Counter-Argumentation and Discourse: A Case Study

Stergos Afantenos
IRIT, Univ. Toulouse
France
stergos.afantenos@irit.fr

Nicholas Asher
IRIT, CNRS,
France
asher@irit.fr

Abstract

Despite the central role that argumentation plays in human communication, the computational linguistics community has paid relatively little attention in proposing a methodology for automatically identifying arguments and their relations in texts. Argumentation is intimately related with discourse structure, since an argument often spans more than one phrase, forming thus an entity with its own coherent internal structure. Moreover, arguments are linked between them either with a support, an attack or a rebuttal relation. Those argumentation relations are often realized via a discourse relation. Unfortunately, most of the discourse representation theories use trees in order to represent discourse, a format which is incapable of representing phenomena such as long distance attachments and crossed dependencies which are crucial for argumentation. A notable exception is Segmented Discourse Representation Theory (SDRT) (Asher and Lascarides, 2003). In this paper we show how SDRT can help identify arguments and their relations. We use counter-argumentation as our case study following Apothéloz (1989) and Amgoud and Prade (2012) showing how the identification of the discourse structure can greatly benefit the identification of the argumentation structure.

1 Introduction

People use arguments to persuade others to adopt a point of view or action they find beneficial to their interests, or alternatively to prevent others from adopting a position or action that they find contrary to their interests. Of course an agent may find it in her interest to convince an interlocutor to adopt a position she herself does not believe; for instance, a seller may want to persuade a buyer that a product is worth more than she believes it is worth. Because argumentation involves an interaction between an arguer and an addressee, it involves game theoretic aspects: it is the means in language for getting an agent to a position of agreement with the position one is advocating, or in game theoretic terms it is an equilibrium in a persuasion game in which the addressee adopts an optimal action based on the conversational history and in which the arguer adopts her conversational strategy based on the addressee's strategy for adopting an action (Glazer and Rubinstein, 2004). Yet, despite its importance in human communication and behavior and despite the fact that textual realizations of arguments and debates are numerous on the web, it is surprising that this area has received very little attention by the Computational Linguistics community.

One domain of research in Computational Linguistics that is of particular interest for argumentation is that of discourse. In a typical argumentation process, which takes the form of a dialogue, every argument has an internal coherence meaning that it can be represented by a discourse graph. Moreover arguments are linked between themselves either with support, attack or rebuttal relations which are realized once again as discourse relations linking either the whole discourse subgraphs representing the arguments or parts of them. Any attempt to automatically extract the argumentation structure from a given text cannot afford to ignore discourse. Our goal in this paper is to show how argumentation is intimately involved with discourse structure. We achieve this by using counter-argumentation (following (Apothéloz, 1989; Amgoud and Prade, 2012)) as a case study.

The remainder of this paper is structured as follows. In section 2 we present the current work in the so-called argumentation mining, the subfield of computational linguistics that deals with the automatic extraction of the argumentation structure from texts. In section 3 we tell a few words on discourse and in section 4 we show how SDRT (Segmented Discourse Representatuin Theory, (Asher and Lascarides, 2003)) can be applied in a case study focused on counter-argumentation. In section 5 we present the future work and we conclude this paper.

2 Argumentation in Computational Linguistics

Despite its general neglect, argumentation has been the focus of some work in Computational Linguistics. Teufel (1999), Teufel and Moens (2002) aim at identifying what they call the argumentative zones of scientific articles. The zones they have used include the aim of the paper, general scientific background, description of the authors' previous work, comparison with other works, etc. They are using a naive bayes model trying to classify each sentence into one of the predefined categories using mostly surface features (position, length, etc) and whether the sentence contains title words or words scoring high in terms of $tf.idf$.

