

 Open access • Journal Article • DOI:10.1007/S10506-010-9104-X

## Argumentation mining — Source link

Raquel Mochales, Marie-Francine Moens

**Institutions:** Katholieke Universiteit Leuven

**Published on:** 01 Mar 2011

**Topics:** Argumentation theory

Related papers:

- [Identifying Argumentative Discourse Structures in Persuasive Essays](#)
- [The uses of argument](#)
- [From Argument Diagrams to Argumentation Mining in Texts: A Survey](#)
- [Classifying arguments by scheme](#)
- [Automatic detection of arguments in legal texts](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/argumentation-mining-lg4tkvq6y6>

# Argumentation Mining

**Raquel Mochales and Marie-Francine Moens**  
Katholieke Universiteit Leuven, Belgium

Received: date / Accepted: date

**Abstract** Argumentation mining aims to automatically detect, classify and structure argumentation in text. Therefore, argumentation mining is an important part of a complete argumentation analysis, i.e. understanding the content of serial arguments, their linguistic structure, the relationship between the preceding and following arguments, recognizing the underlying conceptual beliefs, and understanding within the comprehensive coherence of the specific topic. We present different methods to aid argumentation mining, starting with plain argumentation detection and moving forward to a more structural analysis of the detected argumentation. Different state-of-the-art techniques on machine learning and context free grammars are applied to solve the challenges of argumentation mining. We also highlight fundamental questions found during our research and analyse different issues for future research on argumentation mining.

## 1 Introduction

Argumentation can be defined as a process whereby arguments are constructed, exchanged and evaluated in light of their interactions with other arguments, each of which comprises a set of premises, pieces of evidence, offered in support of a claim. The claim is a proposition, an idea which is either true or false, put forward by somebody as true. The claim of an argument is normally called its conclusion. Argumentation may also involve chains of reasoning, where claims are used as premises for deriving further claims.

People routinely undertake argumentation as an integral part of their daily-life, forming reasons, drawing conclusions and applying them to a case in discussion. For this, they offer facts or assertions as evidence that something is true, i.e. arguments. Arguments are therefore discussions in which reasons (*premises*) are advanced for and against some proposition or proposal (*conclusion*).

Argumentation can be a formal demonstration of the truth of a proposition, e.g. when someone expresses a mathematical proof, but in many areas of human reasoning, including law, this formal demonstration is often not possible. In such domains

---

Address(es) of author(s) should be given

rationality must be based on some more or less informal arguments. Informal arguments are sometimes implicit. That is, the logical structure, the relationship of claims, premises, warrants, relations of implication, and conclusion, is not always spelled out and immediately visible and must sometimes be made explicit by analysis.

The analysis of argumentation focuses on understanding the content of argument chains, as well as analyzing the linguistic structure, determining the relationship between the preceding and following arguments, recognizing the underlying conceptual beliefs, and understanding the arguments' coherence in light of general knowledge of the specific topic. A complete argumentative analysis must depend on several aspects of knowledge: linguistic constraints, domain dependent, conceptual relations, and discourse structure. None of the above aspects are sufficient by themselves for a complete analysis, but they all contribute to complete analysis, especially in the situation where one aspect of information is incomplete. For instance, the beliefs of the speaker may be unknown, unusual or the semantic content may be ambiguous. Argumentation analysis allows a better discourse understanding, a broader knowledge of the speaker's intentions and beliefs, and it highlights interactions between the speakers and their different views of the world.

Argumentation mining focuses on the detection of all the arguments in a text and their relationships with their preceding and following arguments. Argumentation mining does not analyse the validity of the argumentation or its correctness. The aim is to detect those pieces of text which seem to function as argumentative (from a linguistic and semantic point of view) and the relations between them, i.e. their structure. The result is an argumentative structure of the text from the linguistic analysis of its propositions.

The goal of this article is to highlight the challenges and applications of argumentation mining, offering an initial study of how machine learning and other state-of-the-art techniques can help in their accomplishment. The paper is organized as follows. Chap. 2 presents different related research on argumentation inside and outside the legal field. In Chap. 3, we describe the basics of argumentation mining together with formal definitions of the main argumentation aspects involved in argumentation mining. In Chap. 4, we then present a set of methods and models used to achieve some of the tasks of argumentation mining. We then highlight, in Chap. 5, some of the applications of these new methods. We conclude in Chap. 6 with future work and discussion remarks.

## 2 Related Research

The study of argumentation has a long tradition, inside and outside the legal field. Many fields have shown interest in argumentation, e.g. philosophy, logic, psychology or more recently artificial intelligence.

Regarding logics the theories have been divided into two branches: formal logic and informal logic. Formal logic is the study of inference with purely formal content. An inference possesses a purely formal content if it can be expressed as a particular application of a wholly abstract rule, that is, a rule that is not about any particular thing or property. The works of Aristotle contain the earliest known formal study of logic. Mathematical logic, from proof theory [43] to first-order logic [44], is a well-known form of formal logic. Informal logic is the branch of logic whose task is to develop non-formal standards, criteria, and procedures for the analysis, interpretation, evaluation, critique and construction of argumentation in everyday discourse. It arose

in the context of three streams of criticism of the existing formal logic program [37]. First, the pedagogical critique challenged that the tools of logic should be applicable to everyday reasoning and argument of the sorts used in political, social and practical issues. Second, the internal critique challenged the adequacy of existing tools of logic in evaluating everyday argument. Third, the empirical critique challenged the ideas that formal deductive logic can provide a theory of good reasoning, and that the ability to reason well is improved by a knowledge of formal deduction.

Psychology has a long tradition of studies of non-logical aspects of argumentation [45]. For example, studies have shown that simple repetition of an idea is often a more effective method of argumentation than appeals to reason. Empirical studies of communicator credibility and attractiveness, sometimes labeled charisma, have also been tied closely to empirically-occurring arguments. Such studies bring argumentation within the ambit of persuasion theory and practice.

More recent efforts have been made within the field of artificial intelligence, where argumentation is analyzed and performed by computers. Argumentation has been used to provide a proof-theoretic semantics for non-monotonic logic, starting with the influential work of [38]. Computational argumentation systems have found particular application in domains where formal logic and classical decision theory are unable to capture the richness of reasoning, domains such as law and medicine. A comprehensive overview of this area can be found in a recent book [39].

### 3 Argumentation Mining

Argumentation mining moves between natural language processing, argumentation theory and information retrieval. The aim of argumentation mining is to automatically detect the argumentation of a document, i.e. detection of all the arguments involved in the argumentation process, their individual or local structure (rhetorical or argumentative relationships between their propositions), and the interactions between them (the global argumentation structure).

To achieve the aim of argumentation mining an adequate linguistic, formal, and computational study of argumentation is required. Different questions need to be answered:

1. What is the “correct” abstract structure of argumentation? Should we represent argumentation as a tree-structure or is it better to use a graph-structure? What are the constraints that characterize this structure?
2. What are the elementary units of argumentation? And of an individual argument?
3. What are the relations that hold between two arguments and/or argumentation units? Are they grounded into the events and the world that the text describes, or into general principles of rhetoric and linguistics?
4. Can the units of argumentation and/or arguments be determined automatically?
5. Can argumentation structures be determined automatically? If so, how?

