \mathcal{R}_{def}, \mathcal{R}_{sup} \rangle$ is a function $v : \mathcal{A} \rightarrow V$ such that

$$v(A) = g(h_{sup}(v(C_1), \dots, v(C_p)), h_{def}(v(B_1), \dots, v(B_n)))$$

with:

the function h_{def} (resp. h_{sup}): $V^* \rightarrow \mathcal{H}_{def}$ (resp. $V^* \rightarrow \mathcal{H}_{sup}$)⁶ valuating the quality of the defeat (resp. support) on A .

the function g : $\mathcal{H}_{sup} \times \mathcal{H}_{def} \rightarrow V$ with $g(x, y)$ increasing on x and decreasing on y .

The function h , $h = h_{def}$ or h_{sup} , must respect the following constraints:

- if $x_i \geq x'_i$ then $h(x_1, \dots, x_i, \dots, x_n) \geq h(x_1, \dots, x'_i, \dots, x_n)$,
- $h(x_1, \dots, x_n, x_{n+1}) \geq h(x_1, \dots, x_n)$,
- $h() = \alpha \leq h(x_1, \dots, x_n)$, for all x_1, \dots, x_n ⁷,
- $h(x_1, \dots, x_n) \leq \beta$, for all x_1, \dots, x_n ⁸.

Note that the definition 7 produces a generic local gradual valuation. There exist several instances for this generic valuation. One of them is the following:

- $V = [-1, 1]$ interval of reals,
- $\mathcal{H}_{def} = \mathcal{H}_{sup} = [0, \infty[$ interval of reals,
- $h_{def}(x_1, \dots, x_n) = h_{sup}(x_1, \dots, x_n) = \sum_{i=1}^n \frac{x_i+1}{2}$,
- $g(x, y) = \frac{1}{1+y} - \frac{1}{1+x}$.

⁶ V^* denotes the set of the finite sequences of elements of V . \mathcal{H}_{def} and \mathcal{H}_{sup} are ordered sets.

⁷So, α is the minimal value for a defeat (resp. a support) – i.e. there is no defeat (resp. no support) –.

⁸So, β is the maximal value for a defeat (resp. a support) – i.e. for example, if there is an infinity of direct defeaters (resp. supporters) –.

So, we have $\alpha = 0, \beta = \infty$ and $g(\alpha, \alpha) = 0$ ⁹.

On the example 4, we obtain $v(A) = \frac{1}{15}$.

On the example 5, we obtain $v(A) = \frac{-1}{6}$.

Note that the functions h_{def} and h_{sup} represent the two axes of an unipolar bivariate scale¹⁰ Then, with the function g , we obtain a bipolar univariate scale¹¹. So, after using g , we can not distinguish the case “balance between defeats and supports” and the case “no defeat and no support”.

Global valuation in a bipolar argumentation framework

In the global approach, we have also some principles:

Pg1 The value of an argument is a function of all the branches leading to this argument in the bipolar graph.

Pg2 The set of the branches leading to this argument is partitioned in three parts, each part corresponding to a kind of “effect” on the argument (there are only three “effects”, so we have the defeat part, the defence part, the support part).

Pg3 The improvement of the defence or the support parts or the degradation of the defeat part of an argument leads to an increase of the value of this argument.

Pg4 The improvement of the defeat part or the degradation of the defence or the support parts of an argument leads to a decrease of the value of the argument.

Here, the main problem is: how to determine the “effect” of a branch leading to the argument ? There exist different methods and each of them leads to a kind of global valuation. Here, we will present only one of these methods.

Definition 8 (Defeat, defence and support branches for an argument) Let A be an argument.

Defeat (resp. defence) branch for A A branch for A is a defeat branch (resp. defence branch) iff the longest homogeneous path leading to A of this branch is an homogeneous defeat path whose length is an odd integer (resp. even integer).

Support branch for A A branch for A is a support branch iff the longest homogeneous path leading to A of this branch is an homogeneous support path.

⁹Note that $h_{def}(x_1, \dots, x_n, x_{n+1}) \geq h_{def}(x_1, \dots, x_n)$ because $\frac{x_{n+1}+1}{2} \geq 0$ when $x_{n+1} \in [-1, 1]$ (and the same for h_{sup}). We have also $h_{def}() = h_{sup}() = \alpha$, α being the minimal value of $[0, \infty[$, and β being the maximal value of $[0, \infty[$. We can verify also that $g(\alpha, \beta) = g(0, \infty) = -1$ and that $g(\beta, \alpha) = g(\infty, 0) = 1$ (1 and -1 being respectively V_{Min} and V_{Max}).