Palau and Moens (2009) have recently attempted argumentation mining, or the identification of arguments in a text. They assume that an argument consists of a series of premises and a conclusion. Premises and conclusions are represented by propositions in the text. Of course, not all propositions in a given text are part of an argument. In order to tackle the problem of argumentation mining the authors break it into a series of subtasks. Initially they are interested in performing a binary classification of each proposition into either a proposition participating in an argument or not. Propositions that are positively classified are then sent to a second classifier which determines whether it is a premise or a conclusion. For both classification tasks they use a maximum entropy model and the Araucaria corpus[1] as well as a corpus extracted from the european court of human rights. The features they use for the first classifier include surface features ($\{1, 2, 3\}$-grams, punctuation, sentence and word length), POS information (adverbs, verbs and modal auxiliaries) and syntactic parsing. The second classifier uses again surface features, POS tags for the subject and main verb, simple rhetorical and argumentative patterns as well as the results of the first classifier (although no structured prediction is attempted which would probably be more appropriate, given that the two classifiers are not independent). Of course, once one has identified the propositions that are premises and conclusions, one does not yet have the full arguments. In order to get them, the authors create a simple CFG grammar which tries to get the tree structure of arguments. The authors do not attempt to detect the relations (e.g. support, attack, rebuttal) that connect the arguments between each other.

The Araucaria corpus is used by Feng and Hirst (2011) as well but their goal is not performing argumentation mining. Instead they focus on the task of classifying arguments into *argumentation schemes* (Walton et al., 2008). Araucaria arguments contain enthymemes annotated by human subjects which Feng and Hirst (2011) remove. Moreover, each argument is annotated with various argumentation schemes but the authors keep only the ones that are annotated with Walton's schemes. They keep only the 5 more frequent schemes. In total they have 393 arguments which they classify into one of five schemes. Concerning the classification method, they use the C4.5 algorithm implemented in Weka in order to perform either a *one-vs-all* classification or a pairwise classification. The features they use are divided into general ones concerning all schemes (features reflecting textual surface form) or specific ones for each scheme (mostly cue phrases and patterns).

Cabrio and Villata (2012a; 2012b) take a different stance. Their goal is to use Dung's (1995) abstract argumentation framework in order to detect a set of accepted arguments from online debates. They extract arguments from Debatopedia[2] using textual entailment techniques. More precisely, if a sentence T entails another sentence H then they consider that there is a support relation between the two sentences (and thus points of views) otherwise there is an attack relation. They use the open source software package EDITS[3] in order to perform textual entailment. In order then to identify the set of arguments that would be acceptable by a an external observer the authors use Dung's (1995) abstract argumentation framework. In essence an argument belongs to the aforementioned set if all the arguments attacking it are rejected. An argument is rejected if at least one accepted argument attacks it.

3 Discourse

The little prior work on argumentation has ignored discourse structure, and we think this is a mistake. A complete discourse structure of a dialogue will determine how each interlocutor's contribution relates to other contributions, both her own and that of other dialogue participants. This structure already by itself is crucial to determining the structure of an argument—which attacks are directed towards which other contributions. Moreover, an argument is not just a sequence of attacks but a much more complex structure. For one thing, arguments contain support moves as well; a good persuasion strategy is to explain why one's claims are true, but another is to provide background that will enable the addressee to understand one's reasons, and yet another is to provide more details about the claims themselves. All of these "strategies" involve in fact rhetorical moves that are different and that may be appropriate in different situations. A discourse structure makes plain these different types of moves through the use of different discourse relations.

In effect, discourse structure has the promise to give a much more detailed picture of the nature and structure of argumentation. At the moment, we don't know exactly what that picture is. But by pursuing the analysis of dialogues in terms of argument structure and discourse structure we can find out.

4 Counter-Argumentation: A Case Study

To illustrate our point in the previous section, we illustrate how constructed examples of different sorts of

[1] http://araucaria.computing.dundee.ac.uk/

[2] http://debatepedia.idebate.org

[3] http://edits.fbk.eu

arguments given by Apothéloz (1989) look from a discourse structure point of view. Apothéloz (1989) identified four different modes of arguing against a given argument. In this work an argument is simply a pair $\mathcal{C}(x) : \mathcal{R}(y)$ where $\mathcal{R}$ represents the *function of reason* and x its content and $\mathcal{C}$ the *function of conclusion* and x its content. x and y can be either propositions, conclusions or enthymemes. Given the above, Apothéloz (1989) distinguishes between four different modes of arguing against a given argument $\mathcal{C}(x)$:

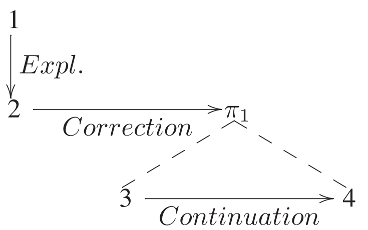
1. disputing the plausibility or the truth of a reason, that is the propositions used in y
2. disputing the completeness of the reason
3. disputing the relevance of the reason with respect to the conclusion, and
4. disputing the argumentative orientation of the reason by showing the reason presented is rather in favor of the conclusion's opposite.

