Making Argumentation More Believable

Anthony Hunter
Department of Computer Science
University College London
Gower Street
London, WC1E 6BT, UK

Abstract

There are a number of frameworks for modelling argumentation in logic. They incorporate a formal representation of individual arguments and techniques for comparing conflicting arguments. A problem with these proposals is that they do not consider the believability of the arguments from the perspective of the intended audience. In this paper, we start by reviewing a logic-based framework for argumentation based on argument trees which provide a way of exhaustively collating arguments and counter-arguments. We then extend this framework to a model-theoretic evaluation of the believability of arguments. This extension assumes that the beliefs of a typical member of the audience for argumentation can be represented by a set of classical formulae (a beliefbase). We compare a beliefbase with each argument to evaluate the empathy (or similarly the antipathy) that an agent has for the argument. We show how we can use empathy and antipathy to define a pre-ordering relation over argument trees that captures how one argument tree is "more believable" than another. We also use these to define criteria for deciding whether an argument at the root of an argument tree is defeated or undefeated given the other arguments in the tree.

Introduction

Argumentation is a vital aspect of intelligent behaviour by humans. There are a number of proposals for logic-based formalisations of argumentation (Prakken & Vreeswijk 2000; Chesnevar, Maguitman, & Loui 2001). These proposals allow for the representation of arguments for and against some conclusion, and for attack or undercut relationships between arguments. In the monological, as opposed to dialectical, approaches we can imagine a knowledgebase Δ as being a set of possibily conflicting pieces of information (each piece of information is represented by a formula) that has been collated by one or more agents, and the role of argumentation is to construct arguments from Δ. However, a problem with the logic-based formalisations is that no account is taken of the audience of the arguments, and in particular, of how believable the arguments may appear to a typical member of the intended audience.

To motivate this need, consider a politician who is giving a speech on a plan by the government to charge car drivers to be able to drive into the city. This requires monological argumentation, since the audience would not be expected to participate in a dialogue *during* the speech. If the audience is a group of commuters who live in the city, then the politician would want to provide arguments that relate to what the audience is likely to be familiar with, perhaps saying that the money raised would be used to buy much-needed new buses, and there would be less pollution for pedestrians. In contrast, if the audience is a group of business executives, then the politician may argue that the cost of the charge to commercial vehicles (e.g. delivery trucks) would be more than offset by the savings made by their vehicles not being stuck in traffic in the city.

The beliefs of these two audiences are unlikely to be the same. They may even be mutually contradictory. The business executives, for example, may live in a different city, and so they might not know whether or not there are enough buses, or whether there is a lot of pollution in that city. And the commuters may be unaware that businesses have greater expenses to pay when their delivery vehicles are stuck in traffic. So the way the politician has to proceed, is to be selective so that the arguments used are likely to be based on assumptions that are already believed, and if not, on assumptions that do not contradict the intended audience's beliefs.

This need is reflected in many professional domains (e.g. medicine, science, law, journalism, etc), and so to deliver better decision-support technology, we first need to develop appropriate formalisms for argumentation that is more believable by the intended audience.

To address this need, we introduce the beliefs of a typical member of the intended audience into logic-based argumentation. These beliefs are represented by a beliefbase (a set of classical formulae). To present this approach, we extend an existing framework to logic-based argumentation by Besnard and Hunter (Besnard & Hunter 2001). In the next section, we review the existing framework, and then in subsequent sections we present our new extension.

Review of Argument Trees

We use $\alpha, \beta, \gamma, \ldots$ to denote formulae and $\Delta, \Phi, \Psi, \ldots$ to denote sets of formulae. Deduction in classical propositional logic is denoted by the symbol $\vdash$ and deductive closure by Th so that $\mathsf{Th}(\Phi) = \{\alpha \mid \Phi \vdash \alpha\}$.

For the following definitions, we first assume a knowledgebase Δ (a finite set of formulae) and use this Δ throughout. We further assume that every subset of Δ is given an enumeration $\langle \alpha_1, \ldots, \alpha_n \rangle$ of its elements, which we call its canonical enumeration. This really is not a demanding constraint: In particular, the constraint is satisfied whenever we impose an arbitrary total ordering over Δ. Importantly, the order has no meaning and is not meant to represent any respective importance of formulae in Δ. It is only a convenient way to indicate the order in which we assume the formulae in any subset of Δ are conjoined to make a formula logically equivalent to that subset.

The paradigm for the approach is a large repository of information, represented by Δ, from which arguments can be constructed for and against arbitrary claims. Apart from information being understood as declarative statements, there is no a priori restriction on the contents, and the pieces of information in the repository can be as complex as possible. Therefore, Δ is not expected to be consistent. It need even not be the case that every single formula in Δ is consistent.

The framework adopts a very common intuitive notion of an argument. Essentially, an argument is a set of relevant formulae that can be used to classically prove some point, together with that point. Each point is represented by a formula.

