Generalizations of Dung Frameworks and Their Role in Formal Argumentation*

Gerhard Brewka[1], Sylwia Polberg[2], and Stefan Woltran[2]

[1] Leipzig University, Informatics Institute, Postfach 100920, 04009 Leipzig, Germany, brewka@informatik.uni-leipzig.de

[2] Vienna University of Technology, Institute of Information Systems, Favoritenstraße 9–11, A-1040 Vienna, Austria, {polberg,woltran}@dbai.tuwien.ac.at

Abstract. The aim of this paper is to provide a short survey of some of the most popular abstract argumentation frameworks available today. We present the general idea of abstract argumentation, highlighting the role of abstract frameworks in the argumentation process and review the original Dung frameworks and their semantics. Then a discussion on generalizations of these frameworks is given, focusing on structures taking preferences and values into account and approaches where not only attack but also support relations can be modeled. Finally we review the concept of abstract dialectical frameworks, one of the most general systems for abstract argumentation providing a flexible, principled representation of arbitrary argument relations.

Keywords: Argumentation, conflict resolution, abstract argumentation frameworks

1 Introduction

Assume you have to make an important decision, for example concerning a new job you have been offered. The conclusion you have to reach will hardly be a simple deductive one. Most probably you will construct arguments in favor of the new job (e.g. better salary, better chances in the future) and arguments against it (further away, bigger workload and thus less time for your family). In the end the pro and con arguments will be weighed against each other and your decision determined.

Given that this kind of argument–based reasoning and decision making is ubiquitous, it is not surprising that argumentation itself has emerged as a scientific field. It studies how to model arguments and their relationships, with the ultimate goal to solve conflicts in the presence of diverging opinions. Thus argumentation has also been referred to as *reasoning tested by doubt*[3]. Argumentation has become a major focus of research in Artificial Intelligence (AI) over the last two decades [14, 16, 59]. It is strongly connected and highly beneficial to various other AI subfields, in particular knowledge representation, nonmonotonic reasoning, and multi–agent systems. It has

* This research has been supported by DFG (projects BR 1817/7-1 and FOR 1513), FWF (project I1102), and the Vienna PhD School of Informatics.

[3] As pointed out by Guillermo R. Simari the credits for this definition belong to David Zarefsky.

Accepted for publication by IEEE [22]

been successfully applied to legal reasoning, which makes use of argumentation principles to formulate legal cases as arguments [15]. Moreover, it has proved valuable in decision support for resolving conflicts between different opinions [4, 33] and in the context of dialogues and persuasion [48, 57]. Finally, argumentation techniques can be found in several expert systems from such different areas as medicine [40,45] or eGovernment [8].

Within argumentation one can distinguish two major lines of research: logic–based and abstract approaches. The former (see for instance [16]) takes the logical structure of arguments into account and defines notions like attack, undercut, defensibility etc. (see e.g. [43]) in terms of logical properties of the chosen argument structures. In contrast, abstract approaches - which will be the topic of this paper - consider arguments as atomic items, focusing entirely on the relations between them. Consequently, it is assumed that the arguments and the relevant relations between them have already been constructed, usually from a given background knowledge base. The obtained argumentation system is then evaluated on the abstract level, yielding possibly alternative sets of (abstract) arguments which may be collectively accepted. Finally, these results are interpreted in terms of the original knowledge base. This three–step creation, evaluation and interpretation process is known as the argumentation process or instantiation–based argumentation [24, 26]. The advantage of this method is that it provides a high degree of modularity, so that the way of solving a given problem is kept on an abstract level, detached from a particular representation in the modeling language used in the knowledge base. In this paper we will provide an overview of the most popular argumentation systems for this abstract level.

The most prominent abstract systems are due to Dung [34]. They come equipped with various types of semantics used for their evaluation. In a nutshell, Dung's frameworks (AFs for short) are directed graphs with the vertices being the abstract arguments and the directed edges corresponding to attacks between them. Conflicts are then resolved using appropriate semantics. The different semantics reflect the different intuitions about what can be considered reasonable, thus providing a suite of *calculi of opposition*. They produce acceptable subsets of the arguments, called extensions, that correspond to various positions one may take in the light of the available arguments. Crucial here are the concepts of conflict–freeness and admissibility. Based on them more advanced semantics have been defined, ranging from Dung's original stable, preferred and grounded semantics to the more recent semi–stable, ideal and cf2–semantics [9].

Recent studies [2, 24] have shown, however, that within the argumentation process the construction of proper argumentation frameworks can cause much more concern than expected. Attention must be given to avoid the risk of violating some natural rationality postulates in the overall instantiation-based argumentation process. Indeed, generating the right argumentation structures is the crucial step here for yielding reasonable — and, in particular, consistent — output.

In general, the more expressive the modeling languages become, the more involved the instantiation step grows. Consequently, the generated AFs may turn out to be far from natural. For instance, recent formal systems like ASPIC+ [51,58] and Carneades [42] provide various useful syntactical features, in particular a separation between strict and defeasible rules, different types of premises and proof standards, preferential information and the like. This makes the modeling language very expressive yet results in a rather complicated instantiation, and it is not always easy to see whether the instantiation "does the right thing".

However, our view is that the general idea underlying the abstract, instantiation–based approach to argumentation is still valid. The modularity of this method provides the flexibility needed in response to changes in the modeling languages. What, as we believe, the mentioned results suggest is that Dung's argumentation frameworks may not necessarily be the best target systems for the instantiation. Indeed, their expressive abilities are limited due to the fact that we have only a binary conflict at hand. This can make modeling e.g. collective [52] or supportive [28] relations unpleasant if not problematic.