Nonetheless, Apothéloz (1989) completely ignores the internal structure that the arguments have. In the following we analyse the different modes of counter-argumentation that Apothéloz (1989) provides, giving examples found in (Amgoud and Prade, 2012). Our goal is to show how discourse analysis can help the field of computational linguistics not only detect relations between arguments but also analyse the internal structure of an argument. In the following, we are using the Segmented Representation Discourse Theory (SDRT) (Asher and Lascarides, 2003). For the sake of representation, discourse is represented as a hypergraph with discourse relations being the edges of the graph and Elementary Discourse Units (EDUs) being nodes containing only one element, while Complex Discourse Units (CDUs) are nodes containing more than one simple elements (Asher et al., 2011).

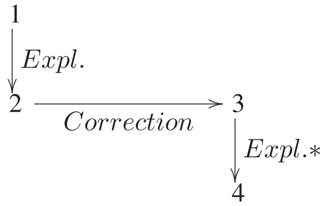
Disputing the plausibility of a reason

When one disputes the plausibility of a reason essentially it amounts to proving that the reason is false. Apothéloz (1989) provides three different ways of showing that; we illustrate them with the following examples.

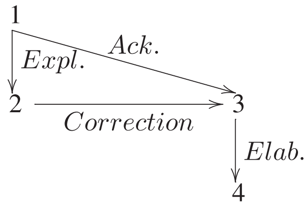
(1) — [Clara will fail her exams.]$_1$ [She did not work hard]$_2$
— [Clara?!]$_3$ [She worked non-stop.]$_4$



(2) — [Clara will fail her exams.]$_1$ [She did not work hard]$_2$
— [No, she worked hard.]$_3$ [Her eyes have bags underneath them.]$_4$



(3) — [Clara works hard]$_1$ [because she is ambitious.]$_2$
— [It is not out of ambition that Clara works hard.]$_3$ [She is not ambitious.]$_4$

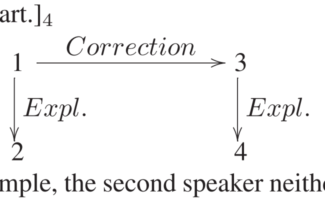


In all three examples, the second speaker does not challenge her interlocutor concerning her conclusion (EDU 1 in all three cases). In fact, in the example (3) the second speaker explicitly acknowledges the content of the conclusion ($Acknowledgment(1, 3)$). Instead the second speaker's disagreement is always with the truth value of the reason behind the conclusion. This takes the form of a $Correction$ relation between the first speaker's EDU representing the reason (EDU 2 in all cases) and the second speaker's counter-argument (EDU 3 for examples (2) and (3) and CDU π_1 for example (1)). For the last two examples the speaker provides additional reason for her beliefs either by means of an $Elaboration$ relation or an $Explanation*$ relation. This last relation signals an explanation of why b said that Clara worked hard. It is an explanation of a speech act and provides epistemic grounds for the content of the assertion. Note that in all the above examples the $Correction$ discourse relation amounts to an attack relation.

Disputing the completeness of a reason

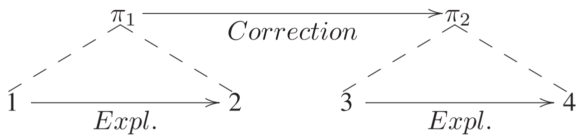
In the second mode of counter-argumentation that Apothéloz (1989) has identified, the second speaker does not attack the truthfulness of the reason but rather its completeness. Here are some examples.

(4) — [Clara will fail her exams.]$_1$ [She did not work hard]$_2$
— [Clara will not fail her exams.]$_3$ [She is very smart.]$_4$



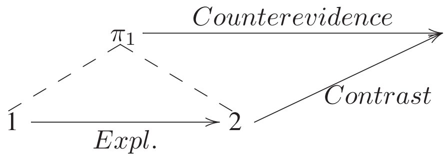
In this example, the second speaker neither affirms neither denies the reason, i.e. the fact that Clara didn't work hard. Instead, she is ignoring it (manifested by the fact that no discourse relation exists between EDUs 2 and 3 or 4). Instead she corrects the conclusion of the first speaker by providing more evidence which lead to the contrary. Again, the *Correction* discourse relation connects two arguments and serves as an attack argumentative relation.

