Why did they post that argument?
Communicative Intentions of Web 2.0 Arguments

Jodi SCHNEIDER, Serena VILLATA and Elena CABRIO

Abstract. The Web is an open platform where users are free to publish their own opinions, to discuss the latest news, to write reviews about a service or product. Given the growing number of online platforms allowing such interactions, it is becoming more and more important to support users in understanding the meaning of such conversations by identifying the context of the discourse, and the evidence and background knowledge required to evaluate the proposed arguments. In this paper, we suggest recording the rough context with communicative intentions to help classify arguments on the Web. We describe five communicative intentions: recreation, information, instruction, discussion, and recommendation. We suggest that this classification can help identify and analyze messages for the Argument Web. In particular, we discuss the suitable combinations of natural language processing techniques and argumentation-based reasoning to support users in understanding the discussions.

Keywords. online argumentation, genre, natural language processing

1. Introduction

In online platforms, people from all over the world share their comments and opinions. Every day billions of people add texts, photos, and videos on a wide range of topics like potholes, Rihanna’s fashion choices, and the Ukraine crisis. Some of these messages argue a point of view.

The goal of the Argument Web is “to encourage debate, to facilitate good argument, and to promote a new online critical literacy” [2]. The Argument Web will “[make] it possible to follow a line of argument...across disparate forums, comments, editorials and multimedia resources” [2]. Today, individuals can explicitly contribute to the Argument Web with bespoke tools, and existing Web content can also be analyzed in order to reconstruct arguments (e.g. [25,21]).

One emerging challenge is making the resulting Argument Web comprehensible and transparent to end-users. To promote online critical literacy, we must
remove barriers to understanding and processing competing opinions. So that people can follow a line of arguments across various different types of media – whether forums, comments, editorials or multimedia resources – we will need to understand which properties of messages are essential and which are superfluous. In this paper we argue that an argument’s communicative intention is a key consideration for incorporating it into a comprehensible Argument Web.

In the following, we first explain why form and dialogue type are not sufficient for distinguishing argumentative Web 2.0 messages (Section 2), leading us to suggest communicative intention as a way to classify arguments on the Web (Section 3). Then we analyze which types of frameworks combining natural language processing techniques and argumentation theory can be exploited to support end users in a deeper understanding of arguments given on the Web, depending on their communicative intention (Section 4). Finally, some conclusions are drawn.

2. Form and dialogue type are not sufficient

Messages on the Web can take various forms: text, image, and video, and multimedia. Merely knowing that a message on the Web is in the form of a text doesn’t describe it very well. Thus more specific categories have been used to analyze argumentative aspects of messages on the Web. These categories mainly focus on the outer form of messages and their containers, for instance Facebook messages [20], Usenet discussion threads [1], tweets [3,13], blogs [5] (as well as particular subgenres such as public issue blogs [14] and academic blogs [12]), Amazon reviews [7,18,25], and hotel reviews [21].

For argumentation, however, knowing what form a message takes is insufficient. Consider for example, the two short messages shown in Figure 1 and Figure 2; both are taken from the same platform (Twitter). These messages share certain characteristics: similar length, having been marked as a favorite and having been retweeted (forwarded/rebroadcasted from another Twitter account), using symbols (the equals sign and the down arrow) as part of the message. They differ in other characteristics: they have different topics and different authors, and only one uses hashtags and only one has replies.

We argue that despite the varying surface characteristics of these two messages, the most important difference between them for argumentation is their communicative intention. Despite extensive attention to the form of Web texts, little attention has been given to communicative intention when conducting argumentative analyses or building the Argument Web.

The closest consideration within the argumentation community is probably Walton’s typology of dialogue [23], shown in Table 1 and first developed in 1995 [24]. In this typology, the type of dialogue can be predicted from the initial situation, the participant’s goal, and the overall goal of the dialogue (which is shared between the participants).

Reviewing our examples (Figure 1 and Figure 2) against this typology, they do not fit: they are not persuasion, negotiation, deliberation, information-seeking, inquiry, discovery, or eristic dialogues. There is no appropriate positioning within these choices. Figure 1 is a humorous argument concluding that chocolate is salad.
Table 1. Walton’s seven types of dialogue \( (23) \) can be organized by the participant’s goal.