Definition 1. *An* **argument** *is a pair $\langle \Phi, \alpha \rangle$ such that: (1) $\Phi \nvdash \bot$; (2) $\Phi \vdash \alpha$; and (3) there is no $\Phi' \subset \Phi$ such that $\Phi' \vdash \alpha$. We say that $\langle \Phi, \alpha \rangle$ is an argument for α. We call α the consequent of the argument and Φ the support of the argument (we also say that Φ is a support for α). For an argument $\langle \Phi, \alpha \rangle$, the support is given by $\mathsf{Support}(\langle \Phi, \alpha \rangle) = \Phi$.*

Example 1. *Let $\Delta = \{\alpha, \alpha \rightarrow \beta, \gamma \rightarrow \neg\beta, \gamma, \delta, \delta \rightarrow \beta, \neg\alpha, \neg\gamma\}$. Some arguments are:*

$$\langle \{\alpha, \alpha \rightarrow \beta\}, \beta \rangle$$
$$\langle \{\neg\alpha\}, \neg\alpha \rangle$$
$$\langle \{\alpha \rightarrow \beta\}, \neg\alpha \vee \beta \rangle$$
$$\langle \{\neg\gamma\}, \delta \rightarrow \neg\gamma \rangle$$

Arguments are not independent. In a sense, some encompass others (possibly up to some form of equivalence). To clarify this requires a few definitions as follows.

Definition 2. *An argument $\langle \Phi, \alpha \rangle$ is* **more conservative** *than an argument $\langle \Psi, \beta \rangle$ iff $\Phi \subseteq \Psi$ and $\beta \vdash \alpha$.*

Example 2. *$\langle \{\alpha\}, \alpha \vee \beta \rangle$ is more conservative than $\langle \{\alpha, \alpha \rightarrow \beta\}, \beta \rangle$.*

Some arguments directly oppose the support of others, which amounts to the notion of an undercut.

Definition 3. *An* **undercut** *for an argument $\langle \Phi, \alpha \rangle$ is an argument $\langle \Psi, \neg(\phi_1 \wedge \ldots \wedge \phi_n) \rangle$ where $\{\phi_1, \ldots, \phi_n\} \subseteq \Phi$.*

Example 3. *Let $\Delta = \{\alpha, \alpha \rightarrow \beta, \gamma, \gamma \rightarrow \neg\alpha\}$. Then, $\langle \{\gamma, \gamma \rightarrow \neg\alpha\}, \neg(\alpha \wedge (\alpha \rightarrow \beta)) \rangle$ is an undercut for $\langle \{\alpha, \alpha \rightarrow \beta\}, \beta \rangle$. A less conservative undercut for $\langle \{\alpha, \alpha \rightarrow \beta\}, \beta \rangle$ is $\langle \{\gamma, \gamma \rightarrow \neg\alpha\}, \neg\alpha \rangle$.*

Definition 4. *$\langle \Psi, \beta \rangle$ is a* **maximally conservative undercut** *of $\langle \Phi, \alpha \rangle$ iff $\langle \Psi, \beta \rangle$ is an undercut of $\langle \Phi, \alpha \rangle$ such that no undercuts of $\langle \Phi, \alpha \rangle$ are strictly more conservative than $\langle \Psi, \beta \rangle$ (that is, for all undercuts $\langle \Psi', \beta' \rangle$ of $\langle \Phi, \alpha \rangle$, if $\Psi' \subseteq \Psi$ and $\beta \vdash \beta'$ then $\Psi \subseteq \Psi'$ and $\beta' \vdash \beta$).*

The value of the following definition of canonical undercut is that we only need to take the canonical undercuts into account. This means we can justifiably ignore the potentially very large number of non-canonical undercuts.

Definition 5. *An argument $\langle \Psi, \neg(\phi_1 \wedge \ldots \wedge \phi_n) \rangle$ is a* **canonical undercut** *for $\langle \Phi, \alpha \rangle$ iff it is a maximally conservative undercut for $\langle \Phi, \alpha \rangle$ and $\langle \phi_1, \ldots, \phi_n \rangle$ is the canonical enumeration of Φ.*

An argument tree describes the various ways an argument can be challenged, as well as how the counter-arguments to the initial argument can themselves be challenged, and so on recursively.

Definition 6. *An* **argument tree** *for α is a tree where the nodes are arguments such that*

1. *The root is an argument for α.*
2. *For no node $\langle \Phi, \beta \rangle$ with ancestor nodes $\langle \Phi_1, \beta_1 \rangle, \ldots, \langle \Phi_n, \beta_n \rangle$ is Φ a subset of $\Phi_1 \cup \cdots \cup \Phi_n$.*
3. *The children nodes of a node N consist of all canonical undercuts for N that obey 2.*

The second condition in Definition 6 ensures that each argument on a branch has to introduce at least one formula in its support that has not already been used by ancestor arguments. As a notational convenience, in examples of argument trees the $\diamond$ symbol is used to denote the consequent of an argument when that argument is a canonical undercut.

Example 4. *Let $\Delta = \{\alpha \vee \beta, \alpha \rightarrow \gamma, \neg\gamma, \neg\beta, \delta \leftrightarrow \beta\}$. For this, two argument trees for the consequent $\alpha \vee \neg\delta$ are given.*

$$\begin{array}{cc} \langle \{\alpha \vee \beta, \neg\beta\}, \alpha \vee \neg\delta \rangle & \langle \{\delta \leftrightarrow \beta, \neg\beta\}, \alpha \vee \neg\delta \rangle \\ \uparrow & \uparrow \\ \langle \{\alpha \rightarrow \gamma, \neg\gamma\}, \diamond \rangle & \langle \{\alpha \vee \beta, \alpha \rightarrow \gamma, \neg\gamma\}, \diamond \rangle \end{array}$$

For an argument tree T, each argument in T is either an **attacking argument** or a **defending argument**. If an argument a_i is a defending argument, then any child a_j of a_i is an attacking argument. If an argument a_j is an attacking argument, then any child a_k of a_j is a defending argument. If an argument a_r is the root, then a_r is the **initiating argument**, and a_r is a defending argument. Finally, for an argument a_i in an argument tree T, let the set of children of a_i be given by $\mathsf{Children}(T, a_i)$.