¹⁰It is the first kind of representation of bipolar informations: each information is independent and corresponds to a specific axe. So, we can distinguish the case “there are no information positive and no information negative” and the case “there are as much information positive as information negative”.

¹¹It is the second kind of representation of bipolar informations: both informations are joined and the result corresponds to only one axe. So, it will be difficult to distinguish the case “there are no information positive and no information negative” and the case “there are as much information positive as information negative”, because generally the result of the join is the same in the both cases.

On the example 4, $E - D - C - B - A$ is a defence branch for A and $F - A$ is a support branch for A . In the example 5, $C - B - A$ is a defence branch for A , $E - D - A$ is a defeat branch for A and $F - A$ is a support branch for A .

Using this definition, we can define the value of the argument under the form of three values (one for each part evoked in the principles – defeat part, defence part, support part of the set of the branches leading to the argument). These values must represent all the branches leading to the argument and having the same effect. So, following the encoding used in (Cayrol & Lagasquie-Schiex 2003b) in the case of a basic argumentation framework, we can consider each value as a tuple of informations (one information for each branch). In (Cayrol & Lagasquie-Schiex 2003b), these informations were only the lengths of branches, but here, it is more complex because of the mix of support edges and defeat edges in a same branch. So, one idea is to encode the description of the branch under the form of a sequence of bits (1 for a defeat edge and 0 for a support edge). With this encoding, it is very easy to distinguish defeat/defence branches and support branches.

In (Cayrol & Lagasquie-Schiex 2004), the reader can find the complete definitions of these three tupled-values and some comparison algorithms on these values respecting the proposed principles.

Note that, although there are three tupled-values per each argument, we have an unipolar bivariate scale, because the defence-value and the defeat-value are used together “against” the support-value.

Note also that, with the global approach, we can distinguish the case “balance between defeats and supports” and the case “no defeat and no support”.

Bipolarity in the selection of the acceptable arguments

Bipolarity appears also when defining the acceptability of arguments, even in the case of a basic argumentation framework¹². In fact, using the Dung’s framework, an argumentation process should return three categories of arguments:

- The class of *acceptable* arguments. Beliefs or goals or decisions supported by such arguments are really justified. In the case of handling inconsistency in knowledge bases, beliefs supported by such arguments will be inferred from the base. Similarly, goals supported by such arguments will be pursued by the agent.
- The class of *rejected* arguments. For example, goals supported only by such arguments will be rejected by the agent even if they can be achieved. Decisions supported by such arguments will be discarded and beliefs supported by such arguments will not be inferred from the knowledge base.
- The class of arguments *in abeyance*. Such arguments are neither acceptable nor rejected.

¹²The definition of acceptability in the case of a bipolar argumentation framework remains to be studied.

Depending on the nature of the arguments and the application which is considered, two kinds of bipolarity can be distinguished when defining the acceptability of arguments. In the first one, the two classes of rejected arguments and arguments in abeyance are defined on the basis of the class of acceptable arguments. For example, in the case of handling inconsistency in knowledge bases, the grounded extension (introduced in the first section) may be used to define the acceptable arguments.

Definition 9 *An argument is acceptable if it belongs to the grounded extension and an argument is rejected if it is attacked by an acceptable argument.*

In the second kind of bipolarity, the classes of acceptable and rejected arguments are defined separately and the class of arguments in abeyance is deduced from these two classes. We illustrate this kind of bipolarity in a particular application where we try to compute the intentions of an agent from its contradictory desires (see (Amgoud 2003) for more details). Let \mathcal{L} be a propositional language, an agent is supposed to be equipped with a base \mathcal{D} of desires, a belief base Σ containing the plans to carry out in order to achieve the desires (we are not interested in the way in which these plans are generated), and finally a base \mathcal{C} of integrity constraints.

- \mathcal{D} contains literals of \mathcal{L} . The elements of \mathcal{D} represent the initial desires of the agent. For example, an agent may have the following desires: *to finish a publication, to go to a dentist, etc...* Note that the set \mathcal{D} may be inconsistent. This means that an agent is allowed to have contradictory desires.
- Σ contains rules having the form $\varphi_1 \wedge \dots \wedge \varphi_n \rightarrow h$ where $\varphi_1, \dots, \varphi_n, h$ are literals of \mathcal{L} . Such a formula means that the agent believes that if he realizes $\varphi_1, \dots, \varphi_n$ then he will be able to achieve h .
- \mathcal{C} contains formulas of \mathcal{L} . They represent a kind of integrity constraints.

