Fostering Explainable Online Review Assessment Through Computational Argumentation

Atefeh Keshavarzi Zafarghandi, Davide Ceolin

Human-Centered Data Analytics, Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

Abstract

Explainable methods have received increased attention within artificial intelligence. Wherever an automated system makes a decision an explanation is required to convince a user about the decision. Furthermore, online information quality assessment is crucial to help users navigate information. However, explaining the assessment of online information had not been clarified well. The current work provides explanations to a user about the assessment of online information and specific, provides explanations for the quality assessments of online reviews. We construct an abstract argumentation framework (AF) based on a set of given reviews. We consider the grounded semantics of AFs to assess each topic. Then, we discuss the question of why a score can be assigned to a topic of a product. Furthermore, we indicate a proper score of a review based on the scores of topics within the review in question. We also collect arguments that can support the chosen score of a review.

Keywords

Online reviews, Abstract Argumentation frameworks, Explainable Artificial Intelligence

1. Introduction

Argumentation is one research area that is frequently mentioned in explainable AI [1]. To our knowledge, the role of argumentation formalisms in the sense of explanation for the assessment of the quality of reviews has not been investigated in depth. In this work we aim to cover this gap, i.e., to use an argumentation formalism not only for assessing the quality of online information but also as a means to explain why a given score is assigned by a system to a topic or a review. This is particularly compelling because online reviews are available in large amounts, but for them to be beneficial to users, their quality needs to be determined. Given the volume of the information at stake here, an automated approach is necessary to address the problem properly. Given that the result of this automated assessment is meant to be used by humans, the ability to explain the assessment process is likewise crucial. Computational argumentation is, in our opinion, a promising methodology in this sense.

Abstract argumentation frameworks (AFs for short), as introduced by Dung [2] is a directed graph in which nodes represent arguments and edges denote attack relation among arguments. In this work, we use AFs as a means of modeling and evaluating reviews and as an explanation for a chosen decision.

1st International Workshop on Argumentation for eXplainable AI (ArgXAI, co-located with COMMA '22), September 12, 2022, Cardiff, UK

akz@cwi.nl (A. Keshavarzi Zafarghandi); davide.ceolin@cwi.nl (D. Ceolin)

0000-0002-5806-1012 (A. Keshavarzi Zafarghandi); 0000-0002-3357-9130 (D. Ceolin)

In the work presented in this paper, we aim to model a set of reviews with AFs and use the grounded semantics of AFs as a means of evaluation and explanation. That is, this model allows reasoning on the 'acceptability of topics within reviews' when reviews are conflicting with each other. Such conflicts are solved by defining a weight for topics within reviews, which extends and generalizes a review weight we defined in previous work [3]. To this end, we first consider each set of reviews containing the same topic with the same score as a single argument. Then, we introduce an attack relation if the weight of arguments is not the same. We first use the grounded semantics of AFs to assess the score of each topic, and secondly to explain why a system chooses a specific score for a topic in question. That is, we aim at providing the final user with an explanation of why a specific score for a topic in question is acceptable (or trustable). In other words, we aim to indicate what the score of a topic about a target/product is based on the set of reviews. Some specific features of our work are as follows:

1. We consider all topics within a review, instead of only picking the most important one. Thus, we do not miss any information presented in a review.
2. By our method, after the assessment, a user can ask about the score (strength) of topics of a product, instead of just asking about the scores of the reviews.
3. Our method is solid enough to explain to a user the reason of assigning a score to a topic. Also, explaining why an argument does not have a roll in the assessment of the score of a topic.
4. We also accumulate the scores of topics within a review to assess a review score. Furthermore, we present an explanation for a review score from a machine point of view, which is a subset of the grounded semantics of AFs. Thus, this explanation is beyond any doubt.
5. For indicating the scores of topics about a product within a set of reviews, based on a system's points of view, we do not need to consider any generalization of AFs. Since the weight of each argument is used to indicate the direction of the attack relation.

The rest of the paper is structured as follows. Section 2 presents related work, and Section 3, introduces abstract argumentation frameworks. Section 4 presents a model to formally represent and reason on reviews. Section 5 concludes and outlines future work directions.

2. Related Work

Explainable artificial intelligence (XAI) has received increased attention to explain decisions of automated systems [4]. Several machine learning methods are used to support decision-making. However, these methods are required to convince a user about the machine decision. In other words, an AI system needs to explain its decision to a user. Argumentation theory can help the process of explanation, (see [1, 5] for a survey). Some argumentation frameworks with respect to their use in support of explainable artificial intelligence (XAI) are presented in [6]. In [7] a new type of argumentation semantics is presented for AFs for capturing explanation. Furthermore, argumentation is used to explain why and/or whether a certain argument can be accepted under certain semantics [8, 9, 10, 11].

