Part 3

Existing Resources and Benchmark Data
Outline

Part 3: Existing Resources and Benchmark Data

3.1 Existing Corpora

3.2 Debate Portals and Community Platforms
Argument component identification (CI)
- Requires presence of non-argumentative text units
- Possible on different granularities (e.g. sentence or clauses)

Argument component classification (CC)
- Commonly claim-premise-schemes
- Other component types include different types of evidence

Structure Identification (SI)
- Argumentative relations between text units (e.g. t1 is-evidence-for t2)
- Note: pair classification systems can be used to identify (counter-)evidence for a given claim
Existing Corpora

Granularity

**Micro-level vs. Macro-level**

Micro-level:
internal structure of arguments

Macro-level:
connections between arguments

**Granularity of argument components**

fine → coarse

- clause-level components
- sentence-level components
- multi-sentence components

---

NLP Approaches to Computational Argumentation – ACL 2016 Tutorial
Existing Corpora

Additional Properties

- Document type
- Language
- Single- vs. multi-documents
- Reliability

κ α
## Existing Corpora

### Overview

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Document types</th>
<th>Language</th>
<th>Tasks</th>
<th>Argument granularity</th>
<th>Component granularity</th>
<th>Size (#documents / #comps)</th>
<th>Single-/multi-docs</th>
<th>Reliability</th>
<th>Specials</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Cabrio and Villata, 2014)</td>
<td>several</td>
<td>en</td>
<td>SI</td>
<td>macro</td>
<td>-</td>
<td>792 argument pairs</td>
<td>-</td>
<td>$\kappa = .71$</td>
<td>multiple sources including wiki revision</td>
</tr>
<tr>
<td>(Eckle-Kohler et al. 2015)</td>
<td>news</td>
<td>de</td>
<td>CI+CC</td>
<td>micro</td>
<td>multi</td>
<td>88 / ~1708</td>
<td>single</td>
<td>$\alpha_u = .40$</td>
<td>-</td>
</tr>
<tr>
<td>(Cabrio and Villata, 2014)</td>
<td>user-generated web</td>
<td>en</td>
<td>CI+CC</td>
<td>micro</td>
<td>multi</td>
<td>340 / ~1k</td>
<td>single</td>
<td>$\alpha_u = .48$</td>
<td>several registers and topics</td>
</tr>
<tr>
<td>(Habernal and Gurevych, 2015)</td>
<td>scientific articles</td>
<td>de</td>
<td>SI</td>
<td>micro</td>
<td>sentence</td>
<td>24 / ~2.7k</td>
<td>single</td>
<td>$\kappa = .43$</td>
<td>-</td>
</tr>
<tr>
<td>(Kirschner et al., 2015)</td>
<td>online comments</td>
<td>en</td>
<td>CI+CC</td>
<td>micro</td>
<td>sentence</td>
<td>119 / 240</td>
<td>single</td>
<td>$\kappa = .62$</td>
<td>several claim types</td>
</tr>
<tr>
<td>(Kwon et al., 2007)</td>
<td>court cases</td>
<td>en</td>
<td>CI+CC</td>
<td>micro</td>
<td>sentence</td>
<td>47 / 1,067</td>
<td>single</td>
<td>$\kappa = .75$</td>
<td>-</td>
</tr>
<tr>
<td>(Mochales-Palau and Moens, 2009)</td>
<td>micro texts</td>
<td>de / en</td>
<td>CI+CC+SI</td>
<td>micro</td>
<td>clause</td>
<td>115 / 576</td>
<td>single</td>
<td>$\kappa = .83$</td>
<td>parallel corpus</td>
</tr>
<tr>
<td>(Reed et al., 2008)</td>
<td>various</td>
<td>en</td>
<td>CI+CC+SI</td>
<td>micro</td>
<td>clause</td>
<td>700 / 2k</td>
<td>single</td>
<td>-</td>
<td>argumentation schemes &amp; enthymeme</td>
</tr>
<tr>
<td>(Rinott et al., 2015)</td>
<td>wiki articles</td>
<td>en</td>
<td>CI+CC+SI</td>
<td>micro</td>
<td>clause</td>
<td>547 / ~7.5k</td>
<td>multi</td>
<td>$\kappa = .39$</td>
<td>different types of evidence</td>
</tr>
<tr>
<td>(Stab et al., 2014)</td>
<td>student essays</td>
<td>en</td>
<td>CI+CC+SI</td>
<td>micro</td>
<td>clause</td>
<td>90 / 1.5k</td>
<td>single</td>
<td>$\alpha_u = .72$</td>
<td>-</td>
</tr>
<tr>
<td>(Walker et al., 2012)</td>
<td>online discussions</td>
<td>en</td>
<td>AT</td>
<td>macro</td>
<td>-</td>
<td>390k / -</td>
<td>-</td>
<td>$\frac{.22}{&lt; \kappa &lt; .62}$</td>
<td>several argument attributions</td>
</tr>
</tbody>
</table>

