Experiments with CST-based Multi-document Summarization

Maria Lucia R. Castro Jorge  
mluciacj@icmc.usp.br

Thiago A.S. Pardo  
taspardo@icmc.usp.br

NLP Group - University of São Paulo
Multi-document Summarization

- One summary from a group of texts on the same topic (Mani, 2001)
  - Relevant information according to user preferences
  - Content selection task
Content Selection Operators
(McKeown and Radev, 1995)

- **Computational artifacts**
  - process a text representation
  - produce a condensed version of it
    - Selects information considering some summarization preference
This work

- **Use of discourse information** for summarization
  - Usefulness
  - Impact in the task
Cross-document Structure Theory (CST)  
(Radev, 2000; Zhang et al., 2003)

- **Discourse relations** among texts

- **Similarities and differences**
  - Content and writing style

- **Graph representation** of the related texts
Fifteen volunteers from the French NGO Action Against Hunger (ACF) were killed in northeastern Sri Lanka today said a spokesman for the organization.

The crimes occurred in the town of Muttur, which during the last two weeks has been experiencing serious conflicts between troops of the Sri Lankan Army and the ones of the Liberation Tigers Tamil Eelam (LTTE).

"We try to send a team to Muttur to look into what is happening, but the soldiers did not allow to enter the city, which is totally blocked," said Director of ACF.

To date, the Sri Lankan authorities did not confirm the deaths or clarified what happens in the town of Muttur.
CST relation typology
(Maziero et al., 2010)
Summarization steps

Source texts → CST parsing (manually done so far) → CST graph → Application of content selection operators → Rank of sentences → Compression rate → Extractive summary
1st step: CST parsing

Source texts

CST Graph

Sentence 1

Historical background

Sentence 4

Attribution

Follow up

Sentence 2

Contradiction

Sentence 5

Sentence 3
2nd step: general CS operator

CST Graph

Sentence 1

Attribution

Sentence 2

Follow up

Sentence 4

Historical background

Contradiction

Sentence 5

Initial rank

(1) Sentence 4
(2) Sentence 2
(3) Sentence 1
(4) Sentence 5
(5) Sentence 3

follow-up

historical background

contradiction

attribution
3rd step: preference CS operators

**Initial rank**

1. Sentence 4
2. Sentence 2
3. **Sentence 1** (highlighted)
4. Sentence 5
5. Sentence 3

**Contextual information CS operator**

**Refined rank**

1. Sentence 4
2. **Sentence 1** (highlighted)
3. Sentence 2
4. Sentence 5
5. Sentence 3
Final step: selection of sentences

Refined rank

1. Sentence 4
2. Sentence 1
3. Sentence 2
4. Sentence 5
5. Sentence 3

contradiction

follow-up

historical background

attribution

Summary
CS operators
(Jorge and Pardo, 2010)

- Each operator represents a particular summarization preference
  - General
  - Redundancy treatment
  - Context information
  - Contradictory information
  - Authorship
  - Events evolution

- Mapping of preferences to CST relations
Hybrid approach

- CST into superficial summarizers

- **Strategy** (Zhang et al., 2002)

  \[ \text{New sentence score} = \text{old sentence score} + \text{number of CST relations} \]

- New rank of sentences
Hybrid approach

- CST into superficial summarizers

  - **MEAD** (Radev et al., 2000)
    - One of the most used systems
      - Criteria: sentence position and size, centroid relatedness

  - **GistSumm** (Pardo et al., 2003)
    - Simple word frequency-based summarizer
Evaluation

- **CSTNews** (Aleixo and Pardo, 2008)
  - 50 clusters of news texts in Brazilian Portuguese, with 3-4 texts per cluster
  - Human summaries, with 30% compression rate
  - Manually annotated according to CST by 4 annotators
    - Good agreement
      - Kappa ≈ 0.5-0.6
      - 80-90 percentage agreement
Automatic Evaluation

- **ROUGE** (Lin, 2004)