A desire is any element of \mathcal{D} . A desire h may have sub-desires. For example, the desire of “going on a journey to central Africa” may have two sub-desires which are: “getting the tickets” and “being vaccinated”. The sub-desire “getting the tickets” may have itself the two following sub-desires: “having a friend who may bring the tickets” and “passing to an agency”.

Definition 10 (Desire/Sub-desire) *Let us consider an agent equipped with the bases $\langle \mathcal{D}, \Sigma, \mathcal{C} \rangle$.*

1. \mathcal{D} is the set of the desires of the agent.
2. $Sub\mathcal{D}$ is the set of the sub-desires of the agent: A literal $h' \in Sub\mathcal{D}$ iff there exists a rule $\varphi_1 \wedge h' \dots \wedge \varphi_n \rightarrow h \in \Sigma$ with $h \in \mathcal{D}$ or $h \in Sub\mathcal{D}$. In that case, h' is a sub-desire of h .

As noted above, an agent may have one or several ways to achieve a given desire. We bring the two notions together in a new notion of *partial plan*.

Definition 11 (Partial plan) *A partial plan for h is a pair $a = \langle h, H \rangle$ such that:*

- h is a desire or a sub-desire.

- $H = \{\varphi_1, \dots, \varphi_n\}$ if there exists a rule $\varphi_1 \wedge \dots \wedge \varphi_n \rightarrow h \in \Sigma$, $H = \emptyset$ otherwise.

The function $Desire(a) = h$ returns the desire or sub-desire of a partial plan a and the function $Support(a) = H$ returns the support of the partial plan. \aleph will gather all the partial plans that can be built from $\langle \mathcal{D}, \Sigma, \mathcal{C} \rangle$.

Note 1 A desire may have several partial plans.

Note 2 Let $a = \langle h, H \rangle$ be a partial plan. Each element of the support H is a sub-desire of h .

Definition 12 A partial plan $a = \langle h, H \rangle$ is elementary iff $H = \emptyset$.

A partial plan shows the actions that should be performed in order to achieve the corresponding desire (or sub-desire). However, the elements of the support of a given partial plan are considered as sub-desires that must be achieved at their turn by another partial plan. The whole way to achieve a given desire is called a *complete plan*. A *complete plan* for a given desire d is an AND tree. Its nodes are partial plans and its arcs represent the sub-desire relationship. The root of the tree is a partial plan for the desire d . It is an AND tree because all the sub-desires of d must be considered. When for the same desire, there are several partial plans to carry it out, only one is considered in a tree. Formally:

Definition 13 (Complete plan) A complete plan g for a desire h is a finite tree such that:

- $h \in \mathcal{D}$ and the root of the tree is a partial plan $\langle h, H \rangle$.
- A node $\langle h', \{\varphi_1, \dots, \varphi_n\} \rangle$ has exactly n children $\langle \varphi_1, H'_1 \rangle, \dots, \langle \varphi_n, H'_n \rangle$ where $\langle \varphi_i, H'_i \rangle$ is a partial plan for φ_i .
- The leaves of the tree are elementary partial plans.

The function $Root(g) = h$ returns the desire of the root. The function $Nodes(g)$ returns the set of all the partial plans of the tree g . \mathcal{G} denotes the set of all the complete plans that can be built from the triple $\langle \mathcal{D}, \Sigma, \mathcal{C} \rangle$. The function $Leaves(g)$ returns the set of the leaves of the tree g .

In (Amgoud 2003), it has been shown that partial plans may be conflicting for several reasons. These different kinds of conflicts are brought together in a unique relation of *conflict* defined as follows:

Definition 14 (Conflict) Let a_1 and a_2 be two partial plans of \aleph . a_1 conflicts with a_2 iff: $\{Desire(a_1), Desire(a_2)\} \cup Support(a_1) \cup Support(a_2) \cup \mathcal{C} \cup \Sigma \vdash \perp$.

More generally, a set of partial plans may be conflicting.

Definition 15 Let $S \subseteq \aleph$. S is conflicting iff $\bigcup_{a \in S} (\{Desire(a)\} \cup Support(a)) \cup \mathcal{C} \cup \Sigma \vdash \perp$.