*(CI = Component identification; CC = Component classification; SI = Structure identification; AT = Argument attribution)*
Resources including arguments

NoDE - Natural language arguments in online Debates

Annotated macro-level relations between arguments (Cabrio and Villata, 2014)

- Three heterogeneous sources
  - Debatepedia
  - Script of a play (“12 Angry Men”)
  - Wikipedia revisions

Pairs of arguments annotated as

- Support and Attack

Inter-annotator agreement

- Between $\kappa = 0.7$ and $\kappa = 0.74$
- (depending on the source)
Resources including arguments

*IAC - Internet Argument Corpus*

Arguments from online debates (Walker et al., 2012)

- Source: 4forums.com
- Includes more than 390k posts

Annotated with several attributes

- Topic
- Degree of agreement
- Cordiality
- Audience-direction (previous poster or a wider audience)
- Combativeness
- Assertiveness
- Emotionality
- Sarcasm

Can be used for argument attribution tasks
Annotated argument components

_ECHR_

Micro argument components at the sentence-level

- Domain: legal (Documents from _European Court of Human Rights_)

Version 1 (Mochales-Palau and Moens, 2008)

- 10 documents (legal cases)
- Inter-annotator agreement: \( \kappa = 0.58 \)

Version 2 (Mochales-Palau and Moens, 2009)

- 47 documents
- Inter-annotator agreement: \( \kappa = 0.75 \)

Can be used for the following tasks

- Identification of argument components (1,067 argumentative and 1,449 non-argumentative sentences)
- Classification of argument components (304 conclusions; 763 premises)
Annotated argument components

Kwon et al. (2007)

Annotation of different claim types (sentence-level)
- Data: Public comments about Environmental Protection Agency (EPA)
  - 119 documents; 240 claims

Claim Types
- Supporting claim (7%)
- Opposing claim (59%)
- Proposing claim (34%)

Reliability
- Identification of claims: $\kappa = 0.62$
- Claim types: $\kappa = 0.80$
Annotated argument components
Eckle-Kohler et al. 2015

Multi-sentence argument components in news articles
• Data: focused crawl (current topics related to educational system in Germany)
  • 88 documents; 3,863 sentences
  • 74 % of the tokens are argumentative

Argument components
• Claim-premise scheme
• 1,708 argument components
• Supporting and attacking premises
• Target claim indicated by annotation scheme (pre, post)

Reliability
• Claim types: $\alpha_U = .402$
Annotated argument components

User-generated Web discourse

Micro argument components at multi-sentence level (Habernal & Gurevych, 2015)

• Domain: Six controversies in education
  – Redshirting; Single-sex education; Prayer in schools; Homeschooling; Mainstreaming; Public schools vs. private schools

• Four Web registers, user-generated content
  – Comments to newswire articles or blogs, discussion forum posts, blog posts, newswire editorials

