How Much Do I Argue Like You? Towards a Metric on Weighted Argumentation Graphs

Markus BRENNEIS [a,1], Maike BEHRENDT [a] Stefan HARMELING [a] and Martin MAUVE [a]
[a] *Department of Computer Science, University of Düsseldorf, Germany*

Abstract. When exchanging arguments with other people, it is interesting to know who of the others has the most similar opinion to oneself. In this paper, we suggest using weighted argumentation graphs that can model the relative importance of arguments and certainty of statements. We present a pseudometric to calculate the distance between two weighted argumentation graphs, which is useful for applications like recommender systems, consensus building, and finding representatives. We propose a list of desiderata which should be fulfilled by a metric for those applications and prove that our pseudometric fulfills these desiderata.

Keywords. argumentation graphs, online argumentation, metric

1. Introduction

In real-world discussions, people exchange arguments on a dedicated issue, such as improving the course of study [11], the distribution of funds [7], or which party to vote for at the next general election. In all those cases, participants discuss positions like "there should be a universal basic income" or "special math courses should be introduced", state their pro and contra arguments and attack other people's arguments.

Each individual participant of an argumentation has a personal view on the arguments and their relative importance: Users can decide for themselves which arguments they consider more convincing, thus which arguments they agree to and how much they agree with a statement. They may consider some positions more important than others.

Based on those individual views, there are useful applications for measuring the similarity or distance between two users: Clustering can be used to find representatives for a group of people with similar argumentation behavior or for finding a consensus. Another application is opinion polling, where one wants to find out why two persons or organizations come to different conclusions. What is more, collaborative filtering, which needs some definition of distance between users, can be used for pre-filtering arguments in applications like Kialo[2].

[1]Corresponding Author: Markus Brenneis, Heinrich-Heine-Universität, Universitätsstraße 1, 40225 Düsseldorf, Germany; E-mail: markus.brenneis@uni-duesseldorf.de.

[2]https://kialo.com/

In this paper, we propose solutions to the two main challenges to achieving the goal of comparing the argumentations of two users: We define weighted argumentation graphs which are a suitable representation of argumentation covering the mentioned aspects, including importance of arguments and agreement with statements. Secondly, we suggest a pseudometric for calculating the distance between two weighted argumentation graphs, which considers the specific structure of argumentation graphs (e.g., opinions deeper in a graph are less important). We contribute a list of useful desiderata for a metric which compares argumentations, and prove that our pseudometric fulfills those properties.

In the following chapter, we present our definition of weighted argumentation graphs. The third chapter introduces our pseudometric and desiderta for a useful metric. Finally, we discuss some limitations of our pseudometric and take a look at related work.

2. Definition of Weighted Argumentation Graphs

To be able to determine the similarity of real-world argumentations, there has to be a suitable representation of them. This representation should be able to capture all aspects mentioned in the introduction, and it should be as simple as possible. Therefore, the following definition is based on the IBIS model [12], which has been successfully tested with users without background in argumentation theory using our D-BAS system [11].

For the application purposes described in the introduction, the model should be able to represent the known opinions and arguments of a person as close as possible. Thus, we use *statements*, not arguments, as atomic elements in our definition, which then can be composed to arguments which can support or attack another statement. Note, though, that this definition can be translated to classical abstract argumentation frameworks based on Dung's definition [5], e.g. using DABASCO [14].

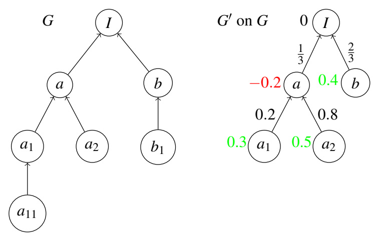
Let S be the (finite) set of all statements, with the special statement $I \in S$. The set $A \subseteq S \setminus \{I\} \times S$ is a set of arguments. For an argument $a = (s_1, s_2) \in A$, s_1 is called premise, s_2 conclusion of a. Let $s \in S$, then $a_{\to s} := \{(t,u) \in A \mid u = s\}$ is the set of arguments with conclusion s.

Note that I is excluded to be a premise since it is the *issue* in the IBIS model. We refer to I as "personal well-being", which allows us to interpret an edge like (b,I) in Figure 1 as "My personal well-being improves, because more wind power plants will be built." The premises of arguments with conclusion I are called positions. Positions are actionable items like "A wall between Austria and Germany should be built," and play an important role in real-world argumentation, e.g. decision-making problems [7].

Definition 1 (Argumentation Graph)**.** *An* argumentation graph *is a directed, weakly connected graph $G = (S,A)$, $A \subseteq S \setminus \{I\} \times S$, where the statements are nodes and the arguments edges, and there is exactly one $I \in S$ which has no outgoing edges.*

Note that this model does not include different relations for attack and support, as known from bipolar argumentation frameworks. Whether an argument is supportive are not, is up to a person's interpretation of the natural language representation of the arguments. The purpose of the model is solely to capture the hierarchy of statement, which we later need for our metric; bipolarity would add unnecessary complexity in this paper.