Pairwise Theory Comparison

When comparing an agent's beliefbase with the support of an argument, we are comparing a pair of theories. To do this we now present a model-theoretic way to compare pairs of theories. Note, we use $\wedge X$ to denote a conjunction of all the formulae in X and we use $\mathsf{Atoms}(X)$ to denote the atom symbols (i.e. propositional letters) used in the formulae in X.

Definition 7. *Let $\mathcal{A}$ be a set of atoms. Each* **interpretation** *w is represented by a subset of $\mathcal{A}$. For each interpretation w, each atom in w is assigned true and each atom in $\mathcal{A} \setminus w$*

is assigned false. Let X be a set of classical propositional formulae. Let $I(X)$ be the set of interpretations of X delineated by the atoms used in X (i.e. $I(X) = \wp(\mathsf{Atoms}(X))$). Let $M(X, Y)$ be the set of **models** *of X that are in $I(Y)$. So $M(X, Y) = \{w \models \wedge X \mid w \in I(Y)\}$ where $\models$ is classical.*

Example 5. *Let $X = \{\alpha\}$ and $Y = \{\beta \wedge \gamma\}$. So $M(X, X \cup Y) = \{\{\alpha, \beta, \gamma\}, \{\alpha, \beta\}, \{\alpha, \gamma\}, \{\alpha\}\}$, and $M(Y, X \cup Y) = \{\{\alpha, \beta, \gamma\}, \{\beta, \gamma\}\}$, where $I(X \cup Y) = \{\{\alpha, \beta, \gamma\}, \{\alpha, \beta\}, \{\alpha, \gamma\}, \{\beta, \gamma\}, \{\alpha\}, \{\beta\}, \{\gamma\}, \{\}\}$.*

Obviously $M(X, X) \subseteq I(X)$. When X is non-empty, we get: (1) $M(X, X) = \emptyset$ iff X is inconsistent; and (2) $I(X) = M(X, X)$ iff X is a set of tautologies. Also $M(X, X \cup Y) \subseteq M(Y, X \cup Y)$ iff $X \vdash \wedge Y$.

The degree of entailment of X for Y is the number of models they have in common divided by the total number of models for X.

Definition 8. *Let X and Y be sets of classical propositional formulae each of which is consistent (i.e. $X \not\vdash \perp$ and $Y \not\vdash \perp$). The* **degree of entailment** *of X for Y, denoted $E(X, Y)$, is defined as follows:*

$$E(X, Y) = \frac{|M(X, X \cup Y) \cap M(Y, X \cup Y)|}{|M(X, X \cup Y)|}$$

Example 6. *$E(\alpha, \alpha \wedge \beta) = 1/2$, $E(\alpha, \alpha \wedge \beta \wedge \gamma) = 1/4$, $E(\alpha, \alpha \wedge \beta \wedge \gamma \wedge \delta) = 1/8$, $E(\alpha \wedge \beta, \alpha \vee \beta) = 1$, $E(\alpha \wedge \beta, \alpha \wedge \epsilon) = 1/2$, $E(\alpha \wedge \beta \wedge \gamma, \alpha \wedge \epsilon) = 1/2$, $E(\alpha \wedge \beta \wedge \gamma \wedge \delta, \alpha \wedge \epsilon) = 1/2$, $E(\alpha \wedge \epsilon, \alpha \wedge \beta \wedge \gamma \wedge \delta) = 1/8$, $E(\alpha \wedge \beta, \alpha \wedge \neg\beta) = 0$.*

Proposition 1. *Let X, Y, and Z be sets of classical propositional formulae: (1) $0 \leq E(X, Y) \leq 1$; (2) $X \vdash \wedge Y$ iff $E(X, Y) = 1$; (3) $X \vdash \neg \wedge Y$ iff $E(X, Y) = 0$; (4) If $E(X, Y) = 1$ then $0 < E(Y, X)$; and (5) $E(X, Y) = 0$ iff $E(Y, X) = 0$.*

We now recall the definition for Dalal distance for comparing pairs of models (Dalal 1988). It is the Hamming distance between two models.

Definition 9. *Let X be a set of classical propositional formulae, and let $w_i, w_j \in I(X)$. The* **Dalal distance** *between w_i and w_j, denoted $\mathsf{Dalal}(w_i, w_j)$, is the difference in the number of atoms assigned true:*

$$\mathsf{Dalal}(w_i, w_j) = |w_i - w_j| + |w_j - w_i|$$

To evaluate the conflict between two theories, we take a pair of models, one for each theory, such that the Dalal distance is minimised. The degree of conflict is this distance divided by the maximum possible Dalal distance between a pair of models in $I(X \cup Y)$ (i.e. $\log_2$ of the total number of models in $I(X \cup Y)$).