Since partial plans may be conflicting, two complete plans may be conflicting too.

Definition 16 (Attack) Let $g_1, g_2 \in \mathcal{G}$. g_1 attacks g_2 iff $\exists a_1 \in Nodes(g_1)$ and $\exists a_2 \in Nodes(g_2)$ such that a_1 conflicts with a_2 .

More generally we are interested in sets of complete plans such that there is no conflict between their nodes. Formally:

Definition 17 (Conflict-free) Let $S \subseteq \mathcal{G}$. S is conflict-free¹³ iff:

$$\left[\bigcup_{g \in S} \left[\bigcup_{a \in Nodes(g)} (Support(a) \cup \{Desire(a)\}) \right] \cup \mathcal{C} \cup \Sigma \not\vdash \perp \right].$$

If $S = \{g\}$, then we say that the complete plan g is conflict-free.

Obviously a desire which has no conflict-free complete plan will be called *unachievable*. This means it is impossible to carry out such a desire.

Definition 18 (Unachievable desire) A desire d is unachievable if $\nexists g \in \mathcal{G}$ s.t $Root(g) = d$ and g is conflict-free.

From the preceding definitions, we can now present the formal system for handling conflicting desires of an agent.

Definition 19 (System for handling desires) Let's consider a triple $\langle \mathcal{D}, \Sigma, \mathcal{C} \rangle$. The pair $\langle \mathcal{G}, Attack \rangle$ will be called a system for handling desires (SHD).

A SHD has the same features as an argumentation framework (Amgoud & Cayrol 2002). Inspired by previous work on argumentation theory, we will define acceptable sets of complete plans. Then we will be able to partition the set \mathcal{G} into three categories thus meeting again the notion of bipolarity in the selection step:

1. The acceptable set(s) of complete plans. They contain the *good plans* to achieve their corresponding desires. These desires will become the intentions of the agent.
2. The class of *rejected complete plans*. These are the *self-attacked* ones.
3. The class of *complete plans in abeyance* which gathers the complete plans which are neither good nor rejected.

We give below the semantics of “acceptable sets of complete plans”.

Definition 20 Let $\langle \mathcal{G}, Attack \rangle$ be a SHD and $S \subseteq \mathcal{G}$. S is an acceptable set of complete plans iff: S is conflict-free and S is maximal (for set inclusion).

Conclusion

This paper presents a comprehensive survey of the use of bipolarity in argumentation frameworks. Indeed, it shows that bipolarity may be used at different levels of an argumentation process. However, this depends broadly on the considered application.

- When constructing arguments, for instance, we have shown that in some situations bipolarity does not appear at all. It is the case in argumentation-based inference systems where arguments are constructed in favour of and against a given conclusion. Even if the two kinds of arguments play different roles, their logical definitions are similar. In argumentation-based decision, things seem different since two kinds of arguments are built and handled differently: arguments in favour of a decision and arguments against a decision. In this case, arguments are bipolar.

¹³Note that this notion is not the same one which is defined by (Dung 1995) and recalled in the first section.

- At a meta level, we can say that we have arguments in favour of other arguments (the support relation) and also arguments against other arguments (the defeat relation). These bipolar relations between arguments may be used in order to define the strengths of arguments. So, in this case, we propose an extension of Dung's framework: an abstract bipolar argumentation framework.
- Finally, the result of an argumentation process may be presented in a bipolar way: acceptable arguments, rejected arguments (the arguments in abeyance being derived from the two previous classes). Here, like in the building of arguments, the bipolarity really appears when we choose acceptable arguments and rejected arguments with two distinct processes.

In all the levels, underlying or introducing bipolarity permit an enrichment and an improvement of the power of the argumentation process and reflects the need of bipolar informations in many real applications (see many discussions about this subject (Boutilier 1994; Tan & Pearl 1994; Lang, Van der Torre, & Weydert 2002; Benferhat *et al.* 2002)).

The future works will be the complete study of:

- the abstract bipolar argumentation framework (in particular, the definition of acceptability in this framework),
- the links between bipolarity in argumentation and other kinds of bipolarity in different domains (in particular in knowledge representation).

