ARGUMENTATION MINING

Marie-Francine Moens joint work with Raquel Mochales Palau and Parisa Kordjamshidi
Language Intelligence and Information Retrieval
Department of Computer Science
KU Leuven, Belgium
FIRE 2013, New Delhi, India
OUTLINE

- Definition of argumentation mining
- Importance of the task
- Current methods and results
- Promising directions to improve the results
- Some applications
ARGUMENTATION MINING

- the detection of an argumentative discourse structure in text or speech, and the detection and the functional classification of its composing components
Argumentation mining = recognition of a rhetorical structure in a discourse

Rhetoric is the art of discourse that aims to improve the capabilities of writers and speakers to inform, persuade or motivate particular audiences in specific situations

ARGUMENTATION

- Is probably as old as mankind
- Has been studied by philosophers throughout the history
From Ancient Greece to the late 19th century: central part of Western education: need to train public speakers and writers to move audiences to action with arguments.

Until the 1950s, the approach of argumentation was based on rhetoric and logic.

Argumentation was/is taught at universities.
SOME HISTORY

■ Highlights:
  ▪ Aristotle’s logical works: *Organon*
  ▪ George Pierce Baker (1895). *The Principles of Argumentation*, 1895
  ▪ Chaïm Perelman describes of techniques of argumentation used by people to obtain the approval of others for their opinions: *Traité de l’argumentation – la nouvelle rhétorique*, 1958
Argumentation in text

One of most the fundamental things we use language for is argument. Arguing means claiming that something is true and trying to persuade other people to agree with your claim by presenting evidence to substantiate it. An argument is statement with three components:

1. A point of view, a claim, something we are arguing in favour of
2. The actual argument, the evidence we are using to argue with
3. A statement that links the initial claim to the argument and ensures that we understand how the argument functions.

The statement that connects the initial claim and the argument is referred to as the warrant. The warrant is thus an argument for the connection between the initial claim and the argument.

We find argumentation in:

- Legal texts and court decisions
- Scientific texts
- Patents
- Reviews
- Debates
- ...

FIRE 2013
In the overload of information users want to find arguments that sustain a certain claim or conclusion.

Argumentation mining refines:
- Search and information retrieval
- Provides the end user with instructive visualizations and summaries of an argumentative structure

Argumentation mining is related to opinion mining, but end user wants to know the underlying grounds and maybe counterarguments.
WHAT IS THE STATE-OF-THE-ART?

- Argumentative zoning
- Argumentation mining of legal cases
= segmentation of a discourse into discourse segments or zones that each play a specific rhetoric role in a text

ARGUMENTATIVE ZONING

BKG: General scientific background (yellow)

OTH: Neutral descriptions of other people’s work (orange)

OWN: Neutral descriptions of the own, new work (blue)

AIM: Statements of the particular aim of the current paper (pink)

TXT: Statements of textual organization of the current paper (in chapter 1, we introduce...) (red)

CTR: Contrastive or comparative statements about other work; explicit mention of weaknesses of other work (green)

BAS: Statements that own work is based on other work (purple)

[PHD thesis of Simone Teufel 2000]
Methods: seen as a classification task: rule based classifier or classifier (e.g., naïve Bayes, support vector machine) is trained with manually annotated examples

[Teufel, S. & Moens, M. *ACL* 1999]
[Teufel, S. & Moens, M. *EMNLP* 2000]
[Hachey, B. & Grover, C. *ICAIL* 2005]
ARGUMENTATION MINING OF LEGAL CASES

- Legal field:
  - Precedent reasoning
  - Search for cases that use a similar type of reasoning, e.g., acceptance of rejection of a claim based on precedent cases
  - Adds an additional dimension to argumentative zoning:
    - Needs detection of the argumentation structure and classification of its components
    - Components or segments are connected with argumentative relationships
Figure 1.1: Reasoning structure of the legal case in Appendix A. Each block is a sentence of the legal case. There are 3 arguments (blue, green and red) that justify the final decision (brown). The contents of each argument and the final decision can be seen in detail in Figures 1.2, 1.3, 1.4 and 1.5.
The Court notes that this complaint is not manifestly ill-founded within the meaning of Article 35 § 3 of the Convention.

It further notes that it is not inadmissible on any other grounds.

It must therefore be declared admissible.

Figure 1.2: Closer view 1st Argument

Article 41 of the Convention provides:
"If the Court finds that there has been a violation of the Convention or the Protocols thereto, and if the internal law of the High Contracting Party concerned allows only partial reparation to be made, the Court shall, if necessary, afford just satisfaction to the injured party."

The applicant has not filed a claim for just satisfaction.

 Accordingly, the Court considers that no award can be made under this provision.

Figure 1.3: Closer view 2nd Argument

FOR THESE REASONS, THE COURT UNANIMATELY
1. Declares the application admissible;
2. Holds that there has been a violation of Article 6 § 1 of the Convention

Figure 1.4: Closer view Final Decision

[PhD thesis Raquel Mochales Palau]
Figure 1.5: Closer view 3rd Argument
Argumentation: a process whereby arguments are constructed, exchanged and evaluated in light of their interactions with other arguments

**Argument**: a set of **premises** - pieces of evidence - in support of a claim

**Claim**: a proposition, 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
For these reasons, the Commission by a majority declares the application admissible, without prejudging the merits.