Evidence for the answers to these questions can come from different disciplines including philosophy, law, linguistics and computer science. Many of these disciplines have a long tradition of argumentation analysis, where they offer many theories, most of them not compatible with the rest. Therefore, adequate and supported answers to the previous questions or even a summary of such answers is a challenging task. The following subsections provide our suggested answers to questions one to four. Then in

---

section 4 we show that questions four and five have positive answers by presenting some results of our research.

### 3.1 Elementary Units of Argumentation

Since it is well-known that argumentation is the process whereby arguments are constructed, exchanged and evaluated in light of their interactions with other arguments, then it is not surprising that all argumentation experts agree that the elementary units of argumentation are *arguments*. However, the definition of an argument is more controversial. Only one thing seems common in all definitions: an argument is always formed by premises and a conclusion.

In free text these premises and conclusion can be implicit (i.e. enthymemes). Therefore, some studies have mentioned that an argument can be presented as a single proposition in its minimal representation. However, even for a human at least two argumentative propositions are needed to have an appropriate certainty when distinguishing arguments from statements. Isolated argumentative propositions are hard to distinguish from simple statements. For example, the isolated sentence: “Councilwoman Radcliffe voted in favour of the tax increase.” does not look like an argumentative sentence. However, when it is placed in the right context: “Councilwoman Radcliffe voted in favour of the tax increase. No one who voted in favour of the tax increase is a desirable candidate. Therefore, Councilwoman Radcliffe is not a desirable candidate.”, it is completely clear that this proposition is part of an argument.

In conclusion, we determine that the elementary units of argumentation are *arguments*, where an argument is *a set of at least two propositions*. However, to complete this definition it is necessary to define the meaning of a proposition.

In linguistics, a proposition is conveyed by a declarative sentence used to make a statement or assertion. But does argumentation mining work on a sentence level or does it need a deeper analysis of smaller text spans? Here we do not constrain, leaving this to a free-choice depending on the type of text at hand. Dialogues or informal text will contain sentences combining information about the conclusion and premise, since those texts are less restrictive in their presentation format. However, more formal texts, such as legal documents, present premises and conclusion in separate sentences, both to clarify the structure and to be able to extend the given information.

### 3.2 Internal Structure of Elementary Units

The definition of an argument proposed is not complete. We need to define the nature and relations that hold between the propositions of an argument. Essentially, most of the argumentation theories assume that the propositions can be classified as premises and conclusions. However, more complex classifications, such as the one of Toulmin [25], [29] or [32], have been presented over the years. We base our work on a theory that studies argumentation schemes, see for example [29], where only premises and conclusions are recognized, assuming that each argument follows a scheme, that defines relations between propositions reflecting reasoning patterns. We base our work on this theory as we believe it should be easier to find patterns in natural text that resemble argumentation schemes, given that they are forms of argument that capture stereotypical patterns of human reasoning, especially defeasible ones. Table 1 presents

Table 1: Argumentation Schemes

Type	Argument Structure	Argument Example
Argument from Example	In this particular case, the individual $a$ has a property $F$ and also property $G$ . $a$ is typical of things that have $F$ and may or may not also have $G$ . Therefore, generally, if $x$ has a property $F$ . then $x$ also has property $G$ .	<b>Conclusion</b> - There will always be a need for elucidation of doubtful points and for adaptation to changing circumstances. <b>Premise</b> - Indeed, in the United Kingdom, as in the other Convention States, the progressive development of the criminal law through judicial law-making is a well entrenched and necessary part of legal tradition.
Argument from Verbal Classification	$a$ has a particular property $F$ . For all $x$ . if $x$ has a property $F$ , then $x$ can be classified as having property $G$ . Therefore, $a$ has property $G$ .	<b>Premise</b> - There will always be a need for elucidation of doubtful points and for adaptation to changing circumstances. <b>Conclusion</b> - Article 7 of the Convention cannot be read as outlawing the gradual clarification of the rules of criminal liability through judicial interpretation from case to case, provided that the resultant development is consistent with the essence of the offence and could reasonably be foreseen .
Argument from Commitment	$a$ is committed to proposition $A$ (generally, or in virtue of what she said in the past). Therefore, in this case, $a$ should support $A$ .	<b>Premise</b> - Contracting Parties do not enjoy an unlimited discretion. <b>Premise</b> - It is for the Court to rule whether, inter alia, the States have gone beyond the "extent strictly required by the exigencies" of the crisis. <b>Conclusion</b> - The domestic margin of appreciation is thus accompanied by a European supervision.
Argument from Position to Know	$a$ is in a position to know whether $A$ is true (false). $a$ asserts that $A$ is true (false). Therefore, $a$ is true (false).	<b>Premise</b> - The Court recalls that it falls to each Contracting State, with its responsibility for "the life of its nation", to determine whether that life is threatened by a "public emergency" and, if so, how far it is necessary to go in attempting to overcome the emergency. <b>Premise</b> - By reason of their direct and continuous contact with the pressing needs of the moment, the national authorities are in principle better placed than the international judge to decide both on the presence of such an emergency and on the nature and scope of the derogations necessary to avert it. <b>Conclusion</b> - This part of the application can therefore also not be rejected.

some examples of argumentation schemes and arguments from real legal texts where it is possible to appreciate how schemes are used in natural language argumentation. Furthermore, Table 2 shows how an argument scheme would be analysed by Toulmin's theory. We observe that the selection of text that created each of its reasoning steps is a hard task, especially where most of the steps are implicit.

Table 2: Argument classified as Verbal Classification seen by Toulmin’s reasoning

Argument of Verbal Classification	Toulmin’s reasoning sequence		
	Class	Reasoning	Text
In voluntary health insurance you generally get a poor return for your money because overhead and profits of the insurance company eat up huge chunks of the premiums you pay. On individual policies these companies spend for overhead and profits an average of about 60% of what you pay them and only 40 cents of your premium dollar goes for benefits to policyholders. Obviously such insurance is a mighty poor buy.	<b>Premise</b>	There is 40% return on health insurance.	<i>only 40 cents of your premium dollar goes for benefits to policyholders</i>
	<b>Warrant</b>	40% can be classified as a poor return.	-
	<b>Qualifier</b>	Unless other companies can do no better.	-
	<b>Conclusion</b>	40% is a poor return.	-

Our definition of argument is established as: “an argument is set of propositions, all of which are premises except, at most, one, which is a conclusion. Any argument follows an argumentation scheme, where the critical questions can be implicit or explicit”.