Every person can have a personal view with personal attitudes on a common argumentation graph G, as depicted in G' in Figure 1. To get an intuition for our next defini-



a	A wall between Austria and Germany should be built.
a_1	The wall stops illegal immigration of cows.
a_{11}	Cows can come via Switzerland.
a_2	The wall is expensive.
b	More wind power plants should be built.
b_1	Wind power is a renewable energy source.

Figure 1. Example for an argumentation graph G and a weighted argumentation graph G' on G with positions a and b and concrete examples for each statement. Edges with weight 0 and nodes with rating 0 are not drawn. Statement ratings are next to nodes, values for relative argument importance next to edges.

tion, let us have a look at what we can deduce about Alice's attitudes from her graph G': Alice strongly accepts position b (rating .4) and is slightly against a (rating $-.2$). She accepts the statement a_2 more than statement a_1 (rating $.5 >$ rating .3). The counterargument (a_2, a) "No wall should be built, because a wall is expensive." is far more important for her than the argument (a_1, a) (relative importance $.8 > .2$).

Furthermore, it makes sense to sort the positions: She considers building more wind power plants (b) more important than building a wall (a, relative importance $\frac{2}{3} > \frac{1}{3}$). So when comparing her attitudes with someone else's attitudes, she would consider a contrary opinion on b more severe than a different opinion on a. For ordinary statements, which are not positions, having an importance does not make sense: One cannot say that "The wall stops illegal immigration of cows" is twice as important as "Cows can come via Switzerland" (important regarding what?); one can only say that the arguments regarding building a wall which are built by those statements are of differing importance.

We will use real numbers to represent those weights and ratings.

Definition 2 (Weighted Argumentation Graph)**.** *Let $G = (S, A)$ be an argumentation graph. A* weighted argumentation graph *G' on G is a quadruple (S, A, r, w) with functions r and w. $r\colon S \to [-0.5, 0.5]$ assigns an agreement score (rating) to every statement, where negative values mean disagreement, 0 no opinion/don't care, and positive values agreement. $w\colon A \to [0,1]$ assigns an importance weight to each argument. The value indicates the importance of that argument relative to other arguments with the same conclusion. The value 0 means that the argument is not used (i.e. has no relevance), and 1 means that the argument is the only relevant argument for the conclusion. The following conditions must hold:*

$$\forall s \in S \sum_{a \in A_{\to s}} w(a) \in \{0, 1\} \tag{1}$$

$$r(I) = 0 \tag{2}$$

Formula 1 means that the sum of weights of arguments with the same conclusion is 1 if there is an argument with positive weight (cf. (a_1, a) and (a_2, a) in Figure 1); the sum is 0 iff no argument for a common conclusion has a weight (cf. $w(b_1, b) = 0$ in Figure 1). This assures that w represents *relative*, not absolute importance. To simplify notation, we write $w(\cdot,\cdot)$ instead of $w((\cdot,\cdot))$. If the underlying argumentation graph G happens to be a directed tree, we call G' a weighted argumentation tree.

w and r can be represented as matrix or vector, respectively, where undefined values are set to the default value 0. For the example in Figure 1, one gets:

$$w = \begin{pmatrix} 0 & 0 & 0\,0\,0\,0\,0 \\ 0.33 & 0 & 0\,0\,0\,0\,0 \\ 0.67 & 0 & 0\,0\,0\,0\,0 \\ 0 & 0.2 & 0\,0\,0\,0\,0 \\ 0 & 0.8 & 0\,0\,0\,0\,0 \\ 0 & 0 & 0\,0\,0\,0\,0 \\ 0 & 0 & 0\,0\,0\,0\,0 \end{pmatrix}, \qquad r = \begin{pmatrix} 0 \\ -0.2 \\ 0.4 \\ 0.3 \\ 0.5 \\ 0 \\ 0 \end{pmatrix} \tag{3}$$

The entry in row i column j of w is the weight of the argument with premise i and conclusion j. The first column and first row must refer to I as premise or conclusion, respectively. Because of Formula 1, the column sum is always 0 or 1.

If we draw or talk about a weighted argumentation graph, "non-existing" edges a are edges with $w(a) = 0$ (argument with no importance), and "non-existing" nodes s are nodes with $r(s) = 0$ (neutral statement). An example is G' shown in Figure 1.

The importance of a position p is represented as weight of the "argument" (p,I) leading to the "personal well-being" I. An application could obtain weights and ratings from a user, for example, by asking them to mark statements which are considered more important, or sorting arguments by relevance, which we do in our *deliberate* system [4].

3. Proposal of a Pseudometric for Weighted Argumentation Graphs

We now propose a pseudometric for calculating a distance between two weighted argumentation graphs, and prove several properties we expect of a function which compares two argumentations. The goal of the metric is to indicated how close the opinions and used arguments of two persons are, considering graph structure and individual assessments of importance; we do *not* want to compare argumentations on abstract levels like consistency, number of arguments used, or if other person's arguments are countered.