(5) — [Paul is in his office $]_1$ [because his car is in the carpark.$]_2$
— [But the car is in the carpark$]_3$ [because it has a mechanical problem and is undriveable.$]_4$



In this case both arguments (as before) are thoroughly supported by an *Explanation* discourse relation. Moreover the second speaker even explicitly agrees with the reason given by the first one ($Acknowledgment(2, 3)$) but she disagrees with the whole argument (note the *Correction* relation between the two CDUs) since she judges that the reason is not enough and provides more evidence (EDU 4) to back her disagreement up.

(6) — [This object is red$]_1$ [since it looks red.$]_2$
— [But the object is illuminated by a red light.$]_3$

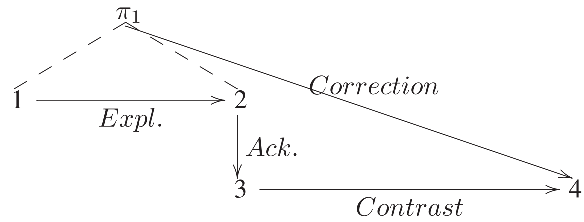


Now, this example is quite more complicated to analyze. There is a contrast between the object's looking red, which generates the expectation that it is red, and the fact that the object is illuminated by a red light, which would tend to put that expectation in doubt. But putting the expectation into doubt also puts into doubt the causal relation supposed by the first speaker between 1 and 2.

Disputing the relevance of a reason

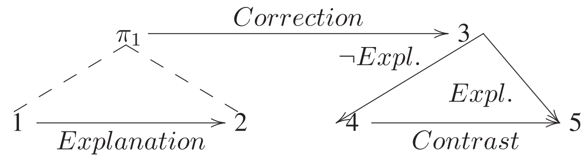
In the third mode of counter-argumentation that Apothéloz (1989) has identified concerns the second speaker does not attack the truthfulness or the completeness of a reason but instead its relevance. Below are some examples of this mode of counter-argumentation.

(7) — [Clara will fail her exams.$]_1$ [She did not work hard$]_2$
— [Indeed, she did not work hard,$]_3$ [but not working hard is not a reason to necessarily fail one's exams.$]_4$



Here the second speaker acknowledges the reason of the first person, as seen by the discourse relation between EDUs 2 and 3, but then shows that there is a contrast between this and her conclusion, disagreeing thus with the whole argument. It is important to note once again that in this example, as the preceding ones, the discourse analysis enables us to clearly pinpoint which elements of the first argument are accepted and which are attacked by the second speaker.

(8) — [Clara will fail her exams.$]_1$ [She did not work hard$]_2$
— [She will not fail her exams$]_3$ [because she did not work hard,$]_4$ [but rather because of the stress.$]_5$

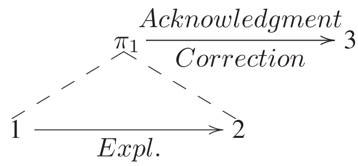


This is a very interesting example. As the discourse analysis shows the undirected cycle that is produced between EDUs 3, 4 and 5 enables the second speaker to explain why she disagrees with the whole of the initial statement.

Disputing the argumentative orientation of a reason

In the final mode of counter-argumentation that Apothéloz (1989) has proposed the second speaker does not dispute neither the reason nor the conclusion. Instead she argues that the reason corroborates towards the opposite of the conclusion. This can be illustrated with the following example.

(9) — [Running a marathon is exhausting.$]_1$ [The whole body undergoes too much stress.$]_2$
— [That's precisely what makes it nice!$]_3$



5 Discussion and Future Work

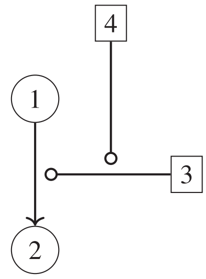
In the previous section we have showed via the use of a case study how the use of a discourse representation theory can help us represent in fine detail the phenomena that take place during argumentation—in this particular case, counter argumentation during a dialogue. In order to represent discourse we have chosen to use the Segmented Discourse Representation Theory (SDRT) of Asher and Lascarides (2003). This choice was made after careful consideration of the phenomena present during argumentation as well as the expressive power of other discourse representation theories.

