Ranking Arguments by Combining Claim Similarity and Argument Quality Dimensions

Lorik Dumani, Ralf Schenkel

Trier University

September 23, 2020

Trier University

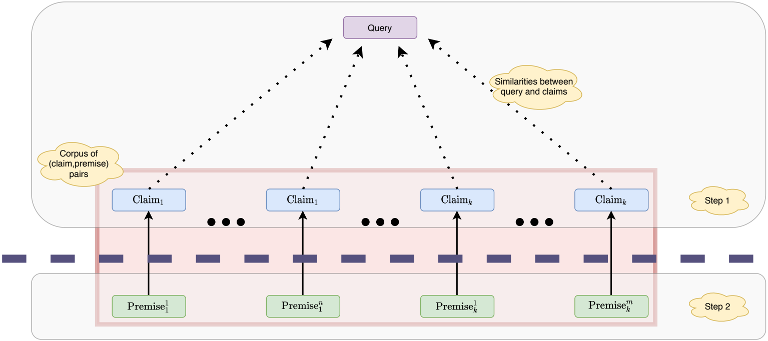
DFG Deutsche Forschungsgemeinschaft

RATIO

CLEF 2020 Thessaloniki

Contents

Draft of the Framework



- Variables: query q, claim c, premise p, various quality aspects Δ.
- $P(p|q,\Delta) = P(c|q) \cdot P(p|c,\Delta)$
- $P(\pi_j|q,\Delta) = \sum_{p \in \pi_j} P(p|q,\Delta)$
  - where π_j is a cluster of premises with the same meaning.
- Now we have to find suitable estimators for $\underline{P(c|q)}$ and $\underline{P(p|c,\Delta)}$.

Step 1: Estimator for Claim Retrieval

Probability	**Description**
$P(c\|q)$	Claim c is relevant to query q.

- Can be estimated with standard text retrieval methods.
- In our implementation we use Divergence from Randomness (DFR) as it yielded promising results in a pre-study.
- In our experiments, DFR was not significantly better than BM25.

Step 2: Estimator for Premise Retrieval

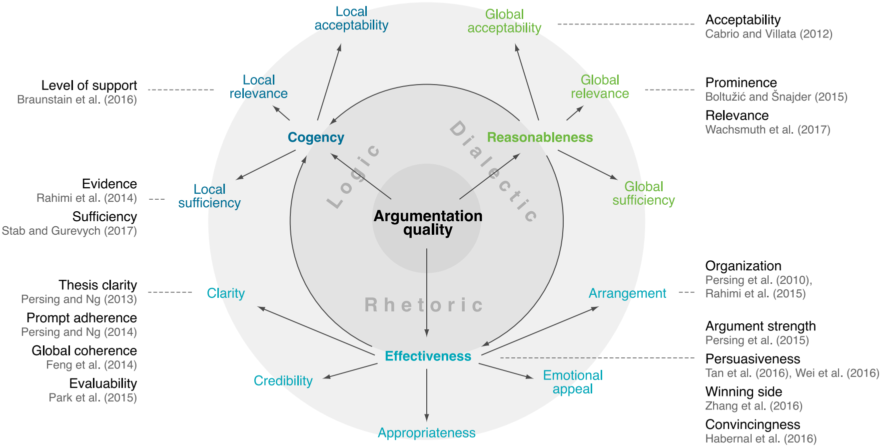
Probability	**Description**
$P(p \mid c, \Delta)$	The user picks a premise p from a claim c, preferring those of high quality in all argument quality dimensions.

- Use and aggregate estimators for various argument quality dimensions.
- Calculate for each premise we the dimension convincing frequency $dcf(p, c, d)$ for a single argument quality dimension d.
  - Count how often a premise p was estimated to be more convincing than all other premises with the same claim c with regard to a dimension d.

$\Rightarrow$ Expressed as probability: $P_{dcf}(p \mid c, d)$

- Multiple dimensions: $P_{dcf}(p \mid c, \Delta) = \prod_{d \in \Delta} P_{dcf}(p \mid c, d)$