3.1. The Pseudometric

We define a distance measure of two weighted argumentation graphs $G_1 = (S,A,r_1,w_1)$, $G_2 = (S,A,r_2,w_2)$ on $G = (S,A)$ as:

$$d_G(G_1,G_2) = (1-\alpha)\sum_{i=1}^{\infty} \alpha^i \|w_1^i[:,1] \odot r_1 - w_2^i[:,1] \odot r_2\|_1 \tag{4}$$

where $\alpha \in (0,1)$ determines the influence of opinions deeper in the graph: A lower α emphasizes opinions on statements $r(s)$, a higher value the similarity of the argumentation underneath a statement s. $w[:,1]$ denotes the first column of the weight matrix w, α^i is the i-th power of the scalar α, w^i the i-th power of the square matrix w, and $\odot$ the Hadamard (entrywise) product. The i-th summand calculates the contribution of the paths with length i ending at I. We drop the index G if the underlying argumentation graph is clear from the context.

The intuition behind this distance measure becomes clearer when rephrasing it for the special case of argumentation *trees* (which have no cycles or re-used statements, thus unique paths from each statement to I). In case of argumentation trees T_1 and T_2 on T, the distance d_G is equivalent to:

$$d_T(T_1,T_2) = (1-\alpha)\sum_{s\in S}\alpha^{\texttt{depth}(s)}\left|r_1(s)\prod_{a\in\rho_{I\to s}}w_1(a) - r_2(s)\prod_{a\in\rho_{I\to s}}w_2(a)\right| \tag{5}$$

$\rho_{s_1\to s_2}$ is the sequence of all arguments (edges) on the path from s_1 to $s_2 \in S$ (where s_2 is deeper in the tree). $\texttt{depth}(s) = |\rho_{I\to s}|$ is the length of the path from I to s, i.e. the number of arguments; $\texttt{depth}(I) = 0$.

The terms in the absolute value measure the similarity of the opinions of a statement s as difference of their ratings, scaled with the product of the "importances" of the arguments leading to s. Hence, statements which are deeper in the argumentation tree get a smaller weight, and the overall relevance of an argumentation branch is limited to its importance.

To see how the calculation works and that the results match intuition, let us calculate the distance between the graphs G' (Figure 1), T_2, and T_3 (Figure 2) for $\alpha = 0.5$. We can expect that T_2 is closer to T_3 than to G', because the opinions on the statements a and b match and only the weights are different. The results confirm the expectation:

$$\begin{aligned} d(T_2,G') =& (1-\alpha)\left(\alpha^1\left|0.5\cdot 0.6-(-0.2)\cdot\frac{1}{3}\right| + \alpha^1\left|0.5\cdot 0.4 - 0.4\cdot\frac{2}{3}\right|\right.\\ &\left. + \alpha^2\left|0\cdot 0\cdot 0.6 - 0.3\cdot 0.2\cdot\frac{1}{3}\right| + \alpha^2\left|0\cdot 0\cdot 0.6 - 0.5\cdot 0.8\cdot\frac{1}{3}\right|\right) = 0.194 \end{aligned} \tag{6}$$

$$d(T_2,T_3) = (1-\alpha)(\alpha^1|0.5\cdot 0.6 - 0.5\cdot 0.3| + \alpha^1|0.5\cdot 0.4 - 0.5\cdot 0.7| = 0.075 \tag{7}$$

Note that the value of d_G and d_T is in $[0,1)$. If d is the depth of T, the maximum value of $d_T(T_1,T_2)$ is $\alpha(1-\alpha^d)$.[3] The maximum value of d_G is $\lim_{n\to\infty}(1-\alpha)\sum_{i=1}^n\alpha^i = \alpha$.

Theorem 1. *Let G be an argumentation graph. d_G is a pseudometric, i.e. has the following properties for all weighted argumentation graphs G_1, G_2, G_3 on G:*

(i) $\quad d_G(G_1,G_1) = 0$

(ii) $\quad d_G(G_1,G_2) = d_G(G_2,G_1)$ *(symmetry)*

(iii) $\quad d_G(G_1,G_3) \le d_G(G_1,G_2) + d_G(G_2,G_3)$ *(triangle inequality)*

Proof. d_G converges: $\sum_i \alpha^i$ is geometric series, which converges for $\alpha \in (0,1)$. The value of the L_1 norm cannot be greater than 1 because for each column sum σ of the w^i, we always have $0 \le \sigma \le 1$.

(i) holds because the same values are subtracted in the L_1 norm.

(ii) is given since the L_1 norm is symmetric.

As each summand fulfills the triangle inequality, d_G also fulfills (iii). □

However, d_G is not a metric, because $d_G(G_1,G_2)$ can be 0 even if G_1 is not equal to G_2: Consider G_1 where all statements are agreed to and every argument weight is 0, and G_2 where all statements have a rating of 0. Because weights and ratings are multiplied, the distance is 0 though the weighted argumentation graphs are different.

[3] Remember that $r_1(\texttt{root}(T)) = r_2(\texttt{root}(T))$ because $\texttt{root}(T) = I$, and $r(I) := 0$.