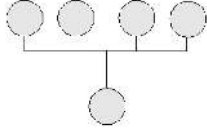
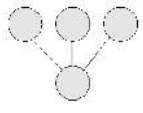
### 3.3 Relations between Elementary Units

Once it is clear which are the elementary units of argumentation, i.e. arguments, it is necessary to define the nature, number and classification of the relations that can hold between these units. Given the aim of argumentation mining to work over free text, the approach of [26], where arguments relate through coordination, a subordination, or form a multiple argumentation relation, seems the most appropriate to us. Table 3 shows examples of this theory. Therefore, our argumentation contains arguments, which are formed by premises and a conclusion following an argumentation scheme, and they relate to other arguments through coordination, subordination or forming a multiple argument.

### 3.4 Argumentation Structure

The previous definitions of the elementary units of argumentation and the possible relations between them facilitate the choice of an argumentation structure. We have decided to use a simple conclusion-premise relation; each premise connects to one conclusion and each conclusion can have one or more premises. Therefore, we can see argumentation as a tree-structure, instead of a more complex graph-structure. This type of structural relation fits perfectly with pragma-dialectic theory [26]. The leaves of the tree are arguments, with premises and conclusion, which converge into other arguments, i.e. premises of higher tree nodes as shown in Figure 1a. The relations between the different arguments of the tree could be subsequently analysed as attacks, supports, etc. However, we are not dealing with such a deep analysis of the argumentation structure. For the moment we only identify a relation between the arguments. We see our tree as a simple representation of a single argumentation structure, where all arguments have a unique relation to another argument of the tree, and finally all

Table 3: Pragma-Dialectics theory and examples

Type	Example
Subordinatively Compound Argumentation 	<i>She won't worry about the exam. She's bound to pass. She's never failed.</i>
Coordinatively Compound Argumentation 	<i>This book has literary qualities: the plot is original, the story is well-told, the dialogues are incredibly natural, and the style is superb.</i>
Multiple Argumentation 	<i>Postal deliveries in Holland are not perfect. You cannot be sure that a letter will be delivered the next day, that it will be delivered to the right address, or that it will be delivered early in the morning.</i>

together support a general conclusion. Some possible extensions of our structure are shown in Figure 1, where 1b shows a visualization of argument schemes while 1c shows a visualization of support-attack relations.

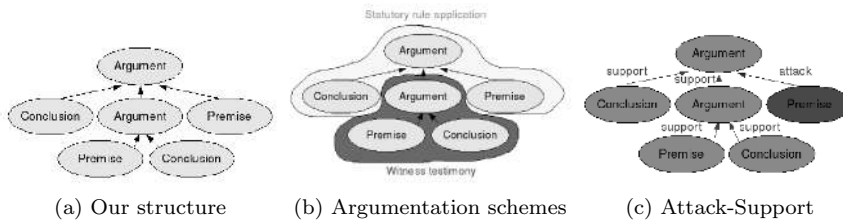


Fig. 1: Argumentation tree-structures

## 4 Experiments on Argumentation Mining

### 4.1 Corpora

The correctness of all the methods proposed in the following sections has been evaluated using two corpora.



### The Argument Text

*The report was commissioned by poker machine operator Tattersalls. Its methodology and the independence of its findings were questioned by InterChurch Gambling Taskforce spokesman Reverend Tim Costello, who said the survey was not reliable because it was based on participants filling out a diary for a fortnight. "People with gambling problems simply don't record their gambling expenditure," he said. He also said that if higher-income earners were really more likely to play on poker machines "there would be as many pokies in Toorak as there are in Footscray".*

### Analysis

- Conclusion: Identifier: A
- *the survey was not reliable*
- Scheme: Argument from Singular Cause (id: 0)
- **Support**
  - Identifier: B
    - *it was based on participants filling out a diary for a fortnight*
    - Scheme: Argument from Singular Cause (id: 0)
  - Identifier: C
    - *People with gambling problems simply don't record their gambling expenditure*
    - Scheme: Argument from Singular Cause (id: 0)
  - Identifier: E (*Enthymeme*)
    - *If the survey was based on participants filling out a diary for a fortnight and people with gambling problems do not record their gambling expenditures, the survey is not reliable.*
    - Scheme: Argument from Singular Cause (id: 0)

Fig. 2: Argument annotated in Araucaria

On one hand we use a general corpus, known as the Araucaria corpus. This corpus comprises two distinct sets of data: a structured set in English collected and analysed according to a specific methodology as a part of a project at the University of Dundee (UK), and an unstructured multi-lingual set of user-contributed analyses. Only the structured data was used for our analysis. The data was collected over a six week period in 2003, during which time a weekly regime of data collection scheduled regular harvests of one argument from 19 newspapers (from the UK, US, India, Australia, South Africa, Germany, China, Russia and Israel, in their English editions where appropriate), 4 parliamentary records (in the UK, US and India), 5 court reports (from the UK, US and Canada), 6 magazines (UK, US and India), and 14 further online discussion boards and “cause” sources such as HUMAN RIGHTS WATCH (HURW) and GlobalWarming.org. Each week, the first argument encountered in each source was identified and analysed by hand. Later on, an amount of non-argumentative information from the same sources was added to the corpus. The corpus is formed by an equal number of sentences that contain an argument and sentences without arguments; see Table 5 for statistics and Table 4 for examples. The sentences are also classified by their text type: newspapers, parliamentary records, legal judgments, weekly magazines, discussion fora, “cause” sources and speeches. All sentences of the corpus are annotated as argumentative or non-argumentative. Moreover, the argumentative sentences are grouped as arguments following argumentation schemes theory, i.e. premises, conclusions, relations between them and type of scheme (see Figure 2).

Table 4: Examples of argumentative and non-argumentative propositions in the Araucaria corpus

Text type	Argument	Non-argument
Discussion fora	On this occasion, however, I shall not vote for any individual or party but will spoil my paper.	I have been voting since 1964 and at one time worked for my chosen party.
Legal judgments	He is aware of the risks involved, and he should bear the risks.	Lest there be any misunderstanding one point should be clarified at the outset.
Newspapers	Labor no longer needs the Liberals in the Upper House.	The independents were a valuable sounding board for Labors reform plans.
Parliamentary records	I have accordingly disallowed the notice of question of privilege.	Copies of the comments of the Ministers have already been made available to Dr. Raghuvansh Prasad Singh.
Weekly magazines	But for anyone who visits Rajasthan's Baran district, the apathy of the district administration and the failure of the Public Distribution System (pds) is clear to see	This time in Rajasthan.

The second corpus we use in our evaluations is the ECHR corpus, a set of documents extracted from legal texts of the European Court of Human Rights (ECHR). The ECHR, over the years, has developed a standard type of reasoning and structure of argumentation. Therefore, its documents are a perfect test set for our argumentation analysis. It provides a set of documents which lack some natural language problems, such as extensive use of metaphor or high appearance of ambiguity, while maintaining enough complexity to be described as real natural language argumentation.

The ECHR corpus annotators identify the different arguments of each document, the schemes they follow and how they interact with the other arguments of the document. To ease the annotation process, different documentation describing argumentation schemes theory and guidelines on the output format were given to the annotators. The overall process took more than a year and included three annotators and one judge to solve disagreements. Once the task was completed, the annotation obtained a 75% agreement between annotators using Cohen's kappa coefficient [35].