On the other hand, a significant amount of research investigates the quality of reviews. In [3] a generalization of abstract argumentation framework is used to assess the quality of the reviews on a product, extending and combining preferred argumentation frameworks [12, 13] and valued-based argumentation frameworks [14, 15] to model and evaluate a set of reviews. Furthermore, to analyze product reviews, different methods for generating probability distributions over constellations of arguments graph have been presented in [16, 17]. In [17] authors assume that an agent(s) specifies a belief in the acceptability status of arguments. However, in [16] the authors propose a scoring method for identifying the probability distribution for a review. After the extraction of support and attack relations between reviews, a set of reviews is modeled by an abstract dialectical framework [18] which generalizes AFs [2].

3. Background: Argumentation Formalisms

We start the preliminaries of our work by recalling the basic notion of Dung's abstract argumentation frameworks (AFs) [2].

Definition 1. [2] An abstract argumentation framework (AF) is a pair (A, R) in which A is a set of arguments and $R \subseteq A \times A$ is a binary relation representing attacks among arguments.

Let $F = (A, R)$ be a given AF. For each $a, b \in A$, the relation $(a, b) \in R$ is used to represent that a is an argument attacking the argument b. An argument $a \in A$ is, on the other hand, defended by a set $S \subseteq A$ of arguments (alternatively, the argument is acceptable with respect to S) (in F) if for each argument $c \in A$, it holds that if $(c, a) \in R$, then there is a $s \in S$ such that $(s, c) \in R$ (s is called a defender of a).

Different extension-based semantics of AFs present which sets of arguments in a given AF can be accepted jointly [2]. We only recall grounded semantics here,[1] because it is proven in [2] that 1. every AF has a unique grounded extension, 2. there is no doubt on the acceptance of the arguments in the grounded extension, 3. in any acyclic finite AF all sets of semantics coincide, 4. and in this work we only have acyclic frameworks.

Set $S \subseteq A$ is called a *conflict-free* set (extension) (in F) if there is no $a, b \in S$ such that $(a, b) \in R$. The *characteristic function* $F : 2^A \mapsto 2^A$ is defined as $F(S) = \{a \mid a \text{ is defended by } S\}$. A set $S \in cf(F)$ is the grounded extension in F if S is a unique fixed point of $F^n(\emptyset)$.

Example 1. *Let $F = (\{a, b, c\}, \{(a, b), (b, c)\})$ be an AF. In F, (a, b) means that argument a attacks b, and (b, c) means that b attacks c. Here, argument c is defended by set $\{a\}$ (alternatively, c is acceptable with respect to $\{a\}$), since a attacks the attacker of c, namely b. The set of conflict-free sets of F is $cf(F) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, c\}\}$. A unique grounded extension of F is $\{a, c\}$. The intuition is that a is not attacked by any argument, thus no one has any doubt about accepting argument a. Argument c is attacked by b, however, it is defended by a which was accepted by everyone.*

[1] The reader interested in semantics of AFs can see [2].

4. Modeling Reviews with Formal Argumentation

Here, we first model reviews in a formal manner by means of abstract argumentation frameworks (AFs), and then show how to use AFs to evaluate the score of a topic of a product based on given reviews (see Section 4.2). To this end, we collect the reviews with the same score containing the same topic and we consider them as a single argument. After indicating the attack relations between arguments, we construct an AF. Then, we use the grounded semantics of AFs for indicating the most trusted arguments and for a reason of explaining why a score is assigned to a topic and a review in our automated method.

4.1. Formal Modeling of Reviews

Let t be a product (target) and let $\{\mathrm{r}_i(t)\}$ be a set of reviews of t (alternatively, $\{\mathrm{r}_{ti}\}$). Each review r_{ti} consists of a numerical score $sc(\mathrm{r}_{ti})$ (e.g., in 5-level Likert scale) and a textual description. The description characterizes the product focusing to specific topics. E.g., in the case of a pair of shoes, the topics can be 'sole', 'upper' or 'comfort'. We represent the list of all topics relevant to the product t as $\mathcal{T}_t$, where $\mathcal{T}_t = \{\phi_1, \ldots, \phi_n\}$.

Each review r_{ti} contains a finite set of topics $\mathcal{T}_{\mathrm{r}_{ti}} \subseteq \mathcal{T}_t$. E.g., review '$\mathrm{r}_{t1}$: the color of the shoes are not my favorite but I use it for a long time and they are still looking good' contains two topics, i.e., $\mathcal{T}_{\mathrm{r}_{t1}} = \{\phi_1 = \text{color}, \phi_3 = \text{quality}\}$. The review score represents the actual reviewer's opinion on the product, while the text aims at motivating such judgment. An agreeing score of two reviews containing the same topic about the same product indicates a support between reviews while disagreeing score indicates a conflict.