Definition 10. *Let X and Y be sets of classical propositional formulae, each of which is consistent, and let $\mathsf{Distances}(X, Y) = \{\mathsf{Dalal}(w_x, w_y) \mid w_x \in M(X, X \cup Y)$ and $w_y \in M(Y, X \cup Y)\}$. The* **degree of conflict** *between X and Y, denoted $C(X, Y)$, is defined as follows:*

$$C(X, Y) = \frac{\mathsf{Min}(\mathsf{Distances}(X, Y))}{\log_2(|I(X \cup Y)|)}$$

Example 7. *$C(\alpha \wedge \beta, \alpha \wedge \neg\beta) = 1/2$, $C(\alpha \wedge \beta, \neg\alpha \vee \neg\beta) = 1/2$, $C(\alpha \wedge \beta, \neg\alpha \wedge \neg\beta) = 1$, $C(\alpha \wedge \beta, \neg\alpha \wedge \beta) = 1/2$, $C(\alpha \wedge \beta \wedge \gamma, \neg(\alpha \wedge \beta) \wedge \neg\gamma \wedge \neg\delta) = 1/2$.*

Proposition 2. *Let X and Y be sets of classical propositional formulae: (1) $0 \leq C(X, Y) \leq 1$; (2) $C(X, Y) = C(Y, X)$; and (3) $C(X, Y) = 0$ iff $X \cup Y \not\vdash \perp$ iff $E(X, Y) \neq 0$.*

In the next section, we use these measures to analyse the believability of arguments.

Believability of Argument Trees

To evaluate the believability of an argument, we use a consistent set of formulae, called a **beliefbase**, that reflects the beliefs of a typical member of the intended audience of the argument. Argumentation for different intended audiences requires different beliefbases.

Definition 11. *Let $\langle \Phi, \alpha \rangle$ be an argument and let Γ be a beliefbase. The* **empathy** *for $\langle \Phi, \alpha \rangle$ is $E(\Gamma, \Phi)$. The* **antipathy** *for $\langle \Phi, \alpha \rangle$ is $C(\Gamma, \Phi)$.*

Example 8. *Let $\Delta = \{\alpha \vee \beta, \alpha \rightarrow \gamma, \neg\gamma, \neg\beta, \delta \leftrightarrow \beta\}$ and consider the following argument tree. If $\Gamma = \{\alpha, \neg\beta\}$, then $E(\Gamma, \{\alpha \vee \beta, \neg\beta\}) = 1$ and $C(\Gamma, \{\alpha \rightarrow \gamma, \neg\gamma\}) = 1/3$.*

$$\begin{array}{c} \langle \{\alpha \vee \beta, \neg\beta\}, \alpha \vee \neg\delta \rangle \\ \uparrow \\ \langle \{\alpha \rightarrow \gamma, \neg\gamma\}, \diamond \rangle \end{array}$$

In order to improve the believability of an argument tree, we want to maximize the empathy for defending arguments, and maximize the antipathy of attacking arguments.

Proposition 3. *Let $\langle \Phi, \alpha \rangle$ be an argument, and let $\langle \Psi, \diamond \rangle$ be a canonical undercut for it, and let Γ be a beliefbase: (1) If $E(\Gamma, \Phi) = 1$, then $E(\Gamma, \Psi) = 0$; and (2) If $C(\Gamma, \Phi) = 1$, then $C(\Gamma, \Psi) \neq 1$.*

Now we provide definitions for the measure of recursive empathy and the measure of recursive antipathy for argument trees. Each of these measures gives a value in the $[-1, 1]$ range. The essential idea is that the empathy at a node can be affected by the nodes below it in the tree. Defending arguments may increase the recursive empathy at a node and attacking arguments may decrease the recursive empathy.

Definition 12. *The* **recursive empathy** *(r.e.) for an argument tree T with the beliefbase Γ, denoted $Re(\Gamma, T)$, is given by $F(a_r)$ where a_r is the root of T and the F function is defined for all nodes a_i in T as follows. Let $k_1 = E(\Gamma, \mathsf{Support}(a_i))$. If $\mathsf{Children}(T, a_i) \neq \emptyset$, then let k_2 be the maximum $F(a_j)$ value for the nodes a_j in $\mathsf{Children}(T, a_i)$, otherwise let $k_2 = 0$. So $F(a_i)$ is:*

1. *If $k_1 > k_2$, then $F(a_i) = k_1$*
2. *If $k_1 = k_2$, then $F(a_i) = 0$*
3. *If $k_1 < k_2$, then $F(a_i) = -k_2$*

For an argument tree T, and a beliefbase Γ, $Re(\Gamma, T) > 0$ when the root has greater non-zero empathy than the r.e. of any of the undercuts, $Re(\Gamma, T) < 0$ when there is an undercut with r.e. greater then the empathy of the root, and

$Re(\Gamma, T) = 0$ when the best undercut (ie. the undercut with maximum r.e.) has r.e. equal to the empathy of the root or when the best undercut has negative r.e. and the root has zero empathy.