3.2. Desiderata for a Metric for Weighted Argumentation Trees

Our pseudometric is only one of many possible metrics for the applications described. We present a list of intuitive desiderata which, as we think, should be fulfilled by any metric comparing two person's argumentations, and thus should be considered when constructing alternative metric proposals. It is, however, hard to capture intuitive properties in graphs which can contain circular references and re-used statements. Therefore, we focus on argumentation trees in this section. Field experiments like [11] have also shown that users seldom create cycles or re-use statements in different branches in real discussions.

After each desideratum, we prove that our pseudometric (Formula 5) fulfills it for weighted argumentation trees. We consider those properties important in many real-world application domains of a metric, albeit not everywhere, as pointed out in Section 5.

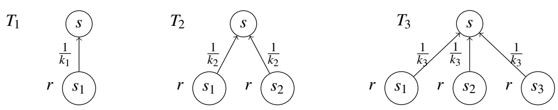
For each desideratum, we indicate why we think it is intuitive. Most desiderata are followed by a visual example making the choice of variable names clearer. Note that each tree in the examples is considered to be part of a bigger weighted argumentation tree, i.e. not all existing nodes and edges are drawn, and irrelevant statement ratings and argument weights are left out.

Desideratum 1 (Proportionally bigger overlap is better)**.** *Consider trees T_1, T_2, T_3, where T_2 is like T_1, but uses one additional argument for a statement s, and T_3 is like T_2, but uses one additional argument for s. Although T_2 and T_1 differ in only one argument, and T_3 and T_2 differ in only one argument, we expect $d(T_1,T_2) > d(T_2,T_3)$ because T_2 and T_3 have a greater overlap regarding the used arguments.*

More formally: For every statement s in a tree T which only has leaves $s_1,\dots,s_n$ ($n > 2$) as premises for s with $r(s_1) = \dots = r(s_n) \neq 0$ and $w(s_1,s) = \dots = w(s_n,s)$ and $\forall a \in \rho_{I\to s} : w(a) \neq 0$, consider the trees T_k, $n \geq k > 0$, which only contain $s_1,\dots,s_k$. Then, given $k_1 < k_2 < k_3$, we want to have $\frac{k_1}{k_2} < \frac{k_2}{k_3} \implies d(T_{k_1},T_{k_2}) > d(T_{k_2},T_{k_3})$, i.e. if the relative overlap of the number of arguments used is greater, the distance is smaller. Likewise, we demand $\frac{k_1}{k_2} > \frac{k_2}{k_3} \implies d(T_{k_1},T_{k_2}) < d(T_{k_2},T_{k_3})$, and $\frac{k_1}{k_2} = \frac{k_2}{k_3} \implies d(T_{k_1},T_{k_2}) = d(T_{k_2},T_{k_3})$.

We require $\forall a \in \rho_{I\to s} : w(a) \neq 0$ (i.e. no argument (edge) with weight 0 along the path from I to s), because if a user gives an argument a weight of 0, they say the premise is not related to the conclusion, thus not related to the topic of the discussion. This means that the user actually does not care about the opinions underneath that argument, which may be treated as if no opinion has been given.

In the following example, we have $k_1 = 1$, $k_2 = 2$, and $k_3 = 3$, thus $\frac{1}{2} < \frac{2}{3} \implies d_T(T_1,T_2) > d_T(T_2,T_3)$:



Proof. Only the argument weights for (s_i,s) (namely $\frac{1}{k_1}$, $\frac{1}{k_2}$, $\frac{1}{k_3}$, respectively, as used in (8) = (9)) are different and contribute to the sum, all other summands are zero. The common weight products and common values for r are summarized as $\overline{w_r}$ for readability and are factored out. For (10) > (11), remember that $\frac{k_1}{k_2} < \frac{k_2}{k_3}$.

$$d_T(T_{k_1},T_{k_2}) = \overline{w_r}(1-\alpha) \sum_{s' \in \{s_1,\dots,s_n\}} \alpha^{\mathtt{depth}(s')} |w_{k_1}(s',s) - w_{k_2}(s',s)| \tag{8}$$

$$= \overline{w_r}(1-\alpha) \sum_{s' \in \{s_1,\dots,s_n\}} \alpha^{\mathtt{depth}(s')} \left(k_1 \cdot \left(\frac{1}{k_1} - \frac{1}{k_2}\right) + (k_2-k_1) \cdot \frac{1}{k_2}\right) \tag{9}$$

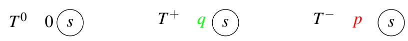
$$= \overline{w_r}(1-\alpha) \sum_{s' \in \{s_1,\dots,s_n\}} \alpha^{\mathtt{depth}(s')} \left(2 - 2\frac{k_1}{k_2}\right) \tag{10}$$

$$> \overline{w_r}(1-\alpha) \sum_{s' \in \{s_1,\dots,s_n\}} \alpha^{\mathtt{depth}(s')} \left(2 - 2\frac{k_2}{k_3}\right) = d_T(T_{k_2},T_{k_3}) \tag{11}$$

The other cases are proven by replacing ">" with "<" or "=", respectively. □

Desideratum 2 (Contrary opinion is worse than no opinion)**.** *Consider trees T_1, T_2, T_3, where all trees are identical, but T_1 has no opinion on a statement s, T_2 a positive opinion on s and T_3 a negative opinion on s. As we definitely know that T_2 and T_3 disagree on s, we want to have $d(T_2,T_3) > d(T_1,T_2)$.*

Formally: For any statement s in a tree T with $\forall a \in \rho_{I\to s} : w(a) \neq 0$, let T^+ be like T but with $r^+(s) = q > 0$, T^- like T with $r^-(s) = p < 0$, and T^0 like T with $r^0(s) = 0$. Then $d(T^+,T^-) > d(T^+,T^0)$.