Our goal is to identify the score of each topic of a product, based on the set of given reviews, i.e., which score is proper for a topic in question. We address this by constructing an AF based on a set of given reviews. Then, we use the grounded semantics of AFs to indicate a score to a topic in question and explain why it is the case. In the end, we further update the scores of reviews based on their topics. We also, prepare an explanation for the updated score of the reviews by using the grounded semantics of AFs.

In review r_{ij} in Table 1, index i indicates the product and j indicates the reviewer. Reviews can support one another if they have the same score and they have a common topic. If a set of reviews has the same score and has a common topic, then it means that these reviews support one another. Thus, we consider them as a single argument. In Table 1, topic 'ϕ : quality', indicated by a topic detection, is a common topic between r_{21} and r_{22}, however, these reviews do not have the same score. Reviewer r_{21} gave a score of 4 out of 5 to this product. The last sentence in r_{21} presents the main reason why the reviewer gave a score of 4 out of 5 to this product because she/he was satisfied with the quality of this product. However, the reviewer of r_{22} is not as satisfied as the reviewer of r_{21} with the quality of this product. Thus, there is an attack relation between reviews r_{21} and r_{22}, because $sc(\mathrm{r}_{21}) \neq sc(\mathrm{r}_{22})$.

Reviews can be classified based on their topics. Assume that ϕ is a topic of a set of reviews $E = \{\mathrm{r}_{t1}, \ldots, \mathrm{r}_{tj}\}$, if there is $E' \subseteq E$, such that for any $\mathrm{r}_{ti}, \mathrm{r}_{tj} \in E'$, it holds that $sc(\mathrm{r}_{ti}) = sc(\mathrm{r}_{tj})$, then we say that reviews in E' support one another and we consider all reviews as a single argument. This leads to the classification of E based on the scores of reviews. Between arguments with different scores, there is a symmetric attack relation. If there is no intersection

Id	Score	Review Text
r_{11}	*****	Comfortable
r_{12}	*	Waaaay too BIG
r_{21}	****	Fit perfectly. I bought dark grey, and they didn't fade
r_{22}	***	They fit great. But they fade bad

Table 1
Example reviews. r_{11} and r_{12} are about the same product, and r_{21} and r_{22} are. The reviews, pairwise, disagree on the scores given. The topics tackled, though, are slightly different.

between topics within r_{ti} and r_{tj}, then there is no relation between these two reviews.

4.2. Modeling Reviews with AFs

Let $\{r_{t1}, \ldots, r_{tn}\}$ be a finite set of reviews on product t. Let $\mathcal{T}_t = \{\phi_i\}$ be the set of topics relevant to the product t. Each review r_{ti} consists of a numerical score $sc(r_{ti})$ (e.g., in 5-level Likert scale) and a textual description. Let $\mathcal{T}_{t,r_{ti}}$ be the set of all topics presented in review r_{ti}. Different topics have different importance in each review, to indicate this importance we introduce a weight function in Definition 2 which shows the initial weight of topic ϕ in review r_{ti} with score $sc(r_{ti})$.

Definition 2. Let r_{ti} be a review, let ϕ be a topic in r_{ti}, and let $sc(r_{ti})$ be a score of r_{ti}. $w(sc(r_{ti}), \phi, r_{ti})$ is called the *initial weight* of ϕ in review r_{ti} and score $sc(r_{ti})$, where w is a value in $[0, \infty]$ such that w is computed by aggregating one or multiple factors meeting the following criteria:

1. at least one of such factors is the result of an abstraction function computed on the review itself. Such abstraction functions should be computable over any review and allow establishing a total order of reviews. Example of such abstraction functions include, for instance, readability scores (e.g., Dale-Chall readability [19]) and complexity measures (e.g., Kolmogorov complexity [20]);

2. optional factors can be computed as abstractions over the combination of the review r_{ti} and/or its topic ϕ.