Definition 13. *The* **recursive antipathy** *(r.a.) for an argument tree T with the beliefbase Γ, denoted $Re(\Gamma, T)$, is given by $A(a_r)$ where a_r is the root of T and the A function is defined for all nodes a_i in T as follows. Let $k_1 = C(\Gamma, \mathsf{Support}(a_i))$. If $\mathsf{Children}(T, a_i) \neq \emptyset$, then let k_2 be the maximum $A(a_j)$ for the nodes a_j in $\mathsf{Children}(T, a_i)$, otherwise let $k_2 = 0$. So $A(a_i)$ is:*

1. *If $k_1 > k_2$, then $A(a_i) = k_1$*
2. *If $k_1 = k_2$, then $A(a_i) = 0$*
3. *If $k_1 < k_2$, then $A(a_i) = -k_2$*

For an argument tree T, and a beliefbase Γ, $Ra(\Gamma, T) > 0$ when the root has greater non-zero antipathy than the r.a. of any of the undercuts, $Ra(\Gamma, T) < 0$ when there is an undercut with greater r.a. than the antipathy of the root, and $Ra(\Gamma, T) = 0$ when the best undercut (ie. the undercut with maximum r.a.) has r.a. that is equal to antipathy of the root or when the best undercut has negative r.a. and the root has zero antipathy.

Example 9. *Consider $\Delta = \{\alpha \leftrightarrow \neg\delta, \beta, \beta \rightarrow \alpha, \gamma \wedge \neg\beta, \neg\gamma, \delta, \neg\delta\}$ giving the argument tree T_1 below.*

$$\langle\{\beta \rightarrow \alpha, \beta\}, \alpha\rangle$$
$$\nearrow \qquad\qquad \nwarrow$$
$$\langle\{\gamma \wedge \neg\beta\}, \diamond\rangle \qquad\qquad \langle\{\delta, \alpha \leftrightarrow \neg\delta\}, \diamond\rangle$$
$$\uparrow \qquad\qquad \uparrow$$
$$\langle\{\neg\gamma\}, \diamond\rangle \qquad\qquad \langle\{\neg\delta\}, \diamond\rangle$$

For $\Gamma = \{\delta, \neg\gamma\}$, the arguments in T_1 are evaluated in the table below, where E is the empathy for the node, F is the r.e. for the node, C is the antipathy for the node, and A is the r.a. for the node, giving $Re(\Gamma, T_1) = -1/2$ and $Ra(\Gamma, T_1) = -1/3$. So r.e. is poor and r.a. is good.

Argument	*E*	*F*	*C*	*A*
$\langle\{\beta \rightarrow \alpha, \beta\}, \alpha\rangle$	0	$-1/2$	0	$-1/3$
$\langle\{\gamma \wedge \neg\beta\}, \diamond\rangle$	0	-1	$1/3$	$1/3$
$\langle\{\delta, \alpha \leftrightarrow \neg\delta\}, \diamond\rangle$	$1/2$	$1/2$	0	-1
$\langle\{\neg\gamma\}, \diamond\rangle$	1	1	0	0
$\langle\{\neg\delta\}, \diamond\rangle$	0	0	1	1

Example 10. *Consider $\Delta = \{\alpha \leftrightarrow \neg\delta, \beta, \beta \rightarrow \alpha, \gamma \wedge \neg\beta, \neg\gamma, \delta, \neg\delta\}$ giving the argument tree T_2 below.*

$$\langle\{\neg\delta, \alpha \leftrightarrow \neg\delta\}, \alpha\rangle$$
$$\uparrow$$
$$\langle\{\delta\}, \diamond\rangle$$

For $\Gamma = \{\neg\delta\}$, T_2 is evaluated in the table below, where E, F, C, and A are as in Example 9, giving $Re(\Gamma, T_2) = 1/2$ and $Ra(\Gamma, T_2) = -1$. So r.e. and r.a. are good.

Argument	*E*	*F*	*C*	*A*
$\langle\{\neg\delta, \alpha \leftrightarrow \neg\delta\}, \alpha\rangle$	$1/2$	$1/2$	0	-1
$\langle\{\delta\}, \diamond\rangle$	0	0	1	1

Example 11. *Consider $\Delta = \{\alpha \leftrightarrow \neg\delta, \beta, \beta \rightarrow \alpha, \gamma \wedge \neg\beta, \neg\gamma, \delta, \neg\delta\}$ giving the argument tree T_3 below.*

$$\langle\{\delta, \alpha \leftrightarrow \neg\delta\}, \neg\alpha\rangle$$
$$\nearrow \qquad\qquad \nwarrow$$
$$\langle\{\beta, \beta \rightarrow \alpha\}, \diamond\rangle \qquad\qquad \langle\{\neg\delta\}, \diamond\rangle$$
$$\uparrow$$
$$\langle\{\gamma \wedge \neg\beta\}, \diamond\rangle$$
$$\uparrow$$
$$\langle\{\neg\gamma\}, \diamond\rangle$$

For $\Gamma = \{\delta, \neg\gamma\}$, T_3 is evaluated in the table below, where E, F, C, and A are as in Example 9, giving $Re(\Gamma, T_3) = 1/2$ and $Ra(\Gamma, T_3) = -1/2$. So r.e. and r.a. are good.

