Argumentation Mining in Persuasive Essays and Scientific Articles from the Discourse Structure Perspective

Christian Stab[†], Christian Kirschner[†‡], Judith Eckle-Kohler[†‡] and Iryna Gurevych[†‡]

[†]Ubiquitous Knowledge Processing Lab (UKP-TUDA)
Department of Computer Science, Technische Universität Darmstadt
[‡]Ubiquitous Knowledge Processing Lab (UKP-DIPF)
German Institute for Educational Research
`www.ukp.tu-darmstadt.de`

Abstract

In this paper, we analyze and discuss approaches to argumentation mining from the discourse structure perspective. We chose persuasive essays and scientific articles as our example domains. By analyzing several example arguments and providing an overview of previous work on argumentation mining, we derive important tasks that are currently not addressed by existing argumentation mining systems, most importantly, the identification of argumentation structures. We discuss the relation of this task to automated discourse analysis and describe preliminary results of two annotation studies focusing on the annotation of argumentation structure. Based on our findings, we derive three challenges for encouraging future research on argumentation mining.

1 Introduction

Argumentation mining is a recent research area which promises novel opportunities not only for information retrieval, educational applications or automated assessment tools but also aims at improving current legal information systems or policy modeling platforms. It focuses on automatically identifying and evaluating arguments in text documents and includes a variety of subtasks like identifying argument components, finding accepted arguments and discovering argumentation structures. Researchers have already investigated argumentation mining in several domains. For instance, Teufel (1999) aims at identifying rhetorical roles of sentences in scientific articles and Mochales-Palau and Moens (2011) identify arguments in legal documents. Also, Feng and Hirst (2011) investigated argumentation schemes in newspapers and court cases and Florou et al. (2013) applied argumentation mining in policy modeling.

However, current approaches mainly focus on the identification of arguments and their components and largely neglect the identification of argumentation structures although an argument consists not only of a set of propositions but also exhibits a certain structure constituted by argumentative relations (Peldszus and Stede, 2013; Sergeant, 2013). We argue in this paper that identifying argumentative relations and the argumentation structure respectively is an important task for argumentation mining. First, identifying argumentative relations between argument components enables the identification of additional reasons for a given claim and thus allows the creation of valuable knowledge bases e.g. for establishing new information retrieval platforms. Second, it is important to recognize which premises belong to a claim, since it is not possible to evaluate arguments without knowing which premises belong to it. Third, automatically identifying the structure of arguments enables novel features of applications, such as providing feedback in *computer-assisted writing* (e.g., recommending reasonable usage of discourse markers, suggesting rearrangements of argument components) or extracting argumentation structures from scientific publications for *automated summarization* systems.

In this paper, we analyze several examples of argumentative discourse from the discourse structure perspective.[1] We outline existing approaches on argumentation mining and discourse analysis and provide an overview of our current work on argumentation structure annotation in scientific articles and persuasive essays. We conclude this paper with a list of challenges for encouraging future

[1]The examples are taken from persuasive essays which are either collected from the writing feedback section of `http://www.essayforum.com` or from the corpus compiled by Stab and Gurevych (2014)

research on argumentation mining.

2 Background

Philosophy and Logic proposed a vast amount of argumentation theories (e.g. Toulmin (1958), Walton et al. (2008), Freeman (2011)).[2] The majority of these theories generally agree that an *argument* consists of several *argument components* which can either be a premise or a claim. The simplest form of an argument includes one claim that is supported by at least one premise (figure 1).

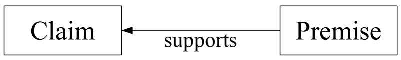

Figure 1: Illustration of a simple argument

The *claim*[3] is the central component of an argument that can either be true or false. Thus, the claim is a statement that should not be accepted by the reader without additional reasons. The second component of an argument, the *premise*[4], underpins the plausibility of the claim. It is usually provided by the proponent (writer) for convincing the reader of the claim. Examples (1) and (2) illustrate two simple arguments, each containing a claim (in bold face) and a single premise (underlined):

> (1) ***"It is more convenient to learn about historical or art items online.*** *With Internet, people do not need to travel long distances to have a real look at a painting or a sculpture, which probably takes a lot of time and travel fees."*
>
> (2) ***"Locker checks should be made mandatory and done frequently*** *because they assure security in schools, make students healthy, and will make students obey school policies."*

These examples illustrate that there exist argument components both on the sentence level and on the clause level.