We aim to use a formalism of argumentation, i.e., abstract argumentation frameworks (AFs), to assess the score of the topics within reviews of a product t. To this end, we construct an AF based on a set of reviews. Our method associates each argument to the set of reviews with the same score about a common topic. That is, we consider all reviews with the same score that presents the same topic as a single argument. Formally, let $[\phi]_k = \{r_{ti} \mid r_{ti}$ contains topic ϕ and $sc(r_{ti}) = k\}$, i.e., $[\phi]_i$ collect all the reviews that contains topic ϕ and have score k. We consider $[\phi]_k$ as a single argument, that contains all the reviews that have topic ϕ and score k. We are interested in evaluating the weight of $[\phi]_k$ as a single argument. To this end, for each k with $k \in \{1, 2, 3, 4, 5\}$, we sum up the initial weights of topic ϕ in review r_{ti} and score k. Note that in Definition 3 if ϕ is not a topic of r_{ti} then $w(k, \phi, r_{ti}) = 0$.

Definition 3. Let ϕ be a topic of a product t, let k be a natural number between 1 and 5. The weight of topic ϕ with respect to score k, alternatively, the weight of argument $[\phi]_k$ is as follows:

$$w([\phi]_k) = \Sigma_{i=1}^{n} w(k, \phi, \mathrm{r}_{ti})$$

where n is the number of reviews about product t, and $w(k, \phi, \mathrm{r}_{ti})$ is the initial weight ϕ in review r_{ti} and score k as introduced in Definition 2.

To construct an AF based on a set of reviews, we consider an attack relation from an argument $[\phi]_i$ to $[\phi]_j$ iff $w([\phi]_i) > w([\phi]_j)$.

Definition 4. Let t be a product, let $\mathcal{T}_t = \{\phi_i\}$ be a set of topics, and let i be a natural number between 1 and 5. For each i, with $i \in \{1.2, 3, 4, 5\}$, and for each $\phi \in \mathcal{T}_t$, we introduce an argument $a_{i,\phi} = [\phi]_i$. Furthermore, $w(a_{i,\phi})$ is the weight of argument $a_{i,\phi}$, equal to $w([\phi]_i)$, as introduced in Definition 3. An AF constructed based on topics is $F = (A, R)$ where,

- $A = \{a_{i,\phi}\}$
- $R = \{(a_{i,\phi}, a_{j,\phi}) \mid a_{i,\phi}, a_{j,\phi} \in A$ and $w(a_{i,\phi}) > w(a_{j,\phi})\}$.

An AF, constructed based on topics of a product, is a directed graph. Each node, indicated by $a_{i,\phi}$ contains all sets of reviews that contain topic ϕ with score i. The reason for collecting all such reviews is that if two reviews contain the same topic ϕ and give the same score to the product, then their content support one another. Thus, we accumulate them in a single node and we consider them as one argument. After indicating the set of arguments from a given set of reviews, we designate attack relations. If two arguments give different scores to product t, but contain the same topic, it means that there may exist a conflict between these two arguments with respect to the topic in question. We consider an attack relation between $a_{i,\phi}$ and $a_{j,\phi}$ when their weights are not the same. In order to indicate the direction of the attack relation between two arguments we consider the weight of the topic with respect to the score, presented in Definition 3. Note that in Definition 4 we only consider relations among arguments with the same topics. Thus, for $\phi, \phi' \in \mathcal{T}_t$, if $\phi \neq \phi'$, then there is no relation between any arguments of $a_{i,\phi}$ and $a_{j,\phi'}$, for $i, j \in \{1, 2, 3, 4, 5\}$. That is, the associated graph to AF F, constructed based on the topics of a product, is a forest, presented formally in Lemma 1. Note that in graph theory a graph is called connected if for every pair of vertices a and b, there is a path between a and b. Furthermore, a connected component is a maximal connected subgraph of an undirected graph.

Definition 5. Let t be a product, and let $\mathcal{T}_t = \{\phi_i\}$ be a set of topics of t. Let $F = (A, R)$ be an AF constructed based on topics. Let $\phi \in \mathcal{T}_t$, set $a_\phi = \{a_{i,\phi}\}$ is called a *component* contains ϕ in F. Furthermore, component a_ϕ is called *connected component* iff for every $a_{i,\phi}, a_{j,\phi} \in a_\phi$ it holds that $(a_{i,\phi}, a_{j,\phi}) \in R$.

Lemma 1. *Let F be an AF, constructed based on topics of product t. If the reviews of the product t contain at least two topics, then the graph associated with F is disconnected. Furthermore, If m is the number of topics presented in the reviews, i.e., $|\mathcal{T}_t| = m$, then the associated graph of F contains at least m connected component.*

Proof. If the reviews contain more than one topic, for instance ϕ and ϕ', then it is clear that there is no relation between arguments containing ϕ and ϕ'. Thus, the associated graph is disconnected. That is, there is no link between the arguments of components a_ϕ and $a_{\phi'}$. Thus, if we have at least two topics in the set of reviews, then the associated graph is disconnected.