Table 5 shows the main characteristics of the ECHR training and test corpora. The dataset deals with different human rights (e.g. child rights, immigration or torture). Figure 3 shows an example of argumentation from a sample ECHR legal case with its corresponding annotation. A more detailed study on the ECHR documents and their argumentation can be found in [19]. The distribution of premises, conclusions and non-argumentative sentences in the corpus shows a clear imbalance between premises and conclusions. The number of premises is 763 while the number of conclusions is just 304 and the number of non-argumentative sentences is 1449. This is a normal characteristic of any real argumentation, where conclusions tend to be justified by many premises to ensure a complete and stable justification of each standpoint.

<p>{ [ <i>SUPPORT</i>: The Court recalls that the rule of exhaustion of domestic remedies referred to in Article x of the Convention art. x obliges those seeking to bring their case against the State before an international judicial or arbitral organ to use first the remedies provided by the national legal system.</p> <p><i>CONCLUSION</i>: Consequently, States are dispensed from answering before an international body for their acts before they have had an opportunity to put matters right through their own legal systems. ]</p> <p>[ <i>SUPPORT</i>: The Court considers that, even if it were accepted that the applicant made no complaint to the public prosecutor of ill-treatment in police custody, the injuries he had sustained must have been clearly visible during their meeting.</p> <p><i>AGAINST</i>: However, the prosecutor chose to make no enquiry as to the nature, extent and cause of these injuries, despite the fact that in Turkish law he was under a duty to investigate see paragraph above.</p> <p><i>SUPPORT</i>: It must be recalled that this omission on the part of the prosecutor took place after Mr Aksoy had been detained in police custody for at least fourteen days without access to legal or medical assistance or support.</p> <p><i>SUPPORT</i>: During this time he had sustained severe injuries requiring hospital treatment see paragraph above.</p> <p><i>CONCLUSION</i>: These circumstances alone would have given him cause to feel vulnerable, powerless and apprehensive of the representatives of the State. ]</p> <p><i>CONCLUSION</i>: The Court therefore concludes that there existed special circumstances which absolved the applicant from his obligation to exhaust domestic remedies. }</p>
--

Fig. 3: An example of legal argumentation with two sub-arguments

Table 5: Characteristics for the Araucaria and the ECHR corpora

Characteristics	Araucaria	ECHR
Number of annotated documents	641	47
Number of words in documents	76970	92190
Number of sentences in documents	3798	2571
Number of annotated arguments	641	257

## 4.2 Argument Detection

The detection of all the arguments presented in a free text is similar to the binary classification of all the propositions of the text as argumentative or non-argumentative. If each proposition of the text can be classified as being part of the argumentation or not, then all units classified as argumentative constitute together all the arguments of the text. However, this approach presents a limitation, as the delimiters of each argument are not defined. Therefore, it is known which information forms the arguments, but not how this information is split into the different arguments. This is known as a segmentation problem.

First, we analyse the classification problem. Following the work of [23] and [9] we studied the use of statistical classifiers, e.g. naïve Bayes, maximum entropy model or support vector machines [12]. First, we represented each sentence as a vector of features. We defined generic features that could easily be extracted from the texts and studied

Table 6: Features used in the classification of argumentative and non-argumentative information

Unigrams	Each word in the sentence.
Bigrams	Each pair of successive words.
Trigrams	Each three successive words.
Adverbs	Detected with a part-of-speech (POS) tagger (e.g. QTag <sup>1</sup> ).
Verbs	Detected with a POS tagger. Only the main verbs (excluding “to be”, “to do” and “to have”) are considered.
Modal auxiliary	Indicates if a modal auxiliary is present using a POS tagger.
Word couples	All possible combinations of two words in the sentence are considered.
Text statistics	Sentence length, average word length and number of punctuation marks.
Punctuation	The sequence of punctuation marks present in the sentence is used as a feature (e.g. “.”). When a punctuation mark occurs more than once in a row, it is considered the same pattern (e.g. two or more successive commas both result in “,+”).
Key words	Keywords refer to 286 words or word sequences obtained from a list of terms indicative for argumentation [13]. Examples from the list are “but”, “consequently”, and “because of”.
Parse features	In the parse tree of each sentence (e.g. Charniak [6]) we used the depth of the tree and the number of subclauses as features.

Table 7: Features for the classification of argumentative information as premise or conclusion

<b>Absolute Location</b>	Position of sentence absolutely in document; 7 segments
<b>Sentence Length</b>	A binary feature, which indicates that the sentence is longer than a threshold number of words (currently 12 words).
<b>Tense of Main Verb</b>	Tense of the verb from the main clause of the sentence; having as nominal values “Present”, “Past” or “NoVerb”.
<b>History</b>	The most probable argumentative category (among the 5 categories) of previous and next sentences (range 1 to 5).
<b>Information 1st Classifier</b>	The sentence has been classified as argumentative or non-argumentative by a first classifier in a preprocessing step.
<b>Rhetorical Patterns</b>	Type of rhetorical pattern occurring on current, previous and next sentences (e.g. “however,”); we distinguish 5 types (Support, Against, Conclusion, Other or None).
<b>Article Reference</b>	A binary feature indicating whether the sentence contains a reference to an article of the law, detected with a POS tagger [27].
<b>Article</b>	A binary feature indicating that the sentence includes the definition of an article detected again with the help of a POS tagger [27] and a collection of legal terminology.
<b>Argumentative Patterns</b>	Type of argumentative pattern occurring in sentence; we have distinguished 5 types of patterns in accordance with our 5 categories (e.g. “see, mutatis mutandis,” “having reached this conclusion”, “by a majority”).
<b>Type of Subject</b>	The agent of the sentence is: the applicant, the defendant, the court or other. The type of agent is detected with a POS tagger. The types of subject were manually drafted taking into account the roles interacting on a ECHR case.
<b>Type of Main Verb</b>	Argumentative type of the main verb of the sentence; we distinguish 4 types (premise, conclusion, final decision or none), manually implemented as a list of corresponding verbs, which are detected in the text with a POS tagger [27].

their contribution in the classification of sentences as argumentative. A classification system was trained with the feature vectors of sentences that were manually annotated.

#### 4.2.1 Argumentative information detection

We used the maximum entropy model, which adheres to the maximum entropy principle [2]. This principle states that, when we make inferences based on incomplete information, we should draw them from that probability distribution that has the maximum entropy permitted by the information we have. In natural language we often deal with incomplete patterns in our training set given the variety of natural language patterns that signal similar content. Hence, this type of classifier is frequently used in information extraction from natural language texts, which motivates our choice of this classifier. Second, the naïve Bayes classifier, specifically a multinomial naïve Bayes classifier [16] is often used. It learns a model of the joint probability of an element  $x$  and its label  $y$ ,  $p(x, y)$ , and makes its predictions by using Bayes rule to calculate  $p(y|x)$  and then selects the most likely label  $y$ . It makes the simplifying (naïve) assumption that the individual features are conditionally independent given the class. The features are typically represented as binary values.