Currently there are two research directions that aim to address such problems and to bridge the gap between the modeling languages and AFs. The first one, called meta–argumentation [17, 50], allows us to stay within the well established setting of Dung. However, it comes at the cost of auxiliary arguments which are required to represent relations other than attack. The second one focuses on extending AFs by equipping them with more expressive concepts to model the aforementioned situations, such as preferences or support relations. Compared to the meta–argumentation approach, the main challenge of the new frameworks is to correspondingly generalize the semantical concepts. They must not only fit the extended frameworks, but also remain intuitive and relatively compatible with the original structures (see [55] for a discussion). The primary objective of this paper is to give an overview of the currently available generalizations of Dung's framework. In the interest of space, we will keep our discussion on a rather informal level but try to raise awareness of the difficulties which arise in this vivid and interesting branch of research in the argumentation community.

This paper is structured as follows. We start with a theoretical background on Dung's systems in Section 2. Then we proceed to describing generalizations that focus on preference and value–based reasoning in Section 3 and ones that study different types of relations between arguments (see Section 4). Please note that we do not aim at a complete survey. Rather, we present what we consider interesting representatives of different kinds of generalizations. In Section 5 we present abstract dialectical frameworks, so far the most general enrichments of AFs. We conclude the paper with a short list of some generalizations we could not describe in detail and pointers for future work.

2 Dung's Argumentation Frameworks

In his seminal paper [34], Phan Minh Dung showed that it is possible to analyze acceptability of arguments in an abstract way, independently of where the arguments come from and how they are generated. Moreover, he aimed at representing different types of nonmonotonic approaches in a uniform setting. To this end, he introduced a surprisingly simple concept: abstract argumentation frameworks.

Definition 1. *A **Dung abstract argumentation framework** (AF) is a pair (A, R), where A is a set of arguments and $R \subseteq A \times A$ represents the **attack** relation. We say x attacks y iff $(x, y) \in R$.*

An AF is thus nothing but a directed graph with a specific intuitive interpretation of nodes and links. As there are no restrictions on the attack relation, cycles, self–attackers, and so on, are all allowed. Arguments do not have any particular structure and the precise conditions for their acceptance are defined by the semantics. In what follows we will present several such semantics, originally defined by Dung.

Let $F = (A, R)$ be an AF. Assume a rational agent accepts a subset S of A. What properties would we expect S to satisfy? First of all, it seems reasonable to require consistency, i.e. that the arguments in S do not attack each other. We say $S \subseteq A$ is *conflict–free* (in F) whenever there are no $x, y \in S$ such that $(x, y) \in R$. However, this is still insufficient. In real life we are often forced to counter the arguments of the opponents that have an opinion conflicting with ours. We say a set of arguments $S \subseteq A$ is *admissible* (in F) whenever S is conflict–free in F and S is able to *defend* itself against outside attacks. By this we mean the following: if there is an $x \in A$ and some $y \in S$ such that $(x, y) \in R$ (an element of A attacks an element of S), then S must contain some element z defending y, that is $(z, x) \in R$ for some $z \in S$.

Admissible sets are good candidates for the sets of arguments a rational agent may adopt. However, not all of them are interesting. The empty set, for instance, is admissible for each AF. Not accepting arguments without any reason can certainly not be viewed as rational. This is why Dung focuses on further notions. A set of arguments $S \subseteq A$ is *complete* iff it is admissible and already contains all arguments defended by S. In other words, each argument whose attackers are attacked by S must be in S. This notion can be further strengthened to *preferred extensions* which simply are the (subset) *maximal* admissible sets. A somewhat stronger notion than preferred extensions are *stable extensions*. Stable extensions are conflict–free subsets of A which attack each argument not in S: a conflict-free set S is a stable extension if for each $x \in A \setminus S$ there is $y \in S$ such that $(y, x) \in R$. Please note that contrary to the other notions discussed here, stable extensions are not guaranteed to exist.

Another interesting semantics is given by the *grounded extension*. It contains all and only arguments whose defense can be traced back to any of the unattacked arguments. Every framework has precisely one such set, even if it is empty. The grounded extension S of an AF can be generated iteratively as follows: starting with the empty set, we first include in the set S all those arguments which are not attacked at all. We then remove them from the framework along with all arguments T attacked by S (together with all attacks between arguments from S and T). We continue like this, adding to S in each step arguments unattacked in the reduced AF and remove them, until we reach a fixpoint, that is, until no further unattacked arguments can be found this way.

Example 1. Here is a short example. Let $F_1 = (A, R)$ with $A = \{a_1, a_2, a_3, a_4, a_5\}$ and $R = \{(a_1, a_2), (a_3, a_2), (a_3, a_4), (a_4, a_3), (a_4, a_5), (a_5, a_5)\}$ (see Figure 1). The conflict-free sets of F_1 are $\{a_1, a_3\}$, $\{a_1, a_4\}$, $\{a_2, a_4\}$ and all their subsets. The admissible sets of F_1 are $\emptyset$, $\{a_1\}$, $\{a_3\}$, $\{a_4\}$ as well as $\{a_1, a_3\}$ and $\{a_1, a_4\}$. Note that, for instance, $\{a_2, a_4\}$ is not admissible since a_2 is not defended by that set against its attacker a_1 (it is however defended against a_3). We have three complete extensions of F_1: $\{a_1\}$, $\{a_1, a_3\}$ and $\{a_1, a_4\}$. The latter two are the preferred extensions, however, only $\{a_1, a_4\}$ is stable. Finally, the grounded extension consists of the single argument a_1.