Argument	*E*	*F*	*C*	*A*
$\langle\{\delta, \alpha \leftrightarrow \neg\delta\}, \neg\alpha\rangle$	$1/2$	$1/2$	0	$-1/2$
$\langle\{\beta, \beta \rightarrow \alpha\}, \diamond\rangle$	0	0	0	$-1/3$
$\langle\{\neg\delta\}, \diamond\rangle$	0	0	$1/2$	$1/2$
$\langle\{\gamma \wedge \neg\beta\}, \diamond\rangle$	0	-1	$1/3$	$1/3$
$\langle\{\neg\gamma\}, \diamond\rangle$	1	1	0	0

There is an intuitive preference for an argument tree with higher r.e. and there is an intuitive preference for an argument tree with lower r.a. For these two dimensions, the best case argument tree has an r.e. of 1 and an r.a of -1, and the worst case argument tree has an r.e. of -1 and an r.a. of 1.

Example 12. *Let T be the following argument tree. If $\Gamma = \{\alpha \wedge \beta\}$, then $Re(\Gamma, T) = 1$ and $Ra(\Gamma, T) = -1$. If $\Gamma' = \{\neg\alpha \wedge \neg\beta\}$, then $Re(\Gamma', T) = -1$ and $Ra(\Gamma', T) = 1$.*

$$\langle\{\alpha \wedge \beta\}, \alpha\rangle$$
$$\uparrow$$
$$\langle\{\neg\alpha \wedge \neg\beta\}, \diamond\rangle$$

In the extreme case where the empathy for the root is 1, then the r.e. for the tree is automatically 1. Similarly, in the (diametrically opposed) extreme case where the antipathy for the root is 1, then the r.a. for the tree is automatically 1.

Proposition 4. *Let $\langle\Phi, \alpha\rangle$ be the root of an argument tree T, and let Γ be a beliefbase: (1) If $E(\Gamma, \Phi) = 1$, then $Re(\Gamma, T) = 1$; and (2) If $C(\Gamma, \Phi) = 1$, then $Ra(\Gamma, T) = 1$.*

The r.e. and r.a. obtained for a tree are coupled. This is a consequence of a beliefbase being unable to have both a non-zero antipathy and a non-zero empathy for an argument.

Proposition 5. *For an argument tree T and a beliefbase Γ, if $Re(\Gamma, T) > 0$ then $Ra(\Gamma, T) \leq 0$.*

So the most important factor in developing believable aruments is for the initiating arguments to have non-zero empathy (ideally unit empathy) and therefore zero antipathy.

Definition 14. *Let T be a argument tree, let Γ be a beliefbase, and let a_i be a node in T. a_i is* **e-disabled** *iff $\exists a_j \in \mathsf{Children}(T, a_i)$ such that $E(\Gamma, \mathsf{Support}(a_i)) \leq F(\Gamma, \mathsf{Support}(a_j))$. a_i is* **c-disabled** *iff $\exists a_j \in \mathsf{Children}(T, a_i)$ such that $C(\Gamma, \mathsf{Support}(a_i)) \leq A(\Gamma, \mathsf{Support}(a_j))$.*

The definition of e-disabled (and similarly c-disabled) give the grounds for empathy (and similarly for antipathy) for a node to not affect the evaluation of r.e. (and similarly r.a) for ancestor nodes.

Definition 15. *Let T be a argument tree, let Γ be a beliefbase, let a_i and a_j be nodes in T.*

a_i **e-voids** a_j *iff* $E(\Gamma, \mathsf{Support}(a_i)) > E(\Gamma, \mathsf{Support}(a_j))$
a_i **c-voids** a_j *iff* $C(\Gamma, \mathsf{Support}(a_i)) > C(\Gamma, \mathsf{Support}(a_j))$

The definition of e-voids (and similarly c-voids) gives the grounds for empathy (and similarly for antipathy) for a node to not be affected by the evaluation of r.e. (and similarly r.a) for an offspring node.

Proposition 6. *Let T be an argument tree, let Γ be a beliefbase, and let a_i be a node in T. (1) a_i is not e-disabled iff $\forall a_j \in \mathsf{Children}(T, a_i)$ a_i e-voids a_j or a_j is e-disabled. (2) a_i is not c-disabled iff $\forall a_j \in \mathsf{Children}(T, a_i)$ a_i c-voids a_j or a_j is c-disabled.*

We will use e-disabled, c-disabled, e-voids, and c-voids to characterise undefeated argument trees.

Improving Believability

Given a knowledgebase Δ, we may be able to generate more than one argument tree for a conclusion such as in Example 4. Using a beliefbase, we can select one of these trees to optimise believability in terms of r.a. and r.e. We can also seek better arguments trees by ignoring some of the formulae in Δ. In other words, we can delete formulae from Δ to give Δ' and then generate argument trees from Δ'.

Example 13. *Continuing Example 9, we could ignore (i.e. delete) δ in Δ. From this, we get argument tree T_4 below.*

$$\langle\{\beta \rightarrow \alpha, \beta\}, \alpha\rangle$$
$$\uparrow$$
$$\langle\{\gamma \wedge \neg\beta\}, \diamond\rangle$$
$$\uparrow$$
$$\langle\{\neg\gamma\}, \diamond\rangle$$

For the beliefbase $\Gamma = \{\delta, \neg\gamma\}$, T_4 is evaluated in the table below, where E, F, C, and A are as in Example 9. So $Re(\Gamma, T_4) = 0$ and $Ra(\Gamma, T_3) = -1/3$ which is an improvement on the r.e. value for T_1.