Argumentative relations are usually directed relations between two argument components and represent the *argumentation structure*. There exist different types like *support* or *attack* (Peldszus and Stede, 2013) which indicate that the source argument component is a reason or a refutation for the target component. For instance, in both of the examples above, an argumentative support relation holds from the premise to the claim. The following example illustrates a more complex argument including one claim and three premises:

> (3) *"**Everybody should study abroad**$_a$. It's an irreplaceable experience if you learn standing on your own feet$_b$ since you learn living without depending on anyone else$_c$. But one who is living overseas will of course struggle with loneliness, living away from family and friends$_d$."*

Figure 2 shows the structure of the argument in (3). In this example, premise$_b$ supports the claim$_a$ whereas premise$_d$ attacks the claim$_a$.

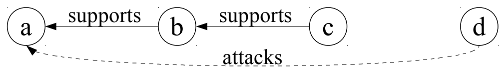

Figure 2: Argumentation structure of example (3).

This example illustrates three important properties of argumentation structures:

1. Argumentative relations can hold between non-adjacent sentence/clauses, e.g. the argumentative attack relation from premise$_d$ to the claim$_a$.

2. Some argumentative relations are signaled by indicators, whereas others are not. For instance, the argumentative attack relation from premise$_d$ to the claim$_a$ is indicated by the discourse marker *'but'*, whereas the argumentative support relation from premise$_b$ to claim$_a$ is not indicated by a discourse marker.

3. Argumentative discourse might exhibit reasoning chains, e.g. the chain constituted between argument components a, b, and c.

3 Argumentation Mining

Previous approaches on argumentation mining cover several subtasks including the separation of argumentative from non-argumentative text units (Moens et al., 2007; Florou et al., 2013), the classification of argument components (with different component classes) (Rooney et al., 2012;

[2] A review of argumentation theory is beyond the scope of this paper. A survey can be found in Bentahar et al. (2010)

[3] also called conclusion (Mochales-Palau and Moens, 2009)

[4] sometimes called support (Besnard and Hunter, 2008) or reason (Anne Britt and Larson, 2003)

Mochales-Palau and Moens, 2009; Teufel, 1999; Feng and Hirst, 2011), and the identification of argumentation structures (Mochales-Palau and Moens, 2009; Wyner et al., 2010).

3.1 Separation of Argumentative from Non-argumentative Text Units

The first step of an argumentation mining pipeline typically focuses on the identification of argumentative text units before analyzing the components or the structure of arguments. This task is usually considered as a binary classification task that labels a given text unit as argumentative or non-argumentative. One of the first approaches was proposed by (Moens et al., 2007). They focus on the identification of argumentative text units in newspaper editorials and legal documents included in the Araucaria corpus (Reed et al., 2008). The annotation scheme utilized in Araucaria is based on a domain-independent argumentation theory proposed by Walton (1996). A similar approach is reported by Florou et al. (2013). In their experiments, they classify text segments crawled with a focused crawler as either containing an argument or not. They focus on the identification of arguments in the policy modeling domain for facilitating decision making. For that purpose, they utilize several discourse markers and features extracted from the tense and mood of verbs.

Although the separation of argumentative from non-argumentative text units is an important step in argumentation mining, it merely enables the detection of text units relevant for argumentation and does not reveal the argumentative role of argument components.

3.2 Classification of Argument Components

The classification of argument components aims at identifying the *argumentative role* (e.g. claims and premises) of argument components.

One of the first approaches to identify argument components is *Argumentative Zoning* proposed by (Teufel, 1999). Each sentence is classified as one of seven rhetorical roles including e.g. claim, result or purpose using structural, lexical and syntactic features. The underlying assumption of this work is that argument components extracted from a scientific article provide a good summary of its content. Rooney et al. (2012) also focus on the identification of argument components but in contrast to the work of Teufel (1999) their scheme is not tailored to a particular genre. In their experiments, they identify claims, premises and non-argumentative text units in the Araucaria corpus. Feng and Hirst (2011) also use the Araucaria corpus for their experiments, but focus on the identification of *argumentation schemes* (Walton, 1996) which are templates for arguments (e.g. argument from example or argument from position to know). Since their approach is based on features extracted from mutual information of claims and premises, it requires that the argument components are reliably identified in advance. Mochales-Palau and Moens (2009) report several experiments for classifying argument components. They solely focus on the legal domain and in particular on legal court cases from the European Court of Human Rights (ECHR). They consider the classification of argument components as two consecutive steps. They utilize a maximum entropy model for identifying argumentative text units before identifying the argumentative role (claim and premise) of the identified components using a Support Vector Machine.