<table>
<thead>
<tr>
<th>Initial situation</th>
<th>Participant’s goal</th>
<th>Goal of dialogue</th>
<th>Type of dialogue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conflict of opinions</td>
<td>Persuade other party</td>
<td>Resolve or clarify issue</td>
<td>Persuasion</td>
</tr>
<tr>
<td>Conflict of interests</td>
<td>Get what you must want</td>
<td>Reasonable settlement both can live with</td>
<td>Negotiation</td>
</tr>
<tr>
<td>Dilemma or practical choice</td>
<td>Coordinate goals and arrange</td>
<td>Decide best viable course of action</td>
<td>Deliberation</td>
</tr>
<tr>
<td>Need information</td>
<td>Acquire or give information</td>
<td>Exchange information</td>
<td>Information-seeking</td>
</tr>
<tr>
<td>Find and verify evidence</td>
<td>Prove/disprove hypothesis</td>
<td>Need to have proof</td>
<td>Inquiry</td>
</tr>
<tr>
<td>Need to find an explanation of facts</td>
<td>Find and defend a suitable hypothesis</td>
<td>Choose best hypothesis for testing</td>
<td>Discovery</td>
</tr>
<tr>
<td>Personal conflict</td>
<td>Verbally hit out at opponent</td>
<td>Reveal deeper basis of conflict</td>
<td>Eristic</td>
</tr>
</tbody>
</table>

Figure 1. A humorous (recreational) argument from the microblog network Twitter.

In fact, it is not a dialogue at all: there are no replies, nor is there any internal bifurcation to be read as a dialogue. Perhaps it is unsurprising that Walton’s dialogical model is not sufficient here: Computer-mediated communication is often seen as midway between writing (a fully monological form) and speaking (a fully dialogical form) \( 9 \). Even though our other example is more dialogical (it at least has replies), Figure 2 is best viewed as a purely informational message that conveys an argument from a scientific study.

3. Communicative Intent

To better categorize these messages, we thus turn to communicative intentions that have been developed for the Web. Sharoff \( 17 \) proposed 5 groups of communicative intentions for Web texts: recreation, information, instruction, discussion, and recommendation. We now discuss each of these in turn.

3.1. Recreation

Recreational messages include humorous messages and texts written for leisure-time reading. Figure 1 is an example of a recreational message. Arguments are used recreationally in two ways: eristic dialogues (for instance discussions to be-
come “King of the hill”[1] and arguments that are themselves amusing (such as Figure[1] which could be read as a parody of argumentation).

For the Argument Web, recreational arguments are problematic. Eristic dialogues are, in most circumstances, uninteresting to others not participating in them. Further, their point is to ‘blow hot air’, and participants are free to sling propositions they would not commit to under other circumstances. These arguments are ephemeral and cannot be relied on. Humorous arguments are also unlikely to form an important part of the Argument Web. Despite their amusement and potential instructional value, from an information perspective, they are noise.

Recreational arguments seem more likely to be found in personal, rather than formal or institutional, situations. They can benefit from an “in crowd” focused around a specific topic, or from strong social bonds. They may be written in rather informal language or in formal language (especially for humorous arguments where parodies can work exceedingly well).

3.2. Information

Informational messages, such as Figure 2, provide information. Arguments appear in informational messages in order to justify or explain the information being provided. Figure 2 argues that produce and probiotics are important to gut health, elaborating on why.

In informational messages, conversation doesn’t form an integral part: typically interaction is secondary to the information content. Conversation around information messages is most relevant when it contains counterexamples, alternative information, or views on the importance of information. We see three replies to the informational message in Figure 2: one reacts to the information, personalizing and joking about it; the second replies to the first; and the third changes the topic, asking for help promoting a video she recommends. We will later introduce the appropriate classifications for these messages: ‘discussion’ for the first two and ‘recommendation’ for the third.

Informational messages are central to the Argument Web, and arguments about the validity and relevance of statements in such arguments are likely to be of key importance. Informational messages can be reused widely to construct other sorts of messages, especially the instruction and recommendation types, as well as discussion and information.

3.3. Instruction

Arguments appear in instructional materials in order to provide the reasons for following certain instructions or to explain their importance. Saint-Dizier and colleagues [15] have worked extensively on identifying arguments in instructional texts. These texts contain advice, warnings, preferences, evaluations, and user
simulations which serve as explanations and arguments. Treating instructional arguments separately focuses the problem of argumentation mining and simplifies it. The constrained domain means that fewer linguistic resources are necessary, and that particular discourse structures can be expected and interpreted. Their corpus contained six types of discourse structures: Instruction, Illustration, Restatement, Purpose, Condition, and Circumstance.