Take for example the Rhetorical Structure Theory (RST, Mann and Thompson (1988)), which is the most widely cited and used discourse representation theory currently. In RST, as in SDRT, the basic units are the same, namely EDUs.[4] In RST *adjacent* EDUs can be linked together with rhetorical relations in order to form what in RST's jargon are called *spans*. Spans can be linked with rhetorical relations either with other *adjacent* EDUs or *adjacent* spans. We keep on emphasizing the word "adjacent" since this constitutes in our opinion (but see also (Peldszus and Stede, 2013)) a limitation of RST since it does not allow this theory to have *long distance dependencies*, a crucial phenomenon in argumentation. SDRT does not have this limitation. Consider example (7). In this simple example the $Correction$ relation—which, incidentally, is the backbone of the second speaker's attack—holds between non-adjacent EDUs. Even if the first speaker's argument was much longer, or if the second speaker elaborated on the fact that Clara did not work hard (and thus we had many EDUs intervening between π_1 and 4) it wouldn't influence the fact that the complex segment π_1 would be attached to EDU 4. Such long distance attachments are impossible with SDRT which requires that each EDU or span is attached to an adjacent EDU or span.

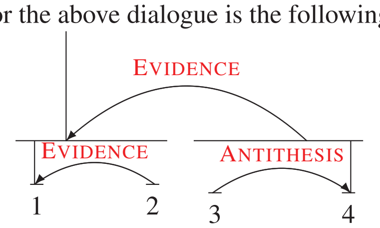
The second problem that RST has as far as the representation of argumentative structures is concerned, is that it cannot correctly represent rebuttals. This is problem that is also reported by Peldszus and Stede (2013) so we are using their example, slightly modified in order to illustrate this point. Consider the following dialogue:

(10) — [We should tear the building down.]$_1$ [It is full of asbestos.]$_2$
— [It is possible to clean it up.]$_3$
— [But that would be forbiddingly expensive!]$_4$

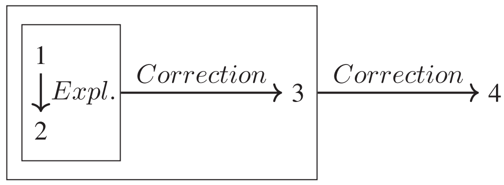
The argumentation graph that results from this dialogue, according to the scheme proposed in (Peldszus and Stede, 2013) is the following:



where edges with arrows denote support relations and edges with circles denote undercuts. The RST graph for the above dialogue is the following:



As we can see, the structural properties of those two graphs are completely different and the use of RST for argumentative analysis does not seem to be a promising path to follow. On the other hand, SDRT neatly follows the argumentation graph (we have used the box representation of SDRT here) making it thus more appropriate for use in argumentative analysis.



At this point we would like to say a few words on the computational extraction of discourse structures. Most of the published work currently is using the RST framework. This is due to two facts. Firstly there are more annotated data available for RST and secondly the problem is computationally less demanding since decisions are always made locally (attachments can be either left or right of a given span) which renders this framework more simple and thus more attractive to researchers. Of course, this implies that all long distance attachements are completely lost, an aspect which is crucial, as we have seen, for argumentation.

Muller et al. (2012) have recently attempted extraction of SDRT structures using data from the ANNODIS corpus (Afantenos et al., 2012), annotated with SDRT structures, with state of the art results. The authors attack the problem of predicting SDRT discourse structures by making some simplifications to the objects that they need to predict, namely they eliminate CDUs by making the assumption that, semantically speaking, attachment to a CDU amounts to attaching to its head—that is the uppermost and leftmost EDU. They have thus structures reminiscent of dependency graphs in syntactic analysis.

