User-focused Multi-document Summarization with Paragraph Clustering and Sentence-type Filtering

Yohei Seki[†], Koji Eguchi[†, ††],and Noriko Kando[†, ††]
The Graduate University for Advanced Studies[†]
National Institute of Informatics[††]

NTCIR Workshop 4 Meeting June 2, 2004

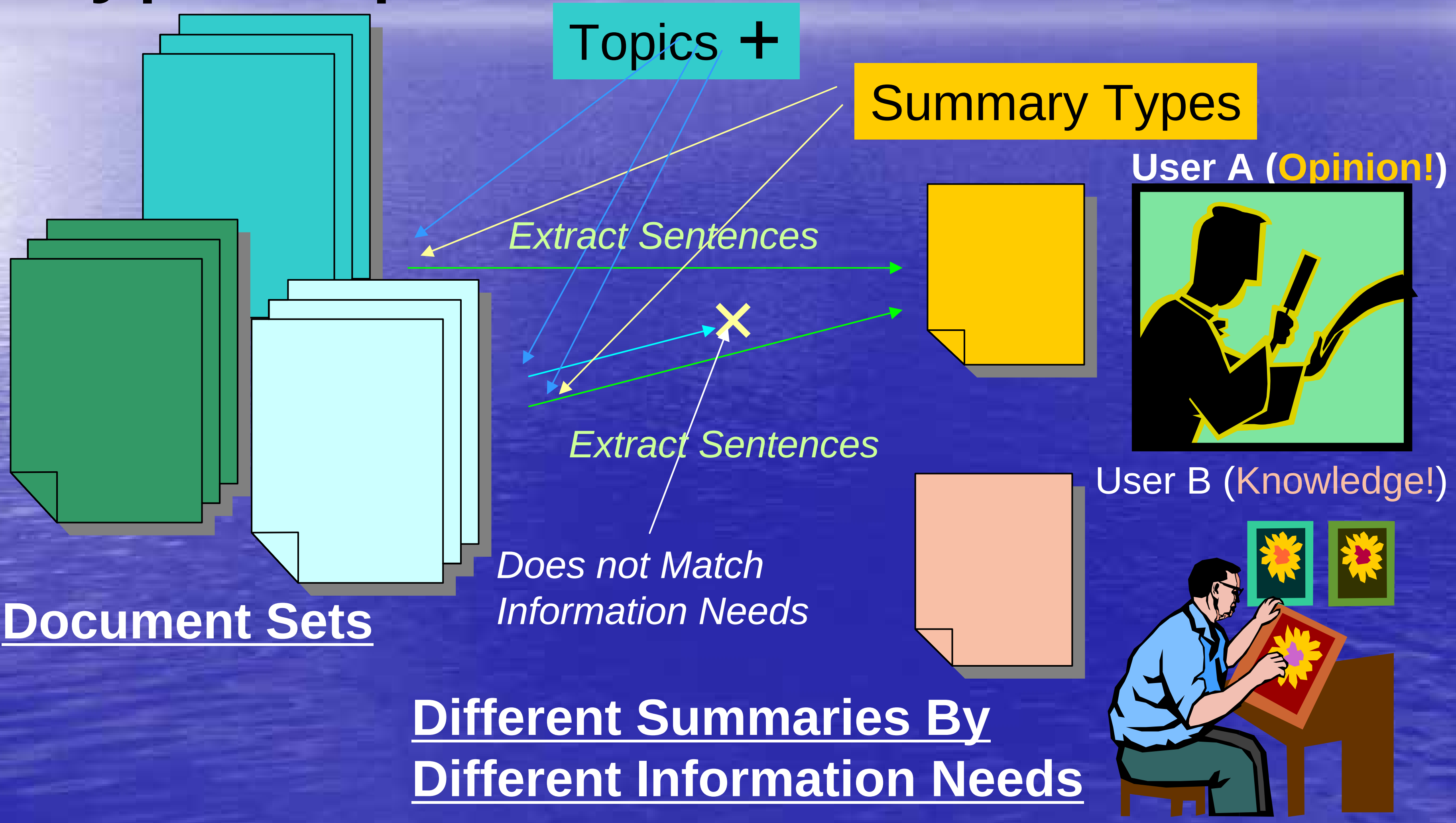
Talk Outline

1. Objective：User-focused Summarization
2. Analysis: Compare Paragraph Clustering-based Summarization Strategies
3. Proposal: Responsiveness Improvement with Sentence-type Filtering for each Cluster
4. Conclusions

Objective：
User-focused Summarization

- Two goals
  1. User-focused interactive summarization for topical requirements
     - Approach: Paragraph Clustering-based Summarization
  2. To produce knowledge-focused summaries (evaluate with question-answering responsiveness)
     - Approach: Sentence-type Filtering

Viewpoint（＝Topic＋Summary Type）-Specified Summarization



Topics ＋

Summary Types

User A (Opinion!)

Extract Sentences

Extract Sentences

User B (Knowledge!)

Does not Match
Information Needs

Document Sets

Different Summaries By
Different Information Needs

Multi-Document Summarization with *Document Clustering*

- "Document clustering techniques" partition a set of objects into clusters
- Closely associated documents tend to be relevant to the same request [cluster hypothesis]
