Summarizing Contrastive Viewpoints

- 2010 U.S. Healthcare Legislation
  - 948 verbatim responses from Gallup opinion phone survey
  - 45% for, 48% against (March 2010)

For:
  “because a lot of people can't afford it [insurance];
  45,000 people die each year because of lack of healthcare.”

Against:
  “everybody should have their own healthcare, and if you can't afford it, you should just die.”

Different viewpoints  |  Same issue
Summarizing Contrastive Viewpoints

• Bitterlemons Corpus
  ○ Editorials about the Israel-Palestine conflict
  ○ Introduced by Lin et al. (2006)
  ○ 312 articles by Israeli authors, 282 articles by Palestinian authors

**Palestinian:** The wall that Israel has been building in the Palestinian occupied territories under the pretext of security, the wall that is being called the apartheid wall by the Palestinian side, has lately drawn a great deal of high-level attention.

**Israeli:** Thus the Palestinian information campaign has succeeded in persuading the world that the fence is a “wall”, even though only a few small segments out of hundreds of kilometers are configured as walls [...].

Saturday, October 9, 2010
Standard Summarization

---

**Generate separate summaries for each viewpoint:**

<table>
<thead>
<tr>
<th>For the healthcare bill</th>
<th>Against the healthcare bill</th>
</tr>
</thead>
<tbody>
<tr>
<td>• there are so many <strong>people</strong> who do not have healthcare and they are in <strong>need</strong> of it.</td>
<td>• just don’t think its going to work out well and will drive the <strong>cost</strong> of healthcare up.</td>
</tr>
<tr>
<td>• because i have poor insurance and i think it might <strong>help</strong> me.</td>
<td>• it’s too much <strong>government</strong>.</td>
</tr>
<tr>
<td>• because there are a lot of <strong>people</strong> out there that don’t go to the doctors</td>
<td>• it’s too <strong>expensive</strong>, it does not provide what it needs to be provided, and the <strong>government</strong> help with catastrophic illnesses. the people pay general routine illnesses. second, it is bankrupting the country.</td>
</tr>
<tr>
<td>because they don’t have enough money.</td>
<td></td>
</tr>
<tr>
<td>• <strong>need</strong> as much as we can because we have so much sickness</td>
<td></td>
</tr>
</tbody>
</table>

• Output based on the *LexRank* algorithm (Erkan & Radev, 2004)
Contrastive Summarization (Macro Level)

**Make the viewpoint summaries more comparable:**
- No alignment of sentences in “macro” summary

<table>
<thead>
<tr>
<th>For the healthcare bill</th>
<th>Against the healthcare bill</th>
</tr>
</thead>
<tbody>
<tr>
<td>• i favor healthcare for who needs it, mostly old <strong>people</strong> who don’t have healthcare. the <strong>government</strong> should <strong>help</strong> the <strong>people</strong> when they are old. they should have that kind of healthcare.</td>
<td>• i think we can’t be responsible for other <strong>people’s</strong> healthcare.</td>
</tr>
<tr>
<td>• i just think something has to be done, the <strong>price</strong> of health is going up.</td>
<td>• doesn’t address things that <strong>need</strong> to be done, addresses things that don’t <strong>need</strong> to be done.</td>
</tr>
<tr>
<td>• [i] pay for private insurance.</td>
<td>• it’s going to increase the <strong>cost</strong> to those insured.</td>
</tr>
<tr>
<td>• bring down <strong>cost</strong>.</td>
<td>• i believe we can’t afford it.</td>
</tr>
<tr>
<td></td>
<td>• way too <strong>expensive</strong>, too intrusive, too much <strong>government</strong> control.</td>
</tr>
</tbody>
</table>

- Output based on our new *Comparative LexRank* algorithm
Contrastive Summarization (Micro Level)

- Explicitly align pairs of contrastive sentences in “micro” summary:

<table>
<thead>
<tr>
<th>For the healthcare bill</th>
<th>Against the healthcare bill</th>
</tr>
</thead>
<tbody>
<tr>
<td>the government already provides half of the healthcare dollars in the united states [...] [they] might as well spend their dollars smarter</td>
<td>government is too much involvement.</td>
</tr>
<tr>
<td>my kids are uninsured.</td>
<td>a lot of people will be getting it that should be getting it on their own, and my kids will be paying a lot of taxes.</td>
</tr>
<tr>
<td>so everybody would have it and afford it.</td>
<td>we cannot afford it.</td>
</tr>
</tbody>
</table>

