

Query Specific Summarization

Information Retrieval Track Group 1

Ashish Gupta, 3rd year UG, IIT Kanpur

Ankit Kumar, 3rd year UG , IIT Kanpur

Kamal Sahni, 3rd year UG, IIT Kanpur

Tarun Kr. Baranwal, 3rd year UG, IIT Kanpur

- **Research Problem**
- Motivation
- Earlier Works
- What people do and what we are doing
- Extracting Keywords from **semantic networks**
- Re-ranking of existing ranked sentences
- Evaluation and our Results.

Outline

- To summarize a single text document in accordance with the **query** specified by the user
- What are the important features of a text summarization system that extracts the words related to query from original documents?

Research Problem

- By just looking at the summary of a document, a user will be able to decide whether the document is of **interest to him/her** without looking at the whole document.
- Although a number of tools like MS AutoSum, Summarist etc. that are available to facilitate the text summarization process automatically, but the summarized text output is still **imprecise** or inaccurate.

Motivation

Author/ Year	Techniques
Luhn, 1958	Word Frequency, Statistical Approach
Baxendale, 1958	Text Positions
Edmunson, 1969	Cue Words and Heading
Miller, 1995	WordNet Lexical Terms
Lin and Hovy, 1997	Sentence Position
Marcu, 1998	Rhetorical Structural Theory
Daume & Marcu, 2002-04	Log Probability & Rhetorical Structural Theory
Kaustubh Patil, 2007	Graph Theory & Node centrality
Bawakid, 2008	Semantic similarity between user query & sentences
Liu, 2009	Correlation Matrix between user queries and sentence

Earlier Work

- **Features Used:**

Cue words, Heading words, Sentence Location, TF-IDF significance, Named Entities etc.

- Sentence **weighting** to **rank** sentence:

$$\mathbf{S_i} = w_1 * F_1 + w_2 * F_2 + w_3 * F_3 + \dots + w_n * F_n$$

- Drawback:

There may be sentences which are not statistically expected but more query oriented.

What people do

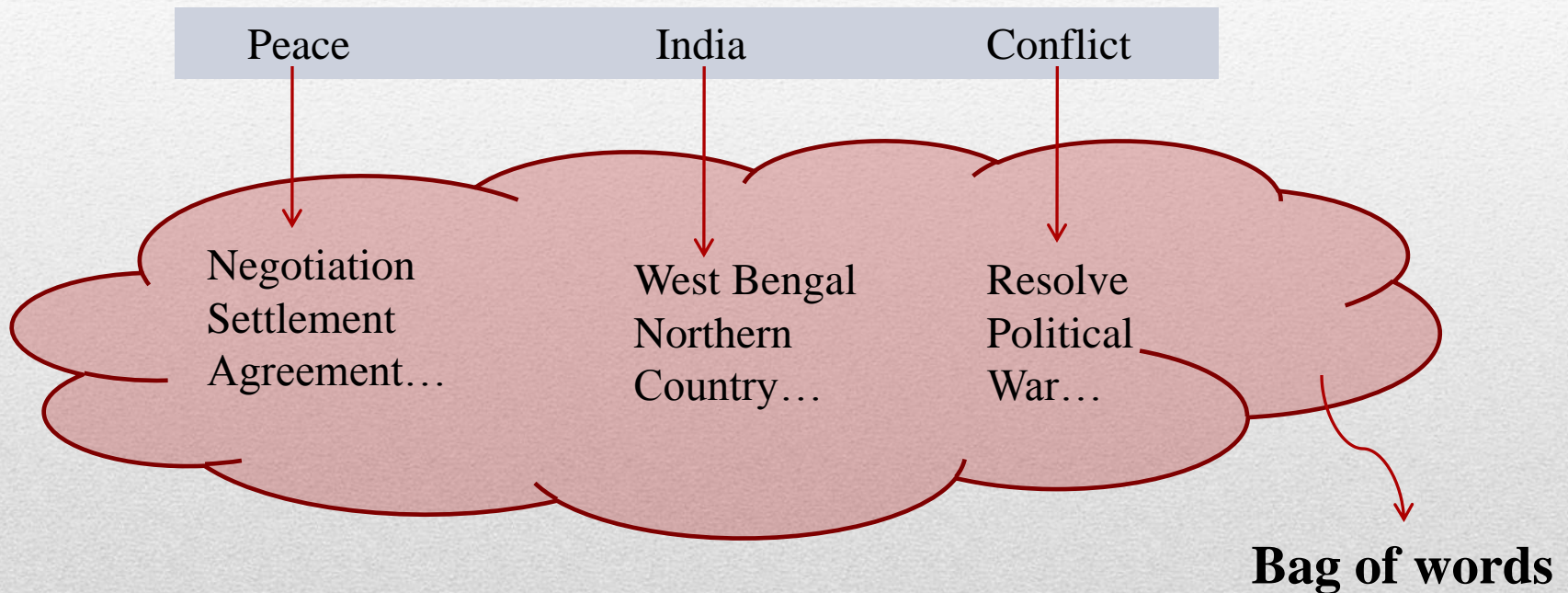
- Sentence weighting to rank sentence :