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>0.5564</td>
<td>0.5303</td>
<td>0.5356</td>
</tr>
<tr>
<td>Redundancy</td>
<td>0.5761</td>
<td>0.5065</td>
<td>0.5297</td>
</tr>
<tr>
<td>Contradiction</td>
<td>0.5503</td>
<td>0.5379</td>
<td>0.5365</td>
</tr>
<tr>
<td>Authorship</td>
<td>0.5563</td>
<td>0.5224</td>
<td>0.5310</td>
</tr>
<tr>
<td>MEAD with CST</td>
<td>0.5599</td>
<td>0.4988</td>
<td>0.5230</td>
</tr>
<tr>
<td>Event Evolution</td>
<td>0.5159</td>
<td>0.5222</td>
<td>0.5140</td>
</tr>
<tr>
<td>Context</td>
<td>0.5196</td>
<td>0.4938</td>
<td>0.4994</td>
</tr>
<tr>
<td>GISTSumm with CST</td>
<td>0.4945</td>
<td>0.5089</td>
<td>0.4994</td>
</tr>
<tr>
<td>MEAD</td>
<td>0.5242</td>
<td>0.4602</td>
<td>0.4869</td>
</tr>
<tr>
<td>GISTSumm</td>
<td>0.3599</td>
<td>0.6643</td>
<td>0.4599</td>
</tr>
</tbody>
</table>
### Automatic Evaluation

- **ROUGE (Lin, 2004)**

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td>0.5564</td>
<td>0.5303</td>
<td><strong>0.5356</strong></td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
<td>0.5761</td>
<td>0.5065</td>
<td><strong>0.5297</strong></td>
</tr>
<tr>
<td><strong>Contradiction</strong></td>
<td>0.5503</td>
<td>0.5379</td>
<td><strong>0.5365</strong></td>
</tr>
<tr>
<td><strong>Authorship</strong></td>
<td>0.5563</td>
<td>0.5224</td>
<td><strong>0.5310</strong></td>
</tr>
<tr>
<td><strong>MEAD with CST</strong></td>
<td>0.5599</td>
<td>0.4988</td>
<td><strong>0.5230</strong></td>
</tr>
<tr>
<td><strong>Event Evolution</strong></td>
<td>0.5159</td>
<td>0.5222</td>
<td><strong>0.5140</strong></td>
</tr>
<tr>
<td><strong>Context</strong></td>
<td>0.5196</td>
<td>0.4938</td>
<td><strong>0.4994</strong></td>
</tr>
<tr>
<td><strong>GISTSumm with CST</strong></td>
<td>0.4945</td>
<td>0.5089</td>
<td><strong>0.4994</strong></td>
</tr>
<tr>
<td><strong>MEAD</strong></td>
<td>0.5242</td>
<td>0.4602</td>
<td><strong>0.4869</strong></td>
</tr>
<tr>
<td><strong>GISTSumm</strong></td>
<td>0.3599</td>
<td>0.6643</td>
<td><strong>0.4599</strong></td>
</tr>
</tbody>
</table>
## Automatic Evaluation

**ROUGE** (Lin, 2004)

<table>
<thead>
<tr>
<th>Method</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>0.5564</td>
<td>0.5303</td>
<td>0.5356</td>
</tr>
<tr>
<td>Redundancy</td>
<td>0.5761</td>
<td>0.5065</td>
<td>0.5297</td>
</tr>
<tr>
<td>Contradiction</td>
<td>0.5503</td>
<td>0.5379</td>
<td>0.5365</td>
</tr>
<tr>
<td>Authorship</td>
<td>0.5563</td>
<td>0.5224</td>
<td>0.5310</td>
</tr>
<tr>
<td>MEAD with CST</td>
<td>0.5599</td>
<td>0.4988</td>
<td>0.5230</td>
</tr>
<tr>
<td>Event Evolution</td>
<td>0.5159</td>
<td>0.5222</td>
<td>0.5140</td>
</tr>
<tr>
<td>Context</td>
<td>0.5196</td>
<td>0.4938</td>
<td>0.4994</td>
</tr>
<tr>
<td>GISTSumm with CST</td>
<td>0.4945</td>
<td>0.5089</td>
<td><strong>0.4994</strong></td>
</tr>
<tr>
<td>MEAD</td>
<td>0.5242</td>
<td>0.4602</td>
<td>0.4869</td>
</tr>
<tr>
<td>GISTSumm</td>
<td>0.3599</td>
<td>0.6643</td>
<td><strong>0.4599</strong></td>
</tr>
</tbody>
</table>
## Automatic Evaluation