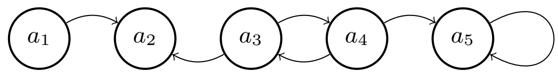

Fig. 1. Sample Dung framework.

Please note that various further alternative definitions of AF semantics have been defined, see [9]. Most prominent are the labeling–based approaches [25] that instead of generating sets of accepted arguments, provide us with three–valued (truth, false, undec) interpretations. However, for the purposes of this paper the notions we introduced are sufficient. Moreover, we do not aim here to discuss computational properties such as complexity or expressiveness. We refer the interested reader to [11, 36, 37]. For an overview of systems for abstract argumentation, see [29].

3 Preferences and Values

In decision making scenarios we are often faced with pro and con arguments, and our decisions are based on preferences among these arguments. For this reason it is quite natural to apply preference handling techniques in argumentation. In fact, this has been done even before abstract argumentation as such has emerged, see for instance [60], where the strength of an argument is measured in terms of the specificity of the underlying information [56]. Including the preference information in argumentation frameworks not only allows us to model the problem more accurately, it also reduces the number of extensions we may obtain.

Generalizations of AFs which include preferences were first introduced in [3,5]. The so–called preference based argumentation frameworks (PAFs) are defined as follows:

Definition 2. *A **preference based argumentation framework** (PAF) is a tuple $(A, R, \geq)$ where A is a set of **arguments**, $R \subseteq A \times A$ is the **attack** relation and $\geq\ \subseteq A \times A$ is a (partial or total) preorder representing **preference**.*[4]

The evaluation of arguments is then based on a simple idea: whenever a strictly less preferred argument a attacks a more preferred argument b (a is strictly less preferred than b, denoted $b > a$, if $b \geq a$ and not $a \geq b$), then the attack $(a, b) \in R$ is simply disregarded. This means that the given PAF $PF = (A, R, \geq)$ is used to generate a standard AF $F = (A, R')$ where $R' = R \setminus \{(a, b) \mid b > a\}$. This AF is then used to define the semantics of the PAF.

Bench-Capon's value based argumentation frameworks (VAFs) [12] are based on similar ideas. However, here it is assumed that arguments promote specific *values*, and the preferences are among these values rather than between the arguments themselves.

[4] Note that depending on the approach, the values/preferences of arguments might form different types of ordering. Although PAFs use a preorder, the mentioned specificity relation is not necessarily a preorder, see [65].

Definition 3. *A **value–based argumentation framework** (VAF) is a tuple* $(A, R, V, val, valprefs)$ *where* A *is a set of **arguments**,* $R \subseteq A \times A$ *is the **attack** relation and* V *is a nonempty set of **values**.* $val : A \to V$ *is a function mapping arguments to values and* $valprefs \subseteq V \times V$ *is a (transitive, irreflexive and asymmetric)* [5] ***preference** relation.*

Again, the evaluation of a VAF is based on the generation of an AF. Here attacks are disregarded whenever the attacked argument promotes a more preferred value than the attacker. VAFs were further generalized to include different audiences which may disagree about the preferences among values. A comparison between PAFs and VAFs can be found in [18, 46].

In many argumentation contexts preferences or values themselves are a matter of debate. It is thus useful to have frameworks in which it is not only possible to reason and argue *with* preferences but also *about* preferences. This observation led to the development of extended argumentation frameworks (EAFs) [49]. Here, reasoning about preferences is modeled by allowing an argument not only to attack other arguments but also other attacks. Like this, an attack (a, b) can be dynamically disabled if, say, an argument is accepted which intuitively expresses that b should be preferred to a.

Definition 4. *An **extended argumentation framework** (EAF) is a tuple* (A, R, D) *where* A *is a set of **arguments**,* $R \subseteq A \times A$ *represents the **attack** relation and* $D \subseteq A \times R$ *is the **defence attack** relation for handling preferences. Moreover, if* $(x, (y, z))$ *and* $(x', (z, y))$ *are in* D*, then* (x, x') *and* (x', x) *are in* R*.*

The main idea behind EAFs is that preferences can be based on arguments themselves and as such they generate a more advanced conflict relation where other attacks can be overridden. The constraint included in the definition states that arguments representing opposite preferences should also be mutually in conflict in the framework. Although the definitions substantially differ from the ones in PAFs and VAFs, the intuition remains similar. Instead of attacks, we focus on *defeats* with respect to some set S: we say that x *defeats*$_S$ y if $(x, y) \in R$ and there is no z in S such that $(z, (x, y)) \in D$. Conflict–free extensions then disregard attacks that are not defeats, assuming that the attack is not symmetric. Intuitively, admissibility is also limited to the successful attacks (i.e., defeats). Defense is now based on defeats, i.e. given a set S, for every y such that y *defeats*$_S$ x where $x \in S$, there exists an argument $z \in S$ s.t. z *defeats*$_S$ y. What is unique in EAFs is that this semantics is later strengthened by ensuring a particular type of defense for the (z, y) defeat itself, referred to as *reinstatement*. Based on this, all further semantics follow naturally and generalize the original ones by Dung; see [49] for further details.