- Output based on our new **Comparative LexRank** algorithm
Previous Work

• Kim and Zhai (2009)
  ○ Micro-contrastive summarization
  ○ Pairs of contradictory sentences
    ▪ e.g., “the battery life is pretty good” vs “battery life sucks”

• Optimizes how well the summary represents the collection as well as the comparability of the sentences in each pair
Lerman and McDonald (2009)
- Macro-contrastive summarization

Summaries are similar to own category but different from opposite category
- e.g. product reviews for two different products; summarize what is unique to each product

Minimize KL-divergence between model of a summary and its viewpoint, but maximize KL-divergence between summary and the opposite viewpoint
Our Complete System

- **Stage 1: Extract viewpoints automatically**
  - *Unsupervised* modeling of viewpoints

- **Stage 2: Summarize the extracted viewpoints**
  - Summarize in a way to highlight *contrast*
  - We’ll describe this stage first
Overview

- **Contrastive summarization algorithm**
  - Comparative LexRank; graph-based approach
- **Summarization evaluation - Supervised**
  - Healthcare corpus
- **Viewpoint modeling and extraction**
  - Unsupervised viewpoint clustering
- **Summarization evaluation - Unsupervised**
  - Bitterlemons corpus
- **Conclusion**
LexRank (Erkan & Radev, 2004)

Line thickness = edge weights = sentence similarity
LexRank (Erkan & Radev, 2004)

\[
P(x_i \rightarrow x_j) = \frac{\text{sim}(x_i, x_j)}{\sum_{j' \in X} \text{sim}(x_i, x_{j'})}
\]
LexRank (Erkan & Radev, 2004)

\[ P(x_i \rightarrow x_j) = \frac{\text{sim}(x_i, x_j)}{\sum_{j' \in X} \text{sim}(x_i, x_{j'})} \]
LexRank (Erkan & Radev, 2004)

\[ P(x_i \to x_j) = \frac{\text{sim}(x_i, x_j)}{\sum_{j' \in X} \text{sim}(x_i, x_j')} \]
LexRank (Erkan & Radev, 2004)

\[ P(x_i \rightarrow x_j) = \frac{\text{sim}(x_i, x_j)}{\sum_{j' \in X} \text{sim}(x_i, x_{j'})} \]
LexRank (Erkan & Radev, 2004)

\[ P(x_i \rightarrow x_j) = \frac{\text{sim}(x_i, x_j)}{\sum_{j' \in X} \text{sim}(x_i, x_{j'})} \]
LexRank (Erkan & Radev, 2004)

This models content **centrality**; stationary distribution $P(X)$ over nodes gives scoring for sentences.
Comparative LexRank

- Sentences belong to viewpoints
- Goal: make viewpoint summaries similar to each other so that they can be directly compared
- Idea: put sentences from all viewpoints into same graph; control which viewpoints the random walker jumps to
Comparative LexRank

Color = viewpoint
Comparative LexRank

Trick: force random walk to move back and forth between views
Comparative LexRank

Trick: force random walk to move back and forth between views
Comparative LexRank

Trick: force random walk to move back and forth between views
Comparative LexRank

Trick: force random walk to move back and forth between views
Comparative LexRank

Trick: force random walk to move back and forth between views
Comparative LexRank

Trick: force random walk to move back and forth between views
Comparative LexRank

Favor sentences with higher inter-viewpoint similarity
Comparative LexRank

- New model: random walker first decides whether to jump to the same or opposite viewpoint according to some probability
  - If $z = 0$, jump to same viewpoint
  - If $z = 1$, jump to opposite viewpoint

- Different transition probabilities conditioned on $z$:

$$P(x_i \rightarrow x_j | z) = \frac{\text{sim}_z(x_i, x_j)}{\sum_{j' \in X} \text{sim}_z(x_i, x_j')}$$

  - Controls which set of nodes can be transitioned to
  - Multiply sim by 0 if between a node you can’t jump to
Comparative LexRank

The transition probability is:

\[ P(x_i \rightarrow x_j) = \lambda P(x_i \rightarrow x_j | z = 0) + (1 - \lambda) P(x_i \rightarrow x_j | z = 1) \]

\( \lambda = P(z = 0) \) controls the level of contrast

- \( \lambda = 1 \) always jump to same viewpoint
  - Equivalent to applying LexRank to viewpoints independently
- \( \lambda = 0.5 \) equal odds of jumping to same or opposite viewpoint
  - Even tradeoff between representation of viewpoint and contrast with opposite viewpoint (2 objectives)
- \( \lambda = 0 \) always jump to opposite viewpoint
  - A viewpoint’s summary will contain sentences that look like the opposite viewpoint
How to score a pair of nodes from opposite viewpoints?