3.3 Identification of Argumentation Structures

Currently, there are only few approaches aiming at the identification of argumentation structures. For instance, the approach proposed by Mochales-Palau and Moens (2011) relies on a manually created context-free grammar (CFG) and on the presence of discourse markers for identifying a tree-like structure between argument components. However, the approach relies on the presence of discourse markers and exploits manually created rules. Therefore, it does not accommodate ill-formatted arguments (Wyner et al., 2010) and is not capable of identifying implicit argumentation structures which are common in argumentative discourse. Indeed, Marcu and Echihabi (2002) found that only 26% of the evidence relations in the RST Discourse Treebank (Carlson et al., 2001) include discourse markers.

Another approach was presented by Cabrio and Villata (2012). They identify relations between arguments of an online debate platform for identifying accepted arguments and to support the interactions in online debates. In contrast to the work of Mochales-Palau and Moens (2011), this approach aims at identifying relations between arguments (macro-level) and not between argument components (micro-level).

4 Argumentation and Discourse Analysis

Discourse analysis aims at identifying discourse relations that hold between adjacent text units with text units being sentences, clauses or nominalizations (Webber et al., 2012). Since text units might be argument components and discourse relations are often closely related to argumentative relations, previous work in automated discourse analysis is highly relevant for argumentation mining.

4.1 Discourse Relations and Argumentative Relations

Most previous work in automated discourse analysis is based on corpora annotated with general discourse relations, most notably the Penn Discourse Treebank (PDTB) (Prasad et al., 2008) and the Rhetorical Structure Theory (RST) Discourse Treebank (Carlson et al., 2003). Whereas RST represents the discourse structure as a tree, the PDTB allows more general graph structure. For the annotation of discourse relations in the PDTB, two different types of discourse relations were distinguished: implicit and explicit relations. Whereas *explicit discourse relations* are indicated by discourse markers, *implicit discourse relations* are not indicated by discourse markers and the identification of those relations requires more sophisticated methods.

Take as an example the argumentation structure discussed in section 2.

> *"Everybody should study abroad$_a$. It's an irreplaceable experience if you learn standing on your own feet$_b$* ***since*** *you learn living without depending on anyone else$_c$.* ***But*** *one who is living overseas will of course struggle with loneliness, living away from family and friends$_d$."*

Whereas the argument components b and c, as well as c and d are related through the discourse marker *'since'* (signalling an explicit CAUSE relation) and *'but'* (signalling an explicit CONTRAST relation), the discourse relation JUSTIFY between a and b is an implicit relation.

Existing approaches of discourse analysis proposed different sets of discourse relations, and there is currently no consensus in the literature about the 'right' set of discourse relations. For instance, the RST (Mann and Thompson, 1988) uses a different set of discourse relations than the PDTB (Prasad et al., 2008).

It is still an open question how the proposed discourse relations relate to argumentative relations. Although, there are preliminary findings that indicate that there are certain similarities (Cabrio et al., 2013), approaches like RST and PDTB aim at identifying general discourse structures and are not tailored to argumentative discourse.

The difference of the relations is best illustrated by the work of Biran and Rambow (2011), which is to the best of our knowledge the only approach that focuses on the identification of distinct argumentative relations. The authors argue that existing definitions of discourse relations are only usable as a building block for argumentation mining and that there are no distinct argumentative relations included in existing approaches. Therefore, they combine 12 relations from the RST Discourse Treebank (Carlson et al., 2001) to a single argumentative support relation for identifying justifications in online discussions.

4.2 Discourse Markers and Indicators of Argumentative Relations

There is a large body of previous research in linguistics on the role of *discourse markers*, signalling discourse relations (e.g.*'because'*, *'therefore'*, *'since'*, etc.) in discourse analysis. Most previous investigations of discourse markers are based on the PDTB (Prasad et al., 2008) and on the RST Discourse Treebank (Carlson et al., 2003).