- Extract one or two representative elements (sentences) from each cluster to produce summaries
- Topical Requirements: Select sentences from clusters in an order similar to queries

Talk Outline

1. Objective：User-focused Summarization
2. Analysis: Compare Paragraph Clustering-based Summarization Strategies
3. Proposal: Responsiveness Improvement with Sentence-type Filtering for each Cluster
4. Conclusions

Comparison: Paragraph Clustering-based Summarization Strategies

- Six clustering options
  1. Cluster units
  2. Features and Cluster Similarities
  3. Clustering algorithm
  4. Cluster size
  5. Sentence extraction clues
  6. Queries

1. Cluster Units: Paragraph

Related Work: Clustering for Summarization

- Stein et al. (1999): Cluster source documents by *single document summaries*
- M. Moens (2000): Cluster source documents by *paragraph* units
- Boros et al. (2001): Cluster source documents by *sentence* units

Our approach (interactive summarization)

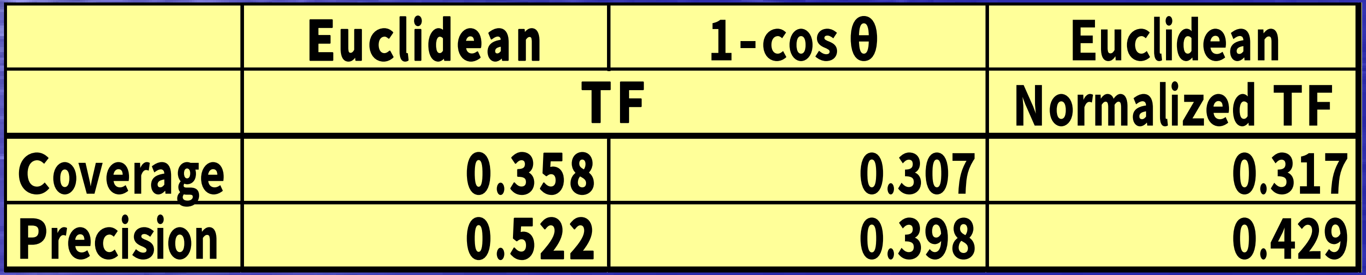
- Sentence features were too sparse to make feature vectors
- Document sizes were too small compared to summary sizes

⇒ Cluster source documents by *paragraph* units

2. Feature and Cluster Distance

Vector-length normalization does not work well for short documents (paragraphs in this research).

