Detecting Argument Components and Structures

Christian Stab and Ivan Habernal

TECHNISCHE UNIVERSITÄT DARMSTADT

Introduction

TECHNISCHE UNIVERSITÄT DARMSTADT

Overview of the talk(s)

- Common Tasks in Argument Detection
- Motivating example and challenges
- Part 1: Argumentation in user-generated web discourse (Ivan Habernal)
- Part 2: Parsing Argumentation Structures in Persuasive Essays (Christian Stab)

Detection of argument structures at the micro-level

- Fine grained analysis of arguments
- Identification of argument components (claims / premises)
- Recognition of argumentation structures (support / attack relations)

UKP

Argument Detection

Common Tasks for Argument Detection

TECHNISCHE UNIVERSITÄT DARMSTADT

Raw input text

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum.

Argument components

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum.

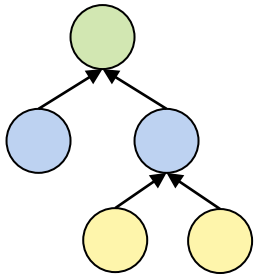
- Separate argumentative from non-argumentative text units
- Identification of argument component boundaries

Component types

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum.

- Argumentative role of argument components
- e.g. conclusions, claims, different types of evidence, etc.

Argument structure



- Identification of relations between argument components
- e.g. support / attack relations

UKP

Argument Detection

Example

TECHNISCHE UNIVERSITÄT DARMSTADT

"*Since researchers at the Roslin Institute in Edinburgh cloned an adult sheep, there is an ongoing debate if cloning technology is morally and ethically right or not. Cloning will be beneficial for many people who are in need of organ transplants. Cloned organs will match perfectly to the blood group and tissue of patients since they can be raised from cloned stem cells of the patient. In addition, it shortens the healing process. Usually, finding an appropriate organ donor takes a long time and by using cloning in order to raise required organs the waiting time can be shortened tremendously*."

UKP

Argument Detection

Example: Identifying Argument Components

TECHNISCHE UNIVERSITÄT DARMSTADT

"*Since researchers at the Roslin Institute in Edinburgh cloned an adult sheep, there is an ongoing debate if cloning technology is morally and ethically right or not. Cloning will be beneficial for many people who are in need of organ transplants. Cloned organs will match perfectly to the blood group and tissue of patients since they can be raised from cloned stem cells of the patient. In addition, it shortens the healing process. Usually, finding an appropriate organ donor takes a long time and by using cloning in order to raise required organs the waiting time can be shortened tremendously*."