However, a critically discussed question in this context is the definition of discourse markers. Are discourse markers in the sense of indicators marking discourse relations just words like *'because'*, *'therefore'*, *'since'*? Taboada (2006) investigates the role of discourse markers in corpora annotated with discourse relations according to the RST. In her discussion of related work on discourse markers in linguistics, she concludes that there are many lexical and linguistic devices signalling discourse relations beyond discourse markers, such as the mood (e.g. indicative or conjunctive) or the modality (e.g. possibility, necessity) of a sentence.

In particular, for *argumentative* discourse, the role of indicators, such as discourse markers, is not well-understood yet, which is due to the lack of corpora annotated with argumentation structures. Recently, Tseronis (2011) summarized intermediate results of a corpus-based analysis of argumentative moves, aiming at the identification of linguistic surface cues that act as *argumentative markers*. According to Tseronis (2011), *any* single or complex lexical expression can act as an argumentative marker, and it can either mark an argumentative relation (i.e., connecting two arguments or argument components) or signal a certain argumentative role, such as a claim or a premise. Moreover, he observed that also sequential patterns of argumentative markers indicate particular argumentative moves, for instance, first stating the common ground (e.g., using the marker *it is understandable* ...) and then presenting an attack to this common ground (e.g., using a marker such as *nevertheless*).

5 Argumentation Structure Annotation

Our research in argumentation mining is motivated by the (1) information access and (2) computer-assisted writing perspective. Currently, we are conducting two annotation studies, focused on analyzing argumentation structures in scientific articles and persuasive essays. In the following subsections we provide an overview of the (preliminary) results.

5.1 Argumentation Structures in Scientific Articles

One of the main goals of any scientific publication is to present new research results to an expert audience. In order to emphasize the novelty and importance of the research findings, scientists usually build up an argumentation structure that provides numerous arguments in favor of their results. The goal of this annotation study is to automatically identify those argumentation structures on a fine-grained level in scientific publications in the educational domain and thereby to improve information access. A potential use case could be an automated summarization system creating a summary of important arguments presented in a scientific article.

Up to now only coarse-grained approaches like Argumentative Zoning (Teufel et al., 2009; Liakata et al., 2012; Yepes et al., 2013) have been developed for argumentation mining in scientific publications. These approaches classify argument components according to their argumentative contribution to the document (see section 3.2) but they do not consider any relations between the argument components. To the best of our knowledge, there is no prior work on identifying argumentation structures on a fine-grained level in scientific full-texts yet (see section 3.3).

Due to the lack of evaluation datasets, we are performing an annotation study with four annotators, two domain experts and two annotators who developed the annotation guidelines. Our dataset consists of about 20 scientific full-texts from the educational domain. For the annotation study, we developed our own Web-based annotation tool (see figure 3 for a screenshot). The annotation tool allows to label argument components directly in the text with different colors and to add different relations (like support or attack) between argument components. The resulting argumentation structure is visualized as a graph (see figure 3).

Next, we plan to develop weakly supervised machine learning methods to automatically annotate scientific publications with argument components and the relations between them. The first step will be to distinguish non-argumentative parts (for example descriptions of the document structure) from argumentative parts (see section 3.1). The second step will be to identify support and attack relations between the argument components. In particular, we will explore lexical features, such as discourse markers (for example *'hence'*, *'so'*, *'for that reason'*, *'but'*, *'however'*, see section 4), and semantic features, such as text similarity or textual entailment.

5.2 Identifying Argumentation Structures for Computer-Assisted Writing

The goal of computer-assisted writing is to provide feedback about written language in order to improve text quality and writing skills of authors respectively. Common approaches are for instance focused on providing feedback about spelling and grammar, whereas more sophisticated approaches also provide feedback about discourse structures (Burstein et al., 2003), readability (Pitler and Nenkova, 2008), style (Burstein and Wolska, 2003) or aim at facilitating second language writing (Chen et al., 2012; Huang et al., 2012).