We used the features presented in Table 6 and we obtained nearly 73% accuracy when detecting argumentation in the Araucaria corpus [20]. Note that we do not detect argumentation schemes, just where the argumentative information is found on the text. The accuracy increases to 80% when the task is performed on the ECHR corpus<sup>2</sup>. These results prove that the classification of sentences as argumentative or non-argumentative is feasible. At this point we want to make clear that we never aimed to study how rhetorical relations and argumentative relations interact, i.e. we did not try to determine for example how many causal relations are also involved in argumentation. Our aim is only to determine which parts of the text are argumentative and how they interact (from an argumentative point of view) using only the information that can be extracted from the words in the text, their positions and linguistic connections.

Table 8: Results from the classification of *Conclusions* in the ECHR

Classifier Combination	Precision	Recall	F-Measure
Maximum Entropy and Support Vector Machine	77.49	60.88	74.07
Context-free Grammar	61.00	75.00	67.27

<sup>1</sup> In a binary classification, accuracy is the proportion of true results (both true positives and true negatives) in the population.

<sup>2</sup> The results presented in [18] were 90%, but the evaluation was done on a previous version of the ECHR corpus. The new version uses the same texts but with an improved human annotation, where a higher agreement between annotators is achieved.

Table 9: Results from the classification of *Premises* in the ECHR

Classifier Combination	Precision	Recall	F-Measure
Maximum Entropy and Support Vector Machine	70.19	66.16	68.12
Context-free Grammar	59.00	71.00	64.03

#### 4.2.2 Argumentative information boundaries

Once the argumentative information has been set apart from the non-argumentative information we need to focus on how to determine the boundaries of each argument, i.e. where does one argument starts and where it does finish.

Some first solutions to this segmentation problem are as follows. First, it is possible to use the structure of the document, i.e. the sections and subsections, to determine where an argument starts or ends. This approach assumes that an argument can not extend between sections or subsections. However, this has obvious limitations, as it is not hard to think of an argument divided in different subsections, one presenting the premises and another the conclusion. Therefore, this option is dependent on the type of text at hand. A second option aims to understand the semantics of the different arguments. For example, one could calculate the semantic distance between the different argumentative units (e.g., sentences), and group sentences in one argument if they discuss content that is semantically related. Besides computing semantic relatedness this method must deal with ambiguity, coreference and pronoun resolution.

We assume that the relatedness of two sentences is a function of the relatedness of their words. There are several approaches for calculating semantic relatedness of words, the most important being ontology- and corpus-based. In the former, the relatedness of words depends on their semantic distances in a lexico-semantic resource such as WordNet [5]. In corpus-based semantic measurement the semantic relatedness is calculated by exploring statistical word correlations. It is assumed that similar words usually occur with the same surrounding words.

### 4.3 Argumentative Proposition Classification

If the detection of the argumentative propositions of a text is possible, then it seems that the classification of these propositions by their argumentative function should also be feasible. Following our formalism, we have studied the classification between premises and conclusions. Our approach is again to work with statistical classifiers.

We first classify the clauses of sentences, obtained using a parsing tool, as being argumentative or not with a maximum entropy classifier (see previous section) using the features discussed in Table 6). In a second step we use a second classifier, a support vector machine for classifying each argumentative clause found into a premise or conclusion. Here, we use more sophisticated features (see Table 7). From the new features, the first three features are selected in accordance with our previous work, as they are based on the general structure of the text and each sentence. The *History* feature models local context; it takes the category of the previous sentence as a feature, as there are often patterns of categories following each other. During the testing, the category of the previous sentence is only probabilistically known, which is why beam search is performed. We define the following novel features. For example, the *Rhetorical Patterns* feature models discursive relations, but it distinguishes the presence of discursive

cues highly related to argumentation, expressed in two types (*premise* or *conclusion*), and the discursive cues (*other*) which are not related to the presence of a premise or conclusion. The *Article* and *Article reference* tend to mark the role of premises, while *Type of Main Verb* signals that verbs such as *conclude* or *decide* have a higher chance of being the main verb of a conclusion than verbs like *recall* or *note*. These verbs are common in argumentative speech in the legal domain, but not restricted to it, and furthermore, they can be easily extended based on linguistic knowledge, e.g. using the verb classes defined in [14].

Table 8 and Table 9 show the best results for clause classification into premise or conclusion on the ECHR corpus, attaining a 68.12 % and 74.07 %  $F_1$  measure respectively.<sup>3</sup>

#### 4.4 Detection of the Argumentation Structure

The detection and classification of argumentative propositions by statistical classifiers has been analysed in the previous section, however this approach does not allow the detection of relations between full arguments. To determine the limits of an argument and the relations it holds with other surrounding arguments is a difficult task. First, there is no limit to the length an argument can take, and we lack any knowledge on what the most probable structures formed by premises and conclusions are. Secondly, even if the argument limits could be detected, how can we know which are the most probable relations between it and other arguments?

Motivated by the work of Marcu [17] on Rhetorical Structure Theory and the research done in sentence parsing, i.e. determining the POS of every word in a sentence, we have studied the possibility of argumentative parsing. There exist different parsing approaches: rule-based (hand-crafted, transformation-based learning) or statistical (Hidden Markov Model, maximum entropy model, memory-based, decision tree, neural network, linear models), but for the time being we have focused on parsing the texts by means of manually derived rules that are grouped into a context-free grammar (CFG).

A CFG defines a formal language, i.e. the set of all sentences (strings of words) that can be derived by the grammar. Sentences in this set are said to be grammatical, while sentences outside this set are said to be ungrammatical. Formally a context-free grammar  $G$  is described as  $G = \langle T, N, S, R \rangle$  where:  $T$  is the set of terminal symbols (represented with non-capital letters), i.e. symbols that form the parts of the statements,  $N$  is the set of non-terminal symbols (represented with capital letters), i.e. symbols that generate statements by substitution of either other nonterminals or terminals or some combination of these,  $S$  is the start symbol and  $R$  are the rules/productions of the form  $X \rightarrow \beta$ , where  $X$  is a non-terminal symbol and  $\beta$  is a sequence of terminal and non-terminal symbols.

Argumentation structure can be seen as a formal language as argumentation is a set of sentences that can be generated from terminal and non-terminal symbols following some rules.

---

<sup>3</sup> The  $F_1$  (or F-measure) is a measure of a test's accuracy. It considers both the precision  $p$  and the recall  $r$  of the test to compute the score:  $p$  is the number of correct results divided by the number of all returned results and  $r$  is the number of correct results divided by the number of results that should have been returned. The  $F_1$  can be interpreted as a weighted average of the precision and recall, where an  $F_1$  score reaches its best value at 1 and worst score at 0.

