

An Annotation Protocol for Collecting User-Generated Counter-Arguments using Crowdsourcing

Paul Reisert^{1,3}, Gisela Vallejo², Naoya Inoue^{3,1}, Iryna Gurevych², and Kentaro Inui^{3,1}

¹RIKEN Advanced Center of Intelligence Project (AIP), Tokyo, Japan

²Ubiquitous Knowledge Processing (UKP), Department of Computer Science, Technische Universität Darmstadt, Darmstadt, Germany

³Graduate School of Information Sciences, Tohoku University, Sendai, Japan



Constructive Feedback

Overview

- Constructive feedback (CF) improves critical thinking skills
- Useful as CF: **counter-arguments (CAs)** and **fallacies**
- Little work on CAs for micro-level (one claim/premise) arguments

Challenge

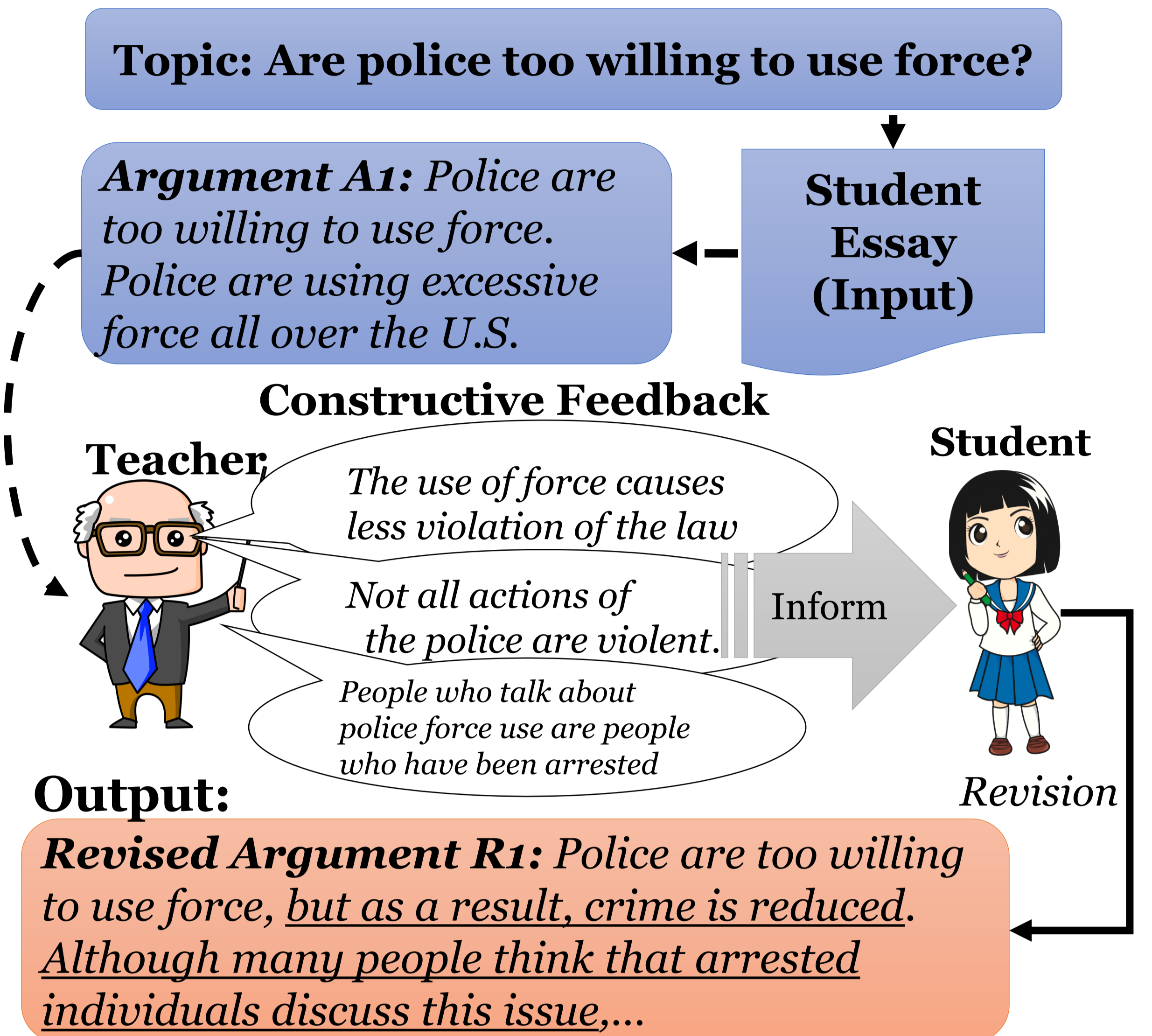
- Collecting large-scale CAs useful for a model is challenging

Solution

- Create an annotation protocol for collecting large-scale CAs

Contribution

- Created protocol for collecting CAs via crowdsourcing
- Conducted qualitative analysis on CA quality



Crowdsourcing

- Useful for collecting large-scale data
- Instructed workers to i) create CA (**CAG**) and ii) create CA after identifying fallacy (**CAG-F**)
- Conducted trials using the Argument Reasoning Comprehension dataset (Habernal+, 2018)
- 5 random fallacy types

Topic: There She Is, Miss America
Argument:
Miss America is good for women because Miss America gives honors and education scholarships.

Does a Hasty Generalization fallacy exist in the argument "Miss America is good for women because Miss America gives honors and education scholarships."?