Argumentative Writing Support is a particular type of computer-assisted writing that aims at providing feedback about argumentation and thus postulates methods for reliably identifying arguments. Besides the recognition of argument components, the identification of the argumentation

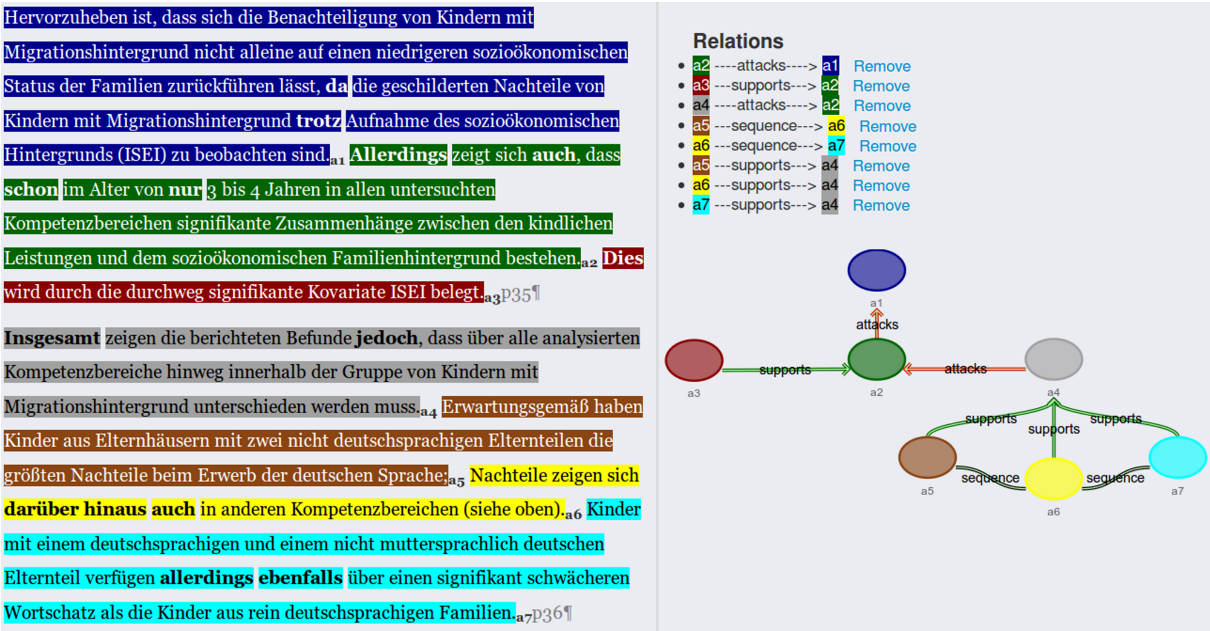

Figure 3: Screenshot of the annotation tool for argumentation structure annotation in scientific full-texts: The left side includes the text of a scientific article and the argument components marked with different colors and labels (a1-a7). The graph visualization on the right side illustrates the argumentation structure. Each node represents an argument component connected with several relations (*'support'*, *'attack'*, *'sequence'*).

structure is crucial for argumentative writing support, since it would open novel possibilities for providing formative feedback about argumentation. On the one hand, an analysis of the argumentation structure would enable the recommendation of more meaningful arrangements of argument components and a reasonable usage of discourse markers. Both have been shown to increase argument comprehension and recall, and thus the quality of the text (Anne Britt and Larson, 2003). On the other hand, by identifying which premises belong to a claim, it would be possible to advice the author to add additional support in her/his argumentation to improve the persuasiveness.

Following this vision, we conducted an annotation study with three annotators to model argument components and the argumentation structure in persuasive essays at the clause-level. The corpus includes 90 persuasive essays which we selected from *essayforum.com*. Our annotation scheme includes three argument components (*major claim*, *claim* and *premise*) and two argumentative relations (*support* and *attack*). For defining the annotation guidelines and the annotation process we conducted a preliminary study on a corpus of 14 short text snippets with five non-trained annotators and found that information about the topic and the author's stance is crucial for annotating arguments. According to these findings, we defined a top-down annotation process starting with the major claim and drilling-down to the claims and the premises so that the annotators are aware of the author's stance and the topic before annotating other components. Using this strategy, we achieved an inter-rater agreement of $\alpha_U = 0.72$[5] for argument components and $\alpha = 0.81$ for argumentative relations indicating that the proposed scheme and annotation process successfully guides annotators to substantial agreement. For more details about this annotation study, we refer the interested reader to (Stab and Gurevych, 2014), which includes a detailed description of the annotation scheme, an analysis of inter-annotator agreements on different granularities and an error analysis. The corpus as well as the annotation guidelines are freely available to encourage future research.[6]

[5] We used Krippendorff's α_U (Krippendorff, 2004) for measuring the agreement since there are no predefined marbles in our study and annotators had also to identify the boundaries of argument components.

