Artificial Intelligence Research and Development
A. Cortés et al. (Eds.)

doi:10.3233/FAIA220354

Combining Support and Attack Interactions for Argumentation Based Discussion Analysis

Teresa ALSINET [1], Josep ARGELICH and Ramón BÉJAR
INSPIRES Research Center – University of Lleida
Jaume II, 69 – 25001 Lleida, SPAIN

Abstract
Online discussion is one of the commonly used tools to engage users to discuss relevant topics for society, where user contributions in the form of posts, comments and votes are essential to their success. The scale, complexity and dynamism of this information leads to a growing interest in understanding what are the major accepted or rejected opinions in different domains by social network users. In this work, we explore how to combine attack and support interactions to extract possible consensus based on some abstract argumentative semantics, and we characterize the set of properties or postulates that this consensus should satisfy.

Keywords. Online discussion, probabilistic weighted interactions, valued arguments, support and attack relations, postulates for consensus analysis.

1. Discussion model

Our goal in this work is to consider a general online discussion platform and to reason, by means of an argumentative approach, about the set of posts that can be accepted as consensus among the participants by combining both the social relevance of posts and the degrees of belief in the answers between them.

Argumentation includes various forms of dialogue such as deliberation and negotiation, which are concerned with collaborative decision-making procedures by which people can express and rationally resolve or at least manage their disagreements. An abstract argument framework, as proposed by Dung [11], is a graph structure in which the nodes denote arguments and the edges denote attacks between the arguments. When we are considering online discussions, another kind of interaction may exist between posts. Indeed, a post can attack another post, but it can also support another one. In addition, it is common for each post to have a degree of popularity, preference or social support, such as the votes it receives throughout the discussion.

To represent the characteristics of online discussions, our proposal is based on graphs extended with weights for both nodes and edges. Each post gives rise to a node in the

[1] Correspondence to: T. Alsinet. INSPIRES Research Centre, University of Lleida. C/Jaume II, 69. Lleida, Spain. Tel.: +34 973702734; E-mail: teresa.alsinet@udl.cat.

graph, and relationships between posts give rise to the edges of the graph expressing answers between them. We say that a post p_1 *answers* a post p_2 whenever p_1 is a direct reply to p_2 or p_1 mentions (refers to) p_2. So, a post can answer many posts. In what follows, Γ denotes a non-empty set of posts, and it is referred to as a discussion from a root (main) post.

Definition 1 *(Weighted Discussion Graph) A weighted discussion graph* (WDisG) *for a discussion* Γ *is a directed graph* $G = \langle N, E, L, W \rangle$*, where*

- *For every post* p_i *in* Γ*, there is a node* n_i *in* N*.*
- *If post* p_1 *answers post* p_2*, there is a directed edge* (n_1, n_2) *in* E*.*
- L *is a labelling function* $L : E \to [0,1]^3$ *for edges in* E*. The labelling function* L *maps an edge* (n_1, n_2) *to a triple of probability values* $(p_a, p_s, p_n) \in [0,1]^3$ *with* $p_a + p_s + p_n = 1$*, which expresses the probability or degree of belief that the answer from post* p_1 *to post* p_2 *can be classified as attack (*p_a*), support (*p_s*) and none (*p_n*), respectively. Attack means that the comment expressed in the post* p_1 *criticizes or disagrees with the claim expressed in the post* p_2*,* $support$ *that* p_1 *agrees with the claim expressed in* p_2 *and* $none$ *that the relation is none of the previous two.*
- W *is a weighting function* $W : N \to [0,1]$ *for nodes in* N*. The weighting function* W *maps the node* n_i *of a post* p_i *in* Γ *to a value in the real interval* $[0,1]$*, which expresses the social (support) acceptance degree that the comment on the post has received throughout the discussion with respect to the rest of comments.*

2. Support and attack interactions

Once we have introduced the formal representation of discussions as Weighted Discussion Graphs, the next key component is the classification of the interactions between nodes according to the labelling function L of edges and the weighting function W of nodes. In our approach, we define two types of interactions between nodes, which we refer to as attack (R_{att}) and support (R_{sup}) relations, and we consider an uncertainty threshold $\alpha \in (0,1]$ which characterizes how much uncertainty about classification we are prepared to tolerate.