Argument	*E*	*F*	*C*	*A*
$\langle\{\beta \rightarrow \alpha, \beta\}, \alpha\rangle$	0	0	0	$-1/3$
$\langle\{\gamma \wedge \neg\beta\}, \diamond\rangle$	0	-1	$1/3$	$1/3$
$\langle\{\neg\gamma\}, \diamond\rangle$	1	1	0	0

Definition 16. *Let T and T' be argument trees and let Γ be a beliefbase. Let $Re(\Gamma, T) = x$, let $Re(\Gamma, T') = x'$, let $Ra(\Gamma, T) = y$, and $Ra(\Gamma, T') = y'$, T is more believable for Γ than T', denoted $T \preceq_\Gamma T'$, iff $x \geq x'$ and $y \leq y'$. T is most believable for Γ iff for all argument trees T', $T \preceq_\Gamma T'$.*

Example 14. *Consider T_1 in Example 9 and T_4 in Example 13. If $\Gamma = \{\delta, \neg\gamma\}$, then $T_4 \preceq_\Gamma T_1$.*

Proposition 7. *Let T be an argument tree and let Γ be a beliefbase. T is most believable for Γ iff $Re(\Gamma, T) = 1$ and $Ra(\Gamma, T) = -1$.*

Obviously, there is not a unique most believable T for Γ since there are many knowledgebases Δ we can consider.

Definition 17. *Let $\mathsf{Trees}(\Delta, \alpha)$ be the set of argument trees constructed from Δ for α (according to Definition 6). Let $\mathsf{Treelets}(\Delta, \alpha) = \bigcup_{\Pi \in \wp(\Delta)} \mathsf{Trees}(\Pi, \alpha)$.*

So if we want to increase the believability of an argument tree $T \in \mathsf{Trees}(\Delta, \alpha)$, we could try to find an argument tree $T' \in \mathsf{Trees}(\Delta, \alpha)$ where $\Delta' \subset \Delta$ and $T' \preceq_\Gamma T$. In other words, we could examine $\mathsf{Treelets}(\Delta, \alpha)$ for a more believable tree. Note, for a given knowledgebase Δ and a given beliefbase Γ, the pre-ordering $(\mathsf{Treelets}(\Delta, \alpha), \preceq_\Gamma)$ does not necessarily have a unique maximal element. Furthermore to maximize believability according to the $\preceq_\Gamma$ pre-ordering may involve a trade-off of increasing r.e. and decreasing r.a.

Winning Argument Trees

Given an argument tree, we want to determine whether the initiating argument wins (i.e. it is undefeated) or whether it looses (i.e. it is defeated). If an initiating argument wins, then we may regard it as an acceptable inference. A simple approach is to consider the binary judge defined next.

Definition 18. *Given an argument tree T, the initiating argument wins according to the* **binary judge**, *denoted $\mathsf{win_b}(T) = yes$, iff for each branch of T, the leaf is a defending argument. Furthermore $\mathsf{win_b}(T) = yes$ iff $\mathsf{win_b}(T) \neq no$.*

So the initiating argument wins whenever all maximal branches of the argument tree have an even length (though not necessarily the same length). This approach captures the essence of a number of approaches to argumentation (see (Besnard & Hunter 2001) for more details).

Now that we have measures of r.a. and r.e. for argument trees, we can take a finer grained approach to evaluating argument trees. Below we define the empathy judge, the antipathy judge, and the combined judge and give examples in Table 1. All the judges we consider are functions that return either yes (when the initiating argument is undefeated) or no (when the initiating argument is defeated).

Using the empathy judge, the initiating argument wins when the empathy for the initiating argument is greater than the r.e. for the undercuts. In other words, the initiating argument is not disabled (by Definition 14). Hence, the initiating argument can win if empathy is sufficient or if the undercuts have lower empathy than a defending argument.

Definition 19. *Given an argument tree T, and beliefbase Γ, the initiating argument wins according to the* **empathy judge**, *denoted $\mathsf{win_e}(\Gamma, T) = yes$, iff $Re(\Gamma, T) > 0$.*

The empathy judge effectively weighs the effect of undercuts, and by recursion undercuts to undercuts, on the initiating argument. This means that the empathy for an undercut can be too low for it to be regarded as sufficient grounds for it to defeat an argument. A consequence of this is that a tree that wins by the binary judge may loose by the empathy judge and vice versa.

Proposition 8. *Let T be a argument tree, let Γ be a beliefbase, and let a_r be the root of argument tree T. $\mathsf{win_e}(\Gamma, T) = yes$ iff for all $a_i \in \mathsf{Children}(T, a_r)$ either a_r e-voids a_i or a_i is e-disabled.*

Using the antipathy judge, the initiating argument wins when the antipathy for the initiating argument is less than the r.a. for the undercuts. Hence, the initiating argument can

	Tree	r.e.	r.a.	$\mathsf{win_b}$	$\mathsf{win_e}$	$\mathsf{win_a}$
Γ_1	T_1	-1	$1/3$	yes	no	no
Γ_1	T_3	1	-1	no	yes	yes
Γ_2	T_1	-1	$1/4$	yes	no	no
Γ_2	T_3	1	-1	no	yes	yes
Γ_3	T_1	$-1/2$	$1/2$	yes	no	no
Γ_3	T_3	0	$-1/2$	no	no	yes
Γ_4	T_1	$1/2$	$-1/2$	yes	yes	yes
Γ_4	T_3	$-1/2$	0	no	no	no

Table 1: Combinations of beliefbases and argument trees are evaluated with the binary judge, the empathy judge, and the antipathy judge, where $\Gamma_1 = \{\delta, \neg\alpha\}$, $\Gamma_2 = \{\delta, \neg\alpha, \gamma\}$, $\Gamma_3 = \{\neg\beta\}$ and $\Gamma_4 = \{\beta\}$ are beliefbases, and T_1 from Example 9 and T_3 from Example 11.

win if antipathy for it is sufficiently low or if undercuts have higher antipathy than a defending argument.