References

- Amgoud, L., and Cayrol, C. 2002. A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence* 34:197–216.
- Amgoud, L., and Prade, H. 2004a. Generation and evaluation of different types of arguments in negotiation. Technical report, Institut de Recherche en Informatique de Toulouse (I.R.I.T.).
- Amgoud, L., and Prade, H. 2004b. Towards argumentation-based decision making: A possibilistic logic approach. Technical report, Institut de Recherche en Informatique de Toulouse (I.R.I.T.).
- Amgoud, L.; Maudet, N.; and Parsons, S. 2000. Arguments, Dialogue and Negotiation. In Horn, W., ed., *Proc. of the 14th ECAI (European Conference of Artificial Intelligence)*, 338–342.
- Amgoud, L. 1999. *Contribution à l'intégration des préférences dans le raisonnement argumentatif*. Ph.D. Dissertation, Université Paul Sabatier, Toulouse.
- Amgoud, L. 2003. A formal framework for handling conflicting desires. In *Proc. of the 7th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU'2003*, 552–563.
- Benferhat, S.; Dubois, D.; Kaci, S.; and Prade, H. 2002. Bipolar representation and fusion of preferences in the possibilistic logic framework. In *Proceedings of the eighth International Conference on Principle of Knowledge Representation and Reasoning (KR'02)*, 158–169.
- Besnard, P., and Hunter, A. 2001. A logic-based theory of deductive arguments. *Artificial Intelligence* 128 (1-2):203–235.
- Boutilier, C. 1994. Towards a logic for qualitative decision theory. In *Proc. of the 4th KR*, 75–86.
- Cayrol, C., and Lagasquie-Schiex, M.-C. 2003a. Gradual acceptability in argumentation systems. In *Proc. of the 3rd CMNA (International workshop on computational models of natural argument)*, 55–58.
- Cayrol, C., and Lagasquie-Schiex, M.-C. 2003b. Gradual handling of contradiction in argumentation frameworks. In Bouchon-Meunier, B.; L.Fouloy; and Yager, R., eds., *Intelligent Systems for Information Processing: From representation to Applications*. Elsevier. chapter Reasoning, 179–190.
- Cayrol, C., and Lagasquie-Schiex, M.-C. 2004. Bipolarité en argumentation. Rapport de recherche 2004-07-R, Institut de Recherche en Informatique de Toulouse (I.R.I.T.), France.
- Dubois, D., and Prade, H. 1995. Possibility theory as a basis for qualitative decision theory. In *14th Inter. Joint Conf. on Artificial Intelligence (IJCAI'95)*, 1924–1930. Montréal: Morgan Kaufmann, San Mateo, CA.
- 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.
- Fox, J., and Parsons, S. 1997. On using arguments for reasoning about values and actions. In *Proc. of AAAI-Symposium on qualitative preferences in deliberation and practical reasoning*, 55–63.
- Gordon, T., and Karacapilidis, N. 1997. The zeno argumentation framework. In *Proc. of the 6th International Conference on Artificial Intelligence and Law*, 10–18. New-York, USA: ACM Press.
- Jakobovits, H., and Vermeir, D. 1999. Robust semantics for argumentation frameworks. *Journal of logic and computation* 9(2):215–261.
- Karacapilidis, N., and Papadias, D. 2001. Computer supported argumentation and collaborative decision making: the HERMES system. *Information systems* 26(4):259–277.
- Krause, P.; Ambler, S.; Elvang, M.; and Fox, J. 1995. A logic of argumentation for reasoning under uncertainty. *Computational Intelligence* 11 (1):113–131.
- Lang, J.; Van der Torre, L.; and Weydert, E. 2002. Utilitarian desires. *Journal of Autonomous Agents and Multi-Agents Systems* 5(3):329–363.
- Parsons, S.; Sierra, C.; and Jennings, N. R. 1998. Agents that reason and negotiate by arguing. *Journal of Logic and Computation* 8(3):261–292.
- Parsons, S. 1997. Normative argumentation and qualitative probability. In *Proc. of the first International Joint Conference on Qualitative and quantitative practical reasoning, ECSQARU-FAPR, LNAI 1244*, 466–480.

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

Tan, S. W., and Pearl, J. 1994. Specification and evaluation of preferences under uncertainty. In *Proc. of the 4th KR*, 530–539.

Verheij, B. 2002. On the existence and multiplicity of extension in dialectical argumentation. In Benferhat, S., and Giunchiglia, E., eds., *Proceedings of the 9th International Workshop on Non-Monotonic Reasoning (NMR'2002)*, 416–425.