• 340 documents, ~ 90,000 tokens, IAA $\alpha_U = 0.48$ (average)

Suitable for the following tasks

• Component identification
  – Free boundaries, mostly aligned to sentences

• Component classification (five types)

Specials

• Modified Toulmin’s scheme, structure is implicitly encoded
• Includes implicit claim annotations
Micro-level argumentation structures

Araucaria

Micro argument structures at the clause level
(Reed et al. 2008)
• Created using a graphical annotation tool
(Reed and Rowe 2004)
• Various document types: e.g. newspaper editorials, parliamentary records, discussions, etc.
• ~700 documents / ~2,000 argument components

Suitable for the following tasks
• Component identification
  – Note that current releases do not include non-argumentative texts
• Component classification
• Argument structures

Specials
• Argumentation schemes (reasoning type)
• Includes implicit argument components (enthymeme)
Micro-level argumentation structures

*Persuasive Essays*

Micro argumentation structures in persuasive essays (Stab and Gurevych, 2014)

- Heterogeneous topics / prompts
- Collected from essayforum.com
- 90 persuasive essay (extended version includes 402 essays)

Usable for the following tasks

- Component identification (e.g. segmentation)
- Component classification (major claim, claim, premise)
- Structure identification (support and attack relations)

Inter-annotator agreement

- Argument components & type: $\alpha_U = .72$
- Argumentative relations $\alpha = .80$

*Extended version available at:* [https://www.ukp.tu-darmstadt.de/data/](https://www.ukp.tu-darmstadt.de/data/)**
Micro-level argumentation structures

*Claims and Evidence from Wikipedia*

Micro argumentation structures over multiple documents (Rinott et al. 2015; Aharoni et al. 2014)

- Argument components at the clause-level
- 58 topics from iDebate.org
- 547 Wikipedia articles
- 2,294 related claims and 4,960 associated evidences from Wikipedia

Different types of evidence

- Study (quantitative analysis),
- Expert (testimony by a person)
- Anecdotal (specific events)

Inter-annotator agreement

- Claims: $\kappa = .39$
- Evidence: $\kappa = .40$
Micro-level argumentation structures

Microtexts

Micro level argument structures (Peldszus and Stede, 2015)
- Controlled complexity (max of 5 argument components, no non-argumentative units)
- 112 documents (one argument per document)
- 576 argument components
- Annotation of opponent or proponent
- Several types of relations: rebuttal, support, undercut

Specials
- Originally created in German
- Professionally translated to English
- First parallel corpus in AM

Inter-annotator agreement
- Three expert annotators
- $\kappa = .83$
Micro-level argumentation structures

Micro argumentation structures at sentence-level (Kirschner et al., 2015)

• 24 German scientific articles
• Educational domain
• Annotations of introduction and discussion sections