1. Feature vector
   - Normalized term frequency vs unnormalized (raw) term frequency
2. Cluster distance measure
   - Euclidean vs cosine



	Euclidean	1-cos θ	Euclidean
	TF		Normalized TF
Coverage	0.358	0.307	0.317
Precision	0.522	0.398	0.429

Unnormalized TF and Euclidean Distance performed well significantly

3. Cluster Algorithm: Ward’s Method

Compare three agglomerative clustering methods: complete-link, group-average, and Ward’s method

	Complete Link	Group Average	Ward's method
Coverage	0.358	0.314	0.364
Precision	0.522	0.499	0.518

The summary resultant with ``Ward’s method” performed better significantly than ``group average method”.

4. Cluster Size

Change cluster size according to number of sentences extracted

Cluster # for Long Summs	× 1	× 1.5	× 2
Cluster # for Short Summs	× 1.5	× 2	× 2.5
Coverage	0.364	0.357	0.353
Precision	0.518	0.543	0.565

Small cluster size performs better, but not significantly improved

5. Sentence Extraction Clues

Compare summarization with three sentence extraction clues:

Title	Yes	Yes	No	Yes
Term Frequency	Yes	Yes	Yes	No
Position	No	Yes	No	No
Coverage	0.339	0.322	0.338	0.315
Precision	0.614	0.606	0.613	0.623

Position weighting did not work well.
Title weighting effect was not clear.
Term Frequency performed well.

6. Queries

Compare cluster ordering using Queries and cluster ordering using Total Frequencies

	Cluster Ordering Similarity	
	to Queries	to Total Frequencies
Coverage	0.364	0.337
Precision	0.518	0.45

With queries, coverage improved 0.02 〜 0.03.

Talk Outline

1. Objective：User-focused Summarization
2. Analysis: Compare Paragraph Clustering-based Summarization Strategies
3. Proposal: Responsiveness Improvement with Sentence-type Filtering for each Cluster
4. Conclusions

Five Sentence-types to Improve User's Requirements

We annotate five sentence-types automatically.

Two Topical Types

- Main Description
- Elaboration

Three Functional Types

- Background
- Opinion
- Prospective

Sentence-type Filtering with Paragraph Clustering-based Summarization

1. The most heavily weighted sentence in each cluster was extracted.
2. For the second/third weighted sentence in each cluster, the sentence-type information was checked.
   - A) The redundancy of sentence-type for the most weighted sentence in the same cluster was checked.
   - B) If the sentence type was not redundant, we extracted it to produce summaries.

Analysis: Which sentence-type improved the responsiveness to Questions?

ID:L/S	Topic	Responsiveness		
		Base	Filtering	Type
310:L	Fossil in Ethiopia	0.2	0.3	Prospective
410:S	Nakata movement	0.273	0.364	Prospective
450:L	Company subsidary move	0.214	0.286	Prospective
510:S	Neutron	0.444	0.556	Prospective
560:L	Mistake in Entrance Examination	0.545	0.636	Prospective
570:S	Space Shuttle	0.308	0.385	Prospective
630:L	Ancient tomb	0.364	0.455	Opinion

``Prospective”-type improved responsiveness for event topics which described forecast in the future

Talk Outline

1. Objective：User-focused Summarization
2. Analysis: Compare Paragraph Clustering-based Summarization Strategies
3. Proposal: Responsiveness Improvement with Sentence-type Filtering for each Cluster
4. Conclusions

Conclusions

For NTCIR-4 TSC3, we focused on multi-document summarization from two different aspects:

1. Paragraph Clustering Techniques for Topical Information Requirements

- Compare Several Parameters:
- Ward's Methods, Unnormalized TF, Euclidean Distance
- Sentences ×1～×1.5 Cluster Size Performed Best

2. Sentence-type Filtering to Improve the Responsiveness to Questions

- To extract the most important sentence and ``Prospective''-type sentence from each cluster improved responsiveness for several topics

Thank you for your attention!