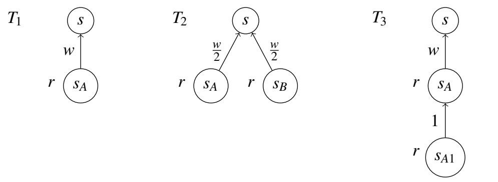
Proof. The only positive summand is the summand for s (which has rating 0, q or p for T^0, T^+, and T^-, respectively). The argument weights and the α term are common to all summands, can be factored out, and are summarized as $\overline{w_\alpha}$.

$$d_T(T^+,T^-) = \overline{w_\alpha} \cdot |q-p| = \overline{w_\alpha} \cdot (q+p) > \overline{w_\alpha} \cdot |q-0| = d_T(T^+,T^0) \tag{12}$$

□

Desideratum 3 (Deviation in deeper parts has less influence than deviation in higher parts)**.** *Consider the trees T_1,T_2,T_3, where T_1 has an argument (s_A,s) with no children and $\forall a \in \rho_{I\to s} : w(a) \neq 0$, T_2 is constructed from T_1 by adding a new statement s_B and argument (s_B,s) with $w_2(s_B,s) = w_2(s_A,s) = \frac{w_1(s_A,s)}{2}$, and T_3 is constructed from T_1 by adding a new statement s_{A1} and argument (s_{A1},s_A) with $w_3(s_{A1},s_A) = 1$. If $r(s_A) = r(s_B) = r(s_{A1}) \neq 0$, then we want to have $d(T_1,T_2) > d(T_1,T_3)$, because arguments deeper in the tree should have a smaller influence since we consider them less important for the overall opinion.*

We require $r(s_A) = r(s_B) = r(s_{A1}) \neq 0$ because adding a statement with "don't care" opinion should not actually change the distance. We want equality because this desideratum should cover only differences in the depth of the statements, not their rating.



Proof. The only differences are the summands including s, s_A and s_{A1}. Let $r := r(s_A) = r(s_B) = r(s_{A1}) \neq 0$ and $w := w_1(s_A, s)$. As before, we summarize common values for weights and α as $\overline{w_\alpha}$.

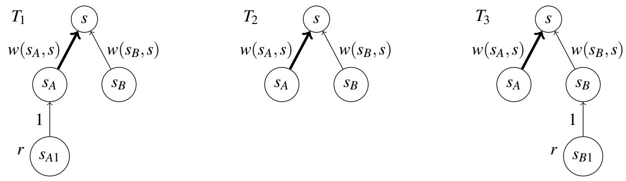
$$d_T(T_1, T_2) = \overline{w_\alpha} \cdot \left(\left| rw - r\frac{w}{2} \right| + \left| 0 - r\frac{w}{2} \right| \right) = \overline{w_\alpha} \cdot |rw| \tag{13}$$

$$> \overline{w_\alpha} \cdot \alpha |rw| = \overline{w_\alpha} \cdot (\alpha \cdot |r \cdot 0 \cdot w - r \cdot 1 \cdot w|) = d_T(T_1, T_3) \tag{14}$$

□

Desideratum 4 (Influence of deeper parts depends on weights in higher parts)**.** *Consider trees T_1, T_2, T_3, where all trees are identical, have statements s_A and s_B with conclusion s, and $w(s_A, s) > w(s_B, s)$, but T_1 has an additional argument for s_A and T_3 has an additional argument for s_B. Although the difference in both cases is only one argument, we expect $d(T_1, T_2) > d(T_3, T_2)$ because (s_A, s) has a larger weight.*

Formally: Let T_2 be a weighted argumentation tree with arguments $(s_A, s), (s_B, s)$ and $w(s_A, s) > w(s_B, s)$, no premises for s_A and s_B and $\forall a \in \rho_{I \to s} : w(a) \neq 0$. T_1 is constructed from T_2 by adding (s_{A1}, s_A) and T_3 from T_2 by adding (s_{B1}, s_B), each with a weight of 1 *and $r_1(s_{A1}) = r_3(s_{B1}) \neq 0$. Then $d_T(T_1, T_2) > d_T(T_3, T_2)$.*



Proof. Let $r := r_1(s_{A1}) = r_3(s_{B1}) \neq 0$. Only the summand which includes s_{A1} or s_{B1}, respectively, contributes a value greater than 0.