UKP

Argument Detection

Example: Identifying Argument Component Types

TECHNISCHE UNIVERSITÄT DARMSTADT

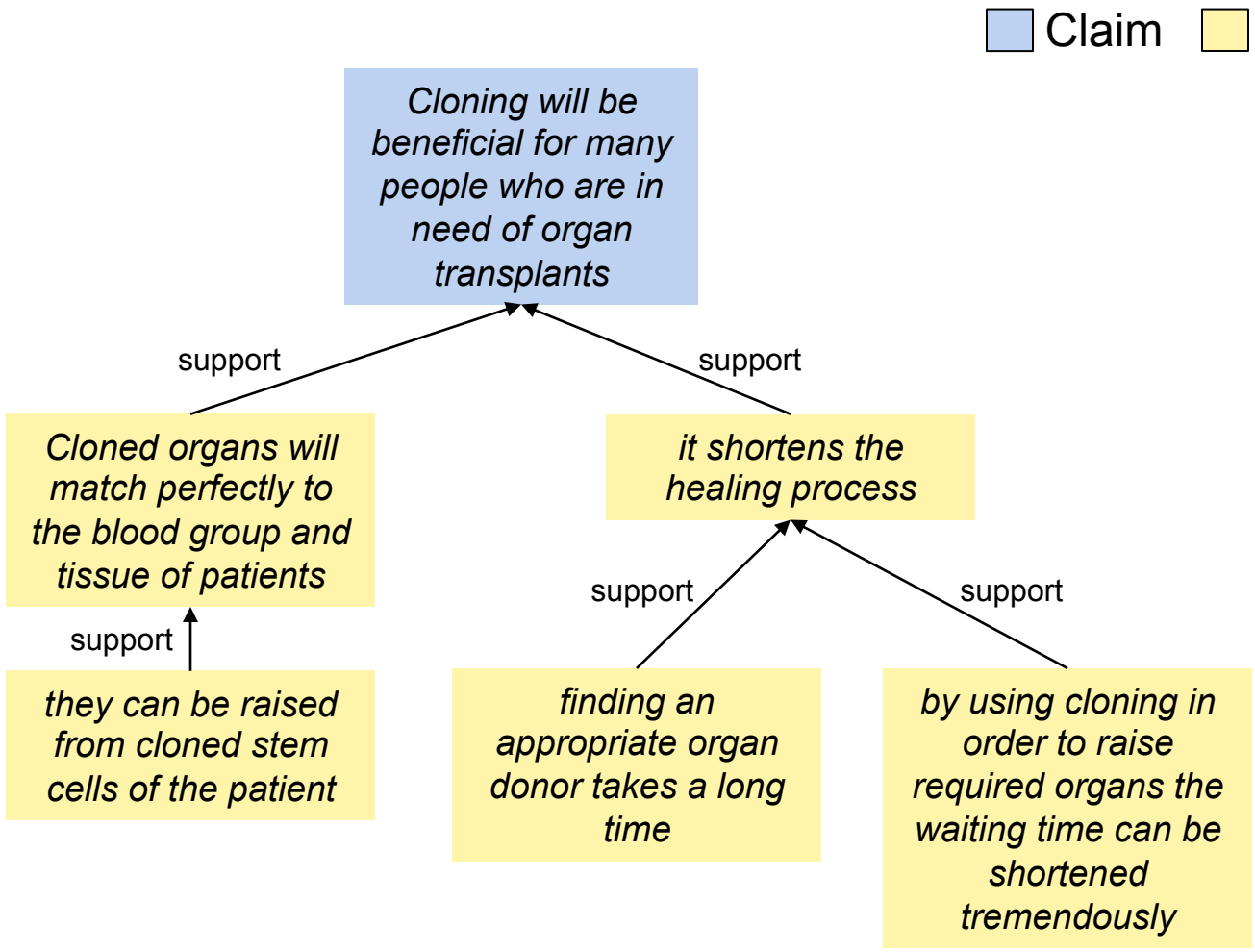
Claim Premise

"*Since researchers at the Roslin Institute in Edinburgh cloned an adult sheep, there is an ongoing debate if cloning technology is morally and ethically right or not. Cloning will be beneficial for many people who are in need of organ transplants. Cloned organs will match perfectly to the blood group and tissue of patients since they can be raised from cloned stem cells of the patient. In addition, it shortens the healing process. Usually, finding an appropriate organ donor takes a long time and by using cloning in order to raise required organs the waiting time can be shortened tremendously*."

UKP

Argument Detection

Example: Identifying of Argumentation Structures



Argument Detection

Challenges and Tasks

TECHNISCHE UNIVERSITÄT DARMSTADT

Non-argumentative information

➔ Separation of argumentative and non-argumentative text units

Several argument components per sentence

➔ Segmentation to identify argument components

Non-adjacent argumentative relations

➔ Consider each argument component pair; bad class distribution

Several arguments in a paragraph

➔ Separation of arguments

UKP

TECHNISCHE UNIVERSITÄT DARMSTADT

Part 1

Argumentation in user-generated web discourse

(Ivan Habernal)

