

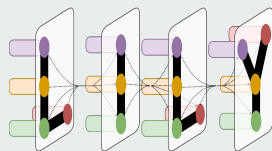


# Surveytree

**Automatic Generation of Survey Structures  
for NLP and AI topics**

Prawat Trairatvorakul, Alexander Fabbri, Dragomir Radev

Yale





# Why is structure important?

## Structure increases readability

Consider:

- (1) Definition, Task, Evaluation, Results
- (2) Results, Task, Definition, Evaluation

## Current systems

- **Domain-specific templates**  
by clustering (Sauper and Barzilay, 2009)
- **HMM** of topic ordering (Jha et al., 2015)

## Surveys cover many subtopics

Each with their own subtopics

Eg. Automatic Summarization

- > Evaluation
  - > ROUGE
  - > BLEU
  - > DUC Quality Questions



# Approach

## Convert document to tree

Nodes are section headings with subheadings as children

## Extract features from heading

Adaptation of tf-idf

- tf: ancestor headings taken into account
- idf: word frequencies from ACL Anthology

## Combine trees into single tree

Nodes merged recursively if similar  
Other nodes are simply added



# Dataset

## Five topics considered

- Linear Algebra
- Statistics and Probability
- Sentiment Analysis
- Automatic Summarization
- Dependency Parsing

4 - 6 input documents from AAN each

## Document Frequencies:

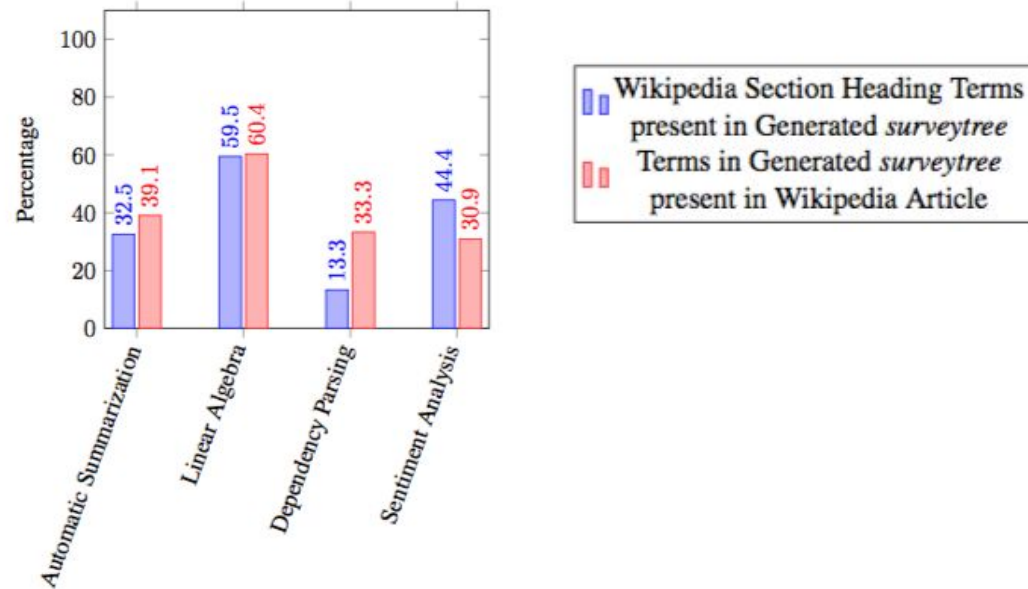
Calculated from papers in ACL Anthology



# Evaluation & Results

- (1) Wikipedia Term Coverage
- (2) Human Evaluation of Node Merges
- (3) Tree Properties
- (4) Inspection of Output

