Semantics for Evidence-Based Argumentation

Nir Oren [a,1] , Timothy J. Norman [b]

[a] *King's College London, London, UK*
[b] *University of Aberdeen, Aberdeen, Scotland*

Abstract. The identification of consistent sets of arguments is one of the most important concerns in the development of computational models of argument. Such *extensions* drive the defeasible reasoning process. In this paper we present a novel framework that offers a semantics for argumentation-based reasoning grounded in an intuitive notion of evidence. In building this model of evidence-based argumentation, we first show how evidential support for and attack against one argument by multiple arguments can be represented. Secondly, we propose a new set of extensions, referred to as the evidential preferred, e-stable, and e-grounded extensions. These extensions are similar to those in Dung's original model, but are modified to cater for the notion of evidence. Our approach involves defining a new notion of attack, and showing how acceptability and admissibility follow on from it. These extensions capture the notion that an argument cannot be accepted unless it is supported by evidence, and will therefore increase the utility of argumentation as part of a reasoning system.

Keywords. Abstract argumentation frameworks, evidence-based reasoning

1. Introduction

Dung's seminal abstract argumentation framework [5] consists of a set of arguments and a binary relation between pairs of arguments. This relation represents the concept of one argument attacking another, and is referred to as the *attacks* relation. For example, consider the following set of arguments: (a) The bridge should be built on solid ground; (b) The bridge should be built at point x, where soft ground exists; (c) The bridge should be built out of wood, not concrete;(d) The bridge should be built out of concrete, not wood; and (e) Building at point x means that concrete should be used, not wood.

These arguments could be represented by the argument framework with arguments $\{a, b, c, d, e\}$ and attack relation $\{(a, b), (b, a), (c, d), (d, c), (e, c)\}$ (i.e. the argument that the bridge should be built on solid ground, a, *attacks* the argument that bridge should be built at point x, b, etc.).

Different modes of reasoning are represented by different *extensions*, allowing one to determine what sets arguments a sceptical, or credulous reasoner would deem acceptable. For example, Dung's preferred extension, representing a credulous reasoner, would

---

[1]Correspondence to: Nir Oren, Department of Computer Science, King's College London, Strand, WC2R 2LS, United Kingdom. Tel.: +44 (0)20 7848 1631 ; Fax: +44 (0)20 7848 2851 ; E-mail: nir.oren@kcl.ac.uk

consist of the arguments $\{d, a, e\}$ and $\{d, b, e\}$; these arguments are, in some sense, "consistent" to a credulous reasoner. A sceptical reasoner on the other hand, may only find $\{d, e\}$ consistent. Deciding what type of extension to use is application dependent. If agents were negotiating, for example, the presence of argument d in all extensions would mean that they all agree to its relevance. The extensions of an argumentation system are often referred to as the system's argumentation semantics. The abstract nature of argument in these frameworks allows them to be applied to many domains through instantiation with different underlying logics. While powerful and elegant, Dung's framework has a number of shortcomings, and various refinements have been proposed.

Looking at the example above, it appears not only as if argument e attacks argument c, but that, in some sense, it supports argument d. This notion of support was recently formalised using bipolar argumentation frameworks (BAFs) [1,2,4].

Another important enhancement of Dung's original work involves allowing multiple arguments to attack one another [6]. For example, consider the two arguments: (f) Financial considerations override any other considerations; and (g) Financial considerations mean that the bridge should be built at point y. It is clear that alone, neither argument is able to attack argument b, but that both arguments, when considered together, combine to create a valid attack against it[1].

In this paper, we draw inspiration from both BAFs and the work of Nielsen and Parsons, with an eye to providing a rich framework that captures our intuitions of evidential reasoning. Evidential reasoning involves determining which arguments are applicable based on some evidence. Such reasoning arises in many fields of human endeavour, and the formalisation of argument for such domains will, we believe, lead to more powerful reasoning mechanisms. Our first extension involves allowing for the use of multiple arguments not only when attacking an argument, but also in providing evidential support for an argument (in the example above, b and e together support argument d). Second, we introduce the notion of support into our framework and present a number of new extensions that allow for an intuitive representation of evidential reasoning. Informally, these extensions only accept arguments that are backed up by a chain of evidence. To allow for this, we redefine the concept of attack so that only supported arguments may attack each other. We then show how supported arguments may be directly, or indirectly attacked.