UKP

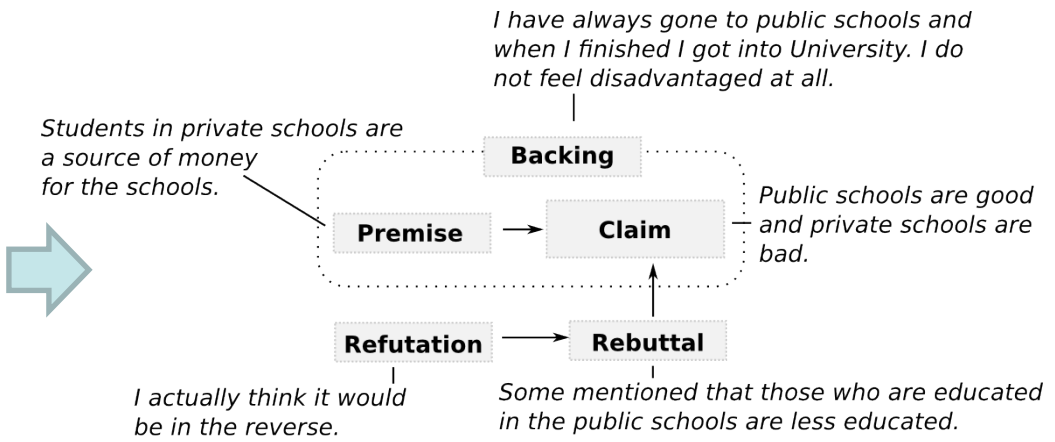
Argumentation in web discourse

Motivation

TECHNISCHE UNIVERSITÄT DARMSTADT

- Example use case: Extracting argument gist

The public schooling system is not as bad as some may think. Some mentioned that those who are educated in the public schools are less educated, well I actually think it would be in the reverse. Student who study in the private sector actually pay a fair amount of fees to do so and I believe that the students actually get let off for a lot more than anyone would in a public school. And its all because of the money. In a private school, a student being expelled or suspended is not just one student out the door, its the rest of that students schooling life fees gone. Whereas in a public school, its just the student gone. I have always gone to public schools and when I finished I got into University. I do not feel disadvantaged at all.



UKP

Argumentation in web discourse

Data

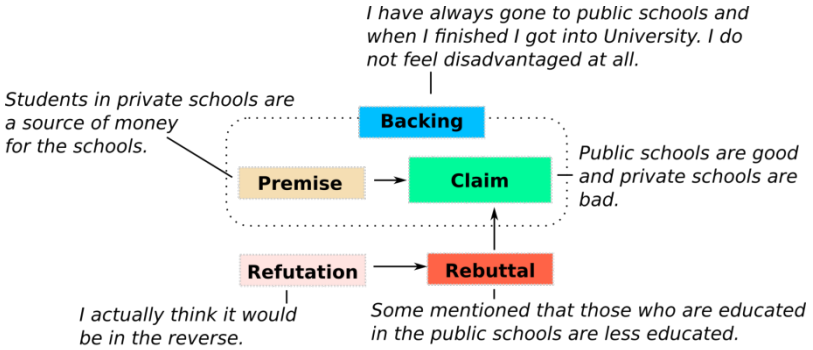
TECHNISCHE UNIVERSITÄT DARMSTADT

Source

- **340** English **documents** (90k tokens)
- **6 topics**: homeschooling, mainstreaming, prayer in schools, public vs. private school, redshirting, single-sex schools
- **4 registers**: comments to articles, forum posts, blogs, newswire opinionated articles

Annotation scheme

- Extended Toulmin's model of argument
- **5 argument component** types
- Implicit relations



Agreement

- 3 annotators, Krippendorff's α_U 0.40-0.65
  - Varies with register, length, and topic

UKP

Argumentation in web discourse

Results

TECHNISCHE UNIVERSITÄT DARMSTADT

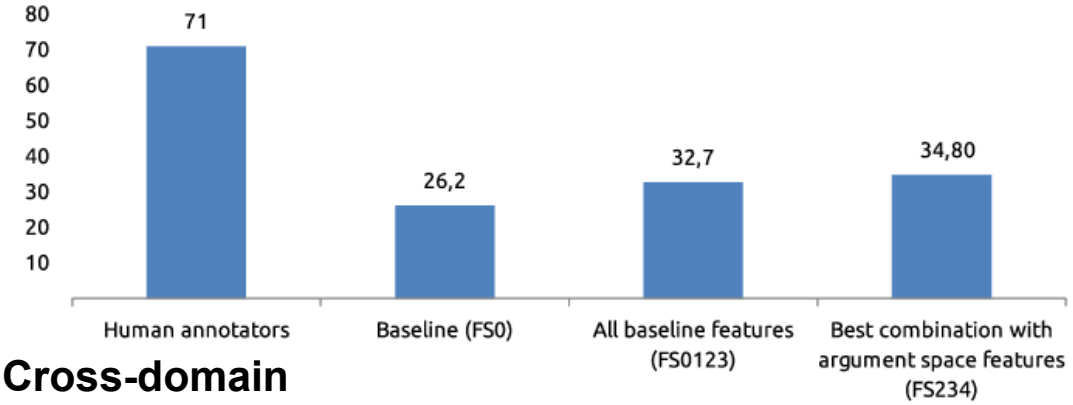
Sequence labeling

- SVM HMM
- Features
  - N-grams baseline (*FS0*)
  - Structure and syntax (*FS1*)
  - Sentiment and topic (*FS2*)
  - Semantic and discourse (*FS3*)
  - Semi-supervised extension – argument space features (FS4)
- BIO tagging, 11 classes

