Addressing Controversial Topics in Search Engines

The Oral Exam of
Yamen Ajjour

To Obtain the Academic Degree of
Dr. rer. nat.

Intelligent Information Systems
Bauhaus-University Weimar
Addressing Controversial Topics in Search Engines

- Motivation
- Identifying argumentative questions
- Topic bias in argument corpora
- Identifying argument frames
- Conclusion
Motivation

Use cases where good arguments are needed.

Student

Lawyer

Politician

Marketing Company
Motivation

Search engines are good at answering factual questions.

California

California was the first state to legalize medical marijuana in 1996. Since then, the medical use of cannabis has been legalized in 40 states and the District of Columbia. The recreational or adult-use of cannabis has been approved in DC and 23 states.
Motivation

Search engines struggle at delivering all perspectives on controversial topics.

PRO

Why Recreational Marijuana Should Stay Illegal - Drug Rehab
Many people think marijuana isn't harmful enough to be illegal, but this is dangerous - See the reasons why recreational marijuana should stay illegal now!

More Reasons States Should Not Legalize Marijuana
von ST Wilkinson · 2013 · Zitiert von: 16 — Medical marijuana should be subject to the same rigorous approval process as other medications prescribed by physicians. Legalizing recreatio... Recreational Marijuana · Myth: Marijuana is Not Addictive · Effects on Cognition

Cons of Legalizing Recreational Marijuana Use
07.03.2022 — Despite the growing acceptance, decriminalizing marijuana should be a federal issue, and it should not be legal to use recreationally.

Petition · Marijuana should remain illegal!
Marijuana should remain illegal. The federal government is planning to legalize marijuana, however, it's policy on marijuana is like an octopus.
Motivation

Argument retrieval systems retrieve pro and con arguments for a query.

Argument retrieval systems promote:

- Transparency
- Explainability
Contributions
Contributions

1. Identifying argumentative questions in web search engines logs
2. Assessing topic bias in argument corpora
3. Frame identification of arguments
Contributions

Goal: Integrating argument retrieval technology in web search engines.

1. Identifying argumentative questions in web search engines logs
2. Assessing topic bias in argument corpora
3. Frame identification of arguments

RQ1. How to identify questions that look for arguments in the query stream of a search engine?
1) Identifying Argumentative Questions

Preparing a Russian Questions Dataset

1. Filter from Yandex logs 4.5 million Russian questions on 19 controversial topics.
   Example topics: Putin, Navalny, Nord Stream, and marijuana.

2. Sample 54,850 questions and annotate them with the annotation scheme:

   - Automatic Topic Filtering
     - Not on Topic
     - On Topic
     - Ill-formed
   - Factual
     - Argumentative
     - Method
1) Identifying Argumentative Questions

Preparing a Russian Questions Dataset

1. Filter from Yandex logs 4.5 million Russian questions on 19 controversial topics
   Example topics: Putin, Navalny, Nord Stream, and marijuana

2. Sample 54,850 questions and annotate them with the annotation scheme:

   ![Automatic Topic Filtering Diagram]

   - On Topic
   - Not on Topic
   - Ill-formed
   - Factual
   - Argumentative
   - Method

Statistics and Examples:

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Percentage</th>
<th>Count</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factual</td>
<td>64%</td>
<td>25,332</td>
<td>Is marijuana legalized in Belgium?</td>
</tr>
<tr>
<td>Argumentative</td>
<td>28%</td>
<td>10,982</td>
<td>Will the president legalize marijuana?</td>
</tr>
<tr>
<td>Method</td>
<td>8%</td>
<td>3,026</td>
<td>How to use medical marijuana?</td>
</tr>
</tbody>
</table>
1) Identifying Argumentative Questions

Analysis of Questions Characteristics

Comparison of argumentative questions with factual and method questions using lexical and syntactical patterns.¹

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Starts with wh-words (except why)</th>
<th>Starts with why</th>
<th>Formed as yes/no</th>
<th>Asks for predictions</th>
<th>Asks for comparisons</th>
<th>Subject is personal pronoun</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factual</td>
<td>65.7%</td>
<td>1.3%</td>
<td>7.2%</td>
<td>3.8%</td>
<td>3.2%</td>
<td>0.3%</td>
<td>18.5%</td>
</tr>
<tr>
<td>Argumentative</td>
<td>41.3%</td>
<td>20.7%</td>
<td>13.8%</td>
<td>8.2%</td>
<td>5.7%</td>
<td>3.8%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Method</td>
<td>93.9%</td>
<td>0.4%</td>
<td>0.0%</td>
<td>0.6%</td>
<td>4.4%</td>
<td>0.4%</td>
<td>0.9%</td>
</tr>
</tbody>
</table>

Finding: Argumentative questions look for predictions and explicitly for reasons.