[6] `http://www.ukp.tu-darmstadt.de/data/argumentation-mining`

6 Challenges

Existing approaches of argumentation mining mainly focus on the identification of argument components (section 3). Based on the examples analyzed in section 2 and on the experience gained in our annotation studies (section 5), we identified the following challenges for future research in argumentation mining that have not been addressed adequately by previous work.

Segmentation: Most of the existing approaches are based on the sentence-level. However, for analyzing arguments, a more fine-grained segmentation is needed (Sergeant, 2013). Apart from the sentence level, in real world data argument components exist on the clause level or can spread over several sentences. For instance, example (4) illustrates that a single sentence can contain multiple argument components (claim in bold face and premise underlined) (see also example (2) in section 2). In example (5) the premise consists of two sentences, because both sentences are needed to represent and support the "different opinions" in the claim.

> (4) *"**Eating apples is healthy** which <u>has to do with substrates which prevent cancer and other diseases</u>."*
> (5) *"**There are different opinions about coffee.** <u>Some people say they need it to stay awake. Other people think it's unhealthy.</u>"*

It is an open question if existing segmentation approaches can be used for reliably identifying the boundaries of argument components. In example (4) we find two times the word "which". This makes it hard for a segmenter to split the sentence correctly in only two parts. On the other hand, the combination of sentences (example (5)) also requires more elaborated techniques that are able to identify sentences that are related and only form in combination the support of a particular claim.

Context Dependence: The context is crucial for identifying arguments, their components and argumentation structures. As illustrated by Stab and Gurevych (2014), it is even a hard task for human annotators to distinguish claims and premises without being aware of the context. For instance, the following three argument components constitute a reasoning chain in which c is a premise for b and b a premise for a:

> (6) *"Random locker checks should be made obligatory.$_a$ Locker checks help students stay both physically and mentally healthy.$_b$ It discourages students from bringing firearms and especially drugs.$_c$"*

In this argumentation structure, a can be classified as a claim. However, without being aware of the argument component a, b becomes a claim which is supported by premise c. The same situation can be found in example (3) in section 2. If we look at the argument components b and c in isolation, we can classify b as claim. However, looking at the whole example, the argument component a is the claim, supported by the premise b. The same holds for the argument components c and a which would be connected by a support relation if they are considered in isolation. Both examples illustrate that the context is crucial for classifying argument components as claims or premises and for identifying the argumentation structure. Although, Stab and Gurevych (2014) proposed an annotation process that facilitates these decisions in manual annotation studies of persuasive essays, it is still an open issue how to model the context in order to improve the performance of automatic argumentation mining methods.

Ambiguity of Argumentation Structures: The most important challenge for identifying argumentation structures is ambiguity, since there are often several possible interpretations of argumentation structures which makes it hard or even impossible to identify one correct interpretation. In previous examples, we have already seen that the classification of argument components depends on the context and the considered argument components respectively. However, even if we consider all components of an argument, there might be several reasonable interpretations of its structure. For instance, the structure of example (6) can be interpreted in three different ways (figure 4). In the first interpretation, the argument component c supports argument component b and argument component b supports argument component a, whereas in the second interpretation argument components b and c both support argument component a. The third interpretation contains all possible argumentative relations from the first and second interpretation combined, and thus represents a graph structure (in contrast to a tree structure).