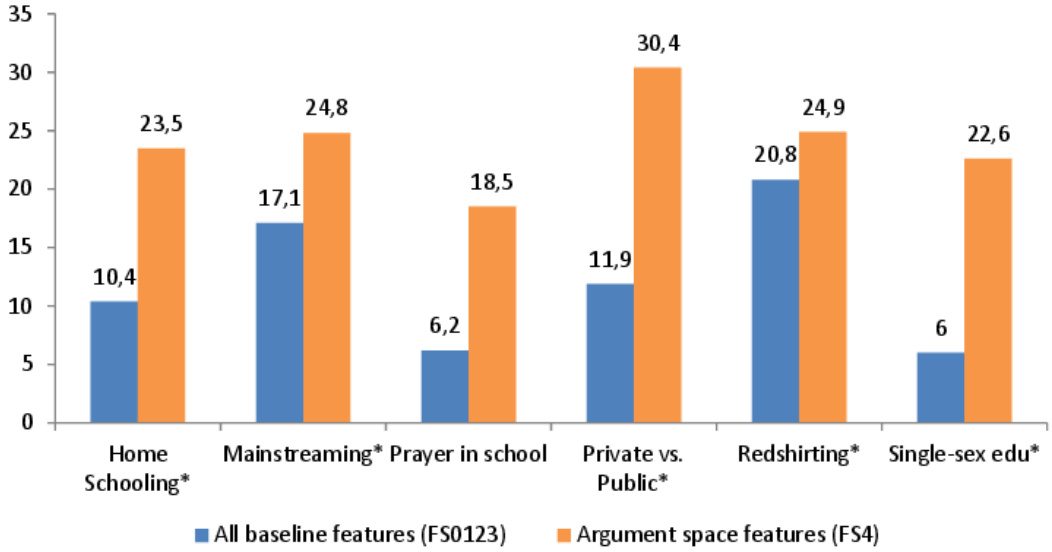
Doc #2224 (forumpost, single-sex-education)

Backing-B *I*	Claim-I *it*
Backing-I *went*	Claim-I *,*
Backing-I *to*	Claim-I *I*
Backing-I *an*	Claim-I *would*
Backing-I *all*	Claim-I *of*
Backing-I *boys*	Claim-I *much*
Backing-I *school*	Claim-I *preferred*
O -	Claim-I *gone*
Claim-B *Can't*	Claim-I *to*
Claim-I *say*	Claim-I *a*
Claim-I *I*	Claim-I *co-ed*
Claim-I *particularly*	Claim-I *.*
Claim-I *liked*	Premise-B *It*
Claim-I *it*	Premise-I *is*
Claim-I *,*	...

Cross-validation



Cross-domain



UKP

TECHNISCHE UNIVERSITÄT DARMSTADT

Part 2

Parsing Argumentation Structures in Persuasive Essays

(Christian Stab)

UKP

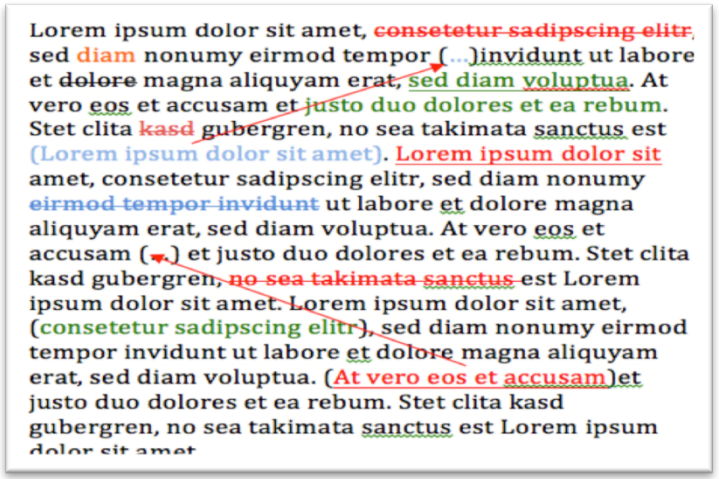
Parsing Argumentation Structures

Motivation

TECHNISCHE UNIVERSITÄT DARMSTADT

Computer-Assisted Writing (CAW)