¹Some question characteristics overlap (e.g., asks for predictions and asks for comparisons.)
1) Identifying Argumentative Questions

Question Type Classification

Developing classifiers to map questions to argumentative, factual or method.

Experimental setting is leave-one-topic-out: test on one topic after training on remaining topics.

F1-score of the three question types and their macro average.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Factual</th>
<th>Argumentative</th>
<th>Method</th>
<th>Macro</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majority Baseline</td>
<td>0.78</td>
<td>0.00</td>
<td>0.00</td>
<td>0.26</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>0.80</td>
<td>0.61</td>
<td>0.52</td>
<td>0.65</td>
</tr>
<tr>
<td>RuBERT</td>
<td>0.85</td>
<td>0.74</td>
<td>0.74</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Finding: Identifying argumentative questions is feasible, even on unseen topics.
Contributions

Goal: Fostering the generalizability of argument mining approaches over topic.

1. Identifying argumentative questions in web search engines logs
2. Assessing topic bias in argument corpora
3. Frame identification of arguments

RQ2. How well do argument corpora represent controversial topics?
2) Topic Bias in Argument Corpora

Survey Regarding Topic Selection

A survey of 59 argument corpora shows that researchers take three approaches:

- Manual selection: choosing a set of topics manually
- Source-driven-greedy: a whole source is exploited
- Source-driven-sample: a source is sampled
2) Topic Bias in Argument Corpora

Trustworthy Topic Ontologies

Topic ontology: a directed graph where

- Nodes are topics
- Edges indicate is part of relation: topics that are part of other topics are called subtopics.
2) Topic Bias in Argument Corpora

Trustworthy Topic Ontologies

Topic ontology: a directed graph where

- Nodes are topics
- Edges indicate is part of relation: topics that are part of other topics are called subtopics.

Three trustworthy topic ontologies with categorized documents

- World Economic Forum (WEF): global issues (mainly economical)
- Debatepedia: biased to western culture
- Wikipedia
2) Topic Bias in Argument Corpora

Trustworthy Topic Ontologies

Topic ontology: a directed graph where

- Nodes are topics
- Edges indicate is part of relation: topics that are part of other topics are called subtopics.

<table>
<thead>
<tr>
<th>Ontology</th>
<th>Topics</th>
<th>Authors</th>
<th>Docs</th>
</tr>
</thead>
<tbody>
<tr>
<td>World Economic Forum Level-1</td>
<td>137</td>
<td>334</td>
<td>940</td>
</tr>
<tr>
<td>World Economic Forum Level-2</td>
<td>822</td>
<td>217</td>
<td>550</td>
</tr>
<tr>
<td>Wikipedia Level-1</td>
<td>14</td>
<td>78,014</td>
<td>68</td>
</tr>
<tr>
<td>Wikipedia Level-2</td>
<td>748</td>
<td>1,930</td>
<td>1</td>
</tr>
<tr>
<td>Debatepedia</td>
<td>89</td>
<td>145</td>
<td>62</td>
</tr>
</tbody>
</table>
2) Topic Bias in Argument Corpora

Units Categorization

The units of 59 corpora are mapped to the three topic ontologies.

- **Manual:**
  mapping the topic labels of a corpus with synonymous or upper topics.

- **Automatic:**
  assessing the similarity between a unit and the documents of a topic.
Findings:

- The topic distribution of existing argument corpora is skewed and concentrated around a small set of topics.
- Argument extractors built on these argument corpora might not be generalizable across topics.
Contributions

Goal: Enable users to select arguments that resonate with their audience.

1. Identifying argumentative questions in web search engines logs
2. Assessing topic bias in argument corpora
3. Frame identification of arguments

RQ3. How to identify the frames of an argument?
3) Frame Identification of Arguments

Introduction

- Framing is to emphasize a specific aspect of a topic while concealing others (Entman et al., 1993).

- A topic like nuclear energy can be framed according to its economical potential or environmental effect among others.
3) Frame Identification of Arguments

Introduction

- Framing is to emphasize a specific aspect of a topic while concealing others (Entman et al., 1993).

- A topic like nuclear energy can be framed according to its economical potential or environmental effect among others.

An argument frames a topic by emphasizing an aspect while rejecting others.