In the next section, we discuss some criticisms that have been levelled at abstract argumentation frameworks that include the notion of support. The following section details our model of evidence-based argumentation, including our notions of acceptability and admissibility, and evidential preferred extensions, evidential stable extensions and evidential grounded extensions. We then use an example in discussing the merits of the model presented in this paper and explore related and future research.

2. The Importance of Evidential Support

Introducing the notion of support between arguments within abstract argumentation frameworks has not been met with universal agreement. The criticisms aired can be summarised as follows: (1) representing support at the argument level leaves too many

---

[1]These enhancements are somewhat controversial, as discussed in the next section, with some arguing that they represent different concepts, and should not be used within abstract argument frameworks.

choices regarding how to represent argument; (2) support represents inference, which should be at the level of logic, not argument; and (3) abstract argument frameworks represent support implicitly through the notion of defence; an argument a supports b by attacking all attackers of b. Similar to (3) is claim (4), namely that abstract argument frameworks represent support by default; i.e. the presence of an argument within a framework means that some support for it exists.

The first argument against support could equally be deemed an argument for its inclusion. While a number of representational choices open up, it is possible to get rid of the redundancy that can appear in attack-only argument frameworks. Similarly, by adopting the maxim "represent arguments in the simplest form possible", one can often eliminate a lot of the choices that may appear. The remaining points are related, as they assume that support is used to somehow infer the case for an argument appearing in either the argument framework, or in its extensions. While this is often the case, particularly in Bipolar Argumentation Frameworks (BAFs), support has another role, namely to allow us to distinguish between *prima facie* and standard arguments. *Prima facie* arguments do not require any sort of support from other arguments to stand, while standard arguments must be linked with at least one *prima facie* argument via a support chain. It is this concept that we formalise in our framework below. Finally, when people reason about argument, they often think of the interactions between support and attack. Therefore, when modelling human argument, it makes sense to have an argument framework that treats the two as equals. BAFs, while allowing for support, use it to model strengthening of conclusions by separate arguments, or occasionally, inference. Arguments may stand with, or without support, meaning that inference in BAFs may take place both internally to the argument, and outside it. The criticisms described above thus apply directly, and we believe legitimately, to Bipolar Argumentation Frameworks. Our notion of support is subtly but importantly different, however; we wish to capture the notion of *evidential support*. We introduce a set of semantics that allow us to reason about arguments that are directly supported by evidence (or are *prima facie* arguments), and arguments that have been inserted into the framework due to a chain of evidential support from a *prima facie* argument. This is a new, and important notion of support.

3. Arguments, Attack and Support

As mentioned in the introduction, an argument is accepted only if it is supported through a chain of arguments, each of which is themselves supported. At the head of this chain of supporting arguments is an argument representing support from the environment (written as support from the special argument η). To represent this notion, we define an evidential argument system as follows:

Definition 1. *(Evidential Argumentation Systems)*
An evidential argumentation system is a tuple (A, R_a, R_e) where A is a set of arguments, R_a is a relation of the form $(2^A \setminus \{\}) \times A$, and R_e is a relation of type $2^A \times A$, such that within the argumentation system, $\nexists x \in 2^A, y \in A$ such that xR_ay and xR_ey. We assume the existance of a "special" argument $\eta \notin A$.

The R_e and R_a relations encode evidential support and attacks between arguments. An element of the evidential support relation of the form $\{\eta\}R_e a$ would represent sup-

port by the environment for the argument a. Within our argument framework, we are interested in seeing which arguments may eventually be considered to hold. Since any argument attacked by the environment will be unconditionally defeated, we believe that it makes no sense to include such arguments, and therefore prohibit the environment from appearing in the attacks relation. Since the environment requires no support, η may not appear as the second element of a member of R_e.

For an argument a, if $\{\eta\}R_e a$, we say (in a slight abuse of the English language) that a has environmental support. Environmental support can be used to model defaults (since an argument that has environmental support is true unless attacked), tautologies (if the argument may never be attacked), and arguments that have incontrovertible evidence in their support. We assume that evidence from the environment that cannot be challenged, such as an observation by an infallible sensor, can be thought of as such evidence.

Unlike Bipolar Argumentation Frameworks (BAFs), we allow for attacks and supports by sets of arguments. A set S is said to attack an argument a iff there is a $T \subseteq S$ such that $T R_a a$. S is a minimal attack on a if there is no subset of T that also attacks a. We say that a set S attacks a set S' if S attacks one of the members of T. Similar notions exist for support, but are complicated by the fact that an argument is not supported by simply appearing in the R_e relation. In an evidence based approach, an argument is only acceptable if it is supported by another supported argument, or by evidence from the environment.