Yes
 No
 Unsure

Please write a counter-argument using the fallacy pattern:

- Avoid copy and paste as much as possible.
- Keep the answer as brief as you can
- Avoid writing your own knowledge as much as possible.

It's too hasty to assume that...

Enter your text here.

CAG-F

Annotation Study

- Annotate quality of collected CAs
- Observe **inter-annotator agreement (IAA)**
- Annotators: two fluent-English speakers
- 100 instances for each task (CAG & CAG-F)

Stage	IAA	#Agreed
CAG	0.29	54/100
CAG-F	0.37	74/100

Corpus of large-scale CAs

Large-scale CA corpus

- Created large-scale corpus after submission
- Includes 4 fallacy types

Fallacy Type	Definition
Hasty Generalization (HG)	Someone generally always assumes something based on a few instances.
Begging the Question (BQ)	The truth of the premise is already assumed by the claim.
Questionable Cause (QC)	The cause of an effect is questionable.
Red Herring (RH)	Someone reverts attention away from the original claim by changing the topic.

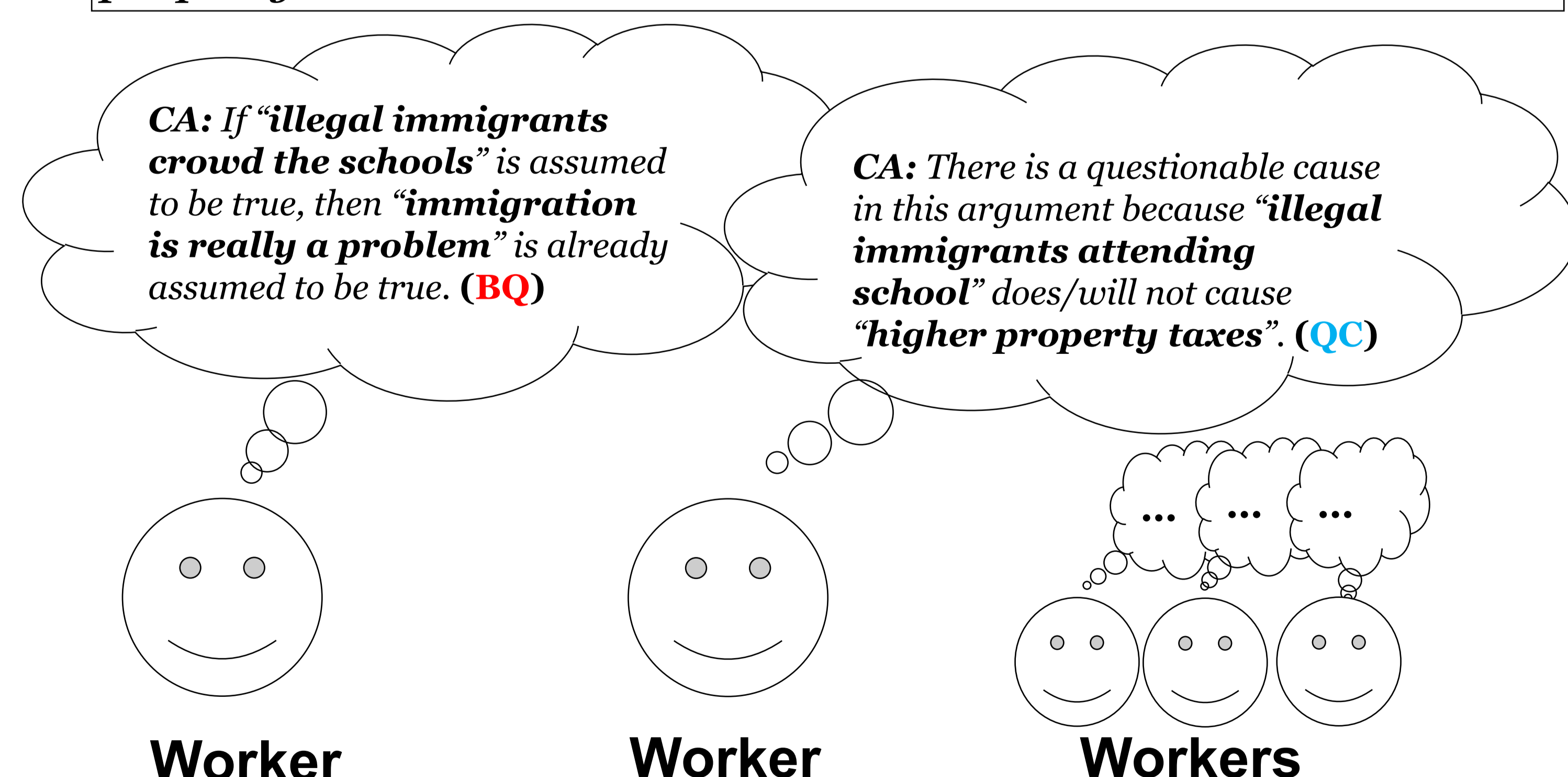
Statistics

Criteria	HG	BQ	QC	RH	Total
Unsure	2,043	315	2,136	1,879	6,373
CAs (FS)	3,365	3,818	2,121	1,772	11,076
CAs (O)	907	2,182	2,058	2,664	7,811
CAs (total)	4,272	6,000	4,179	4,436	18,887

CA Examples

- Collected around 18K CAs from workers

Topic: Is immigration really a problem?
Argument: Immigration is really a problem because illegal immigrants crowd the schools which in turn causes higher property taxes.



Contact

- Contact author: Paul Reisert, paul.reisert@riken.jp

Future Work

- Create a model for automatically identifying fallacies and generating informative CAs