$$S_i = (w_1 * F_1 + w_2 * F_2 + w_3 * F_3 + \dots + w_n * F_n) + w_s * S_{fi}$$

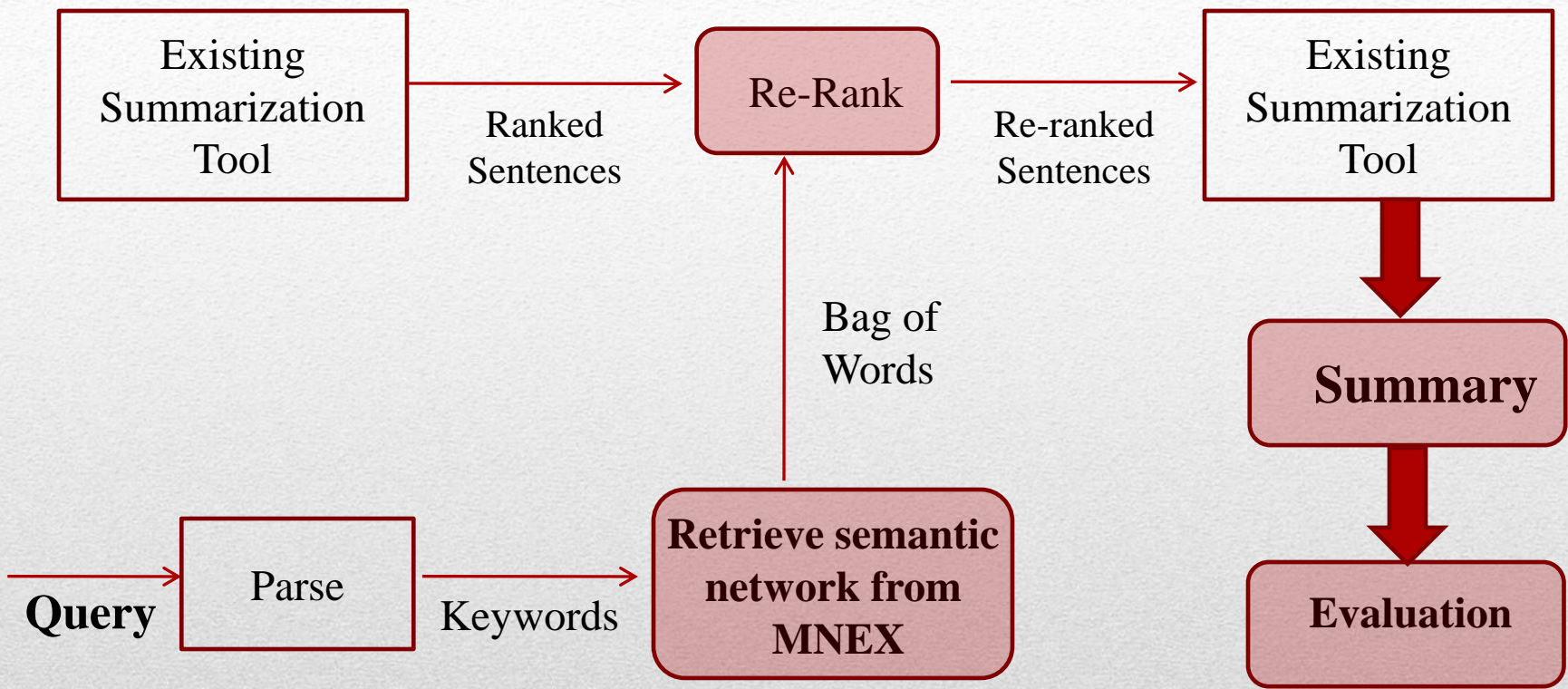
- S_{fi} are few **extra query based keywords** that we are introducing in existing model.
- These weights should be calculated simultaneously with **regression** models that we couldn't have done in given time line.
- Solution: **Re-ranking** of existing ranked sentences.