Definition 2. *(Evidential Support) An argument a is e-supported by a set S iff*

1. *$S R_e a$ where $S = \{\eta\}$ or*
2. *$\exists T \subset S$ such that $T R_e a$ and $\forall x \in T$, x is supported by $S \setminus \{x\}$*

S is a minimum support for a if there is no $T \subset S$ such that a is supported by T.

If a is e-supported by S, we may say that S e-supports a.

This notion of support is stronger than the one present in BAFs, and, in our opinion, offers a far stronger justification for the inclusion of a support relation within abstract argumentation systems. Our notion of support requires evidence at the start of a chain of support (i.e. $\{\eta\}R_e x$ for some argument $x \in S$) which leads, through various arguments to a, before the argument a may be used. With this notion, we may define the notion of an evidence-supported (or e-supported) attack:

Definition 3. *(Evidence-Supported Attack) A set S carries out an evidence-supported attack on an argument a if*

- *$X R_a a$ where $X \subseteq S$, and,*
- *All elements $x \in X$ are supported by S.*

A supported attack by a set S is minimal iff there is no $T \subset S$ such that T carries out an evidence-supported attack on a.

An e-supported attack attempts to represent the notion of an attack backed up by evidence or facts, and is necessary for our definition of acceptability.

Support for an argument is clearly one requirement for acceptability. However it is not enough. Following Dung, an argument should be acceptable (with respect to some set) if it is defended from attack by that set. The question arises whether all attacks should
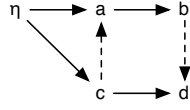
Figure 1. An evidential argument system with four arguments; dashed arrows represent attacks relations and solid arrows represent supports relations. Evidence nodes are those supported by the special argument η.

be defended against, or only e-supported attacks. In this work, we choose the latter, and, by doing so, we allow an argument to be defended from attack by having the attack itself attacked or by having some means of support for the argument attacked by the defending set.

Definition 4. *(Acceptability) An argument a is acceptable with respect to a set S iff*

1. *S e-supports a, and*
2. *Given a minimal evidence-supported attack $X \subseteq 2^A$ against a, $\exists T \in S$ such that $T R_a x$, where $x \in X$ so that $X \setminus \{x\}$ is no longer a supported attack on a.*

An argument is thus acceptable with respect to a set of arguments S if any argument that e-support attacks it is itself attacked (either directly or by being rendered unsupported) by a member of S. The set S must also e-support the acceptable argument. It is clear from this definition of acceptability that an asymmetry arises with regards to acceptability, as we ignore elements of S that might themselves be attacked by other arguments, which would prevent S from supporting a. To overcome this problem, the notion of admissibility is required, and for this, we must first define two other notions.

Definition 5. *(Conflict free Sets) A set of arguments S is conflict free iff $\forall y \in S$, $\nexists X \subseteq S$ such that $X R_a y$.*

Definition 6. *(Self Supporting Sets) A set of arguments S is self supporting iff $\forall x \in S$, S e-supports x.*

It should be noted that this definition of a conflict free set is very strong. A weaker definition is possible by defining a conflict free set as one not containing any e- supported attacks on itself. Since we intend to use the concept of a conflict free set in our definition of admissibility, we will show in Lemma 1 that, for our purposes, these two definitions of a conflict free set coincide.

It is clear (from Definition 2) that if a set is a minimum e-support for an argument, then the set is self supporting. We may thus define the concept of admissibility:

Definition 7. *(Admissible Set of Arguments) A set of arguments S is said to be admissible iff*

1. *All elements of S are acceptable with respect to S.*
2. *The set S is conflict free.*

Figure 1 illustrates these concepts. Here, a, b, c, d are all supported arguments. The set of arguments $\{a, b\}$ e-support attacks d, while c e-support attacks a. Argument d is acceptable with respect to $\{c\}$, since c prevents b from being e-supported. Finally, the set $\{c, d\}$ is an admissible set.

We may then show a number of useful results (proofs for the lemmas can be found in [7]):

Lemma 1. *An admissible set contains no supported attacks on itself.*

Lemma 2. *An admissible set is self supporting.*

Lemma 3. *The empty set is admissible.*

Lemma 4. *Given an admissible set S, and two arguments x, y which are acceptable with respect to S,*

1. *$T = S \cup x$ is admissible,*
2. *y is acceptable with respect to T.*

Given this groundwork, we may now define modified versions of a number of Dung's extensions.