The initial idea behind evaluating the preference (value–based) frameworks was to disregard a conflict if the attacked argument was stronger than its attacker (or if the attack itself was attacked in case of EAFs) and focus on the remaining relations. This treatment can result in extensions that do not follow the traditional intuition behind conflict–freeness and depending on how we understand preferences, they can be seen as intended [49] or not [5]. In the latter case, the cited work also proposes a new approach towards semantics for PAFs that is meant to fix such issues and as such presents

[5] We keep the original definition here, even though some of the properties of $valprefs$ imply one another.

a different type of reasoning than in the Dung's frameworks. Instead of being driven by acceptability, it makes use of a so-called dominance relation to establish whether a set of arguments (candidate for an extension) is better than another. The initial relation is described in terms of postulates that incorporate conflict–freeness and the behavior of attacks in the PAF setting. It is then further adjusted depending on the (Dung–style) semantics that one wants to generalize. The extensions are represented by the maximal elements of the dominance relation. While preferences assign different levels of importance to arguments, the relation between arguments remains purely conflicting. In the next section, we take a look at systems which relax this restriction.

4 Relationships Beyond Attack

Dung's frameworks consider only a single relationship among arguments, namely attack. In various contexts it is natural to go beyond conflict. In particular, the ability to model various notions of support appears useful. This observation has led to the development of so-called *bipolar* argumentation frameworks [27, 28] where both of the relations are modeled.

Definition 5. *A* ***bipolar argumentation framework*** *(BAF) is a tuple* (A, R_{att}, R_{sup}) *where* A *is a set of* ***arguments****,* $R_{att} \subseteq A \times A$ *represents the* ***attack*** *relation and* $R_{sup} \subseteq A \times A$ *represents the* ***support*** *relation. We require that* $R_{att} \cap R_{sup} = \emptyset$*.*

Including a new type of relation requires a careful adaptation of the existing semantics. The combination of attack and support leads to indirect attacks, referred to as complex attacks in [28]. For instance, there is a *supported attack* from a_1 to b if there is a sequence of support links from a_1 to a_n and an attack from a_n to b. There is a *secondary attack* from b to a_n if there is a sequence of support links from a_1 to a_n and an attack from b to a_1. A *mediated attack* from b to a_1 takes place if there is a sequence of support links a_1 to a_n and an attack from b to a_n. For the generalization of Dung–style semantics to BAFs these indirect, complex notions of attack then need to be taken into account adequately. There are various ways of doing this, and what is adequate depends on the specific interpretation of the support.

Example 2. We will now show some of the complex attacks. Let $BF_1 = (A, R_{att}, R_{sup})$ with $A = \{a_1, a_2, a_3, a_4, a_5, a_6\}$, $R_{att} = \{(a_3, a_4), (a_6, a_5)\}$ and $R_{sup} = \{(a_1, a_2),$ (a_2, a_3), $(a_4, a_5)\}$ (see Figure 2). The normal attacks are depicted with normal lines, supports with dashed and complex attacks with dotted.

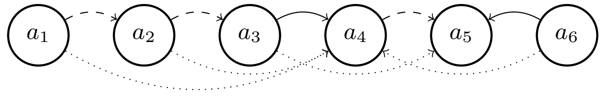

Fig. 2. Sample BAF.

Here, a_2 supports a_3 which attacks a_4, creating a supported attack from a_2 to a_4. Another supported attack on a_4 comes from a_1, a supporter of a_2. There is a secondary attack from a_3 to a_5 and a mediated one from a_6 to a_4.

In [28] an analysis of three different types of support is carried out: *deductive* support, *necessary* support and *evidential* support. The former two are quite strong notions and, as shown in that paper, dual to each other. If a deductively supports b, then b must be accepted whenever a is accepted. Necessary support between a and b means that accepting a is a necessary (but not necessarily sufficient) precondition for accepting b.

The notion of evidential support was first developed in evidential argumentation systems (EASs) [54]. This approach builds on a generalization of Dung AFs in [52] where sets of arguments rather than single arguments may be needed to attack another argument. In addition, the frameworks introduce a distinction between so–called *prima facie* arguments and ordinary arguments. The former can be accepted without further requirements. The support of the latter needs to be rooted in such special arguments to be considered valid. Otherwise it not only cannot be accepted, it is not even considered a valid attacker. Thus the resulting semantics are stronger versions of the Dung semantics that impose a type of a grounding on the arguments.

We refer the reader to the original papers for further details and to [30] for a survey on various types of support in argumentation. To summarize this section, research on bipolar argumentation frameworks has certainly demonstrated that going beyond the attack relation is interesting and useful. On the other hand, it is apparent from the literature that there is no such thing as a single interpretation of support. One can easily think of further interpretations, not among the three we briefly discussed. For instance, an argument may strengthen another one without guaranteeing its acceptance as required for deductive support. One can also think of situations where different notions need to be combined in flexible ways. Finally, in all bipolar frameworks so far, there is the hidden assumption that conflict is stronger than support. This means that no matter the support an argument receives, it still has to be defended from incoming attacks. It thus appears useful to have frameworks that — rather than being built on a fixed interpretation of support — make it possible to specify exactly, for each argument, in what way support and attack interact. This is exactly the functionality provided by abstract dialectical frameworks which we will discuss in the next section.