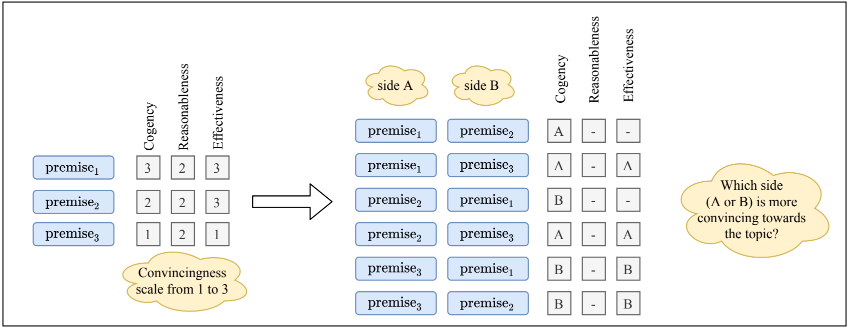
Argument Quality Dimensions



Wachsmuth et al., EACL 2017

Preprocessing of the Data

- Train classifiers for predicting the argument quality with the dataset DAGSTUHL-15512 ARGQUALITY CORPUS.
- It consists of 32 (issue,stance) pairs with 10 premises each (320 arguments) with labels between 1 (low) and 3 (high).
- Transform the dataset to ($premise_1$, $premise_2$) pairs with labels A, B.
- Learn which argument (1 or 2) is better with regard to dimension d.



Example

Issue: *is the school uniform a good or bad idea*
Stance: *bad*

Premise$_1$: *i thik thier bad because i think ushould be free with out nobody telling u wat to do*
⇒ Cogency: 1
⇒ Reasonableness: 1
⇒ Effectiveness: 1

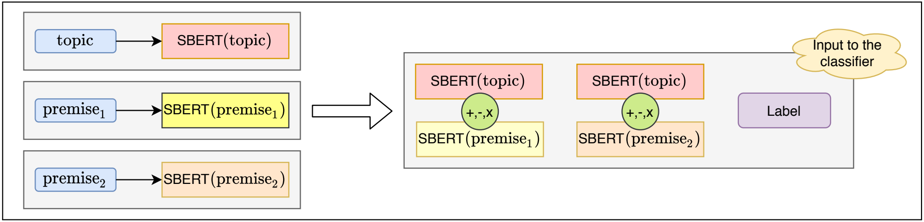
Premise$_2$: *The school my mother works at, plus the school district my cousin's 3 children are in, are utilizing school uniforms. One reason is to "reduce bullying", which in reality, doesn't even address the problem concerning bullying. The only good it does is that it gets rid of or reduces students being bullied because they aren't wearing a specific clothing label that they dictate is the IN thing to wear. While it's a problem, all it does is sweep the one basic type of bullying under the rug. Kids will find other reasons to bully others . It also infringes upon their basic rights to be individuals and to express their individuality .*

⇒ Cogency : 2.667
⇒ Reasonableness: 2.667
⇒ Effectiveness : 2.667

Classifiers for Predicting Argument Quality

- Calculated embeddings by applying Sentence-BERT (SBERT).
- Calculate (1) the sum, (2) the difference, and (3) the product of each dimension of the two premises to the topic pointwise.
- Concatenate the two premise vectors and add a label.

  ⇒ Input to the classifier.

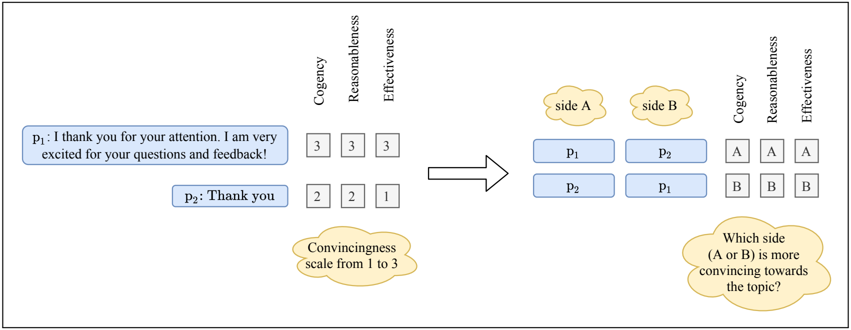


Evaluation of the Classifiers

- Evaluation of seven standard classifiers with leave-one-out cross-validation (32 folds).
- Logistic Regression and Random Forest are significantly better (tested with Tukey's HSD test) than the other classifiers (except Stochastic Gradient Descent) for the three dimensions.

	Accuracy		
Classifier	**Cogency**	**Reasonableness**	**Effectiveness**
Random Forest	**.971**	.972	**.977**
Logistic Regression	.958	**.976**	.97
Stochastic Gradient Descent	.951	.964	.965
Gradient Boosting	.932	.942	.952
Support Vector Machine	.918	.917	.922
K Nearest Neighbours	.887	.89	.902
Naive Bayes	.792	.784	.778

Thank you for your kind attention!



contact: dumani@uni-trier.de