Definition 8. *(**Evidential Preferred Extensions**) An admissible set S is an evidential preferred extension if it is maximal with respect to set inclusion. That is, there is no admissible set T such that $S \subset T$.*

The evidential preferred (or e-preferred) extension is analogous to Dung's preferred extension, and we may prove the following result:

Lemma 5. *Any argumentation system has at least one e-preferred extension.*

It should be noted that Dung's preferred extension can be captured in our framework by having support exist only between the environment (η) and all other arguments in the system.

There is no trivial transformation between BAF d/s/c-preferred semantics and our e-preferred semantics (or, as will be seen when other semantics for the evidential framework are described, between those and any BAF semantics). The reason for this is that the type of consistency BAF semantics describe is completely different to the consistency we are interested in. A set of arguments is deemed "compatible", and thus part of an extension in one of the BAF semantics if, in some sense, none of the arguments in the extension both support, and attack some argument. Two types of compatibility were identified, namely internal, and external compatibility. Internal compatibility occurs when an argument is not accepted into the set if it would lead to the defeat of a member of the set. External compatibility on the other hand means that an argument would not be accepted into a set because its acceptance would mean that an argument outside the set would be both directly, or indirectly, supported and attacked by members of the set. We do not see this simultaneous attack and defence as a problem, additional evidence may allow us to still have a consistent set of arguments in this case (as shown by the semantics introduced here).

We may also introduce analogies to Dung's stable and grounded extensions. Informally, an e-stable extension is one that ensures that nothing but the e-stable set can be derived. This can be achieved by attacking arguments not in the set, or by attacking the support for those arguments.
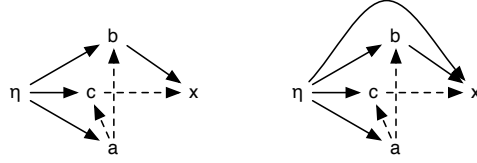
Figure 2. Two sets of arguments and their interactions.

Definition 9. *E-stable Extensions An e-stable extension of an argument framework A is a conflict free and self supporting set S such that for any supported argument $a \notin S$,*

1. *S (support) attacks a or*
2. *$\forall T$ such that T minimally supports a, S (support) attacks T.*

Lemma 6. *A set S is an e-stable extension iff $S = \{a|a$ is not support attacked by S and is supported by $S\}$, and $S \neq \{\{\}\}$.*

To introduce the evidential grounded extension, we first introduce the characteristic function of an argument framework:

Definition 10. *(The Characteristic Function) The characteristic function of an argumentation framework A is the function*

$$F_A : 2^A \to 2^A$$

$$F_A(S) = \{a|a \text{ is acceptable with respect to } S\}$$

The following lemmas then hold:

Lemma 7. *A conflict free set S of arguments is admissible iff $S \subseteq F_A(S)$*

Lemma 8. *F_A is monotonic with respect to set inclusion.*

From this, we can define the evidential grounded extension:

Definition 11. *(E-grounded Extensions) An evidential grounded extension of a finitary argument framework A is the least fixed point of F_A.*

4. Discussion and Future Work

Evidential extensions are able to represent situations other frameworks are unable to handle. For example, consider the following scenarios:

1. The lefthand diagram in Figure 2 represents the following set of arguments: (a) it was dark, when it is dark, a witness statement could be wrong; (b) witness b said he saw a man fly by waving his arms. A witness statement can be viewed as evidence for a claim; (c) witness c said he saw a man waving his arms, but the man didn't fly. A witness' statement can be viewed as evidence for a claim; and (x) From the evidence, we can conclude that the man flew.

2. The righthand diagram in Figure 2 represents the following set of arguments: (a) It was dark, when it is dark, a witness statement could be wrong; (b) Witness b made the statement that the bird flew. A witness statement can be viewed as evidence; (c) Witness c made the statement that the bird did not fly. A witness statement can be viewed as evidence; and (x) We know that birds can normally fly, and thus given some evidence, we may claim that birds fly.

Given the arguments outlined in scenario 2, we wouldwant our extension to contain the arguments $\{a, x\}$; i.e. we would consider the arguments that it was dark and when it is dark a witness statement could be wrong, and we know birds can normally fly, and thus given some evidence we may claim that birds fly. In scenario 1, we would, intuitively, want to accept argument a; the removal of some background knowledge (birds, by default, fly; humans do not) means that, unlike for the original argument x, we have no evidence one way or another that will allow us to determine the status of x.