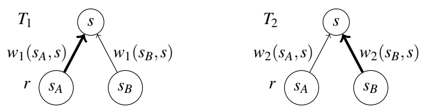
$$d_T(T_1, T_2) = \overline{w_\alpha} \cdot |w(s_A, s) \cdot r - 0| > \overline{w_\alpha} \cdot |0 - w(s_B, s) \cdot r| = d_T(T_2, T_3) \tag{15}$$

□

Desideratum 5 (Weights of arguments have influence even if they are the only difference)**.** *Consider trees T_1, T_2, where all trees are identical and have the arguments (s_A, s) and (s_B, s), but the weights are different: $w_1(s_A, s) \neq w_2(s_A, s)$ and $w_1(s_B, s) \neq w_2(s_B, s)$. We want to have $d(T_1, T_2) > 0$ if there exists a statement s' below s_A (or s_A itself) with $r_1(s') = r_2(s') \neq 0$ and $\forall a \in \rho_{I \to s'} : w(a) \neq 0$.*

We demand $r_1(s') = r_2(s') \neq 0$ because it makes sense if weights leading only to statements which are rated as “don’t care” are ignored.

In this example, we have $s_A = s'$:



Proof. It is enough to show that there is at least one summand greater than 0.

$$|w_1(s_A,s) - w_2(s_A,s)| > 0 \tag{16}$$

$$\implies \alpha^{\texttt{depth}(s')}(1-\alpha)|w_1(s_A,s) - w_2(s_A,s)| > 0 \tag{17}$$

$$\implies \alpha^{\texttt{depth}(s')}(1-\alpha) r_1(s') \prod_{s_{A'} \in \rho_{s' \to I} \setminus (s_A,s)} w_1(s_{A'}) |w_1(s_A,s) - w_2(s_A,s)| > 0 \tag{18}$$

$$\implies \alpha^{\texttt{depth}(s')}(1-\alpha) \left| r_1(s') \prod_{s_{A'} \in \rho_{s' \to I}} w_1(s_{A'}) - r_2(s') \prod_{s_{A'} \in \rho_{s' \to I}} w_2(s_{A'}) \right| > 0 \tag{19}$$

For (18) $\implies$ (19), remember that all weights in $\rho_{s' \to I} \setminus (s_A,s)$ and all ratings are the same for T_1 and T_2, e.g. $r_1(s') = r_2(s')$. □

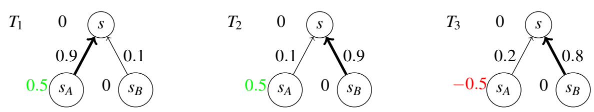
Desideratum 6 (Symmetry regarding negation of opinion)**.** *Let T_1, T_2 be any weighted argumentation trees, and T_3, T_4, respectively, the same trees, but the opinion for each statement is negated, i.e. $r_3(s) = -r_1(s)$ and $r_4(s) = -r_2(s)$ for all $s \in S$. We expect that a metric is symmetric regarding negation, i.e. $d(T_1,T_2) = d(T_3,T_4)$.*

Proof. This holds because $|r_1(s) - r_2(s)| = |(-r_3(s)) - (-r_4(s))| = |r_3(s) - r_4(s)|$. □

Desideratum 7 (Trade-off between argument weights and agreement)**.** *Consider trees T_1, T_2, T_3 which are nearly identical and have leaf statements s_A and s_B with common conclusion s and $\forall a \in \rho_{I \to s} : w(a) \neq 0$. We have $r_1(s_A) = r_2(s_A) = 0.5$, $r_3(s_A) = -0.5$, $r_1(s_B) = r_2(s_B) = r_3(s_B) = 0$, and $w_1(s_A,s) > w_2(s_A,s)$. Furthermore, $w_2(s_B,s) = w_1(s_B,s) + w_1(s_A,s) - w_2(s_A,s)$, $w_3(s_B,s) = w_1(s_B,s) + w_1(s_A,s) - w_3(s_A,s)$, i.e. s_B is neutral and "collects" remaining weight such that the sum is 1.*

If $w_1(s_A,s) - w_2(s_A,s) > w_2(s_A,s) + w_3(s_A,s)$, although T_1 and T_2 have the same opinion on s_A, we want to have $d(T_1,T_2) > d(T_2,T_3)$, because both T_2 and T_3 do not care much about their (different) opinions on s_A. Likewise, if $w_1(s_A,s) - w_2(s_A,s) < w_2(s_A,s) + w_3(s_A,s)$, we expect $d(T_1,T_2) < d(T_2,T_3)$ because the weights w_1 and w_2 are closer to each other and give a greater weight for opposing opinions on s_A.

The following example trees depict the first case with concrete weight values. Because the different opinions on statement s_A are underneath an argument edge with small weight, we want to have $d(T_1,T_2) > d(T_2,T_3)$.