The ambiguity of argumentation structures rep-

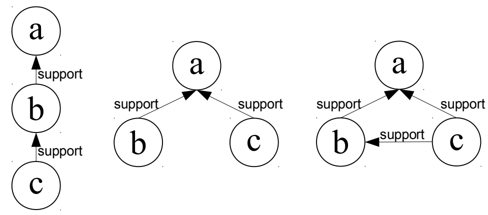

Figure 4: Several interpretations of the argumentation structure of example (6).

resents a major challenge for argument annotation studies and consequently the creation of reliable gold standards for argumentation mining. In all annotation studies we know, exactly one annotation is considered to be correct which means that other possibly correct interpretations are considered as incorrect and therefore downgrade the results for the inter annotator agreement and the performance of automatic classifiers. Consequently, it might be interesting to explore different evaluation methods. For instance, evaluation schemes used in automatic text summarization could be considered as an alternative. In text summarization, inter annotator agreement for human-generated summaries is particularly low, and hence, each human-generated summary is considered valid for evaluating an automatic summarization system (Nenkova and McKeown, 2012).

7 Conclusion

In this paper, we showed that existing approaches to argumentation mining mainly focus on the identification of argument components and largely neglect the identification of argumentation structures, although this task is crucial for many promising applications, e.g., for building novel argument related knowledge bases. By examining several examples, we derived characteristic properties of argumentation structures. We discussed the relation of discourse analysis and argumentation structure and showed that previous works in discourse analysis are not capable of identifying argumentation structures, because discourse relations do not cover all argumentative relations and are limited to relations between adjacent text units. Based on our observations, we derived three challenges for encouraging future research, i.e., (i) identifying the boundaries of argument components, (ii) modeling the context of argument components and argumentative relations, and (iii) addressing the problem of ambiguous argumentation structures. In particular, the ambiguity of argumentation structure poses an important issue for future work.

References

M. Anne Britt and Aaron A. Larson. 2003. Constructing representations of arguments. *Journal of Memory and Language*, 48(4):794–810.

Jamal Bentahar, Bernard Moulin, and Micheline Bélanger. 2010. A taxonomy of argumentation models used for knowledge representation. *Artificial Intelligence Review*, 33(3):211–259.

Philippe Besnard and Anthony Hunter. 2008. *Elements of argumentation*, volume 47. MIT press Cambridge.

Or Biran and Owen Rambow. 2011. Identifying justifications in written dialogs by classifying text as argumentative. *International Journal of Semantic Computing*, 05(04):363–381.

Jill Burstein and Magdalena Wolska. 2003. Toward Evaluation of Writing Style: Finding Overly Repetitive Word Use in Student Essays. In *Proceedings of the Tenth Conference on European Chapter of the Association for Computational Linguistics - Volume 1*, EACL '03, pages 35–42, Budapest, Hungary.

Jill Burstein, Daniel Marcu, and Kevin Knight. 2003. Finding the WRITE Stuff: Automatic Identification of Discourse Structure in Student Essays. *IEEE Intelligent Systems*, 18(1):32–39.

Elena Cabrio and Serena Villata. 2012. Natural language arguments: A combined approach. In *Proceedings of the 20th European Conference on Artificial Intelligence*, ECAI '12, pages 205–210, Montpellier, France.

Elena Cabrio, Sara Tonelli, and Serena Villata. 2013. From discourse analysis to argumentation schemes and back: Relations and differences. In João Leite, TranCao Son, Paolo Torroni, Leon Torre, and Stefan Woltran, editors, *Computational Logic in Multi-Agent Systems*, volume 8143 of *Lecture Notes in Computer Science*, pages 1–17. Springer Berlin Heidelberg.

Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski. 2001. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of the Second SIGdial Workshop on Discourse and Dialogue - Volume 16*, SIGDIAL '01, pages 1–10, Aalborg, Denmark.

Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski, 2003. *Building a discourse-tagged corpus in the framework of rhetorical structure theory*, chapter 5, pages 85–112. Springer.

Mei-Hua Chen, Shih-Ting Huang, Hung-Ting Hsieh, Ting-Hui Kao, and Jason S. Chang. 2012. FLOW: A First-language-oriented Writing Assistant System. In *Proceedings of the ACL 2012 System Demonstrations*, ACL '12, pages 157–162, Jeju Island, Korea.

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 - Volume 1*, HLT '11, pages 987–996, Stroudsburg, PA, USA. Association for Computational Linguistics.

Eirini Florou, Stasinos Konstantopoulos, Antonis Koukourikos, and Pythagoras Karampiperis. 2013. Argument extraction for supporting public policy formulation. In *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, pages 49–54, Sofia, Bulgaria, August. Association for Computational Linguistics.