The authors perform structured prediction on the dependency graphs they produced which can be broken

[4]There is a big difference as far as EDUs are concerned between the two theories. In SDRT EDUs can be embedded the one within the other whilst RST does not allow it.

down into two steps. Initially they learn local probability distributions for attaching and labeling EDUs, based on naive bayes and logistic regression models. They effectively thus create a complete graph where each node represents an EDU and each arc a probability of attachment. The authors then move on to the decoding phase where the goal is to extract the graph that approaches the reference object. They use two decoding approaches based on $A*$ and Maximum Spanning Tree (MST) algorithms.

Closing this paper we would like to state that one of the main reasons that extraction of argumentative structures has not been more widely explored by the computational linguistics community is due to the fact that few annotated corpora exist. We believe that a project with the goal of jointly annotating argumentative and discourse structures is crucial for the advancement of this field, as well as other fields such as automatic summarization (Afantenos et al., 2008), question answering, etc.

References

Stergos D. Afantenos, Vangelis Karkaletsis, Panagiotis Stamatopoulos, and Constantin Halatsis. 2008. Using Synchronic and Diachronic Relations for Summarizing Multiple Documents Describing Evolving Events. *Journal of Intelligent Information Systems*, 30(3):183–226, June.

Stergos Afantenos, Nicholas Asher, Farah Benamara, Myriam Bras, Cecile Fabre, Mai Ho-Dac, Anne Le Draoulec, Philippe Muller, Marie-Paul Pery-Woodley, Laurent Prevot, Josette Rebeyrolles, Ludovic Tanguy, Marianne Vergez-Couret, and Laure Vieu. 2012. An empirical resource for discovering cognitive principles of discourse organisation: the annodis corpus. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey, may. European Language Resources Association (ELRA).

Leila Amgoud and Henri Prade. 2012. Can ai models capture natural language argumentation? *International Journal of Cognitive Informatics and Natural Intelligence*, 6(3):19–32, July.

Denis Apothéloz. 1989. Esquisse d'un catalogue des formes de la contre-argumentation. *Travaux du Centre de Recherches Sémiologiques*, 57:69–86.

Nicholas Asher and Alex Lascarides. 2003. *Logics of Conversation*. Studies in Natural Language Processing. Cambridge University Press, Cambridge, UK.

Nicholas Asher, Antoine Venant, Philippe Muller, and Stergos D. Afantenos. 2011. Complex discourse units and their semantics. In *Contstraints in Discourse (CID 2011)*, Agay-Roches Rouges, France.

Elena Cabrio and Serena Villata. 2012a. Combining textual entailment and argumentation theory for supporting online debates interactions. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 208–212, Jeju Island, Korea, July. Association for Computational Linguistics.

Elena Cabrio and Serena Villata. 2012b. Natural Language Arguments: A Combined Approach. In *20th European Conference on Artificial Intelligence (ECAI 2012)*, Montpellier, France.

Phan Minh Dung. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357.

Vanessa Wei Feng and Graeme Hirst. 2011. Classifying arguments by scheme. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 987–996, Portland, Oregon, USA, June. Association for Computational Linguistics.

Jacob Glazer and Ariel Rubinstein. 2004. On optimal rules of persuasion. *Econometrica*, 72(6):1715–1736, November.

William C. Mann and Sandra A. Thompson. 1988. Rhetorical Structure Theory: Towards a Functional Theory of Text Organization. *Text*, 8(3):243–281.

Philippe Muller, Stergos Afantenos, Pascal Denis, and Nicholas Asher. 2012. Constrained decoding for text-level discourse parsing. In *Proceedings of COLING 2012*, pages 1883–1900, Mumbai, India, December. The COLING 2012 Organizing Committee.

Raquel Mochales Palau and Marie-Francine Moens. 2009. Argumentation mining: The detection, classification and structure of arguments in text. In *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, ICAIL '09, pages 98–107, New York, NY, USA. ACM.

Andreas Peldszus and Manfred Stede. 2013. From argument diagrams to argumentation mining in texts: A survey. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 7(1):1–31.

Simone Teufel and Marc Moens. 2002. Summarizing Scientific Articles: Experiments with Relevance and Rhetorical Status. *Computational Linguistics*, 28(4):409–445, December.

Simone Teufel. 1999. *Argumentative Zoning: Information Extraction from Scientific Articles*. PhD Thesis, University of Edinburgh.

Douglas Walton, Chris Reed, and Fabrizio Macagno. 2008. *Argumentation Schemes*. Cambridge University Press.