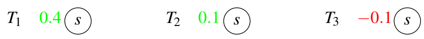
Proof. We proof the first case. For (20) = (21), remember that $w_1(s_A,s) > w_2(s_A,s)$.

$$\begin{aligned} d(T_1,T_2) &= \overline{w_\alpha} \cdot |0.5 \cdot w_1(s_A,s) - 0.5 \cdot w_2(s_A,s)| && (20) \\ &= \overline{w_\alpha} \cdot 0.5 \cdot (w_1(s_A,s) - w_2(s_A,s)) && (21) \\ &> \overline{w_\alpha} \cdot 0.5 \cdot (w_2(s_A,s) + w_3(s_A,s)) && (22) \\ &= \overline{w_\alpha} \cdot |0.5 \cdot w_2(s_A,s) - (-0.5) \cdot w_3(s_A,s)| = d(T_2,T_3) && (23) \end{aligned}$$

The other case follows by replacing "$>$" with "$<$". $\square$

Desideratum 8 (Trade-off between statement ratings and agreement)**.** *Consider trees T_1, T_2, T_3 which are nearly identical and have a statement s and $\forall a \in \rho_{I \to s} : w(a) \neq 0$. We have $r_1(s) > r_2(s) > 0 > r_3(s)$ such that $|r_1(s) - r_2(s)| > |r_2(s) - r_3(s)|$. Although T_1 and T_2 have the same positive opinion on s, we want to have $d(T_1, T_2) > d(T_2, T_3)$, because both T_2 and T_3 have a weak opinion on s. Likewise, if $|r_1(s) - r_2(s)| < |r_2(s) - r_3(s)|$, we expect $d(T_1, T_2) < d(T_2, T_3)$ because the ratings $r_1(s)$ and $r_2(s)$ are closer to each other than $r_2(s)$ and $r_3(s)$.*

The first case, $d(T_1, T_2) > d(T_2, T_3)$, is shown in the following example:



Proof. Only the summand for s contributes to the distance, all other summands are 0. Remember that all weights on the path to s are positive.

$$d(T_1, T_2) = \overline{w_\alpha}|r_1(s) - r_2(s)| > \overline{w_\alpha}|r_2(s) - r_3(s)| = d(T_2, T_3) \tag{24}$$

The other case follows by replacing "$>$" with "$<$". $\square$

Desideratum 9 (Weights limit the influence of a path)**.** *Consider graphs T_1, T_2 which are nearly identical, have an argument (s_A, s) with $w = w_1(s_A, s) = w_2(s_A, s)$ and only the ratings and weights below (and including) s_A may differ in any way. No matter how those values are chosen, we want to have $d(T_1, T_2) \leq w$, i.e. the maximum influence of the differences below s_A is limited by the weight of s_A to its conclusion.*

Proof. It is enough to consider paths which include a, since the summands for all other parts are 0. Let S_a be the set of all statements which have an argument leading to a, including a itself. We abbreviate $\alpha^{\texttt{depth}(s')}(1 - \alpha)$ as $\overline{\alpha}(s')$.

$$d_T(T_1, T_2) = \sum_{s' \in S_a} \overline{\alpha}(s') \left| r_1(s') \prod_{s_{A'} \in \rho_{S \to I}} w_1(s_{A'}) - r_2(s') \prod_{s_{A'} \in \rho_{S \to I}} w_2(s_{A'}) \right| \tag{25}$$

$$= w \sum_{s' \in S_a} \overline{\alpha}(s') \left| r_1(s') \prod_{s_{A'} \in \rho_{S \to I} \setminus (s_A, s)} w_1(s_{A'}) - r_2(s') \prod_{s_{A'} \in \rho_{S \to I} \setminus (s_A, s)} w_2(s_{A'}) \right| \tag{26}$$

As the factor after w is in $[0, 1]$, we get $d_T(T_1, T_2) \leq w$. $\square$

4. Limitations

Although the proposed pseudometric fulfills several intuitive desiderata, there are some limitations which we will discuss in the following.

If a weight of an argument is 0, the proposed pseudometric ignores all weights which are underneath this argument. When comparing how similar people argue for or against

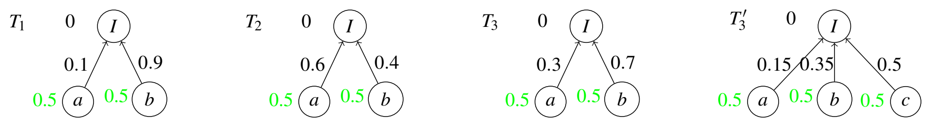


Figure 2. Possible unexpected change in order if an unrelated opinion is added: $0.05 = d(T_1,T_3) < d(T_2,T_3) = 0.075$, but $0.1375 = d(T_1,T_3') > d(T_2,T_3') = 0.125$ with $\alpha = 0.5$

the top-level positions, this is okay, but if also the way how arguments which are not supported are attacked should influence the distance, the metric has to be extended.

Moreover, ordering can be changed by adding an unrelated opinion. Consider trees T_1, T_2, T_3 with $d(T_1,T_3) < d(T_2,T_3)$. At first sight, it might seem unexpected that this order can be changed to $d(T_1,T_3') > d(T_2,T_3')$ by adding a new position c to T_3 and keeping the relative weights of the other positions. Depending on the application context, e.g. a voting advice application (VAA), this might be unwanted. An example is depicted in Figure 2. This is due to the normalization of the argument weights.