The evidence-based argumentation framework presented in this paper neatly captures the distinction between these two cases; in scenario 1 the e-preferred extension consists of $\{a\}$, whereas in scenario 2 the e-preferred extension consists of $\{a, x\}$. Both extensions, therefore, correspond to our intuitions.

In contrast, it is not possible to capture this distinction within BAFsThis is simply because the notion of support within BAFs is not intended to capture the grounding of a chain of arguments in evidence. Support within BAFs captures a form of inference, which, as discussed in Section 2, should be captured at the level of logic, not argument. Our notion of *evidential support* from a *prima facie* argument is importantly distinct from the notion of support within BAFs, and, we believe, provides the first clear justification for including a notion of support within abstract argumentation frameworks.

Dung's extensions are clearly unable to handle the notion of support, and cannot represent either scenario, while the various extensions described by BAFs cannot distinguish between the two scenarios.

This added expressiveness is useful wherever reasoning with evidence, or with default and *prima facie* arguments, appears. For example, when performing contract monitoring, an agent could determine which clauses definitely hold by computing the e-grounded extension based on what it knows about the environment. In a contract enforcement setting, where two agents disagree as to what the contract state is, they could advance arguments, and the shared elements of the e-preferred extensions could again be used to determine which clauses are in force.

Our extensions reduce to Dung's original extensions when support links only exist between the empty argument and all other arguments. Since d-preferred extensions are identical to Dung's preferred extension, they can be similarly encapsulated by our framework. S-preferred and c-preferred extensions overlap with our extensions only for certain configurations of support.

As in BAFs, attack in our framework is more powerful than support. That is, if an attack succeeds, no amount of support will be able to negate it. One way of overcoming this weakness involves the concept of valuation of arguments [3], and we are actively pursuing this avenue of research.

Many relations between extensions are known to exist in Dung's framework as well as in BAFs. We intend to investigate which of these properties extend to our framework. Other extensions have been proposed for argumentation frameworks [2], and we intend to see how these translate into our more expressive model.

5. Conclusions

In this paper we have presented a novel abstract argumentation framework that captures the concepts of evidence-based support and support by multiple arguments. We have introduced a number of extensions allowing us to reason about sets of arguments in various ways. These extensions are based on the concepts of *evidence* and *support*; that is, an argument is only considered part of an extension if it is supported by a chain of arguments rooted by evidence, and is defended from attack by other arguments which are in turn supported by evidence. We have been able to prove many of the results shown by Dung for his original extensions, but are able to support more complicated arguments. Furthermore, our extensions are able to provide more intuitive results than Bipolar Argumentation Frameworks in cases where evidential reasoning is required; we would argue that this would include most, if not all, realistic domains. Finally, we believe that this research provides clear justification for the inclusion of the notion of support within abstract argumentation frameworks.

6. Acknowledgements

References

[1] L. Amgoud, C. Cayrol, and M.-C. Lagasquie-Schiex. On the bipolarity in argumentation frameworks. In *Proceedings of the 10th International Workshop on Non-monotonic Reasoning*, pages 1–9, Whistler, Canada, 2004.

[2] C. Cayrol, C. Devred, and M.-C. Lagasquie-Schiex. Handling controversial arguments in bipolar argumentation systems. In *Proceedings of the 2006 Conference on Computational Models of Argument*, pages 261–272, 2006.

[3] C. Cayrol and M.-C. Lagasquie-Schiex. Graduality in argumentation. *Journal of Artificial Intelligence Research*, 23:245–297, 2005.

[4] C. Cayrol and M.-C. Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In *Pro. of the Eighth European Conference on Symbolic and Quantitative Approaches to Reasoning With Uncertainty*, volume 3571 of *LNAI*, pages 378–389. Springer-Verlag, 2005.

[5] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–357, 1995.

[6] S. H. Nielsen and S. Parsons. A generalization of Dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In *Proceedings of the Third International Workshop on Argumentation in Multi-Agent Systems*, pages 7–19, Hakodate, Japan, 2006.

[7] N. Oren. *An Argumentation Framework Supporting Evidential Reasoning with Applications to Contract Monitoring*. Phd thesis, University of Aberdeen, Aberdeen, Scotland, 2007.