We now show that the associated graph contains at least m connected components if the set of reviews contains m number of different topics, by induction on m.

Base case: assume that $m = 1$, that is all reviews contain the single topic ϕ. If all reviews give the same score to the product, then we only have one argument in the AF constructed based on these reviews. Thus, we have exactly one component. Note that if we have different arguments with the same weight, i.e., $|a_\phi| > 1$ and $w(a_{i,\phi}) = w(a_{j,\phi})$ for every $a_{i,\phi}, a_{j,\phi} \in a_\phi$, then there are more than one connected component.

Inductive hypothesis: Assume that $m = j$, then the associated graph contains at least j connected components.

Inductive step: Assume that $m = j + 1$, then we have to show that the associated graph contains at least $j + 1$ components. Let K be j different topics of t, i.e., $K \subseteq \mathcal{T}_t$ and $|K| = j$. By the inductive hypothesis, the associated graph contains at least j connected components. Let ϕ be a topic such that $\phi \notin K$, and let $\phi' \in K$. Thus, there is no relation between arguments of a_ϕ and $a_{\phi'}$. Thus, if $m = j + 1$ the associated graph has at least $j + 1$ connected component. □

An AF $F = (A, R)$ is called *acyclic* (or well-founded) if there is no infinite sequence of arguments $a_1, \ldots, a_i, \ldots$ such that $(a_{i+1}, a_i) \in R$. By Definition 4, in an AF $F = (A, R)$, it holds that $(a_{i,\phi}, a_{j,\phi}) \in R$ iff $w(a_{i,\phi}) > w(a_{j,\phi})$. Furthermore, because of the transitive property of relation $<$ in real numbers, any AF constructed based on a set of reviews is acyclic. It is proven in [2] that in any acyclic AF all sets of semantics coincide. Thus, in the following of this work, to assess the topics, presented in reviews, we focus on the grounded extension of the constructed AF based on a given set of reviews.

Corollary 1. *Let F be an AF, constructed based on topics of the product t. Every connected component in the associated graph of F is acyclic.*

Proof. This corollary is the direct result of the fact that F is acyclic. Thus, any connected component of F is also acyclic. □

Since, by Corollary 1 every connected component in F is acyclic, every component has an initial argument, i.e., an argument that does not have any parents. Proposition 1 is the direct result of Definition 4.

Proposition 1. *Let F be an AF constructed based on a set of topics of the product t. Let $a_{i,\phi}$ be an initial argument of F. It holds that $w(a_{i,\phi})$ is maximum among the weights of other arguments.*

Lemma 2. *Let $F = (A, R)$ be an AF constructed based on a set of topics of the product t. Let I be the set of initial arguments of F, i.e., $I = \{b \mid \text{ there is no } a \in A \text{ such that } (a, b) \in R\}$. The grounded extension of F is none empty and it is equal with I.*

Proof. Since every AF constructed based on topics is acyclic and since each acyclic AF has a none empty grounded extension, F has a none empty grounded extension. By the definition of grounded semantics, every initial argument is in the grounded extension, i.e., $I \subseteq grd(F)$. It remains to show that $grd(F) \subseteq I$. Toward a contradiction, assume that $grd(F) \nsubseteq I$. That is, there exists an argument in $grd(F)$ which is not an initial argument. Assume that $a_{i,\phi}$ is an argument such that $a_{i,\phi} \in grd(F)$ but $a_{i,\phi} \notin I$. Since $grd(F) \neq \emptyset$ and $I \subseteq grd(F)$, there exists $a_{j,\phi} \in grd(F) \cap I$. By Proposition 1, $w(a_{j,\phi})$ is the maximum among the weights of other arguments, in specific, $w(a_{j,\phi}) > w(a_{i,\phi})$. Thus, by Definition 4, $(a_{j,\phi}, a_{i,\phi}) \in R$. That is, $a_{i,\phi}$ is attacked by $a_{j,\phi}$ in F. This is a contradiction of the assumption that $a_{i,\phi} \in grd(F)$. Thus, the assumption that there exists an argument in the grounded extension which is not an initial argument is wrong. Hence, $grd(F) = I$.

□

4.3. What is an AI system explanation of the score of a topic?

In this section, we explain how AFs as an AI system can be used to indicate the score of each topic of the product t, based on the set of reviews.