“because i have no insurance and i need it.”

“because i have health insurance.”
Comparative LexRank

\[ P(x_i) P(x_i \rightarrow x_j | z = 1) + P(x_j) P(x_j \rightarrow x_i | z = 1) \]

“because i have no insurance and i need it.”

“because i have health insurance.”

Saturday, October 9, 2010
Overview

- Contrastive summarization algorithm
  - Comparative LexRank; graph-based approach
- **Summarization evaluation - Supervised**
  - Healthcare corpus
- Viewpoint modeling and extraction
  - Unsupervised viewpoint clustering
- **Summarization evaluation - Unsupervised**
  - Bitterlemons corpus
- Conclusion
Evaluation Setup (Healthcare Corpus)

- Gold standard summaries for each viewpoint
  - Prominent reasons found in data as analyzed by humans

For:

<table>
<thead>
<tr>
<th>Reason</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>People need health insurance/Too many uninsured</td>
<td>29%</td>
</tr>
<tr>
<td>System is broken/Needs to be fixed</td>
<td>18%</td>
</tr>
<tr>
<td>Costs are out of control/Would help control costs</td>
<td>12%</td>
</tr>
<tr>
<td>Moral responsibility to provide/Obligation/Fair</td>
<td>12%</td>
</tr>
<tr>
<td>Would make healthcare more affordable</td>
<td>10%</td>
</tr>
<tr>
<td>Don’t trust insurance companies</td>
<td>5%</td>
</tr>
<tr>
<td>Respondent or family member currently lacks coverage</td>
<td>4%</td>
</tr>
<tr>
<td>To help senior citizens</td>
<td>4%</td>
</tr>
<tr>
<td>To help the poor</td>
<td>3%</td>
</tr>
</tbody>
</table>
Evaluation Setup

- **ROUGE**
  - Recall-based evaluation metric compares against gold summary
  - Modification: scale term counts by prominence in data

- **Against:**
  - Will raise costs of insurance/Make it less affordable: 20%
  - Does not address real problems: 19%
  - Need more information/clarity on how system would work: 8%
  - Against big government/Too much government involvement (general): 8%
  - Government should not be involved in healthcare: 7%
  - Healthcare is a privilege, not an entitlement: 6%
  - Would cost government too much/Too much spending/Increase the deficit: 5%
  - People should not be required to buy health insurance: 5%
  - Will affect respondent’s current health insurance: 4%
  - Socialism/Socialized medicine: 4%
  - Oppose the “public option” proposal: 3%
  - Rushing it through process/Should take more time: 3%
  - Would hurt senior citizens/Medicare: 3%
  - Would pay for abortions: 2%
  - Has not worked in other countries: 1%
Baseline Approach

- Compare against non-comparative LexRank
- Analogous to $\lambda = 1$!
  - Always jump to same viewpoint

Remember:

$$P(x_i \rightarrow x_j) = \lambda P(x_i \rightarrow x_j | z = 0) + (1 - \lambda) P(x_i \rightarrow x_j | x = 1)$$
Evaluation Results (Healthcare Corpus)

- Evaluate summaries against the **opposite** viewpoint:
Evaluation Results (Healthcare Corpus)

- Evaluate summaries against their own viewpoint:

![Graph showing ROUGE scores for different Lambda values]

Saturday, October 9, 2010
Overview

- **Contrastive summarization algorithm**
  - Comparative LexRank; graph-based approach

- **Summarization evaluation - Supervised**
  - Healthcare corpus

- **Viewpoint modeling and extraction**
  - Unsupervised viewpoint clustering

- **Summarization evaluation - Unsupervised**
  - Bitterlemons corpus

- **Conclusion**
Comparative LexRank

- So far we’ve assumed that there is a way to partition the data into viewpoints
- Question: how do we know if the nodes are red or blue?
- Viewpoint membership might be probabilistic
- Viewpoint membership might not be labeled
Comparative LexRank

Sentences may represent viewpoints to varying degrees. Intuition: assign higher scores to more representative sentences.
Comparative LexRank

