Structure or Content? Towards assessing Argument Relevance

Marc Feger, Jan Steimann and Christian Meter

Heinrich-Heine-University Düsseldorf
Institute of Computer Science
Department for Computer-Networks and Communication-Systems

September 9, 2020

Motivation

What would you like to search for?
What would you like to know?

Motivation

Question: ***Why should peanuts be banned on board aircraft?***

(a_1): *Peanut reactions can be **life threatening**. An individual **doesn't have to consume** the product to have a life threatening reaction. They can have **contact or inhalation reactions.***

(a_2): *Providing **buffer zones** to avoid contact with peanuts is a **thoughtful gesture**. But from a practical point of view, it does not work.*

(a_3): *With so **many food choices** available, **why are peanuts a necessary choice?***

(a_4): ***Restricting** the ban of peanut products to certain flights is **not enough**.*

Possible ranking

$a_1 > a_2 \geqslant a_3 > a_4$?

About this work

This paper covers the following topics:

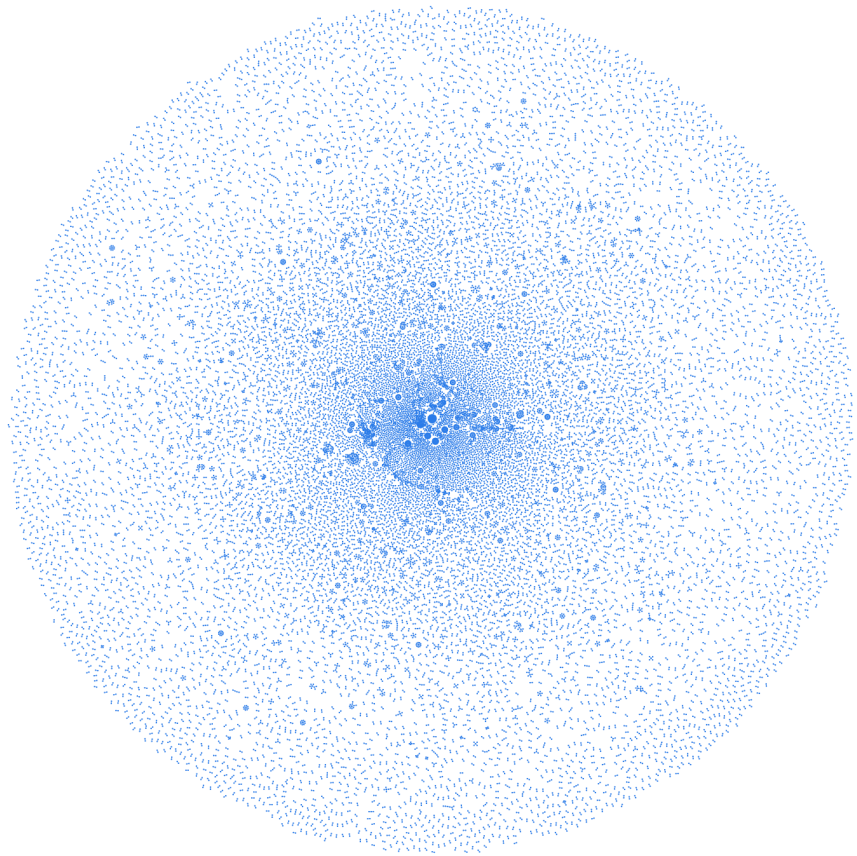
- Follow-up of Wachsmuth et al. [2017]
- Evaluation of methods for determining relevant arguments
  - Analysis of PageRank
  - Intuitive content- and knowledge-based methods

Results

- PageRank is not entirely sufficient
- Better results with content- and knowledge-based methods

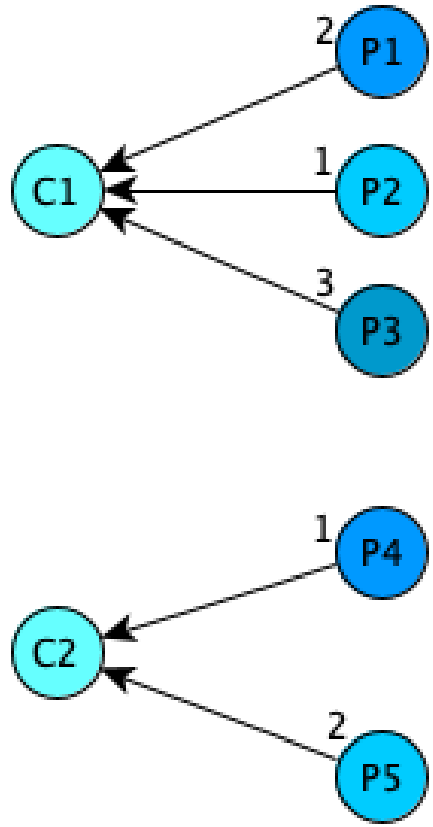
Corpus

- Webis-ArgRank2017 dataset by Wachsmuth et al. [2017].
- Ground-truth argument graph
- Benchmark for argument ranking