Proposition 2. *Let $F = (A, R)$ be the constructed AF based on topics of product t. Let ϕ be a topic of product t. There exists $i \in \{1, 2, 3, 4, 5\}$ such that $a_{i,\phi}$ is in the grounded extension of F.*

Proof. Let F be an AF, constructed based on topics of the product t. Let ϕ be a topic and let a_ϕ be the connected component of ϕ, as introduced in Definition 5. By Corollary 1, every connected component in the associated graph of F is acyclic. Thus, for each ϕ, a_ϕ contains an initial argument. By Lemma 2, the grounded extension of F and the set of initial arguments of F coincide. Thus, for each ϕ an argument of a_ϕ is in the grounded extension. Hence, for any ϕ, there exists i such that $a_{i,\phi}$ is in the grounded extension of F. □

We now define basic explanation in terms of functions. The function $sc_{AI}(-)$ is a unary function that takes a topic as an input and returns an appropriate score to that topic by considering all the reviews containing that topic, presented in Definition 6.

Definition 6. (Score of a topic) Let $F = (A, R)$ be an AF constructed based on topics of product t. Let ϕ be a topic of product t. The score ϕ based on an AI system is denoted by $sc_{AI}(\phi)$, defined as follows:

$$sc_{AI}(\phi) = \text{round}\left(\frac{\Sigma_{a_{i,\phi} \in grd(F)}\, i}{|\{i \mid a_{i,\phi} \in grd(F)\}|}\right)$$

In $sc_{AI}(\phi)$ the output of the function round is the nearest integer to $\frac{\Sigma_{a_{i,\phi} \in grd(F)}\, i}{|\{i \mid a_{i,\phi} \in grd(F)\}|}$.

Note that by Proposition 2 for each topic ϕ there exists at least an i such that $a_{i,\phi} \in grd(F)$. Hence, Definition 6 is well-defined. Note that for each ϕ, if component a_ϕ is connected, then there exists exactly one i such that $a_{i,\phi} \in grd(F)$. Intuitively, for a ϕ, if component a_ϕ is connected and $a_{j,\phi}$ is an initial argument, then it holds that $sc_{AI}(\phi) = j$. In this case, the choice of machine, i.e., $sc_{AI}(\phi) = j$ can be explained that the score of topic ϕ is j because argument $a_{j,\phi}$ is an initial argument, i.e., $w(a_{j,\phi})$ is maximum among all other arguments in component a_ϕ. The highest weight of $a_{j,\phi}$ among the arguments of a_ϕ means that sum of the weights of reviews containing ϕ with score j is the highest weight.

Furthermore, in the following, we try to explain to a user why from a machine point of view topic ϕ has the score $\mathrm{sc}_{AI}(\phi)$. To this end, we define an explanation function, denoted by $\mathrm{Exp}(\phi, \mathrm{sc}_{AI}(\phi))$ in Definition 7. Function $Exp(\phi, \mathrm{sc}_{AI}(\phi))$ is a binary function that takes a topic ϕ and its machine score $\mathrm{sc}_{AI}(\phi)$ as inputs and returns the set of arguments that their scores have a role in the machine decision in Definition 6.

Definition 7. (Explanation of a score of a topic) Let F be an AF, constructed based on topics $\{\phi_i\}$ of product t. Let ϕ be a topic and let $\mathrm{sc}_{AI}(\phi)$ be the score of ϕ from a machine point of view, as introduced in Definition 6. An explanation of why the score of topic ϕ is $\mathrm{sc}_{AI}(\phi)$ is as follows:

$$\mathrm{Exp}(\phi, \mathrm{sc}_{AI}(\phi)) = \{a_{i,\phi} \mid a_{i,\phi} \in \mathit{grd}(F)\}$$

Let ϕ be a topic, $\mathrm{Exp}(\phi, \mathrm{sc}_{AI}(\phi))$ collects all $a_{i,\phi}$ that is in the grounded extension of F, i.e., the set of arguments, with respect to topic ϕ, that the acceptance of them are beyond of any doubts. Assume that $\mathrm{sc}_{AI}(\phi) = i$, the notation $\neg\mathrm{sc}_{AI}(\phi)$ is a number j such that $j \neq i$.

Furthermore, the function $\mathrm{NotDef}(\phi)$ collects all the arguments that contain topic ϕ but do not have any role in the computation of the score of topic ϕ, from a machine point of view. In other words, $\mathrm{NotDef}(\phi)$ contains all arguments which are attacked by $\mathrm{Exp}(\phi, \mathrm{sc}_{AI}(\phi))$.