Examples:

Frame 1: Environment

<table>
<thead>
<tr>
<th>Premise</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>We should keep nuclear energy</td>
<td>It produces zero carbon emissions</td>
</tr>
</tbody>
</table>

Frame 2: Economy

<table>
<thead>
<tr>
<th>Premise</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>We should invest in nuclear energy</td>
<td>It is the lowest-cost form of power</td>
</tr>
</tbody>
</table>
3) Frame Identification of Arguments

Generic vs Topic-specific Frames

Examples of topic-specific frames:

Bill Clinton is a bad president
Lewinksy scandal lowered his credibility

Bill Clinton is a good president
NAFTA led to thousands of jobs

Frame 1: Lewinksy Scandal
Frame 2: NAFTA

First argument frames dataset covering 467 topics.

<table>
<thead>
<tr>
<th>Frame Type</th>
<th>Count of Frames</th>
<th>Count of Arguments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic</td>
<td>330</td>
<td>7,052</td>
</tr>
<tr>
<td>Topic-specific</td>
<td>1,293</td>
<td>5,274</td>
</tr>
<tr>
<td>All</td>
<td>1,623</td>
<td>12,326</td>
</tr>
</tbody>
</table>
3) Frame Identification of Arguments

Approach

a) Topic Clustering

b) Topic Removal

c) Frame Clustering
3) Frame Identification of Arguments

Approach: a) Topic Clustering

Semantic Spaces:

- TF-IDF

Clustering algorithm: K-means with euclidean distance.
3) Frame Identification of Arguments

Approach: b) Topic Removal

Two models for topic removal:

- **Content-based removal:**
  Remove tokens with high TF-IDF values in each topic cluster.

- **Structure-based removal:**
  Remove the conclusion of an argument.
3) Frame Identification of Arguments

Approach: c) Frame Clustering

Input
A set of arguments

<table>
<thead>
<tr>
<th>Args with topic 1</th>
<th>Topic-free arg's 1</th>
<th>Args with frame 1</th>
<th>Generic frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF</td>
<td>TF</td>
<td>TF</td>
<td>F</td>
</tr>
<tr>
<td>TF</td>
<td>TF</td>
<td>TF</td>
<td>F</td>
</tr>
<tr>
<td>TF</td>
<td>TF</td>
<td>TF</td>
<td>F</td>
</tr>
<tr>
<td>TF</td>
<td>TF</td>
<td>TF</td>
<td>F</td>
</tr>
<tr>
<td>TF</td>
<td>TF</td>
<td>TF</td>
<td>F</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Args with topic m</th>
<th>Topic-free arg's m</th>
<th>Args with frame 2</th>
<th>Topic-specific frames (from one topic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF</td>
<td>TF</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>TF</td>
<td>TF</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>TF</td>
<td>TF</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>TF</td>
<td>TF</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>TF</td>
<td>TF</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Args with frame 3</th>
<th>Args with frame 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

Legend: TF argument | TT topics | FFFF frames

Semantic Spaces:

- TF-IDF
- LSA

Clustering algorithm: K-means with euclidean distance.
3) Frame Identification of Arguments

Experiments

- Topic Clustering
- Frame Clustering
3) Frame Identification of Arguments

Experiment results: Generic Frame Clustering

Clustering effectiveness in bcubed F1-score.

<table>
<thead>
<tr>
<th>Semantic Space</th>
<th>Topic Removal</th>
<th>Topic Clustering</th>
<th>Frame Clustering</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF-IDF</td>
<td>No-removal</td>
<td>0.45</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>Content-based</td>
<td>0.42</td>
<td><strong>0.28</strong></td>
</tr>
<tr>
<td></td>
<td>Structure-based</td>
<td>0.17</td>
<td>0.26</td>
</tr>
<tr>
<td>LSA</td>
<td>No-removal</td>
<td>0.44</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>Content-based</td>
<td>0.40</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>Structure-based</td>
<td>0.25</td>
<td><strong>0.20</strong></td>
</tr>
</tbody>
</table>

- Removing topic-specific information helps identifying generic frames.
- Structure-based argument removal models is more effective at removing topic-information.
3) Frame Identification of Arguments

Experiment Results: Topic-specific Frame Clustering

Clustering effectiveness in bcubed F1-score.