A simple example of a CFG to parse single sentences, e.g. “*The house is big*” or “*The car is old*” could be as follows:

$$\begin{aligned}
S &\Rightarrow NP VP \\
NP &\Rightarrow Det N \\
VP &\Rightarrow V Adj \\
Det &= \{the, a\} \\
N &= \{house, car, birds\} \\
V &= \{is, are\} \\
Adj &= \{big, old\}
\end{aligned}$$

Where  $S$  stands for sentence,  $NP$  for Noun Phrase,  $VP$  for Verb Phrase,  $Det$  for Determiner,  $N$  for Noun,  $V$  for Verb and  $Adj$  for Adjective. The sentence “*The house is big*” is produced from the following rules sequence:  $S \Rightarrow NP VP \Rightarrow Det N VP \Rightarrow the N VP \Rightarrow the house VP \Rightarrow the house V Adj \Rightarrow the house is Adj \Rightarrow the house is big$ .

Argumentative parsing is a difficult task, therefore we focus our efforts on proving that it is a promising approach and we restrict our research to a limited complexity. Our approach is for the moment only related to the legal domain, which makes the task easier, at least when drafting the rules manually.

$$\begin{aligned}
T &\Rightarrow A^+ D \\
A &\Rightarrow \{A^+ C | A^* C n P^+ | C n s | A^* s r_c C | P^+\} \\
D &\Rightarrow r_c f \{v_c s | \cdot\}^+ \\
P &\Rightarrow \{P_{verbP} | P_{art} | P P_{sup} | P P_{ag} | s P_{sup} | s P_{ag}\} \\
P_{verbP} &= s v_p s \\
P_{art} &= s r_{art} s \\
P_{sup} &= \{r_s\} \{s | P_{verbP} | P_{art} | P_{sup} | P_{ag}\} \\
P_{ag} &= \{r_a\} \{s | P_{verbP} | P_{art} | P_{sup} | P_{ag}\} \\
C &= \{r_c | r_s\} \{s | P_{verbP} | C\} \\
C &= s^* v_c s
\end{aligned}$$

Fig. 4: Context-free grammar used for argumentation structure detection and proposition classification

Using information extracted from 10 ECHR documents, which are excluded from the evaluation set, we define the context-free grammar shown in Figure 4 using the terminal and non-terminal symbols defined in Table 10. We implement the grammar using Java and JSCC<sup>5</sup>.

<sup>5</sup> <http://jsc.c.jmksf.com/>



Table 10: Terminal and non-terminal symbols from the context-free grammar used in the argumentation structure detection

$T$	General argumentative structure of legal case.
$A$	Argumentative structure that leads to a final decision of the factfinder <sup>4</sup> $A = \{a_1, \dots, a_n\}$ , each $a_i$ is an argument from the argumentative structure.
$D$	The final decision of the factfinder $D = \{d_1, \dots, d_n\}$ , each $d_i$ is a sentence of the final decision.
$P$	One or more premises $P = \{p_1, \dots, p_n\}$ , each $p_i$ is a sentence classified as premise.
$P_{ag}$	Premise with at least one contrast rhetorical marker.
$P_{art}$	Premise with at least one article rhetorical marker.
$P_{sup}$	Premise with at least one support rhetorical marker.
$P_{verbP}$	Premise with at least one verb related to a premise.
$C$	Sentence with a conclusive meaning.
$n$	Sentence, clause or word that indicates one or more premises will follow.
$s$	Sentence, clause or word neither classified as a conclusion nor as a premise ( $s! = \{C P\}$ ).
$r_c$	Conclusive rhetorical marker (e.g. therefore, thus, ...).
$r_s$	Support rhetorical marker (e.g. moreover, furthermore, also, ...).
$r_a$	Contrast rhetorical marker (e.g. however, although, ...).
$r_{art}$	Article reference (e.g. terms of article, art. para. ....).
$v_p$	Verb related to a premise (e.g. note, recall, state,...).
$v_c$	Verb related to a conclusion (e.g. reject, dismiss, declare, ...).
$f$	The entity providing the argumentation (e.g. court, jury, commission, ...).

To determine which terminal symbols compose each sentence we focus on common expressions encountered in the legal documents, such as “*For these reasons*”, “*in the light of all the material*” or “*see mutatis mutandis*”, and rhetorical markers, such as “*However*” or “*Furthermore*”.

The grammar can be overviewed as follows: each document has a tree-structure ( $T$ ). Each tree-structure ( $T$ ) is formed by an argument ( $A$ ) and a decision ( $D$ ). The decision ( $D$ ) can be form by one or more sentences that contain: (1) conclusive rhetorical marker and reference to the entity providing the argument or (2) the previous and conclusive verb. The argument ( $A$ ) can be formed by:

- other arguments and a conclusion ( $A^+C$ )
- possible arguments, a conclusion, a nexus and one or more premises ( $A^*CnP^+$ )
- a conclusion, a nexus and a sentence, clause or word neither classified as a conclusion nor as a premise ( $Cns$ )
- possible arguments, a sentence, clause or word neither classified as a conclusion nor as a premise and a conclusion ( $A^*sr_cC$ )
- just one or more premises ( $P^+$ )

A premise ( $P$ ) can be:

- a premise with an occurrence of a verb of premise ( $P_{verbP}$ )
- a premise with a reference to an article ( $P_{art}$ )
- one or more premises followed by a premise with a support reference ( $PP_{sup}$ )
- one or more premises followed by a premise with a contrast reference ( $PP_{ag}$ )
- a sentence, clause or word neither classified as a conclusion nor as a premise followed by a premise with a support reference ( $sP_{sup}$ )
- a sentence, clause or word neither classified as a conclusion nor as a premise followed by a premise with a contrast reference ( $sP_{ag}$ )

A conclusion ( $C$ ) can be:

- a sentence, clause or word neither classified as a conclusion nor as a premise containing a conclusive verb ( $s^*v_{cs}$ )
- a conclusive rhetorical marker followed by a sentence, clause or word neither classified as a conclusion nor as a premise ( $r_{cs}$ )
- a conclusive rhetorical marker followed by a premise with a verb of premise ( $r_cP_{verbP}$ )
- a support rhetorical marker followed by a conclusion ( $r_sC$ )