5 Abstract Dialectical Frameworks

The abstract dialectical frameworks (ADFs) have been proposed by Brewka and Woltran [23] as a generalization of Dung-style AFs. The idea behind them was to allow not only abstract arguments, but also highly flexible and abstract relations. This is achieved by adding to each argument a a specific acceptance condition C_a. More formally, an abstract dialectical framework is a directed graph whose nodes represent arguments, statements or positions which can be accepted or not. The links represent dependencies: the status of a node s only depends on the status of its parents (denoted $par(s)$), that is, the nodes with a direct link to s. In addition, each node s has an associated acceptance condition C_s specifying the exact conditions under which s is accepted. C_s is a function assigning to each subset of $par(s)$ one of the truth values $\mathtt{t}$, $\mathtt{f}$. Intuitively, if for some

$R \subseteq par(s)$ we have $C_s(R) = \mathbf{t}$, then s will be accepted provided the nodes in R are accepted and those in $par(s) \setminus R$ are not accepted.

Definition 6. *An **abstract dialectical framework** is a tuple $D = (S, L, C)$ where S is a set of **statements** (positions, nodes), $L \subseteq S \times S$ is a set of **links** and $C = \{C_s\}_{s \in S}$ is a set of total functions $C_s : 2^{par(s)} \to \{\mathbf{t}, \mathbf{f}\}$, one for each statement s. C_s is called the **acceptance condition** of s.*

In many cases it is convenient to represent the acceptance conditions as propositional formulas, as described in [38]. For this reason we frequently use a logical representation of ADFs (S, L, C) where C is a collection $\{\varphi_s\}_{s \in S}$ of formulas expressing the Boolean functions from Definition 6.

Acceptance conditions can specify arbitrary relationships between arguments and their parents, thus allowing us to model complex interactions. A case where an argument a can only be successfully attacked if two attacking arguments b and c are jointly accepted can be easily expressed with a condition $\varphi_a = \neg b \vee \neg c$. Dung's standard AFs can be recovered as the special case where the acceptance condition of an argument, say a, is defined as the formula $\neg c_1 \wedge \cdots \wedge \neg c_n$ where $c_1, \ldots, c_n$ are all arguments attacking a.

Example 3. A representation of the Dung framework from Figure 1 in terms of an ADF is depicted in Figure 3 and formally given as $ADF_1 = (S, L, C)$ with $S = \{a_1, a_2, a_3, a_4, a_5\}$, $L = \{(a_1, a_2), (a_3, a_2), (a_3, a_4), (a_4, a_3), (a_4, a_5), (a_5, a_5)\}$ and $C = \{a_1 : \top,\ a_2 : \neg a_1 \wedge \neg a_3,\ a_3 : \neg a_4,\ a_4 : \neg a_3,\ a_5 : \neg a_4 \wedge \neg a_5\}$.

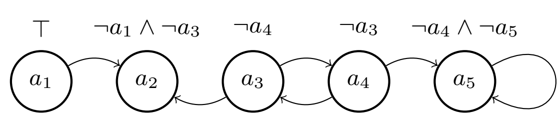

Fig. 3. Dung–style ADF

As the definition of acceptance conditions suggests, ADFs can naturally express much more. Consider the framework depicted in Figure 4: $ADF_2 = (S, L, C)$ with $S = \{a_1, a_2, a_3, a_4, a_5, a_6\}$, $L = \{(a_1, a_2), (a_3, a_2), (a_3, a_4), (a_4, a_3), (a_5, a_4), (a_6, a_5), (a_6, a_6)\}$ and $C = \{a_1 : \top, a_2 : a_1 \vee a_3, a_3 : \neg a_4, a_4 : \neg a_3 \vee \neg a_5, a_5 : a_6, a_6 : \neg a_6\}$.

The relation between a_2 and its parents is an example of a group support similar to the one presented in EAS [54]. This means that accepting either a_1 or a_3 is sufficient for a_2, we do not have to commit to both. Attack by a set of arguments, i.e. where more arguments are needed for a successful attack [52] is carried out by a_3 and a_5 on a_4. Finally a_5 is (necessary) supported by a_6.

The semantics of ADFs in the original paper [23] have later been generalized and further developed in various directions. The initial approach and the one described

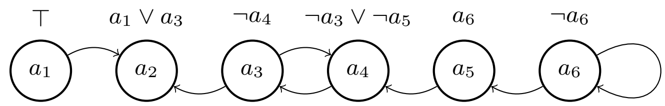

Fig. 4. Sample ADF

in [55] fall into the extension–based category, while the works in [20, 61] are labeling–based. In the first approach, a conflict–free extension is a set of arguments having their acceptance condition satisfied. Admissibility generalizes the original intuition from Dung by making sure that the extension has the power to discard undesired arguments. In the labeling–based approach, the definition of ADF semantics is based on the notion of a model. A (two-valued) interpretation v – a mapping from arguments to the truth values $\mathbf{t}$ and $\mathbf{f}$ – is a model of an ADF whenever v maps exactly those statements to $\mathbf{t}$ whose acceptance conditions are satisfied under v. The definition of grounded, complete, preferred and stable semantics is then derived from an analysis in terms of three–valued interpretations where an additional truth value $\mathbf{u}$ (undefined) is used. A key notion is the following *consensus operator*: for an ADF D and a three-valued interpretation v, the operator Γ_D returns the (three-valued) interpretation $\Gamma_D(v)$ which assigns to each statement s the consensus truth value for its acceptance formula φ_s, where the consensus takes into account all possible two-valued interpretations w that extend the input valuation v. The relevant semantical notions are derived from this operator. For instance, the grounded model of D is defined as the least fixpoint of Γ_D. For further technical details we refer the reader to [20, 61]. A comparison between the extension and labeling–based approaches can be found in [55]. However, we want to emphasize that the ADF semantics are actually proper generalizations of the original Dung semantics in the sense that they treat ADFs corresponding to AFs in exactly the same way as defined by Dung.