Our approach

- Query: **Efforts** made toward **peace** between **India** and **Pakistan** over **Kashmir conflict**.



Extra query based keywords



Methodology

- Initially **ranked** sentences are taken from an **existing** summarization tool.
- Re-ranking of sentences done as
 S_i = initial score + **extra score.**
- Re-ranked sentences were input back into the existing summarization tool to generate summary.

What is Re ranking ?

- Number of **word overlap** between sentence and semantic bag of words is found : **S_i**
- Method-1:
All words obtained from semantic network assigned **equal weighting**
$$S_i = (\text{initial score}) + W_s * (\text{no of words overlap})$$
- Method-2:
Summary generated with **Jaccard Indexing**:
$$S_i = (\text{initial score}) + W_s * (\text{no_of_words overlap}) / (\text{no_of_words in_sentence})$$

Extra Score ?

- Measures similarity between our generated summary and **gold set** summary
- **Gold set** Summary is available from TAC 2009 dataset
- Evaluation: comparison of Rouge-N Scores of **existing** query based summarization baseline model of (IIT-H) and **our** generated summary

Evaluation with ROUGE Scores

ROUGE-1			ROUGE-2			ROUGE-SU4		
R	P	F	R	P	F	R	P	F
Baseline Summaries								
0.3573	0.3551	0.3560	0.0827	0.0820	0.0823	0.1236	0.1229	.1232
Weight=0.02								
0.3557	0.3523	0.3539	0.0875	0.0870	0.0872	0.1238	0.1229	0.1233
Weight=0.06								
0.3278	0.3305	0.3290	0.0670	0.0678	0.0674	0.0603	0.1068	0.1062

R : Recall

P : Precision

F : F- measure

Rouge Scores w/o Jaccard Indexing

ROUGE-1			ROUGE-2			ROUGE-SU4		
R	P	F	R	P	F	R	P	F
Baseline Summary								
0.3573	0.3551	0.3560	0.0827	0.0820	0.0823	0.1236	0.1229	0.1232
Weight=0.1								
0.3577	0.3554	0.3564	0.0863	0.0854	0.0858	0.1238	0.1229	0.1232
Weight=1								
0.3493	0.3467	0.3479	0.0816	0.0809	0.0812	0.1204	0.1197	0.1200

R : Recall

P : Precision

F : F- measure

Rouge Scores with Jaccard Indexing

- These ROUGE scores do not appear to improve the existing system.
- Possible reasons
 - The Gold summary set for this dataset is 100 words only and is also an **abstractive** one
 - The Gold summary is **diverse** while our summary is more **query focused**. Hence, the low ROUGE scores.
- New evaluation technique required.