Figure 3 shows an argument from WikiHow about how to ‘seal the deal’ when being persuasive. It suggests a reason for using language carefully: that using words that de-emphasize separateness and emphasize similar interests helps you to be seen as part of a single unit with the person you wish to persuade.

The Instruction type contains arguments mostly about how or why to do something. These might perhaps form a kind of Argument Web of Instructional How-Tos, for instructional arguments. In instructional arguments, certain kinds of counterarguments are interesting: arguments that something should be done in a significantly different way, and explaining why to do them differently. There may also be arguments about whether to take a subsidiary action needed to achieve the larger goal.

3.4. Discussion

For messages in a discussion, the communicative intention is to discuss a state of affairs. Discussion can appear in a number of contexts: Sharoff [17] also considers newspaper articles, academic papers, and travel stories as discussions. The discussion may support an operational procedure or the decision making of an online group, such as in bug reports [10] or open collaboration [16]. This means that the discussion class is rather overloaded and needs further refinement with appropriate subcategories, which might come in part from Walton’s dialogue types.

One example discussion is the bug report given in Figure 4. This is about establishing a fault in a technical standard, HTML5, and determining the priorities for fixing it. Bug reports typically involve information-seeking (what is the fault, why is it important) as well as deliberation (should the problem be fixed), from Walton’s typology. In this case the fault is that HTML5 does not allow 13-month calendars.

The context of the bug report helps disambiguate the messages it contains. For example, one reply (‘Comment 1’) asks ‘What’s the use case?’. This can, at first sight, be read as a request for information. Yet it is not only that: the implication is also that, without a specific example, the problem pointed out in the first message may not be worth solving. In other words, in the context of the bug report, Comment 1 can be read as a counterargument to the claim that HTML5 should allow 13-month calendars.

Discussions seem to be the most dialogical form of argument on the Web, because the discussion may contribute an additional, overall argument which is not stated in any individual message. In our example from Figure 4, we see that the arguments made by one participant may be strengthened by the reasons proffered by other participants. In effect, claims and rationales may be split across messages. For example, the third message (‘Comment 2’) provides a use case. This strengthens the first commenter’s claim that HTML5 should allow 13-month calendars.
calendars by showing a practical need. The argument is elicited iteratively in an interaction between multiple participants. Consequently, to take account of the argumentation in the bug report, the message is not the appropriate granularity: the whole dialogue needs to be presented.

Discussions present an interesting challenge for the Argument Web in that it may be difficult to present discussions to end-users since their messages often depend on one another. This raises questions such as: In what cases is it appropriate to display individual messages from a longer discussion thread? Under what conditions will a link to the rest of the thread suffice? Discussion-type messages may also belong to a special context (such as bug reporting inside a community making decisions), making them relevant only within a certain scope.

3.5. Recommendation

Recommendation messages are closely related to Walton’s persuasion dialogues. They suggest certain actions in a persuasive manner. We consider reviews as recommendations. In these messages, the form and purpose are closely aligned. Consequently, the rhetorical function of statements can be predicted and used. This has led to great success in studying the argumentative aspects of this type of message.

Figure 5 shows a recommendation in the form of an Amazon product review. There is one reply which provides a counterargument. This is an interesting type of reply for the Argumentation Web: counterarguments are important to display along with recommendation messages.

Biased and sponsored messages may appear as recommendations, in addition to messages with strong opinions. For Sharoff [17] the recommendation category
includes propaganda and advertising. For the Argument Web, it would be useful to make the source of an argument more obvious. Trusted sources may vary from person to person, and people often want to make their own determinations about which sources to trust.

4. Considerations for argument identification and reconstruction

In the previous section, the main communicative intentions of arguments published on the Web were identified. Starting from the challenge of making the Argument Web transparent and understandable to end users, there is a need to develop support tools addressing this issue. Two points have to be highlighted: (i) the arguments presented on the Web, independently from the communicative intention they have, are expressed in natural language meaning that Natural Language Processing (NLP) techniques have to be adopted to process the text and return information to help users in their understanding, and (ii) such information returned by the NLP phase are exploited in different kinds of argumentation frameworks, to address the reasoning step that is necessary to actually support end users in dealing with evidence, background knowledge, context, etc. Furthermore, the information extracted with a combination of NLP techniques and argumentative reasoning should be exploited and translated into standard Web languages like RDF and OWL, in order to process such information. In the remainder of this section, we will investigate, for each communicative intention presented in Section 3, which kind of information should be retrieved to support users in a better understanding of such arguments, which NLP techniques can be adopted to achieve that task, and finally what argumentation framework should be exploited to reason over the retrieved information, highlighting the pros and cons of existing argumentation systems with respect to the specified requirements.