Fuzzy argumentation frameworks deal with the uncertainty of arguments or attacks caused by incompleteness or vagueness, and its semantics have been studied in various ways. In [12] the authors propose a fuzzy argumentation approach by enriching the expressive power of the classical argumentation model proposed by Dung [11], by allowing to represent the relative strength of the attack relationships between arguments, as well as the degree to which arguments are accepted. They define extensions as fuzzy sets by aggregating the strength of the attack with the degree of acceptance of arguments. Although our approach is not based on fuzzy sets semantics, we also propose to use an aggregation operation to define the attack and support relations.

Definition 2 *(Support and Attack Interactions) Let* $G = \langle N, E, L, W \rangle$ *be a* WDisG *for a discussion* Γ *and let* $\alpha \in (0,1]$ *be a threshold on the distances of the probability values. Moreover, let* $\wedge : [0,1] \times [0,1] \to [0,1]$ *be a t-norm; i.e., a binary aggregation operation satisfying: monotone, associative, commutative, and* $\forall x \in [0,1], 1 \wedge x = x \wedge 1 = x$*. We define two binary relations on* N *as follows:*

- *A binary relation R_{sup} on N called the* support relation*:*
  $R_{sup} = \{(n_1, n_2) \in E \mid L(n_1, n_2) = (p_a, p_s, p_n)$ *with*
  $p_s \geq \max(p_a, p_n) + \alpha$ *and* $W(n1) \wedge p_s \geq W(n_2)\}$.
- *A binary relation R_{att} on N called the* attack relation*:*
  $R_{att} = \{(n_1, n_2) \in E \mid L(n_1, n_2) = (p_a, p_s, p_n)$ *with*
  $p_a \geq \max(p_s, p_n) + \alpha$ *and* $W(n1) \wedge p_a \geq W(n_2)\}$.

Notice that the support and attack relations verify that $R_{sup} \cap R_{att} = \emptyset$ and $R_{sup} \cup R_{att} \subseteq E$; i.e., the answer between two posts is neither a support nor an attack interaction, or, if it is, it is either a support or an attack interaction, but not both.

Finally, from R_{sup} and R_{att}, we recursively define the set of posts that reinforce (*support*) or contradict (*attack*) a post $n_i \in N$ in the following way:

- $\mathit{support}(n_i) = \{n_i\} \cup \{n_1 \mid (n_1, n_2) \in R_{sup} \cap (N \times \mathit{support}(n_i))\}$
- $\mathit{attack}(n_i) = \{n_1 \mid (n_1, n_2) \in R_{att} \cap (N \times \mathit{support}(n_i))\}$

Then, we say that a discussion Γ is *coherent* whenever $\mathit{support}(n_i) \cap \mathit{attack}(n_i) = \emptyset$, for all node $n_i \in N$.

3. Rationality postulates

The basic idea of argumentation theory is to construct arguments in favour and against a statement (in our approach, a post), to select the "acceptable" ones, accepted as consensus among the participants, and, finally, to determine whether the original statement (root post) can be accepted or not.

In [1] we defined a reasoning system for analysis of discussions on Twitter, where each tweet is represented by an argument and the notion of acceptability is based on Value-based Abstract Argumentation. Value-based Abstract Argumentation [6], attach arguments with social values, and makes the semantics dependent on a particular preference order over values, representing a particular audience. While our reasoning system is useful in applications as user profile analysis [2], it can lead to some unintuitive results, since the acceptability semantics does not consider support relationships between arguments. In [1] the authors take into account support interactions by reinforcing the social values of arguments propagating support relationships between tweets.

In order to consider a more intuitive acceptability semantics for online discussions, in this preliminary work, we are interested in defining principles, as the rational postulates proposed by Caminada and Amgoud [8] for rule-based argumentation systems, [7] for non-monotonic reasoning with strict and defeasible rules, and [4] for weighted bipolar settings.