• Assign a probability of viewpoint membership to each sentence

Recall: 

\[ P(x_i \rightarrow x_j | z) = \frac{\text{sim}_z(x_i, x_j)}{\sum_{j' \in X} \text{sim}_z(x_i, x_j')} \]

• Multiple sim by the \textbf{probability} that \((i, j)\) belong to the same viewpoint (if \(z = 0\)) or that they belong to the opposite viewpoint (if \(z = 1\)).
Probabilistic Topic Modeling

- Topic models
  - Latent Dirichlet Allocation (LDA)
- Idea: use LDA with 2 “topics” to discover viewpoints
- 2 improvements:
  - Use better features than “bag of words”
    - “bag of features”
    - Dependency information, also negation/polarity
  - Use a better model than LDA
Imagine a set of **product reviews**

- Each word might depend on the viewpoint/sentiment as well as the topic/aspect being discussed

<table>
<thead>
<tr>
<th>View/</th>
<th>Usability</th>
<th>Service</th>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>easy intuitive</td>
<td>friendly helpful</td>
<td>sleek durable</td>
</tr>
<tr>
<td>Negative</td>
<td>confusing difficult</td>
<td>rude slow</td>
<td>flimsy ugly</td>
</tr>
</tbody>
</table>

**TAM**: each document is both a mixture of *topics* and a separate mixture of *viewpoints*

- Words may depend on both, one or the other, or neither
Clustering Results

- Measured accuracy by comparing cluster assignments to gold labels
  - Dependency features make a big difference!

- Healthcare corpus:
  - Median clustering accuracy (200 trials):
    - Bag of words: 61.0%
    - Best feature set: 70.7%

- Bitterlemons corpus:
  - Median clustering accuracy (50 trials):
    - Bag of words: 69.3%
    - Best feature set: 88.1%
Overview

- Contrastive summarization algorithm
  - Comparative LexRank; graph-based approach
- Summarization evaluation - Supervised
  - Healthcare corpus
- Viewpoint modeling and extraction
  - Unsupervised viewpoint clustering
- **Summarization evaluation - Unsupervised**
  - Bitterlemons corpus
- Conclusion
Evaluation Setup (Bitterlemons Corpus)

- Unsupervised viewpoint summarization
- Run TAM on document collection
  - Use dependency features
  - Repeat 10 times, take model with best data likelihood
- Generate macro-level summaries for 2 viewpoints
  - $\lambda = 0.5$ (even balance)
  - Summary length = 6 sentences
- Ask humans to label each summary as the “Israeli” or “Palestinian” viewpoint
  - Measures clustering accuracy and summarization salience
  - Randomly partition each summary in half for each judge
Evaluation Results (Bitterlemons Corpus)

- 2 viewpoints x 6 sentences = 12 sentences
  - 11 of 12 sentences clustered correctly by TAM

- 8 human judges given 4 summaries
  - correctly labeled 78% of the summaries

- ROUGE scores on the healthcare set were similarly degraded when using the unsupervised output
  - More contrast (smaller lambda) worsens this
Conclusion

- Unsupervised viewpoint modeling
  - Achieved large gains in clustering accuracy by using simple but rich syntactic features
  - Showed that rich feature sets can be used with topic models simply by using a Naïve Bayes-like “bag of features” approach

- Contrastive multi-viewpoint summarization
  - Introduced *Comparative LexRank* algorithm
  - Same algorithm can be used for macro-level and micro-level contrastive summaries, and can generalize to >2 viewpoints
  - Our random walk formulation based on class membership could generalize to other tasks beyond summarization
Greedy Summary Generation

- Partition sentences into their viewpoints
- Choose sentences that have high scores but are not redundant with one another
  - We don’t care about the order of the sentences
  - Simple approach:
    - At each step, add the sentence with the highest score as long as $\text{sim}(\text{sentence}, S) < \delta$
    - Repeat until S exceeds user-specified length limit
Evaluation Results (Healthcare Corpus)

- Scores for the micro-contrastive summaries (summaries with explicitly aligned pairs)
  - Created gold summary by having annotators identity contrastive pairs in the gold summaries

![Graph showing ROUGE scores for Unigram and Bigram models with varying Lambda values.](image)
### Israeli viewpoint

- The American war on Iraq, however problematic for much of the world, is for most of us in Israel a welcome attempt by a friend and ally to deal with a strategic danger that we have been struggling to cope with on our own for decades.

### Palestinian viewpoint

- If the Israelis do that, in line with the Americans and the international community, I believe that after the end of the occupation, we could start real negotiations on the other issues.