- Feedback about written text
- Support of different writing tasks
- ***Examples***: Spell checking and grammar checking



Argumentative Writing Support (AWS)

- Feedback about written arguments
- Identify flaws in argumentation
- ***Goal***: Improve argumentation skills and text quality

Argumentation Mining in writing environments

- Enables students to inspect their texts for plausibility
- Improves argumentation quality
- Tailored feedback about argumentation

UKP

Parsing Argumentation Structures

Data

TECHNISCHE UNIVERSITÄT DARMSTADT

Persuasive essays

- Written for e.g. IELTS, TOEFL, etc.
- Collected from www.essayforum.com
- 402 essays; 7,116 sentences; 147,271 tokens

Annotation scheme

- Argumentative structure as tree structure
- Argument components: Major Claim (751), Claim (1,506) and Premise (3,832)
- Argumentative relations: Support (3,613) and Attack (219)

*Inter-Annotator Agreement**

- Argument components: α_U =.767
- Argumentative relations: α = .723 (avg. of support & attack)

*determined among three annotators on a subset of 80 essays

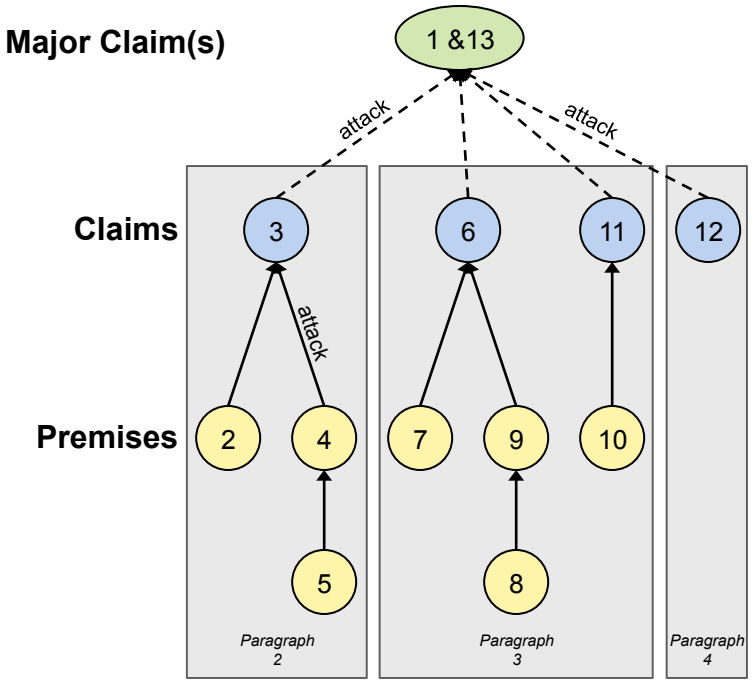
UKP

Parsing Argumentation Structures

Example argumentation structure

TECHNISCHE UNIVERSITÄT DARMSTADT

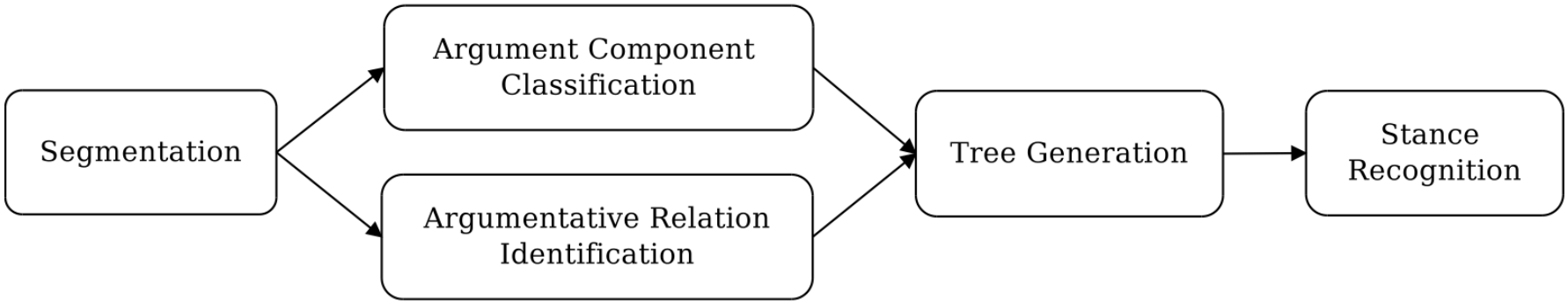
Argumentation Structure:



UKP

Argumentation Structure Parsing

Pipeline



Argument component classification

- Classify each argument component as major claim, claim or premise

Argumentative Relation Identification

- Classify argument component pairs as argumentatively related or not

Problem: Result is an arbitrary graph NOT a tree

Solution: Joint Modeling (Tree generation)

Approach
Joint Modeling

TECHNISCHE UNIVERSITÄT DARMSTADT

Component types and argumentative relations share mutual information

Component Type	*Argumentative Relation*
Claim	No outgoing relations (root node)
Premise	Exhibits outgoing relations
Claim	More incoming relations
Premise	Less incoming relations

Idea:

Jointly model **argument component types** and **argumentative relations** to find an optimal tree

ILP-based joint model

- Finds the tree structure which optimizes previous analysis results
- Allows to find several trees (arguments) in a paragraph

Argumentation Structure Parsing

Joint Modeling

TECHNISCHE UNIVERSITÄT DARMSTADT

		component classification					*relation identification*			***statistics***		
		F1	MC	Cl	Pr		F1	NoLi	Link	Cl → Pr	Pr → Cl	Trees
Baseline heuristics		.724	.740	.560	.870		.660	.885	.436	-	-	100%
Base classifier	†	.773	**.865**	.592	.861	†	.736	**.917**	.547	-	-	20.9%
IncBaseline	†	.776	**.865**	.601	.861	†	.739	**.917**	.555	206	1,144	24.2%
ILP Joint Model	†‡	**.823**	**.865**	**.701**	**.904**	†	**.752**	.913	**.591**	297	283	100%

Baseline Heuristic

- Last component in introduction and first component in conclusion as major claim
- First component in paragraph as claim, remaining as premise
- Link all premises to first component in paragraph

IncBasline

- Incorporates baseline in base classifiers if both base classifiers fail to predict claims or relations in a paragraph

† significant improvement over baseline heuristic; ‡ significant improvement over base classifier

UKP

Conclusion

TECHNISCHE UNIVERSITÄT DARMSTADT

Introduction of common tasks of argument detection

- Argument component identification
- Classifying argument component types
- Identifying argumentation structures

Argumentation mining in Web discourse

- Our modified Toulmin's model offers a trade-off between its expressiveness and annotation reliability
  - But there are aspects of online argumentation that lack their established theoretical counterparts, such as rhetorical questions, figurative language, or narratives
- Semi-supervised approach using „argument space“ features improves perfomance up to 90% in cross-domain and cross-register evaluation