Definition 8. Let F be an AF, constructed based on topics $\{\phi_i\}$ of product t. Let ϕ be a topic and let $\mathrm{sc}_{AI}(\phi)$ be the score of ϕ from a machine point of view, as introduced in Definition 6. Function $\mathrm{NotDef}(\phi)$ is a unary function that takes a topic as an input and returns the set of arguments that do not have any role in the evaluation of $\mathrm{sc}_{AI}(\phi)$.

$$\mathrm{NotDef}(\phi) = \{a_{i,\phi} \mid a_{i,\phi} \notin \mathit{grd}(F)\}$$

For each topic ϕ, $\mathrm{Exp}(\phi, \mathrm{sc}_{AI}(\phi))$ and $\mathrm{NotDef}(\phi)$ are disjoint functions, i.e., $\mathrm{Exp}(\phi, \mathrm{sc}_{AI}(\phi)) \cup \mathrm{NotDef}(\phi) = a_{\phi}$. For each ϕ, if $a_{i,\phi} \notin \mathrm{Exp}(\phi, \mathrm{sc}_{AI}(\phi))$, then $a_{j,\phi} \notin \mathit{grd}(F)$, that is, there exists $a_{i,\phi} \in \mathit{grd}(F)$ such that $a_{i,\phi}$ attack $a_{j,\phi}$. That is, $w(a_{i,\phi}) > w(a_{j,\phi})$, thus, it is reasonable that the machine neither consider the score of $a_{j,\phi}$ in the evaluation of $\mathrm{sc}_{AI}(\phi)$ nor argument $a_{j,\phi}$ in $\mathrm{Exp}(\phi, \mathrm{sc}_{AI}(\phi))$. Furthermore, the function $\mathrm{SC}_{AI}(-)$ gives a score to a review r_{ti} based on the score of each topic within r_{ti}. That is, SC_{AI} revise the score of each review given by a reviewer.

Definition 9. (Score of a review) Let r_{ti} be a review of product t. Let $\mathcal{T}_{t,\mathrm{r}_{ti}}$ be a set of topics presented in review r_{ti}. The score of the review r_{ti} based on a machine assessment is the output of function $\mathrm{SC}_{AI}(\mathrm{r}_{ti})$, evaluated as follows:

$$\mathrm{SC}_{AI}(\mathrm{r}_{ti}) = \mathrm{round}\left(\frac{\Sigma_{\phi \in \mathcal{T}_{t,\mathrm{r}_{ti}}} \mathrm{sc}_{AI}(\phi)}{|\mathcal{T}_{t,\mathrm{r}_{ti}}|}\right)$$

The score of review r_{ti}, from a machine point of view, is the round of the average of scores of the topics presented in r_{ti}, where the score of each topic from a machine point of view is presented in Definition 6.

In Definition 10 grounded semantics of AFs are implicitly used to explain the reasons for the score given to a review from a machine's point of view.

Definition 10. (Explanation of a score of a review) Let r_{ti} be a review of product t. An explanation of why the score $\mathrm{SC}(\mathrm{r}_{ti})$ is given to a review r_{ti} is as follows:

$$\mathrm{Exp}(\mathrm{r}_{ti}, \mathrm{SC}(\mathrm{r}_{ti})) = \bigcup_{\phi \in \mathcal{T}_{t,\mathrm{r}_{ti}}} \mathrm{Exp}(\phi, \mathrm{sc}_{AI}(\phi))$$

In Definition 10, function $\mathrm{Exp}(\mathrm{r}_{ti}, \mathrm{SC}(\mathrm{r}_{ti}))$ collects all the explanations for all the topics within r_{ti}. In other words, this function collects all the initial arguments of components a_ϕ, where $\phi \in \mathcal{T}_{t,\mathrm{r}_{ti}}$. That is, $\mathrm{Exp}(\mathrm{r}_{ti}, \mathrm{SC}(\mathrm{r}_{ti})) \subseteq grd(F)$, thus, there is no doubt on the acceptance of these arguments.

5. Conclusion and Future Work

We present a method to consider reviews containing a comment topic with the same score as a single argument. We aggregate the initial weights of the reviews in an argument to indicate the main weight of the arguments. We introduce an attack relation between arguments by considering the weight of arguments. Then, we use the notion of grounded semantics of AFs to evaluate the most trusted arguments. Since each argument $a_{i,\phi}$ in the constructed AF is attached by $a_{j,\phi}$ if $j \neq i$ and $w(a_{j,\phi}) > w(a_{i,\phi})$, we evaluate the score of topics in the grounded extension. Next, we introduce a function to explain the reason for choosing the associated score of a topic by a system. Then, we present a function to accumulate the scores of topics within a review to assign a score to a review from a machine point of view. In the next step, as an explanation method, we also introduce a function to give a review and its score from a system point of view as inputs and returns all the arguments that have an effect on the assessment of a review in question from a system point of view.