Conclusion with Rouge Scores

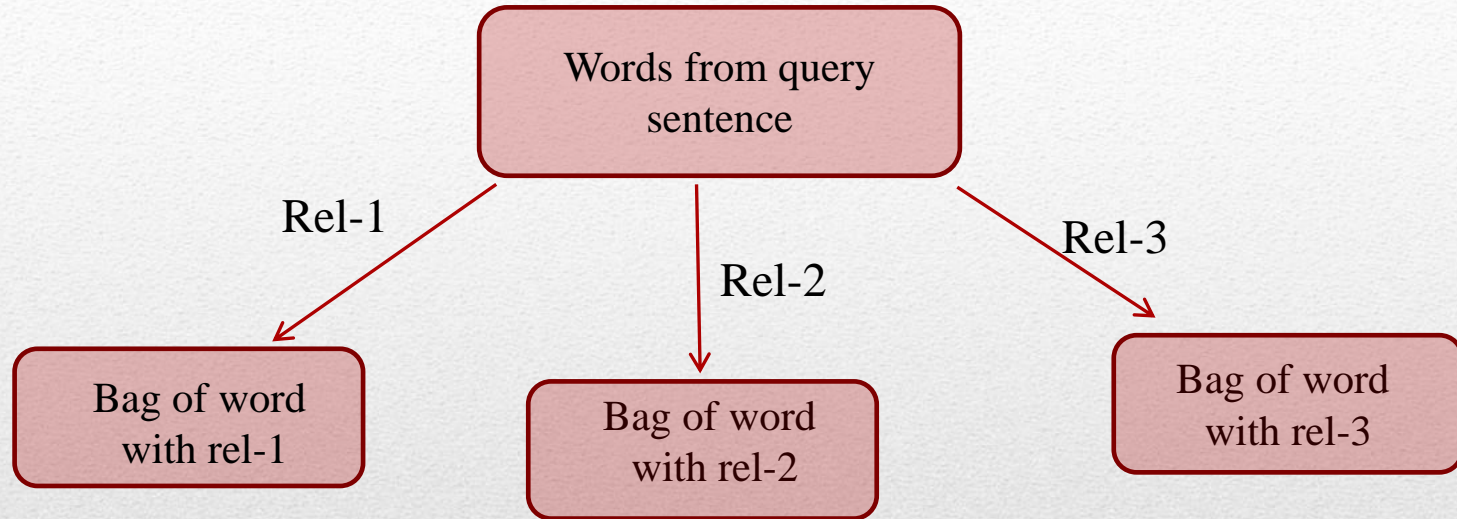
- Manually ranked sentences in binary (i.e. sentence can be relevant or irrelevant) are generated.
- Generated with **compress ratio** of 5% , 10%, and 20%
- This is used as reference summary sentences and we evaluated how close are our ranked sentences to this reference summary sentences.
- A correlation **Score** is calculated:
 - If (score > 0) : Our summary is better
 - If (score < 0) : Baseline summary is better

Evaluation with correlation

Document Id	Compress Ratio =5%	Compress Ratio =10%	Compress Ratio =20%
D0902A	21.1875	14.45161	6.966101
D0905A	2.84210	2.19512	-0.05405
D0901A	-1.23076	7.95	13.4705
D0910B	5.25	4.3333	3.5789
D0908B	-2.4285	-4.2142	5.2413
Average	5.124468	6.628846	5.8405502

Correlation score

Preferential Weighting with different relations



- This removes noise introduced in the system by semantic networks
- Expected to improve Re ranking of sentences.

Future work

Thanks!!!

Acknowledgement

- Dr. Carolyn Rose
- Dr. Vasudeva Varma
- Elijah Mayfield

Special thanks to..

- Rohit Bharadawaj, MS, IIIT-H
- Sudheer kovelamudi, MS, IIIT-H

Q/A



- Assigning different weights depending upon relation of word in semantic with query word.

$$S_i = (\text{initial weight}) + W_{r1} * (\text{no_of_Sfi_r1}) + W_{r2} * (\text{no_of_Sfi_r1}) + \dots \\ \dots + W_{rn} * (\text{no_of_Sfi_r1}).$$

Preferential weightings

- **Parsing** query sentences, **stemming** and deleting stop words from query sentence to generate additional keywords
- Retrieving **Semantic** network with **nodes** and **relations** for keywords obtained above using online tools viz. Microsoft Research: **MNEX**

Extracting Keywords

■ Semantic Networks:

- Using Microsoft Research: **MNEX**, the online MindNet explorer: [With_step-1\screenshot2.jpg](#)

- Bag of Keywords:

[With_step-1\bag of words kashmir.txt](#)

Getting extra keywords (sfi)

□ With_step-1\existing Sumarry_IND_PAK.docx

- With weights, $W_s = 0.1$, assigned to each word of our bag of words are quite higher, So summary will be **biased** to our bag of words

- This partial generated summary seems more query oriented

With_step-1\OurSumarry_IND_Pak_wt=0.1.docx

- With weights, $W_s = 0.02$, assigned to each word of our bag of word, we got summary i.e. more biased toward initial summarization.

- Very much similar to initial one that we got.

With_step-1\OurSumarry_IND_PAK_wt=0.02.docx

With different weighting...

- Cue words
- Heading words
- Sentence Location
- Sentence Length
- Presence of uppercase words
- TF-IDF significance of sentence
- Named Entities in sentence
- Dates in sentence
- Quotation marks in sentence
- Pronouns in sentence
- Numbers in sentence

General Features used to weight sentences