# Wikipedia Term Coverage



# Human Evaluations of Node Merges

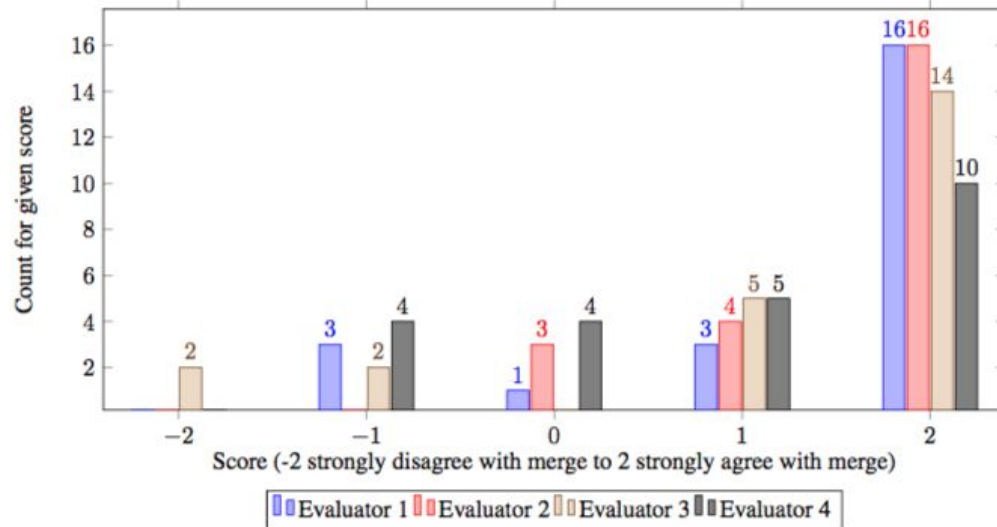


Figure 3: Evaluation of Heading Merges for Automatic Summarization

# Human Evaluations of Node Merges

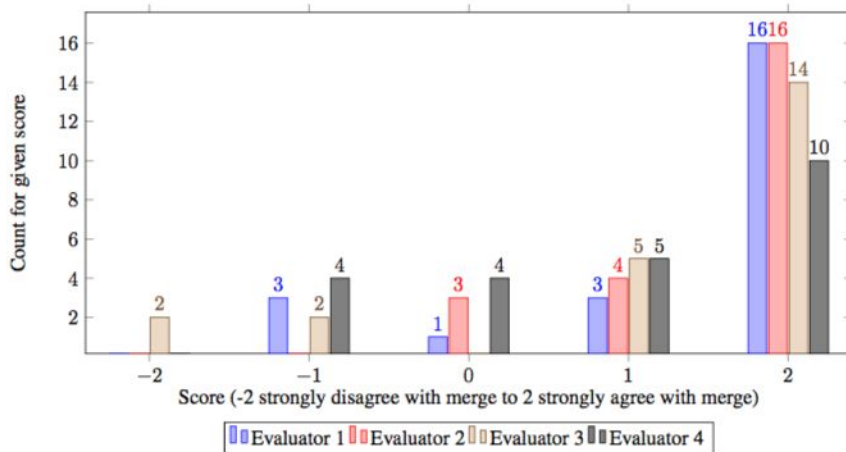


Figure 3: Evaluation of Heading Merges for Automatic Summarization

Topic	Krippendorff's $\alpha$
Automatic Summarization	0.73
Linear Algebra	0.54
Statistics & Probability	0.48
Dependency Parsing	0.38
Sentiment Analysis	0.39

Table 2: Inter-Evaluator Agreement

Thanks Alex, Michi, and Ryan!



# Human Evaluations of Node Merges

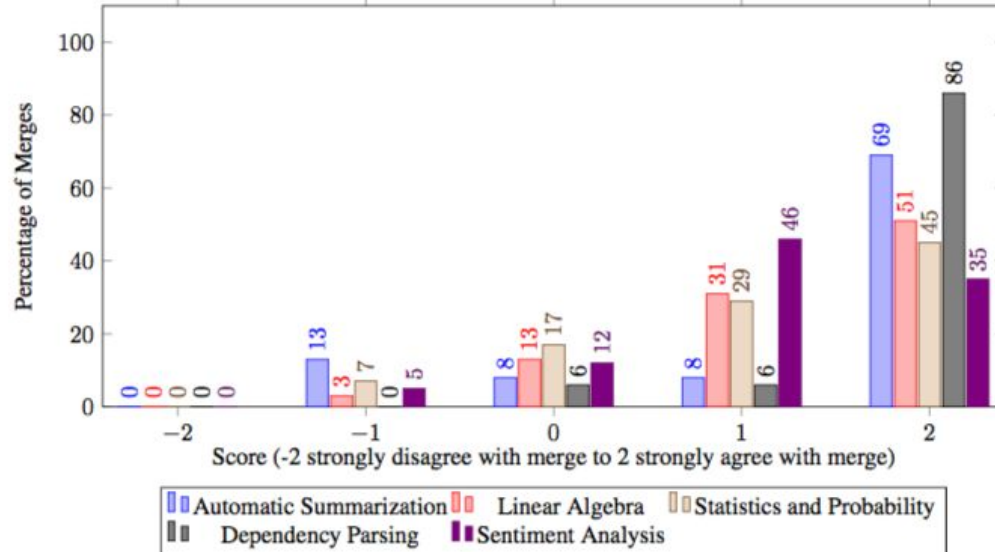


Figure 2: Mean Merging Agreement Score Distributions for each topic



# Tree Properties

	<i>surveytree</i>	<i>dt 1</i>	<i>dt 2</i>	<i>dt 3</i>	<i>dt 4</i>	<i>dt 5</i>	sum	mean
# Nodes	123	90	73	56	51	28	298	59.6
# Nodes in <i>surveytree</i>	123	90	41	49	41	18	239	47.8
Max Depth	4	3	2	2	2	2	N/A	2.2
Mean Depth	2.8	2.7	2.0	0.8	2.0	2.0	N/A	1.9
Max # Children	19	11	12	25	8	13	N/A	13.8
Mean # Children	3.9	3.6	6.5	3.2	5.6	5.4	N/A	4.9

Table 3: Tree Structure Properties for Linear Algebra



# Selections from Generated Survey Trees

```
(1)Sentiment Analysis and Opinion Mining
├── (1)Sentiment Analysis: A Fascinating Problem
│   └── ...
├── (2)The Problem of Sentiment Analysis
│   └── ...
├── (1)Document Sentiment Classification
│   └── ...
├── (2)Sentence Subjectivity and Sentiment Classification
│   ├── (3)Subjectivity Classification
│   │   └── (3)Classification Based on Supervised Learning
│   ├── (1)Dealing with Conditional Sentences
│   ├── (1)Dealing with Sarcastic Sentences
│   └── ...
├── ...
└── (1)Concluding Remarks
```

Figure 4: Selection of nodes from the generated *surveytree* of Sentiment Analysis

```
(1)Introduction to Bayesian Statistics
├── ...
├── (2)Hypothesis Testing and Model Selection
│   ├── (1)An Example Hypothesis Test
│   │   ├── (1)The "Testing" Prior
│   │   ├── (1)Some Terminology
│   │   ├── (1)Hypothesis Testing and the Marginal Likelihood
│   │   └── (3)The hypothesis-testing framework
│   └── ...
├── ...
├── (3)Probability
│   ├── (1)Random Variables
│   │   └── (3)Discrete random variables
│   └── (5)Continuous Random variables
├── ...
└── ...
```

Figure 5: Selection of nodes from the generated *surveytree* of Statistics and Probability



## Conclusion & Future Work

- This approach has potential
- Heading comparison works fairly well
- Also has a lot of room to improve
- Explore possibility of using vector magnitudes
- We need more reliably parsed input documents