ADFs also provide a new handle on the treatment of preferences [1, 6] and values, respectively audiences [13]. As shown in [23], preferences on links between statements can directly be used in the definition of acceptance conditions. In [20], a treatment of preferences among arguments in the style of PAFs has been introduced. A prioritized ADF (PADF) consists of a prioritized set of arguments, a set of support links and a set of attack links. A PADF is then compiled to a standard ADF. The approach is shown to be a proper generalization of PAFs. Similar generalizations of VAFs are straightforward.

The same paper also proposes a new approach to argumentation with dynamic preferences which, rather than being given in advance, are a matter of debate themselves. In a nutshell, dynamic preferences are handled as follows. We first guess a (stable, preferred, grounded) extension M. We assume some nodes in M carry preference information. We extract this information and check whether M can be reconstructed under this preference information, thus verifying that the preferences represented in the model itself were taken into account adequately to construct the model.

Brewka and Woltran have shown how legal proof standards - which play an essential role in legal reasoning - can be modeled using ADFs [23]. Moreover, in [21]

ADFs were used to provide a reconstruction of the much-cited Carneades system [42]. The reconstruction not only puts Carneades on a safe formal ground. Even more importantly, it allows the somewhat unrealistic restriction of the original system to acyclic argumentation scenarios to be relaxed. This shows the potential of ADFs as systems for generalizing not just abstract frameworks, but also more logic–based approaches.

6 Concluding Remarks

In this paper we have described the field of abstract argumentation and provided an overview of the currently available frameworks that extend Dung's initial system by incorporating preferences and relations beyond attack. In what follows we give a few more pointers to the literature; a more detailed discussion of these works had to be omitted due to space constraints. Concerning preferences, such further works include audience specific argumentation frameworks in which arguments can promote multiple values [46], uniform argumentation frameworks [7], and multi-contextual preference based argumentation frameworks [18]. Other generalizations incorporate probabilities [44, 63] and certain forms of weights [32, 35, 47]. Concerning the frameworks focusing on generalizing the relations between arguments, we mention here three further representatives. First, the SETAF [52] approach formalizes the concept of *sets* of attacking arguments. Second, in the AFRA (argumentation frameworks with recursive attacks) approach [10], not only arguments but also attacks can be attacked (similar to the EAF approach discussed in Section 3 but without any limitations and motivated from a different angle). Finally, there exists another type of bipolar structures – abstract frameworks with necessities – that focus on necessary support [53].

As we have seen ADFs go even further trying to capture many of the aforementioned generalizations at once. Closely related to ADFs are constrained argumentation frameworks (CAFs) [31], Weydert's hyperframeworks [64] and equational argumentation networks (EANs) [41]. The main difference between CAFs and ADFs lies in the fact that the former use the relations to filter the extensions rather than exploiting them during the extension computation, as it is done in ADFs. The motivations behind the hyperframeworks and EANs are close to the one behind ADFs. However, they leave a lot of detail to the user and lack the combination of simplicity and flexibility that ADFs provide.

Coming back to the entire argumentation process as described in the introduction, let us mention that research on instantiations in the context of ADFs [62] has recently emerged. Also efficient implementations [39] are now available making ADFs a valuable tool to experiment with and also to better understand "weaker" generalizations of Dung's AFs. ADFs, as we believe, do provide a highly useful interface between modeling languages on one hand and AF-based implementation techniques on the other. In other words, ADFs act as argumentation "middleware" bridging the gap between highly complex argumentation applications and the core abstract frameworks of Dung (as higher order programming languages are used to implement algorithms but are by no means intended to replace machine code). This view is also backed up by polynomial compilations from ADFs to AFs which preserve at least some of the ADF semantics [19]. As in the meta-argumentation approach, additional arguments are needed for this purpose. Even if one is willing to stick to AF-based implementation techniques, ADFs thus provide a useful interface between the more complex non–abstact modeling languages on one hand and the purely Dung layer on the other.

As to future work, it seems fruitful to continue the recent research on ADF semantics [20, 55] with a particular focus on the limits of expressiveness. Another interesting, so far unexplored aspect is to switch the language of the acceptance conditions. So far, these conditions describe relations between arguments in terms of *standard propositional logic*. Moving to temporal, modal, or even nonmonotonic logic for the interpretation of the acceptance conditions offers new exciting research perspectives in the area of formal argumentation.