Definition 3 *(Postulates) Let R_{sup} and R_{att} be the support and attack relations, respectively, for a* WDisG $G = \langle N, E, L, W \rangle$ *of a coherent discussion* Γ. *An acceptable consensus $A \subseteq N$ of Γ should satisfy:*

- *Consistency, meaning that* $\neg\exists\ (n_1, n_2) \in A \times A : \exists\ n_3 \in \mathit{attack}(n_2)$ *and* $n_1 \in \mathit{support}(n_3)$.
- *Closure, meaning that* $\neg\exists\ n_i \in A : \mathit{support}(n_i) \not\subseteq A$.
- *Self-defence, meaning that* $\forall\ (n_1, n_2) \in R_{att} \cap ((N \backslash A) \times A) : \exists\ (n_3, n_4) \in R_{att} \cap (A \times \mathit{support}(n_1))$.

Notice that *consistency* ensures both *direct* (i.e., $\neg\exists\ (n_1, n_2) \in A \times A : n_1 \in \mathit{attack}(n_2)$) and *indirect* consistency (i.e., $\neg\exists\ (n_1, n_2) \in A \times A : \exists\ n_3 \in \mathit{attack}(n_2)\ \mathit{and}\ n_1 \in \mathit{support}(n_3)$).

Finally, another feature or postulate that should satisfy an acceptable consensus $A \subseteq N$ is that of maximality, meaning that $\neg\exists\ n_i \in (N \backslash A) : A \cup \{n_i\}$ satisfies consistency, closure and self-defence.

In the literature, we find different approaches to incorporate support between arguments in the context of Abstract Argumentation Frameworks. In [9], the authors introduce Bipolar Abstract Argumentation, extending the defeat relation with indirect attacks, and in [10] acceptability semantics are proposed based on enforcing coherence of the admissible sets and taking into account attack and support interactions for proposing gradual labelling for the arguments. On the other hand, weighted semantics [3,5] focus on the evaluation of individual arguments rather than sets of arguments and assign a weight to each argument, allowing for a fine-grained classification of the acceptance and rejection of arguments. As future work, we intend to explore these approaches in order to extract an order of preference for the possible acceptable consensus sets of a discussion.

Acknowledgments The authors would like to thank the anonymous reviewers for providing helpful comments. This work was partially funded by Spanish Project PID2019-111544GB-C22, by the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement 723596 and Grant Agreement 768824, and by 2017 SGR 1537.

References

[1] Alsinet, T., Argelich, J., Béjar, R., Fernández, C., Mateu, C., Planes, J.: An argumentative approach for discovering relevant opinions in twitter with probabilistic valued relationships. Pattern Recognition Letters 105, 191–199 (2018)

[2] Alsinet, T., Argelich, J., Béjar, R., Martínez, S.: User profile analysis in reddit debates. In: Artificial Intelligence Research and Development, CCIA 2019. vol. 319, pp. 275–284. IOS Press (2019)

[3] Amgoud, L., Ben-Naim, J.: Ranking-based semantics for argumentation frameworks. In: SUM 2013. Lecture Notes in Computer Science, vol. 8078, pp. 134–147. Springer (2013)

[4] Amgoud, L., Ben-Naim, J.: Evaluation of arguments in weighted bipolar graphs. Int. J. Approx. Reason. 99, 39–55 (2018)

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

[6] Bench-Capon, T.J.M.: Persuasion in practical argument using value-based argumentation frameworks. Journal of Logic and Computation 13(3), 429–448 (2003)

[7] Caminada, M.: Rationality postulates: Applying argumentation theory for non-monotonic reasoning. FLAP 4(8) (2017)

[8] Caminada, M., Amgoud, L.: On the evaluation of argumentation formalisms. Artif. Intell. 171(5-6), 286–310 (2007)

[9] Cayrol, C., Lagasquie-Schiex, M.: On the acceptability of arguments in bipolar argumentation frameworks. In: Proceedings of ECSQARU 2005. pp. 378–389 (2005)

[10] Cayrol, C., Lagasquie-Schiex, M.: Bipolar abstract argumentation systems. In: Argumentation in Artificial Intelligence, pp. 65–84. Springer (2009)

[11] 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)

[12] Janssen, J., Cock, M.D., Vermeir, D.: Fuzzy argumentation frameworks. In: Information Processing and Management of Uncernainty in Knowledge-Based Systems, IPMU'2008 (2008)