Benchmark Argument Ranking

- 110 arguments
- 32 conclusions
- 7 annotators
- Agreement of Kendall's $\tau \approx 0.36$
- $\tau = -1$ (no agreement)
- $\tau = 1$ (absolute agreement)

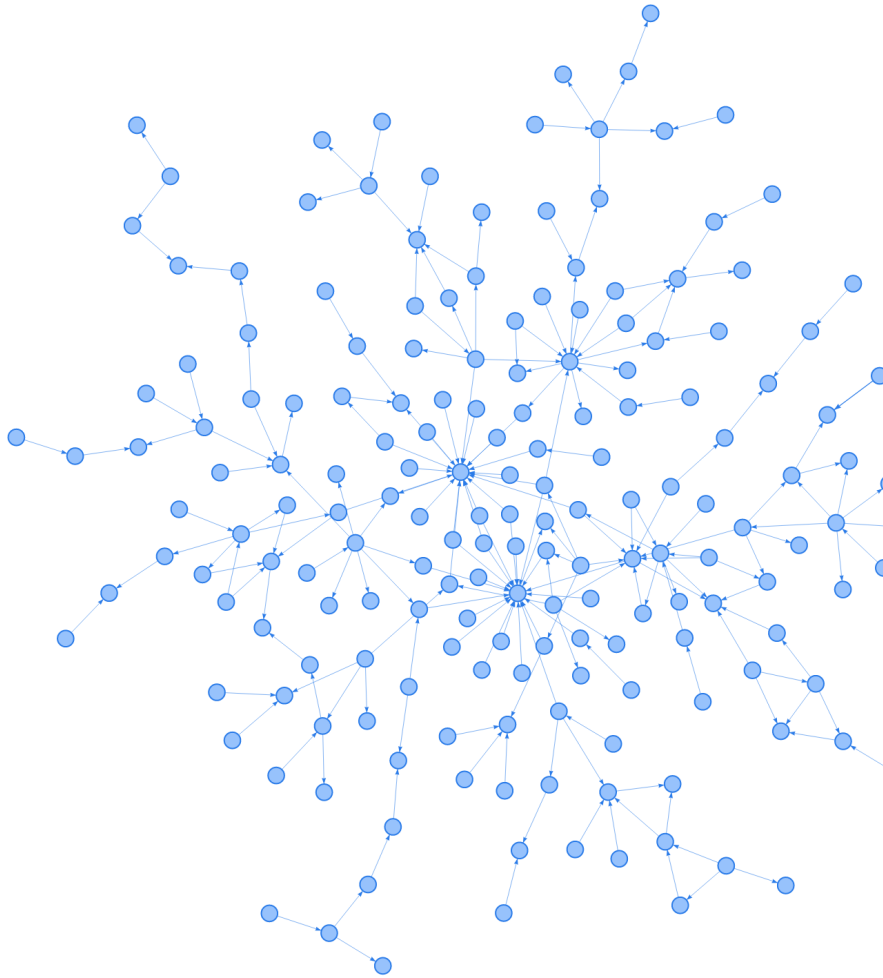


PageRank

- PageRank originally used for websites
- Websites replaced by arguments

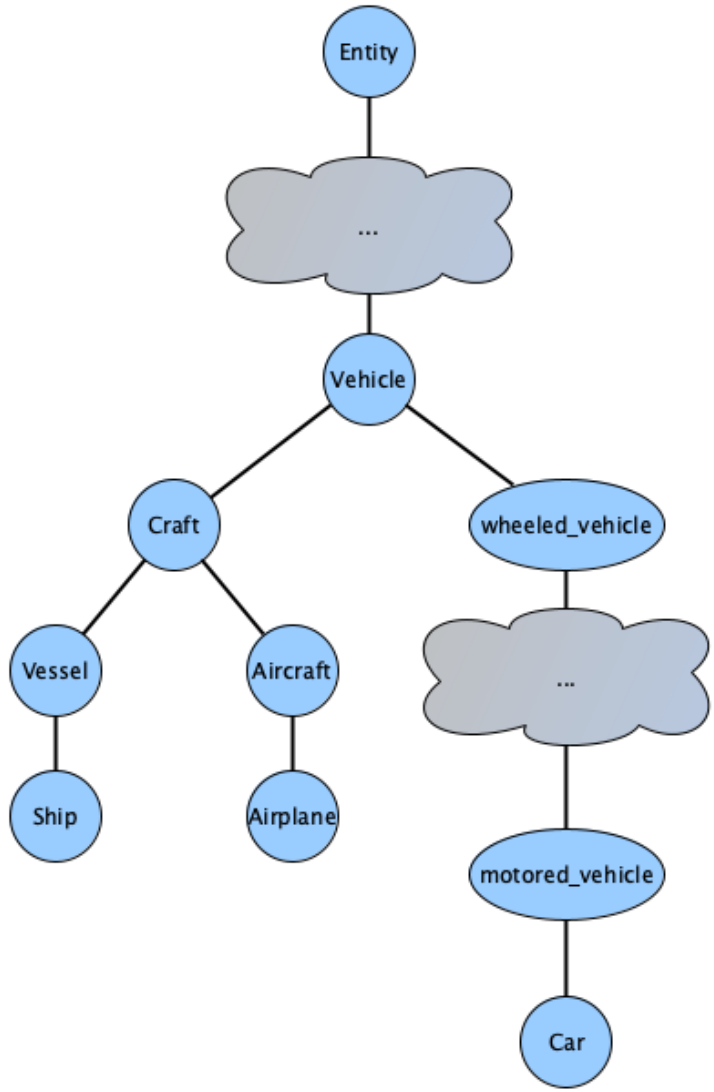
Custom-made PageRank

$$p_t(c_i) = \begin{cases} (1-\alpha)G_{rel} + \alpha L_{rel} & : t > 0 \\ G_{rel} & : t = 0 \end{cases}$$



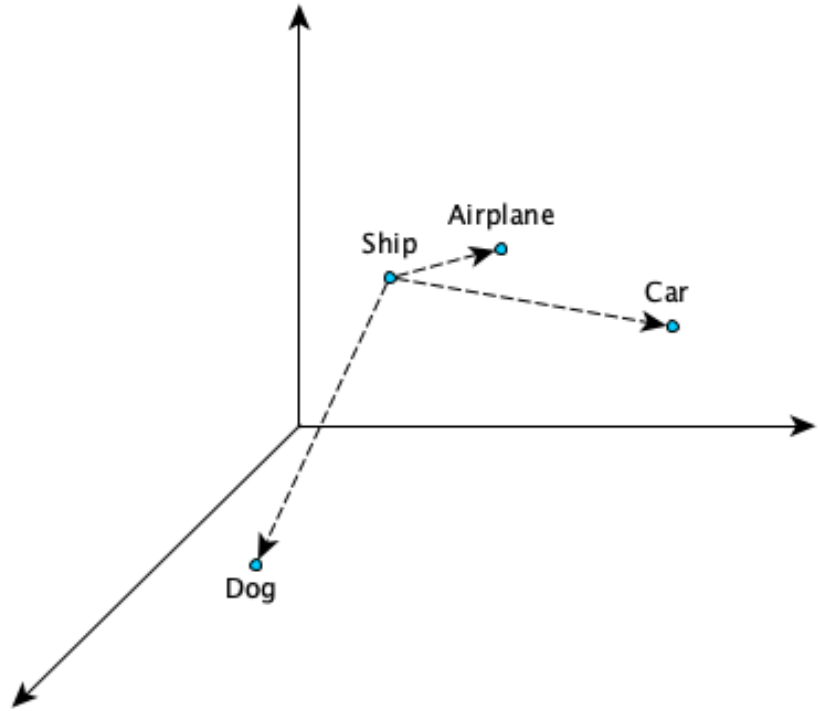
WordNet

- Knowledge-based method
- Conceptual similarity between conclusion and premise



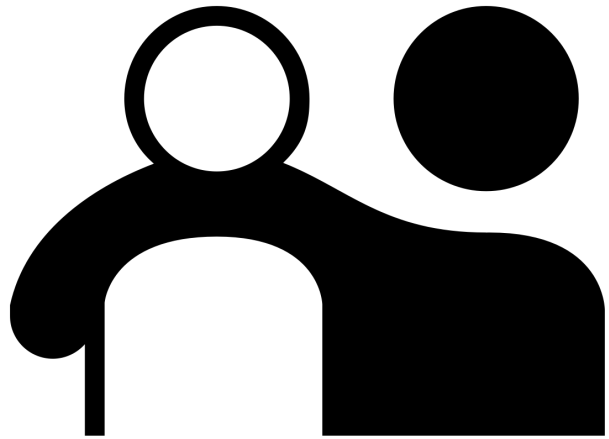
Similarity

- Semantic similarity
- Different embeddings
- BERT, ELMo and GloVe
- Similarity over $Cos(C, P)$



Sentiment

- Positivity of premise
- Captures the constructivity
- Neuronal network based on FastText



Results

PageRank performs best for $\alpha \approx 0$ indicating nearly no interconnection

First Result

- PageRank is not satisfying yet
- Reader not able to judge about interconnection

Results

Simple content- and knowledge-based methods better then PageRank

Second Result

- Relevance captured at word-level
- Reader judges on context at word-level
- Interconnection of secondary importance

Thank you!