Using the context-free grammar for parsing the texts from the ECHR corpus we obtain around 60 % accuracy in detecting the argumentation structures. This is measured manually comparing the full structures given by our tool and the structures given by the annotators. We check if all the argumentative information is included in the structure, if the individual arguments are well-constructed and if the connections between arguments are correct. In most cases the argumentative information is all included in the structures and most of the problems involve connecting subordinate arguments as coordinate and vice versa. The resulting structures maintain around 70%  $F_1$  on the classification between premises and conclusions.

```

T
|--D
| |--: For these reasons, the Commission by a majority declares the application admissible,
|       without prejudging the merits.
|--A
| |--A
| | |--C
| | | |--: It follows that the application cannot be dismissed as manifestly ill-founded.
| | |--A
| | | |--P
| | | | |--: It considers that the applicant 's complaints raise serious issues of fact
| | | | and law under the convention, the determination of which should depend on
| | | | an examination of the merits.
| | |--P
| | | |--: The Commission has taken cognizance of the submissions of the parties.
|--A
| |--C
| | |--: In these circumstances, the Commission finds that the application cannot be
|         declared inadmissible for non-exhaustion of domestic remedies.
|--A
| |--P
| | |--: The Commission recalls that article art. x of the convention only requires
|         the exhaustion of such remedies which relate to the breaches of the
|         convention alleged and at the same time can provide effective and sufficient
|         redress.
|--P
| | |--: The Commission notes that in the context of the section powers the
|         secretary of state has a very wide discretion.
|--P
| | |--: The Commission recalls that in the case of temple v. the united kingdom
|         no. x dec. d.r. p.
|--P
| | |--: The Commission held that recourse to a purely discretionary power on
|         the part of the secretary of state did not constitute an effective
|         domestic remedy.
| |--: The Commission finds that the suggested application for discretionary
|         relief in the instant case cannot do so either.

```

Fig. 5: Output of the automatic system: small fragment of the argumentation tree-structure of a document

Figure 5 is a small fragment of the tree-structure obtained in a test of our automatic system without human intervention. It contains the top-level of the tree, i.e. the decision (D) and the main argumentation (A), which contains six arguments. The figure shows only two of them, which are the ones closer to the end of the document, i.e. closer to the final decision. The first is a single conclusion (C) and two premises (P). The second is a single conclusion (C) and four premises. The first three premises are single and were found thanks to the verbs *notes* and *recalls*; therefore internally they are  $P_{verbP}$ . Note that the first one could have been detected by the article reference, but as the CFG applies rules by order the first choice is  $P_{verbP}$ . The last premise contains two premises, the first a sentence, clause or word neither classified as a conclusion nor as a premise (s) and the second a support premise ( $P_{sup}$ ) due to the word *either* and the verb *finds*.

## 5 Applications of Argumentation Mining in Law

Argumentation has been recognised as a key area of importance in Artificial Intelligence (AI) and Law for over a decade. The study of argumentation in general AI is receiving increasing attention as AI comes to recognise that a persuasive or convincing argument is not the same as a formal mathematical proof. However, little research addressing this theme has yet found its way to practice.

Arguments, as found in case files and judicial decisions, can often be rather complex, so that understanding the web of relationships becomes difficult. Argumentation mining addresses this need in research on argument structuring, which has uses in areas where the clear presentation of the argument is of prime importance, such as preliminary fact investigation, teaching or case management. When used for case management, a software using argumentation mining and other natural language and information retrieval techniques could allow the user to structure a collection of case-related documents in terms of the argumentation structure of a case. The structure would capture: the main issues; the main positions and arguments taken by the parties with respect to the issues; the available evidence related to them; and so on. Documents (statutes, case law, journal articles, testimony, letters, and so on) could be indexed following the argumentative structure, i.e. relating them to the relevant argumentation parts of the case, making it easier to find the evidence that supports or attacks that part of the argumentation.

In case-based reasoning systems, such as [33], [34], [46], [47], [48] or [49], where expertise is embodied in a library of past cases and the current problem is matched against those past cases to retrieve only the similar cases, argumentation mining offers a new way to compare cases, i.e. by their argumentation structure. Lawyers could compare cases by the amount of arguments that were necessary to prove a common conclusion. Then they would find which argumentation was more efficient and complete, i.e. the minimal set of arguments that can not be defeated. In another more complex application a lawyer could be interested in searching for all the previous cases which used a type of argument to attack another; here we refer to for example the use of an argument of expert opinion to attack a previous argument. The lawyer's intention could be to find examples where it yields a favourable result and examples where the factfinder (e.g. court) was not convinced. Then the lawyer would analyze how those arguments diverge, and therefore he would find the most effective or persuasive presentation of the facts. The current state of argumentation mining research would not allow this.

However, we are confident that once the argumentation structure is well determined a classification of its arguments would be possible.

Furthermore, formal or computational models of legal argument have been designed and applied to knowledge bases with formalised knowledge to produce automatic decision systems [42]. This research finds its main obstacle in the knowledge engineering bottleneck, that makes their practical applicability restricted. However, the use of argumentation mining can offer an easy way to obtain arguments to update the bases, opening new avenues for reasoning system development.

Another important research area in argumentation has focused on argument visualisation tools, such as Araucaria [40]. Most of these tools are stand-alone tools for visualising the structure of a single document or a single argument. However some of them are starting to incorporate analysis of different sources, e.g. finding arguments in different webpages and their relations. Extending any of these visualisation tools with argumentation mining would open new possibilities for the users of these tools. For example, the system could offer a suggestion of argumentation for the given text allowing the user to modify it if needed. Another option would be that the system automatically creates an argumentation structure for the selected text by the user.

Finally, in areas such as management or finances, argumentation mining can provide a better way to maintain on-going discussions or large and long meetings. It presents a more structured visualization of the discussed points and the arguments used by the different parties, keeping track of which points have already been accepted or rejected, and under which arguments or lack of arguments that happened. Therefore, new members in the discussion will be provided with a better understanding of the current state of the discussion and if the discussion has to re-start after a long break restatements of arguments or reopening the discussion of previously well argued topics could be avoided.

## 6 Conclusions and future work

Research on argumentation mining is still limited and leaves many interesting challenges. However, our experiments prove that it is possible to automatically detect written argumentation and to structure it with general AI methods. We have presented a framework that follows state-of-the-art argumentation theories enabling an automatic approach to argumentation detection. Our experiments on distinguishing between argumentative and non-argumentative information in text show better than expected results, specially for the legal texts.

The future of argumentation mining depends on three main aspects: (a) the study and creation of new corpora to capture all possible types of argumentation in all scenarios, (b) the implementation of new techniques to detect more complex argumentation structures than simple trees and (c) an extended evaluation of all the methods by legal experts and users.

## References

1. J. Allen. *Natural Language Understanding (2nd ed.)*. Benjamin-Cummings Publishing Co., Inc., Redwood City, CA, USA, 1995.
2. A. L. Berger, V. J. D. Pietra, and S. A. D. Pietra. A maximum entropy approach to natural language processing. *Computational Linguistics*, 22(1):39–71, 1996.