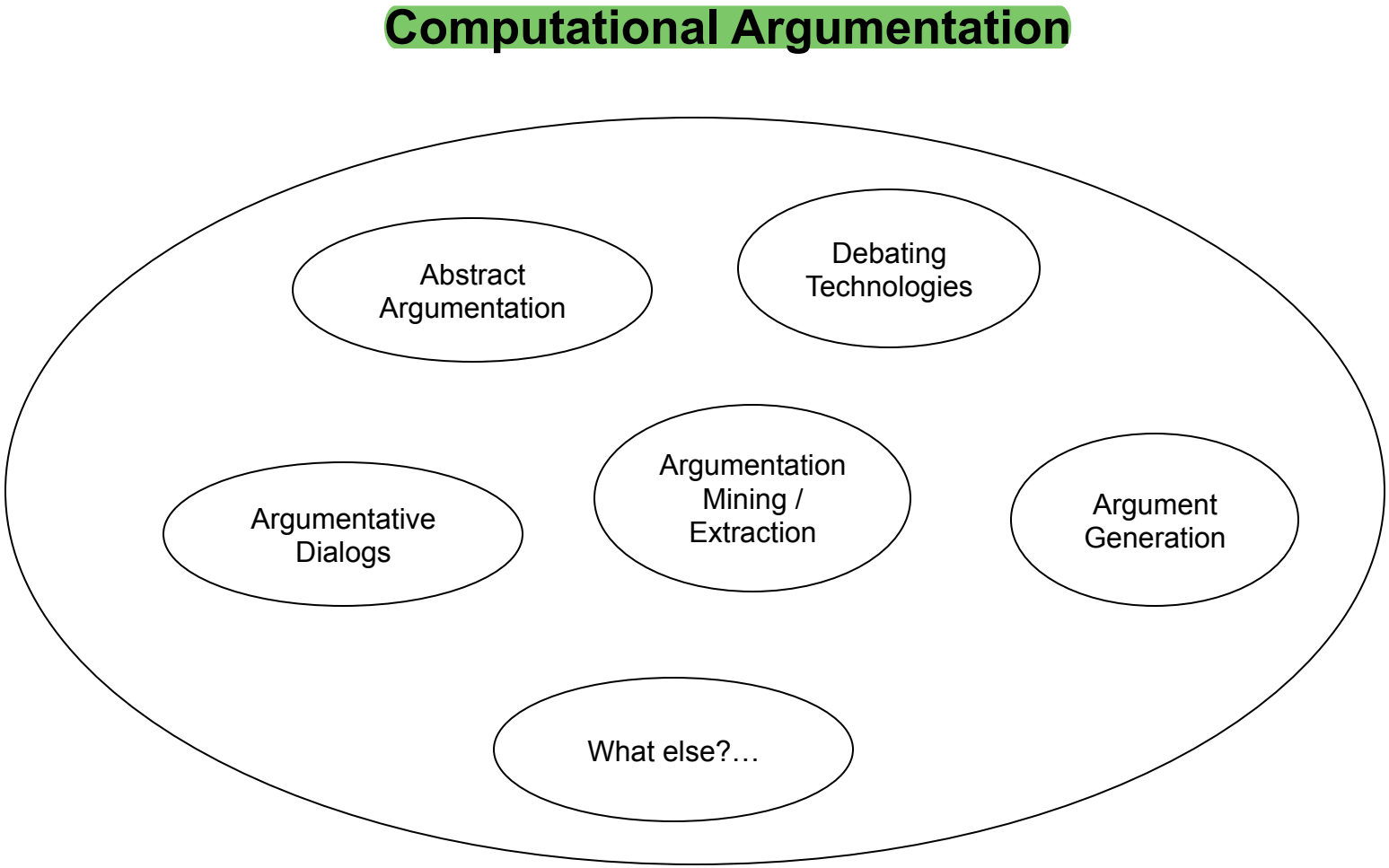
Jointly modeling argument component structures

- Works well for finding argumentation structures (see also Peldszus2015)
- Simultaneously improves results of upstream tasks

UKP

Thank you for your attention!
Questions?

TECHNISCHE UNIVERSITÄT DARMSTADT



UKP

TECHNISCHE UNIVERSITÄT DARMSTADT

Contact

Christian Stab
Technische Universität Darmstadt
Ubiquitous Knowledge Processing Lab

Hochschulstr. 10, 64289 Darmstadt, Germany
+49 (0)6151 16 - 4197
+49 (0)6151 16 - 5455
stab (at) ukp.informatik.tu-darmstadt.de

Dr. Ivan Habernal
Technische Universität Darmstadt
Ubiquitous Knowledge Processing Lab

Hochschulstr. 10, 64289 Darmstadt, Germany
+49 (0)6151 16 - 5420
+49 (0)6151 16 - 5455
habernal (at) ukp.informatik.tu-darmstadt.de

UKP

References

Habernal, I., & Gurevych, I. (2015). Exploiting Debate Portals for Semi-Supervised Argumentation Mining in User-Generated Web Discourse. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (pp. 2127–2137). Lisbon, Portugal: Association for Computational Linguistics.

Habernal, I., & Gurevych, I. 2016. Argumentation Mining in User-Generated Web Discourse. Under review in *Computational* Linguistics

Peldszus, Andreas and Manfred Stede. 2015. Joint prediction in mst-style discourse parsing for argumentation mining. In Conference on Empirical Methods in Natural Language Processing (EMNLP 2015), page 938-948 , Lisbon, Portugal.

Stab, Christian and Iryna Gurevych. 2014a. Annotating argument components and relations in persuasive essays. In Proceedings of the 25th International Conference on Computational Linguistics (COLING 2014), pages 1501–1510, Dublin, Ireland, August.

Stab, Christian and Iryna Gurevych. 2014b. Identifying argumentative discourse structures in persuasive essays. In Conference on Empirical Methods in Natural Language Processing (EMNLP 2014), pages 46–56, Doha, Qatar.

Stab, Christian and Iryna Gurevych. 2016. Parsing Argumentation Structures in Persuasive Essays. Under review in Computational Linguistics