<table>
<thead>
<tr>
<th>Semantic Space</th>
<th>Topic Removal</th>
<th>Topic Clustering</th>
<th>Frame Clustering</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF-IDF</td>
<td>No-removal</td>
<td>0.45</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Content-based</td>
<td>0.42</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>Structure-based</td>
<td>0.17</td>
<td>0.45</td>
</tr>
<tr>
<td>LSA</td>
<td>No-removal</td>
<td>0.44</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>Content-based</td>
<td>0.40</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>Structure-based</td>
<td>0.25</td>
<td>0.46</td>
</tr>
</tbody>
</table>

- Removing topic-specific information helps identifying frames only in LSA space.
- Using TF-IDF semantic space without topic removal performs the best.
Conclusion: Research Questions

1. User Perspective
   - Topic

2. Mining Perspective
   - Argument Retrieval System
   - Argument Index

3. Analysis Perspective
   - pro
   - con

1. How to identify questions that look for arguments in the query stream of a search engine?
2. How well do argument corpora represent controversial topics?
3. How to identify the frames of an argument?
Conclusion

Contributions:

- Enabling search engines to identify and respond to questions that pertain to controversial topics and those that look for arguments.
- Method to quantify topic bias in argument corpora and resources to help researchers sample topics in a more representative way.
- A model and an approach for frames in argumentation.

Findings:

- Argumentative questions ask for predictions or reasons.
- The topic distribution of existing argument corpora is skewed and concentrated around a small set of topics.
- Identifying the topic of an argument and removing it helps identifying its frames.
Future Work

- **User Perspective**
  1. Exploiting session information (i.e., not only one question but a series of questions)
  2. Know more about the user intent (e.g., use case, audience, types of arguments).

- **Mining Perspective**
  1. Developing a unified topic ontology.
  2. Developing topic sampling strategies.
  3. Assessing topic-robustness of argument extractors.

- **Analysis Perspective**
  1. Detecting effective frames sequence from news articles.
  2. Generating frame labels based on argument clusters.
A survey of 59 argument corpora shows that researchers take three approaches:

- **Manual selection**: choosing a set of topics manually
- **Source-driven-greedy**: a whole source is exploited
- **Source-driven-sample**: a source is sampled
Automatic Corpora Unit Categorization

About a third of argument corpora do not provide corpora topic labels and hence is not included in the previous analysis.

**Approach**: Semantic indexing calculates the cosine similarity between a corpus unit and the documents categorized under an ontology topic.

**Evaluation**: Pooled evaluation for 104 corpora units with a depth of five ontology topics.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Wikipedia</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level-1</td>
<td>Level-2</td>
<td>Level-1</td>
<td>Level-2</td>
</tr>
<tr>
<td>Direct match</td>
<td>0.06</td>
<td>0.40</td>
<td>0.29</td>
<td>0.19</td>
</tr>
<tr>
<td>Semantic Indexing</td>
<td>0.43</td>
<td><strong>0.59</strong></td>
<td><strong>0.34</strong></td>
<td><strong>0.33</strong></td>
</tr>
<tr>
<td>Text2vec-SI $^{BERT}$</td>
<td><strong>0.47</strong></td>
<td>0.31</td>
<td>0.28</td>
<td>0.23</td>
</tr>
</tbody>
</table>

F1-score of the approaches
Dataset Construction from Debatepedia.org

<table>
<thead>
<tr>
<th>Aspect 1</th>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aspect 2</th>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aspect 3</th>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Dataset Construction from Debatepedia.org
## Dataset Construction from Debatepedia.org

<table>
<thead>
<tr>
<th>Count of Topics</th>
<th>Count of Frames</th>
<th>Count of Arguments</th>
</tr>
</thead>
<tbody>
<tr>
<td>467</td>
<td>1623</td>
<td>12,326</td>
</tr>
</tbody>
</table>

### Count of Generic Frames

- **1-50**: 11%
- **51-100**: 25%
- **101-150**: 5%
- **151-200**: 1%

### Count of Topic-Specific Frames

- **1**: 80%
- **2**: 3%
- **3**: 2%
- **4**: 4%
- **>4**: 11%

### Count of Arguments

- **1-5**: 92%
- **6-10**: 5%
- **11-15**: 2%
- **16-20**: 1%
### Experiment Results: Topic Clustering

<table>
<thead>
<tr>
<th>Semantic Space</th>
<th># Topics</th>
<th>Bcubed F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSA Debate</td>
<td>310</td>
<td>0.52</td>
</tr>
<tr>
<td>TF-IDF</td>
<td>260</td>
<td>0.45</td>
</tr>
<tr>
<td>LSA</td>
<td>280</td>
<td>0.44</td>
</tr>
</tbody>
</table>

LSA Debate is the best semantic space to model the topic of arguments.