References

1. Leila Amgoud. Using preferences to select acceptable arguments. In *Proc. ECAI*, pages 43–44. John Wiley and Sons, 1998.
2. Leila Amgoud and Philippe Besnard. Bridging the gap between abstract argumentation systems and logic. In *Proc. SUM*, volume 5785 of *LNCS*, pages 12–27. Springer, 2009.
3. Leila Amgoud and Claudette Cayrol. A reasoning model based on the production of acceptable arguments. *Ann. Math. Artif. Intell*, 34(1-3):197–215, 2002.
4. Leila Amgoud and Henri Prade. Using arguments for making and explaining decisions. *Artif. Intell.*, 173(3-4):413–436, 2009.
5. Leila Amgoud and Srdjan Vesic. A new approach for preference-based argumentation frameworks. *Ann. Math. Artif. Intell*, 63:149–183, 2011.
6. Leila Amgoud and Srdjan Vesic. Two roles of preferences in argumentation frameworks. In *Proc. ECSQARU*, volume 6717 of *LNCS*, pages 86–97. Springer, 2011.
7. Katie Atkinson, Trevor J. M. Bench-Capon, and Paul E. Dunne. Uniform argumentation frameworks. In *Proc. COMMA*, volume 245 of *Frontiers in Artificial Intelligence and Applications*, pages 165–176. IOS Press, 2012.
8. Katie Atkinson, Trevor J. M. Bench-Capon, and Peter McBurney. PARMENIDES: Facilitating deliberation in democracies. *Artif. Intell. Law*, 14(4):261–275, 2006.
9. Pietro Baroni, Martin Caminada, and Massimiliano Giacomin. An introduction to argumentation semantics. *Knowledge Eng. Review*, 26(4):365–410, 2011.
10. Pietro Baroni, Federico Cerutti, Massimiliano Giacomin, and Giovanni Guida. AFRA: Argumentation framework with recursive attacks. *Int. J. Approx. Reasoning*, 52(1):19–37, 2011.
11. Pietro Baroni and Massimiliano Giacomin. On principle-based evaluation of extension-based argumentation semantics. *Artif. Intell.*, 171(10-15):675–700, 2007.
12. Trevor J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *J. Log. Comput.*, 13(3):429–448, 2003.
13. Trevor J. M. Bench-Capon. Try to see it my way: Modelling persuasion in legal discourse. *Artif. Intell. Law*, 11(4):271–287, 2003.
14. Trevor J. M. Bench-Capon and Paul E. Dunne. Argumentation in artificial intelligence. *Artif. Intell.*, 171(10-15):619–641, 2007.
15. Trevor J. M. Bench-Capon, Henry Prakken, and Giovanni Sartor. Argumentation in legal reasoning. In Guillermo Simari and Iyad Rahwan, editors, *Argumentation in Artificial Intelligence*, pages 363–382. Springer, 2009.
16. Philippe Besnard and Anthony Hunter. *Elements of Argumentation*. MIT Press, 2008.
17. Guido Boella, Dov M. Gabbay, Leendert van der Torre, and Serena Villata. Meta-argumentation modelling I: Methodology and techniques. *Studia Logica*, 93(2-3):297–355, 2009.

18. Jean-Rémi Bourguet, Leila Amgoud, and Rallou Thomopoulos. Towards a unified model of preference-based argumentation. In *Proc. FoIKS*, volume 5956 of *LNCS*, pages 326–344. Springer, 2010.
19. Gerhard Brewka, Paul E. Dunne, and Stefan Woltran. Relating the semantics of abstract dialectical frameworks and standard AFs. In *Proc. IJCAI*, pages 780–785. IJCAI/AAAI, 2011.
20. Gerhard Brewka, Stefan Ellmauthaler, Hannes Strass, Johannes Peter Wallner, and Stefan Woltran. Abstract dialectical frameworks revisited. In *Proc. IJCAI*, pages 803–809. IJCAI/AAAI, 2013.
21. Gerhard Brewka and Thomas F. Gordon. Carneades and abstract dialectical frameworks: A reconstruction. In *Proc. COMMA*, volume 216 of *Frontiers in Artificial Intelligence and Applications*, pages 3–12. IOS Press, 2010.
22. Gerhard Brewka, Sylwia Polberg, and Stefan Woltran. Generalizations of Dung frameworks and their role in formal argumentation. *Intelligent Systems, IEEE*, PP(99), 2013. In Press.
23. Gerhard Brewka and Stefan Woltran. Abstract dialectical frameworks. In *Proc. KR*, pages 102–111. AAAI Press, 2010.
24. Martin Caminada and Leila Amgoud. On the evaluation of argumentation formalisms. *Artif. Intell.*, 171(5-6):286–310, 2007.
25. Martin Caminada and Dov M. Gabbay. A logical account of formal argumentation. *Studia Logica*, 93(2):109–145, 2009.
26. Martin Caminada and Yining Wu. On the limitations of abstract argumentation. In *Proc. BNAIC*, pages 59–66, 2011.
27. Claudette Cayrol and Marie-Christine Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In *Proc. ECSQARU*, volume 3571 of *LNCS*, pages 378–389. Springer, 2005.
28. Claudette Cayrol and Marie-Christine Lagasquie-Schiex. Bipolarity in argumentation graphs: Towards a better understanding. *Int. J. Approx. Reasoning*, 54(7):876–899, 2013.
29. Günther Charwat, Wolfgang Dvořák, Sarah A. Gaggl, Johannes P. Wallner, and Stefan Woltran. Implementing Abstract Argumentation – A Survey. Technical Report DBAI-TR-2013-82, Vienna University of Technology, 2013.
30. Andrea Cohen, Sebastian Gottifredi, Alejandro J. García, and Guillermo R. Simari. A survey of different approaches to support in argumentation systems. *The Knowledge Engineering Review*, FirstView:1–38, 2013.
31. Sylvie Coste-Marquis, Caroline Devred, and Pierre Marquis. Constrained argumentation frameworks. In *Proc. KR*, pages 112–122. AAAI Press, 2006.
32. Sylvie Coste-Marquis, Sébastien Konieczny, Pierre Marquis, and Mohand Akli Ouali. Weighted attacks in argumentation frameworks. In *Proc. KR*, pages 593–597. AAAI Press, 2012.
33. Yannis Dimopoulos, Pavlos Moraitis, and Leila Amgoud. Extending argumentation to make good decisions. In *Proc. ADT*, volume 5783 of *LNCS*, pages 225–236. Springer, 2009.
34. Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.*, 77:321–357, 1995.
35. Paul E. Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artif. Intell.*, 175(2):457–486, 2011.
36. Wolfgang Dvořák. *Computational Aspects of Abstract Argumentation*. PhD thesis, Vienna University of Technology, 2012.
37. Wolfgang Dvořák and Stefan Woltran. On the intertranslatability of argumentation semantics. *J. Artif. Int. Res.*, 41(2):445–475, 2011.
38. Stefan Ellmauthaler. Abstract dialectical frameworks: properties, complexity, and implementation. Master's thesis, Vienna University of Technology, 2012.