It follows that the application cannot be dismissed as manifestly ill-founded.

The Commission has taken cognizance of the submissions of the parties.

In these circumstances, the Commission finds that the application cannot be declared inadmissible for non-exhaustion of domestic remedies.

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.

The Commission notes that in the context of the section powers the secretary of state has a very wide discretion.

The Commission recalls that in the case of temple v. the united kingdom no. x dec. d.r 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. 6: Output of the automatic system: small fragment of the argumentation tree-structure of a document.
Experiments with decisions of the European Court of Human Rights (ECHR)

[Mochales & Moens, AI & Law 2011]
Features of classifier: Clauses described by unigrams, bigrams, adverbs, legal keywords, word couples over adjacent clauses, ...

Context free grammar allows also to recognize the full argumentation structure: accuracy: 60%

[Mochales & Moens, AI & Law 2011]
FUTURE WORK

- Joint recognition of a claim and its composing arguments
- Learning of event relationships
- Joint recognition with latent variables
- Integration in retrieval and visualization models
Promising structured learning approaches: e.g., segmenting and jointly classifying the argumentation components

Can be expanded to the joint recognition of nested arguments as found in legal cases

Or to the Toulmin model or the many different argumentation schemes discussed in Douglas Walton (1996). *Argumentation Schemes for Presumptive Reasoning*. Mahwah, New Jersey: Lawrence Erlbaum Associates
Structured learning: modeling of interdependence among output labels:

- Probabilistic graphical models [Koller and Friedman 2009]
- Generalized linear models, e.g., structured support vector machines and structured perceptrons [Tsochantaridis et al. JMLR 2006]

The interdependencies between output labels and other background knowledge can be imposed using constraint optimization techniques during prediction and training.
Considering the interdependencies and structural constraints over the output space easily leads to intractable training and prediction situations:

- Models for decomposition, communicative inference, message passing, ...
- [PhD of Parisa Kordjamshidi 2013] [Kordjamshidi & Moens NIPS workshop 2013]
The discourse structure is often signaled by typical keywords (e.g., in conclusion, however, ...), but often this is not the case.

Humans who understand the meaning of the text can infer whether a claim is a plausible conclusion given a set of premises, or a claim rebuts another claim.

=> Background or domain knowledge that an argumentation mining tool should also acquire: how?

Work on event causality: [Xuan Do et al. EMNLP 2011]
Semi-supervised induction of discourse parse grammars: e.g., by means of inside outside algorithm

Warrant as a latent variable?
Visualization: e.g., work of Chris Reed [Reed & Rowe IJAIT 2004]: the recognized argumentation scheme can be easily visualized

Retrieval: need for search tools that take into account argumentative reasoning
OPINION MINING: FINDING ARGUMENTS AND COUNTER ARGUMENTS FOR AN OPINING EXPRESSED:

- Find support for the opinion, explain the opinion

An opinion, whether it is grounded in fact or completely unsupportable, is an idea that an individual or group holds to be true. An opinion does not necessarily have to be supportable or based on anything but one's own personal feelings, or what one has been taught. An argument is an assertion or claim that is supported with concrete, real-world evidence.

[http://wiki.answers.com]
POSSIBLE APPLICATIONS

- Mining of the supporting evidence of claims in scientific publications and patents and their visualization for easy access

[http://undsci.berkeley.edu/article/howscienceworks_07]
Digital humanities: finding and comparing the arguments that politicians use in their speeches:

- Then that little man in black there, he says women can't have as much rights as men, 'cause Christ wasn't a woman! Where did your Christ come from? Where did your Christ come from? From God and a woman! Man had nothing to do with Him. [Sojourner Truth (1797-1883): Ain't I A Woman?, Delivered 1851, Women's Convention, Akron, Ohio]
The Araucaria corpus (constructed by Chris Reed at the University of Dundee, 2003)

Sources:
- 19 newspapers (from the UK, US, India, Australia, South Africa, Germany, China, Russia and Israel, in their English editions)
- 4 parliamentary records (in the UK, US and India)
- 5 court reports (from the UK, US and Canada)
- 6 magazines (UK, US and India)
- 14 further online discussion boards such as Human Rights Watch and GlobalWarming.org

The annotation by experts of the Araucaria collection follows Walton’s classification and argumentation scheme
The **ECHR corpus** annotated by legal experts in 2006 under supervision of Raquel Mochales Palau:

- 25 legal cases
- 29 admissibility reports
- 12,904 sentences, 10,133 non-argumentative and 2,771 argumentative, 2,355 premises and 416 conclusions
CONCLUSIONS

- Argumentation mining: novel and promising research domain
- Potential of structured learning integrating known interdependencies between the structural components in the argumentation and expert knowledge
- Several interesting applications of the technology