3. P. Besnard and A. Hunter. *Elements of Argumentation*. The MIT Press, 2008.
4. L. Birnbaum. Argument molecules: a functional representation of argument structures. In *Proceedings of the Third National Conference on Artificial Intelligence (AAAI-82)*, pages 63–65, Pittsburgh, PA, August 18–20 1982.
5. E. Budanitsky and G. Hirst. Evaluating WordNet-based measures of lexical semantic relatedness. *Computational Linguistics*, 32:13–47, 2006.
6. E. Charniak. A Maximum-Entropy-Inspired Parser. Technical Report CS-99-12, 1999.
7. R. Cohen. Analyzing the Structure of Argumentative Discourse. *Computational Linguistics*, 13:11–24, 1987.
8. L. Fontan and P. Saint-Dizier. Constructing a know-how repository of advices and warnings from procedural texts. In *DocEng '08: proceedings of the Eighth ACM Symposium on Document Engineering*, pages 249–252, New York, NY, USA, ACM, 2008.
9. B. Hachey and C. Grover. Automatic legal text summarisation: experiments with summary structuring. In *ICAIL '05: proceedings of the 10th International Conference on Artificial Intelligence and Law*, pages 75–84, New York, NY, USA, ACM, 2005.
10. H. Horacek and M. Wolska. Interpreting semi-formal utterances in dialogs about mathematical proofs. *Data Knowledge Engineering*, 58(1):90–106, 2006.
11. E. H. Hovy. Automated discourse generation using discourse structure relations. *Artificial Intelligence*, 63(1-2):341–385, 1993.
12. D. Jurafsky and J.H. Martin. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. Prentice Hall (Series in Artificial Intelligence), 2 edition, February 2009.
13. A. Knott and R. Dale. Using linguistic phenomena to motivate a set of rhetorical relations. Technical report, Discourse Processes, 1993.
14. B. Levin. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: The University of Chicago, 1993.
15. W. C. Mann and S. A. Thompson. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281, 1988.
16. C. D. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing*. The MIT Press, June 1999.
17. D. Marcu. The rhetorical parsing of unrestricted texts: a surface-based approach. *Computational Linguistics*, 26(3):395–448, September 2000.
18. R. Mochales and M.-F. Moens. Study on Sentence Relations in the Automatic Detection of Argumentation in Legal Cases. In *Legal Knowledge and Information Systems. Jurix 2007*. IOS Press, 2007.
19. R. Mochales and M.-F. Moens. Study on the Structure of Argumentation in Case Law. In *Legal Knowledge and Information Systems. Jurix 2008*. IOS Press, 2008.
20. M.-F. Moens, E. Boiy, R. Mochales, and C. Reed. Automatic detection of arguments in legal texts. In *ICAIL '07: Proceedings of the 11th International Conference on Artificial Intelligence and Law*, pages 225–230, New York, NY, USA, ACM Press, 2007.
21. M. Moser and J. D. Moore. Toward a synthesis of two accounts of discourse structure. *Computational Linguistics*, 22:22–3, 1996.
22. T. J. Sanders. Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse Processes*, 24:119–147, 1997.
23. S. Teufel. *Argumentative Zoning: Information Extraction from Scientific Text*. PhD thesis, University of Edinburgh, 1999.
24. S. Teufel and M. Moens. Sentence Extraction as a Classification Task. In *Workshop 'Intelligent and scalable Text summarization', ACL/EACL*, 1997.
25. S. E. Toulmin. *The Uses of Argument*. Cambridge University Press, July 1958.
26. F. H. Van Eemeren and Grootendorst. *A Systematic Theory of Argumentation: The pragma-dialectic approach*. Cambridge University Press, 2004.
27. J. Wagner, D. Seddah, J. Foster, and J. van Genabith. C-structures and F-structures for the British National Corpus. In M. Butt and T. H. King, editors, *The Proceedings of the LFG '07 Conference*, pages 418–438, University of Stanford, California, USA, 2007.
28. D. N. Walton. *Argumentation Schemes for Presumptive Reasoning*. Erlbaum, Mahwah NJ, USA, 1996.
29. D. N. Walton. *The New Dialectic, Conversational Contexts of Argument*. University of Toronto Press, 1998.
30. D. N. Walton. *Legal Argumentation and Evidence*. Penn State Press, 2002.
31. D. N. Walton, C. Reed, and F. Macagno. *Argumentation Schemes*. Cambridge University Press, 2008.

- 
32. J. H. Wigmore. *The Principles of Judicial Proof*. Little, Brown & Co, 1931.
  33. V. Aleven and K. D. Ashley. *Teaching Case-Based Argumentation through a Model and Examples Empirical Evaluation of an Intelligent Learning Environment*. 1997.
  34. K. D. Ashley. *Reasoning with cases and Hypotheticals in HYPO*. *Int. J. ManMachine Studies*, 1991, Vol. 34, pp. 753–796.
  35. J. Cohen. *A Coefficient of Agreement for Nominal Scales*. *Educational and Psychological Measurement*, Vol. 20, No. 1. (1 April 1960), pp. 37-46.
  36. F. Wolf and E. Gibson. *Coherence in Natural Language: Data Structures and Applications*. MIT Press, June 2006.
  37. R. H. Johnson and J. A. Blair. Informal logic and the reconfiguration of logic. Dov M. Gabbay, Ralph H. Johnson, Hans Jrgen Ohlbach and John Woods (eds.), *Handbook of the Logic of Argument and Inference: Turn Towards the Practical, Amsterdam and New York: Elsevier*, pp. 340-396. 2002.
  38. P. M. Dung. *On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games*. *Artificial Intelligence*, 77: 321-357, 1995.
  39. I. Rahwan and G. Simari *Argumentation in Artificial Intelligence*. X. 494 p. 100 illus., Hardcover, 2009.
  40. C. Reed and G. Rowe. *ARAUCARIA: ARGUMENT DIAGRAMMING AND XML, Araucaria: Software for Puzzles in Argument Diagramming and XML*.
  41. P. A. Kirschner, S. J. Buckingham Shum and C. S. Carr. *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*. Springer-Verlag: London, 2003.
  42. T. J. Bench-Capon and K. Atkinson and P. McBurney. Altruism and agents: an argumentation based approach to designing agent decision mechanisms. *AAMAS (2), 2009*, pp. 1073-1080.
  43. S. Negri and J. von Plato. *Structural Proof Theory*. Cambridge, 2001.
  44. J. Barwise. An introduction to first-order logic. *Handbook of Mathematical Logic, Studies in Logic and the Foundations of Mathematics*, Amsterdam: North-Holland, 1977.
  45. Edited by James F. Voss and Julie Van Dyke. *Argumentation in Psychology*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 2002.
  46. K. Ashley. *Modeling Legal Argument: Reasoning with Cases and Hypotheticals*. Cambridge, MA: MIT Press, 1990.
  47. K. Ashley. Case-based reasoning. In A. Oskamp en A. Lodder (Eds.), *Information Technology and Lawyers* (pp. 21-45). Berlin: Springer, 2006.
  48. L. K. Branting. *Reasoning with Rules and Precedents: A Computational Model of Legal Analysis*. Boston, MA: Kluwer Academic Publishers, 2000.
  49. E. Rissland and D. Skalak. CABARET: Statutory interpretation in a hybrid architecture. *International Journal of Man-Machine Studies*, 34, 1991.
  50. S. Brninghaus and K. D. Ashley. *Predicting Outcomes of Case-Based Legal Arguments*. In ICAIL '03: proceedings of the 9th International Conference on Artificial Intelligence and Law, 233-242, 2003.