39. Stefan Ellmauthaler and Hannes Strass. The DIAMOND system for argumentation: Preliminary report. In Michael Fink and Yuliya Lierler, editors, *Proc. ASPOCP*, 2013.
40. John Fox, David Glasspool, Vivek Patkar, Mark Austin, Liz Black, Matthew South, Dave Robertson, and Charles Vincent. Delivering clinical decision support services: There is nothing as practical as a good theory. *Journal of Biomedical Informatics*, 43(5):831–843, 2010.
41. Dov M. Gabbay. Introducing equational semantics for argumentation networks. In *Proc. ECSQARU*, volume 6717 of *LNCS*, pages 19–35. Springer, 2011.
42. Thomas F. Gordon, Henry Prakken, and Douglas Walton. The Carneades model of argument and burden of proof. *Artif. Intell.*, 171(10-15):875–896, 2007.
43. Nikos Gorogiannis and Anthony Hunter. Instantiating abstract argumentation with classical logic arguments: Postulates and properties. *Artif. Intell.*, 175(9-10):1479–1497, 2011.
44. Anthony Hunter. A probabilistic approach to modelling uncertain logical arguments. *International Journal of Approximate Reasoning*, 54(1):47 – 81, 2013.
45. Anthony Hunter and Matthew Williams. Aggregating evidence about the positive and negative effects of treatments. *Artificial Intelligence in Medicine*, 56(3):173–190, 2012.
46. Souhila Kaci and Leendert van der Torre. Preference-based argumentation: Arguments supporting multiple values. *Int. J. Approx. Reasoning*, 48(3):730–751, 2008.
47. Diego C. Martínez, Alejandro Javier García, and Guillermo Ricardo Simari. An abstract argumentation framework with varied-strength attacks. In *Proc. KR*, pages 135–144. AAAI Press, 2008.
48. Peter McBurney and Simon Parsons. Dialogue games for agent argumentation. In Guillermo Simari and Iyad Rahwan, editors, *Argumentation in Artificial Intelligence*, pages 261–280. Springer, 2009.
49. Sanjay Modgil. Reasoning about preferences in argumentation frameworks. *Artif. Intell.*, 173(9-10):901–934, 2009.
50. Sanjay Modgil and Trevor J. M. Bench-Capon. Metalevel argumentation. *J. Log. Comput.*, 21(6):959–1003, 2011.
51. Sanjay Modgil and Henry Prakken. A general account of argumentation with preferences. *Artif. Intell.*, 195:361–397, 2013.
52. Søren Nielsen and Simon Parsons. A generalization of Dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In *Proc. ArgMAS*, volume 4766 of *LNCS*, pages 54–73. Springer, 2007.
53. Farid Nouioua and Vincent Risch. Argumentation frameworks with necessities. In *Proc. SUM*, volume 6929 of *LNCS*, pages 163–176. Springer, 2011.
54. Nir Oren and Timothy J. Norman. Semantics for evidence-based argumentation. In *Proc. COMMA*, volume 172 of *Frontiers in Artificial Intelligence and Applications*, pages 276–284. IOS Press, 2008.
55. Sylwia Polberg, Johannes Peter Wallner, and Stefan Woltran. Admissibility in the abstract dialectical framework. In *Proc. CLIMA*, volume 8143 of *LNCS*, pages 102–118. Springer, 2013.
56. David L. Poole. On the comparison of theories: Preferring the most specific explanation. In *Proc. IJCAI*, pages 144–147. Morgan Kaufmann, 1985.
57. Henry Prakken. Models of persuasion dialogue. In Guillermo Simari and Iyad Rahwan, editors, *Argumentation in Artificial Intelligence*, pages 281–301. Springer, 2009.
58. Henry Prakken. An abstract framework for argumentation with structured arguments. *Argument & Computation*, 1(2):93–124, 2010.
59. Iyad Rahwan and Guillermo R. Simari. *Argumentation in Artificial Intelligence*. Springer, 1st edition, 2009.
60. Guillermo R. Simari and Ronald P. Loui. A mathematical treatment of defeasible reasoning and its implementation. *Artif. Intell.*, 53(23):125 – 157, 1992.

61. Hannes Strass. Approximating operators and semantics for abstract dialectical frameworks. *Artificial Intelligence*, 205:39 – 70, 2013.
62. Hannes Strass. Instantiating knowledge bases in abstract dialectical frameworks. In *Proc. CLIMA*, volume 8143 of *LNCS*, pages 86–101. Springer, 2013.
63. Matthias Thimm. A probabilistic semantics for abstract argumentation. In *Proc. ECAI*, volume 242 of *Frontiers in Artificial Intelligence and Applications*, pages 750–755. IOS Press, 2012.
64. Emil Weydert. Semi-stable extensions for infinite frameworks. In *Proc. BNAIC*, pages 336–343, 2011.
65. Claus-Peter Wirth and Frieder Stolzenburg. David poole's specifity revised. *CoRR*, abs/1308.4943, 2013.