Definition 20. *Given an argument tree T, and beliefbase Γ, the initiating argument wins according to the* **antipathy judge***, denoted* $\mathsf{win_a}(\Gamma, T) = yes$*, iff* $Ra(\Gamma, T) < 0$.

The antipathy judge effectively weighs the effect of undercuts, and by recursion undercuts to undercuts, on the initiating argument. This means that the antipathy for an undercut can be too high for it to be regarded as sufficient grounds for it to defeat an argument. Furthermore, this antipathy to the undercut could be sufficient for the intiating argument to win because the undercut has "alienated" the intended audience irrespective of the other undercuts. A consequence of this is that a tree that wins by the binary judge may loose by the antipathy judge and vice versa.

Proposition 9. *Let T be a argument tree, let Γ be a beliefbase, and let a_r be the root of argument tree T.* $\mathsf{win_c}(\Gamma, T) = yes$ *iff there is an* $a_i \in \mathsf{Children}(T, a_r)$ *such that a_i c-voids a_r and a_i is not c-disabled.*

The combined judge classifies a tree as undefeated iff both the empathy judge and the antipathy judge classify the tree as undefeated.

Definition 21. *Given an argument tree T, and beliefbase Γ, the initiating argument wins according to the* **combined judge***, denoted* $\mathsf{win_c}(\Gamma, T) = yes$*, iff* $\mathsf{win_e}(\Gamma, T) = yes$ *and* $\mathsf{win_a}(\Gamma, T) = yes$.

Argument trees with initiating arguments that win by the combined judge are more believable than those that do not.

Proposition 10. *Let* $T_1, T_2 \in \mathsf{Trees}(\Delta, \alpha)$ *and let Γ be a beliefbase. If* $\mathsf{win_c}(\Gamma, T_1) = yes$ *and* $\mathsf{win_c}(\Gamma, T_2) = no$*, then* $T_2 \not\preceq_\Gamma T_1$.

A necessary condition for $\mathsf{win_e}(\Gamma, T) = yes$ or $\mathsf{win_c}(\Gamma, T) = yes$ is for $\Gamma \neq \emptyset$. Furthermore, the empathy Γ has for the root of T has to be greater than 0. These are natural expectations to have about believable argumentation. If the audience has zero empathy for the initiating argument, then the consequence of the initiating argument should not be regarded as an acceptable inference.

Discussion

The proposal for evaluating believability of arguments is a way for argumentation systems to interact intelligently with users. Given a beliefbase that reflects the intended audience, the values for empathy/antipathy provide an intuitive ranking over arguments. The definitions for r.a. and r.a. extend this ranking to argument trees. This may then be used to optimize the presentation of argumentation for audiences as part of decision-support technology. It may also be used to define judges that capture important criteria for deciding when an initiating argument should be an inference.

Whilst the presentation is based on a particular approach to logic-based argumentation, the proposal could be adapted for a range of other logic-based approaches to argumentation. It also seems that the proposal in this paper could be generalised for dialectical argumentation where arguments are generated by different agents in a debate or dialogue in a multi-agent system. Furthermore, the beliefbase of an audience could be constructed dynamically as part of a dialogue involving an agent asking questions to determine some of the beliefs of the audience before presenting arguments.

Other approaches to evaluating potentially inconsistent information, such as information-theoretic measures (Lozinskii 1994), possibilistic measures (Dubois, Lang, & Prade 1994), epistemic action measures (Konieczny, Lang, & Marquis 2003), and measures based on four valued models (Hunter 2003), may give interesting alternatives to the empathy and antipathy definitions.

References

Besnard, P., and Hunter, A. 2001. A logic-based theory of deductive arguments. *Artificial Intelligence* 128:203–235.

Chesnevar, C.; Maguitman, A.; and Loui, R. 2001. Logical models of argument. *ACM Computing Surveys* 32:337–383.

Dalal, M. 1988. Investigations into a theory of knowledge base revision: Preliminary report. In *Proceedings of the 7th National Conference on Artificial Intelligence (AAAI'88)*, 3–7. MIT Press.

Dubois, D.; Lang, J.; and Prade, H. 1994. Possibilistic logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3. Oxford University Press. 439–513.

Hunter, A. 2003. Evaluating the significance of inconsistency. In *Proceedings of the International Joint Conference on AI (IJCAI'03)*, 468–473. Morgan Kaufmann.

Konieczny, S.; Lang, J.; and Marquis, P. 2003. Quantifying information and contradiction in propositional logic through epistemic actions. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI'03)*, 106–111.

Lozinskii, E. 1994. Information and evidence in logic systems. *Journal of Experimental and Theoretical Artificial Intelligence* 6:163–193.

Prakken, H., and Vreeswijk, G. 2000. Logical systems for defeasible argumentation. In Gabbay, D., ed., *Handbook of Philosophical Logic*. Kluwer.