**ROUGE** (Lin, 2004)

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>0.5564</td>
<td>0.5303</td>
<td>0.5356</td>
</tr>
<tr>
<td>Redundancy</td>
<td>0.5761</td>
<td>0.5065</td>
<td>0.5297</td>
</tr>
<tr>
<td>Contradiction</td>
<td>0.5503</td>
<td>0.5379</td>
<td>0.5365</td>
</tr>
<tr>
<td>Authorship</td>
<td>0.5563</td>
<td>0.5224</td>
<td>0.5310</td>
</tr>
<tr>
<td>MEAD with CST</td>
<td>0.5599</td>
<td>0.4988</td>
<td><strong>0.5230</strong></td>
</tr>
<tr>
<td>Event Evolution</td>
<td>0.5159</td>
<td>0.5222</td>
<td>0.5140</td>
</tr>
<tr>
<td>Context</td>
<td>0.5196</td>
<td>0.4938</td>
<td>0.4994</td>
</tr>
<tr>
<td>GISTSumm with CST</td>
<td>0.4945</td>
<td>0.5089</td>
<td>0.4994</td>
</tr>
<tr>
<td>MEAD</td>
<td>0.5242</td>
<td>0.4602</td>
<td><strong>0.4869</strong></td>
</tr>
<tr>
<td>GISTSumm</td>
<td>0.3599</td>
<td>0.6643</td>
<td>0.4599</td>
</tr>
</tbody>
</table>
## Automatic Evaluation

- **ROUGE** (Lin, 2004)

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contradiction</td>
<td>0.5503</td>
<td>0.5379</td>
<td>0.5365</td>
</tr>
<tr>
<td>Authorship</td>
<td>0.5563</td>
<td>0.5224</td>
<td>0.5310</td>
</tr>
<tr>
<td>MEAD with CST</td>
<td>0.5599</td>
<td>0.4988</td>
<td>0.5230</td>
</tr>
<tr>
<td>Event Evolution</td>
<td>0.5159</td>
<td>0.5222</td>
<td>0.5140</td>
</tr>
<tr>
<td>Context</td>
<td>0.5196</td>
<td>0.4938</td>
<td>0.4994</td>
</tr>
<tr>
<td>GISTSumm with CST</td>
<td>0.4945</td>
<td>0.5089</td>
<td>0.4994</td>
</tr>
<tr>
<td>MEAD</td>
<td>0.5242</td>
<td>0.4602</td>
<td>0.4869</td>
</tr>
<tr>
<td>GISTSumm</td>
<td>0.3599</td>
<td>0.6643</td>
<td>0.4599</td>
</tr>
</tbody>
</table>

Statistically significant differences with 95% confidence
Human Evaluation

- User satisfaction
  - Random sample of texts
  - 6 evaluators
    - Coherence, cohesion, informativity and redundancy
  - Grades to the summaries
    - 0: Unacceptable
    - 1: Bad
    - 2: Regular
    - 3: Good
    - 4: Excellent
Human Evaluation

- 0: Unacceptable
- 1: Bad
- 2: Regular
- 3: Good
- 4: Excellent

<table>
<thead>
<tr>
<th>Preference</th>
<th>Coherence</th>
<th>Cohesion</th>
<th>Informativity</th>
<th>Redundancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>3.6</td>
<td>3.2</td>
<td>3.6</td>
<td>1.8</td>
</tr>
<tr>
<td>Context</td>
<td>2.1</td>
<td>2.7</td>
<td>2.2</td>
<td>3.6</td>
</tr>
<tr>
<td>Authorship</td>
<td>3.3</td>
<td>2.4</td>
<td>3.0</td>
<td>2.8</td>
</tr>
<tr>
<td>Contradiction</td>
<td>2.4</td>
<td>2.7</td>
<td>3.7</td>
<td>2.5</td>
</tr>
<tr>
<td>Events evolution</td>
<td>2.1</td>
<td>2.5</td>
<td>3.2</td>
<td>2.6</td>
</tr>
</tbody>
</table>
Conclusions

- Use of CST allows to explore user preferences
- CST improves superficial methods
Current work

- 3 main lines
  - Automatic CST parsing
  - Machine learning of good summary CST configuration
  - Combination of CST-based CS operators with traditional summarization strategies
Experiments with CST-based Multi-document Summarization

Thank you!

www.nilc.icmc.usp.br