Although even end-user friendly systems like D-BAS support for undercuts, i.e. arguments that have an argument as conclusion, undercuts are currently not explicitly modeled in our model. This is no big problem, because in many applications, for instance, in a VAA, arguments can be preselected such that no undercut attack is necessary (since the arguments make sense), or rephrased such that the premise is attacked. For example, consider the argument “We should build more nuclear power plants because cats are cute” and the attack “Cats and nuclear power plants are unrelated”. Though this is technically an undercut, a user interface may present this as an attack on “Cats are cute”.

5. Related Work

Calculating the distance between argumentation graphs to compare how similar the attitudes of two agents are has already been used in other systems. The Carneades opinion formation and polling tool presented in [9] is able to compare one’s argumentation with the argumentation of other entities like organizations. This comparison is simply done by counting the number of statements where the agreement/disagreement is the same. This approach is much simpler than our proposal, but uses neither weights nor ratings and violates i.a., Desideratum 3.

The mobile application described in [1] also bases on IBIS and extends it with an agreement value for each argument in the argumentation tree. This information is used, for opinion prediction using collaborative filtering. In contrast to our work, the idea of relative argument importance in combination with statement rating is not present.

Another application of calculating the similarity of weighted trees is match-making of agents, which are represented by weighted trees. In [3], a recursive similarity measure for this application is proposed. Its parameter N serves a similar purpose to our parameter α. They also give examples which are similar to our desiderata, e.g. Example 4 is like our Desideratum 2. Nodes, however, do not have a weight, and some desiderata are not fulfilled; for instance, Desideratum 5 is explicitly not demanded in their Example 2.

There are already other definitions of weighted argumentation graphs based upon Dung’s definition of argumentation frameworks, but many lack the differentiation between argument relevance and statement ratings (e.g. [2,8,10,6]), which we think is important since argument weights limit how much a branch of an argumentation is relevant, whereas statement ratings are only relevant for the single statement. Furthermore, in most cases, there is a global assignment of values in the graph and no user-specific views, which is why a strength of 0 for attacks would be meaningless in the model presented in [13]. Note that most related work in this field is concerned about evaluating consistency or calculating extensions, whereas our main goal is comparing the attitudes of agents, not caring about whether an agents' attitude is logically consistent or not.

6. Conclusion and Future Work

In this paper we proposed a pseudometric to calculate the distance between two argumentation graphs representing the attitudes of different persons, and several desiderata which should be considered when proposing other metrics for the same purpose, and are fulfilled by our pseudometric. Possible next steps include developing other sensible metrics and comparing them regarding theoretical properties and practicality.

In future work, we want to check if the desiderata are not only intuitive for experts in argumentation theory, but are following the intuition of untrained humans. We also want to test the metric in a VAA to compare the argumentation of voters with those of parties. Thereby we see whether the results of the metric are accepted in an application context.

References

[1] Althuniyan, N., Sirrianni, J.W., Rahman, M.M., Liu, X.F.: Design of mobile service of intelligent large-scale cyber argumentation for analysis and prediction of collective opinions. In: International Conference on AI and Mobile Services. pp. 135–149. Springer (2019)

[2] Amgoud, L., Ben-Naim, J., Doder, D., Vesic, S.: Acceptability semantics for weighted argumentation frameworks. In: IJCAI. vol. 2017, pp. 56–62 (2017)

[3] Bhavsar, V.C., Boley, H., Yang, L.: A weighted-tree similarity algorithm for multi-agent systems in e-business environments. Computational Intelligence **20**(4), 584–602 (2004)

[4] Brenneis, M., Mauve, M.: deliberate – online argumentation with collaborative filtering. In: COMMA (2020), to appear

[5] Dung, P.M.: 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] Dunne, P.E., Hunter, A., McBurney, P., Parsons, S., Wooldridge, M.: Weighted argument systems: Basic definitions, algorithms, and complexity results. Artificial Intelligence **175**(2), 457–486 (2011)

[7] Ebbinghaus, B.: Decision making with argumentation graphs. arXiv preprint arXiv:1908.03357 (2019)

[8] Gordon, T.F., Walton, D.: Formalizing balancing arguments. In: COMMA. pp. 327–338 (2016)

[9] Gordon, T.F.: Structured consultation with argument graphs. From Knowledge Representation to Argumentation in AI. A Festschrift in Honour of Trevor Bench-Capon on the Occasion of his 60th Birthday pp. 115–133 (2013)

[10] Hunter, A.: A probabilistic approach to modelling uncertain logical arguments. International Journal of Approximate Reasoning **54**(1), 47–81 (2013)

[11] Krauthoff, T., Meter, C., Mauve, M.: Dialog-Based Online Argumentation: Findings from a Field Experiment. In: Proceedings of the 1st Workshop on Advances in Argumentation in Artificial Intelligence. pp. 85–99 (November 2017)

[12] Kunz, W., Rittel, H.W.J.: Issues as elements of information systems, vol. 131. Citeseer (1970)

[13] Martınez, D.C., Garcıa, A.J., Simari, G.R.: An abstract argumentation framework with varied-strength attacks. In: Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning (KR'08). pp. 135–144 (2008)

[14] Neugebauer, D.: Dabasco: Generating af, adf, and aspic+ instances from real-world discussions. In: COMMA. pp. 469–470 (2018)