Recreation. Given that recreational messages include humorous messages and texts written for leisure-time reading, the language adopted is often deprecated
(e.g. slang words, acronyms, SMS-like). Moreover, it can contain metaphors, puns, or incomplete sentences. This language is difficult to process using common NLP tools, that are normally trained on formal-style (e.g. news) documents. However, several approaches have recently addressed Twitter-style language, either to annotate Named Entities e.g. [11], or to normalize such language so that it can be automatically processed [6]. From the argumentation viewpoint, eristic dialogues – having just a leisure purpose – cannot be formally represented by standard frameworks modeling persuasion or argument games. There is no point into modeling them into Dung-like argumentation, given that identifying conflicts between arguments is not the goal of the dialogue. However, it could be interesting to investigate new frameworks where relations among arguments reflect the joke-style of the interaction.

Information. The language style tends to be more formal (e.g. news-like, scientific discussions), even if the medium can bias the style in which messages are written (e.g. in Twitter, as can be noticed in Figure 2). Relations among natural language arguments can be detected applying systems built to detect semantic inferences between textual fragments (e.g. Recognizing Textual Entailment systems [3]). Opinion mining techniques can also be applied to understand people’s opinions, appraisals, attitudes, and emotions toward entities, individuals, issues. From the argumentation viewpoint, in this kind of interactions it is important to detect the evidence put forward in favor of or against a given argument, to detect at the end the winning arguments applying argumentation semantics. Moreover, it is also very important to weight the information depending on the source proposing them, taking into account the competences of the source with respect to the topic of the dialogue (several existing formal frameworks model sources and trustworthiness, e.g. [22], [19]).

Instruction. This kind of argument appears in informational messages to justify or explain the information provided. Given that they often assume the form of a question answering interaction, the question interpretation phase is the key point in understanding the message. In general, this kind of message is not part of dialogical interactions, even if the same question can correspond to several proposed answers. From the argumentation viewpoint, the aim is to translate such arguments in a logical format, so that the internal argument coherence can be assessed. Attempto Controlled English (ACE) [4] can suit this purpose.

Discussion. In order to help users to understand online discussions, natural language techniques can be combined with Dung-like abstract argumentation to identify and generate the arguments from natural language text, and then to evaluate this set of arguments to know which are the accepted ones. Starting from the participants’ opinions, the arguments implying or contradicting, even indirectly, the issue of the debate should be detected – applying e.g. the Textual Entailment (TE) approach. TE is a generic framework for applied semantics, where linguistic objects are mapped by means of semantic inferences at a textual level. TE automatically identifies, from a natural language text, the arguments. Bipolar

argumentation can then be applied to reason over the set of generated arguments with the aim of deciding which are the accepted ones.

**Recommendation.** Given that recommendation messages have persuasive goals, sentiment analysis approaches including natural language processing and text analysis techniques can be applied to identify and extract subjective information in source materials. Formal models of persuasive argumentation are starting to be proposed in the community, even if no standard assessed model exists. However, also in this kind of interaction, it is important to automatically retrieve the relations among the reviews and the associated comments, which can be addressed by applying the same approaches explained for Discussion interactions. Moreover, also in this case it is important to evaluate the evidence in favor or against a certain recommendation.

5. Conclusions

This paper proposes communicative intention as a relevant classification for Web 2.0 arguments, which can help identify which arguments are appropriate for the Argument Web, and can provide rough context for interpreting the messages. We introduced five communicative intentions and showed messages of each type: recreation, information, instruction, discussion, and recommendation. We argued that the communicative intention is different from both the dialogue type and the form of the message. We also suggested that this classification can also guide the choice of natural language processing algorithms used to identify and analyze messages for the Argument Web.

In future work we plan to test whether communicative intention impacts the choice of natural language processing algorithms. We expect that considering the communicative intention will make it easier to accurately identify and reconstruct arguments with natural language processing.

Further, we will test whether other aspects of messages should be taken into consideration for the Argument Web. Herring suggested that the best way to handle the fast-changing pace of computer-mediated communication was use a faceted classification [9]. Communicative intention, form, and dialogue type are among the relevant facets for argumentation in social media, but further study is needed to develop a full faceted genre description of social media appropriate for the Argument Web.

References