James B. Freeman. 2011. *Argument Structure: Representation and Theory*, volume 18 of *Argumentation Library*. Springer.

Chung-chi Huang, Ping-che Yang, Mei-hua Chen, Hung-ting Hsieh, Ting-hui Kao, and Jason S. Chang. 2012. TransAhead: A Writing Assistant for CAT and CALL. In *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, EACL '12, pages 16–19, Avignon, France.

Klaus Krippendorff. 2004. Measuring the Reliability of Qualitative Text Analysis Data. *Quality & Quantity*, 38(6):787–800.

Maria Liakata, Shyamasree Saha, Simon Dobnik, Colin Batchelor, and Dietrich Rebholz-Schuhmann. 2012. Automatic recognition of conceptualization zones in scientific articles and two life science applications. *Bioinformatics*, 28(7):991–1000.

William C. Mann and Sandra A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.

Daniel Marcu and Abdessamad Echihabi. 2002. An unsupervised approach to recognizing discourse relations. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL '02, pages 368–375.

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.

Raquel Mochales-Palau and Marie-Francine Moens. 2011. Argumentation mining. *Artificial Intelligence and Law*, 19(1):1–22.

Marie-Francine Moens, Erik Boiy, Raquel Mochales Palau, and Chris Reed. 2007. Automatic detection of arguments in legal texts. In *Proceedings of the 11th International Conference on Artificial Intelligence and Law*, ICAIL '07, pages 225–230, New York, NY, USA. ACM.

Ani Nenkova and Kathleen McKeown. 2012. A survey of text summarization techniques. In *Mining Text Data*, pages 43–76.

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.

Emily Pitler and Ani Nenkova. 2008. Revisiting readability: A unified framework for predicting text quality. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '08, pages 186–195.

Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. 2008. The Penn Discourse TreeBank 2.0. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, pages 28–30, Marrakech, Morocco.

Chris Reed, Raquel Mochales-Palau, Glenn Rowe, and Marie-Francine Moens. 2008. Language resources for studying argument. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation*, LREC '08, pages 2613–2618, Marrakech, Morocco.

Niall Rooney, Hui Wang, and Fiona Browne. 2012. Applying kernel methods to argumentation mining. In *Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Society Conference*, FLAIRS '12, pages 272–275, Marco Island, FL, USA.

Alan Sergeant. 2013. Automatic argumentation extraction. In *Proceedings of the 10th European Semantic Web Conference*, ESWC '13, pages 656–660, Montpellier, France.

Christian Stab and Iryna Gurevych. 2014. Annotating argument components and relations in persuasive essays. In *Proceedings of the 25th International Conference on Computational Linguistics (COLING 2014)*, page (to appear), Dublin, Ireland, August.

Maite Taboada. 2006. Discourse Markers as Signals (or Not) of Rhetorical Relationsteu. *Journal of Pragmatics*, 38(4):567–592.

Simone Teufel, Advaith Siddharthan, and Colin Batchelor. 2009. Towards discipline-independent argumentative zoning: evidence from chemistry and computational linguistics. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 1493–1502.

Simone Teufel. 1999. *Argumentative Zoning: Information Extraction from Scientific Text*. Ph.D. thesis, University of Edinburgh.

Stephen E. Toulmin. 1958. *The uses of Argument*. Cambridge University Press.

Assimakis Tseronis. 2011. From connectives to argumentative markers: A quest for markers of argumentative moves and of related aspects of argumentative discourse. *Argumentation*, 25(4):427–447.

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

Douglas N Walton. 1996. *Argumentation schemes for presumptive reasoning*. Routledge.

Bonnie Webber, Mark Egg, and Valia Kordoni. 2012. Discourse structure and language technology. *Natural Language Engineering*, 18:437–490, 10.

Adam Wyner, Raquel Mochales Palau, Marie-Francine Moens, and David Milward. 2010. Approaches to text mining arguments from legal cases. In *Semantic Processing of Legal Texts*, volume 6036 of *Lecture Notes in Computer Science*, pages 60–79. Springer.

Antonio Jimeno Yepes, James G. Mork, and Alan R. Aronson. 2013. Using the argumentative structure of scientific literature to improve information access. In *Proceedings of the 2013 Workshop on Biomedical Natural Language Processing*, pages 102–110.