In our approach, we identify an attack relation between arguments if they are about the same topic but their weights are different. As future work, we are eager to present relations among arguments that do not have a common topic (we are eager to consider some other features of arguments for presenting relations among them). Furthermore, we are interested in using the non-monotonic nature of argumentation theory for working on a temporal way of reasoning instead of considering a fixed set of reviews. That is, we aim to show how we can evaluate the quality of reviews if we consider the order of time of presenting of reviews. To have further human-machine interaction we aim to consider user preferences over the topics of a product to evaluate the score of a review. It is a possible topic for future work to use the generalizations of AFs, namely, valued-based AFs [14, 15] or ADFs [21], for modeling and assessing the quality of the set of the reviews.

Note that the tasks of argument mining from a given KR and automatically constructing of the associated AF are not within the scope of this work because we focus on studying the use of the grounded semantics of AFs as a means to assess product reviews. Given that this shows to be a promising direction, future work will focus on optimizing the task of AF construction by combining human and automated computation.

Acknowledgments This research has been supported by the Centrum Wiskunde & Informatica (CWI) and Supported by the Netherlands eScience Center project "The Eye of the Beholder" (project nr. $027.020.G15$).

References

[1] K. Cyras, A. Rago, E. Albini, P. Baroni, F. Toni, Argumentative XAI: A survey, in: IJCAI, ijcai.org, 2021, pp. 4392–4399.

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

[3] D. Ceolin, G. Primiero, J. Wielemaker, M. Soprano, Assessing the quality of online reviews using formal argumentation theory, in: ICWE, Springer, 2021, pp. 71–87.

[4] C. Lacave, F. J. Díez, A review of explanation methods for heuristic expert systems, Knowl. Eng. Rev. 19 (2004) 133–146.

[5] A. Vassiliades, N. Bassiliades, T. Patkos, Argumentation and explainable artificial intelligence: a survey, The Knowledge Engineering Review 36 (2021).

[6] M. Demollin, Q. Shaheen, K. Budzynska, C. Sierra, Argumentation theoretical frameworks for explainable artificial intelligence, in: 2nd Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence, NL4XAI@INGL 2020, Dublin, Ireland, December 18, 2020, Association for Computational Linguistics, 2020, pp. 44–49.

[7] X. Fan, F. Toni, On computing explanations in abstract argumentation, in: ECAI 2014 - 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic - Including Prestigious Applications of Intelligent Systems (PAIS 2014), volume 263 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2014, pp. 1005–1006.

[8] A. Borg, F. Bex, A basic framework for explanations in argumentation, IEEE Intell. Syst. 36 (2021) 25–35.

[9] Z. G. Saribatur, J. P. Wallner, S. Woltran, Explaining non-acceptability in abstract argumentation, in: ECAI, volume 325 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2020, pp. 881–888.

[10] X. Fan, F. Toni, On explanations for non-acceptable arguments, in: TAFA, volume 9524 of *Lecture Notes in Computer Science*, Springer, 2015, pp. 112–127.

[11] X. Fan, F. Toni, On computing explanations in argumentation, in: AAAI, AAAI Press, 2015, pp. 1496–1502.

[12] L. Amgoud, S. Vesic, Two roles of preferences in argumentation frameworks, in: ECSQARU, volume 6717, Springer, 2011, pp. 86–97.

[13] S. Modgil, Reasoning about preferences in argumentation frameworks, Artif. Intell. 173 (2009) 901–934.

[14] T. J. M. Bench-Capon, Value-based argumentation frameworks, in: NMR, 2002, pp. 443–454.

[15] T. J. M. Bench-Capon, Persuasion in practical argument using value-based argumentation frameworks, J. Log. Comput. 13 (2003) 429–448.

[16] K. Noor, A. Hunter, Analysing product reviews using probabilistic argumentation, in: COMMA, volume 326, IOS Press, 2020, pp. 295–306.

[17] A. Hunter, K. Noor, Aggregation of perspectives using the constellations approach to probabilistic argumentation, in: AAAI, AAAI Press, 2020, pp. 2846–2853.

[18] G. Brewka, S. Ellmauthaler, H. Strass, J. P. Wallner, S. Woltran, Abstract dialectical frameworks: An overview, in: Handbook of Formal Argumentation, College Publications, 2018, pp. 237–285.

[19] J. S. Chall, E. Dale, Readability revisited: The new Dale-Chall readability formula, Brookline Books, 1995.

[20] M. Li, P. Vitányi, et al., An introduction to Kolmogorov complexity and its applications, volume 3, Springer, 2008.

[21] G. Brewka, S. Woltran, Abstract dialectical frameworks, in: Proceedings of the Twelfth International Conference on the Principles of Knowledge Representation and Reasoning (KR 2010), 2010, pp. 102–111.