Several argumentative relations

• Support
• Attack
• Detail
• Sequence

Inter-Annotator Agreement

• $\kappa = .43$ among four annotators
# Existing Corpora

## Overview

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Document types</th>
<th>Language</th>
<th>Tasks</th>
<th>Argument granularity</th>
<th>Component granularity</th>
<th>Size (#documents/#comps)</th>
<th>Single-/multi-docs</th>
<th>Reliability</th>
<th>Specials</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Cabrio and Villata, 2014)</td>
<td>several</td>
<td>en</td>
<td>SI</td>
<td>macro</td>
<td>-</td>
<td>792 argument pairs</td>
<td>-</td>
<td>κ = .71</td>
<td>multiple sources including wiki revision</td>
</tr>
<tr>
<td>(Eckle-Kohler et al. 2015)</td>
<td>news</td>
<td>de</td>
<td>CI+CC</td>
<td>micro</td>
<td>multi</td>
<td>88 / ~1708</td>
<td>single</td>
<td>α_U = .40</td>
<td>-</td>
</tr>
<tr>
<td>(Habernal and Gurevych, 2015)</td>
<td>user-generated web</td>
<td>en</td>
<td>CI+CC</td>
<td>micro</td>
<td>multi</td>
<td>340 / ~1k</td>
<td>single</td>
<td>α_U = .48</td>
<td>several registers and topics</td>
</tr>
<tr>
<td>(Kirschner et al., 2015)</td>
<td>scientific articles</td>
<td>de</td>
<td>SI</td>
<td>micro</td>
<td>sentence</td>
<td>24 / ~2.7k</td>
<td>single</td>
<td>κ = .43</td>
<td>-</td>
</tr>
<tr>
<td>(Kwon et al., 2007)</td>
<td>online comments</td>
<td>en</td>
<td>CI+CC</td>
<td>micro</td>
<td>sentence</td>
<td>119 / 240</td>
<td>single</td>
<td>κ = .62</td>
<td>several claim types</td>
</tr>
<tr>
<td>(Mochales-Palau and Moens, 2009)</td>
<td>court cases</td>
<td>en</td>
<td>CI+CC</td>
<td>micro</td>
<td>sentence</td>
<td>47 / 1,067</td>
<td>single</td>
<td>κ = .75</td>
<td>-</td>
</tr>
<tr>
<td>(Peldszus and Stede, 2015)</td>
<td>micro texts</td>
<td>de / en</td>
<td>CC+SI</td>
<td>micro</td>
<td>clause</td>
<td>115 / 576</td>
<td>single</td>
<td>κ = .83</td>
<td>parallel corpus</td>
</tr>
<tr>
<td>(Reed et al., 2008)</td>
<td>various</td>
<td>en</td>
<td>CI+CC+SI</td>
<td>micro</td>
<td>clause</td>
<td>700 / 2k</td>
<td>single</td>
<td>-</td>
<td>argumentation schemes &amp; enthymeme</td>
</tr>
<tr>
<td>(Rinott et al., 2015)</td>
<td>wiki articles</td>
<td>en</td>
<td>CI+CC+SI</td>
<td>micro</td>
<td>clause</td>
<td>547 / ~7.5k</td>
<td>multi</td>
<td>κ = .39</td>
<td>different types of evidence</td>
</tr>
<tr>
<td>(Stab et al., 2014)</td>
<td>student essays</td>
<td>en</td>
<td>CI+CC+SI</td>
<td>micro</td>
<td>clause</td>
<td>90 / 1.5k</td>
<td>single</td>
<td>α_U = .72</td>
<td>-</td>
</tr>
<tr>
<td>(Walker et al., 2012)</td>
<td>online discussions</td>
<td>en</td>
<td>AT</td>
<td>macro</td>
<td>-</td>
<td>390k / -</td>
<td>-</td>
<td>.22 &lt; κ &lt; .62</td>
<td>several argument attributions</td>
</tr>
</tbody>
</table>

(CI = Component Identification; CC = Component Classification; SI = Structure Identification; AT = Argument Attribution)
DKPro Argumentation

Unified type system for modeling argumentation
- Based on UIMA
- Easy to expand

Enables cross-resource-experiments
- Seamless integration in DKPro and DKPro-TC
- Several data sets using it

Supports all subtasks of AM
- Component identification on diff. granularities
- Component classification
- Relation identification

Available here: https://github.com/dkpro/dkpro-argumentation
Existing Resources

Summary

Existing resources are heterogeneous
  • No one-fits-it-all-resource
  • Different, tasks, genres, component types, granularities, etc.
  • However, “claim-and-premise-scheme” seems to be common

Our taxonomy attempts to understand the differences w.r.t.:
  • Argument detection tasks
  • Argument granularities (micro- vs. macro-level)
  • Argument component granularities (clause vs. sentence vs. multi-sentence)
    ➔ Facilitates the selection of existing benchmark resources
    ➔ Helps to define the focus of future annotation studies
References (1/2)


References (2/2)


