Text Summarisation for Evidence Based Medicine

Diego Mollá

Centre for Language Technology,
Macquarie University

IIT Patna, 16 December 2012
Contents

Evidence Based Medicine
  What is Evidence Based Medicine?
  EBM and NLP
  A Corpus for Summarisation

Text Summarisation
  Sentence Extraction
  Cohesion Check
  Balance and Coverage

Proposals for Text Summarisation
  Single-document Summarisation
  Optimisation and Summarisation
About us: Macquarie University
About us: Centre for Language Technology

http://www.clt.mq.edu.au

Core Staff (* involved in the AISRF project)

- Prof. Robert Dale
- * Prof. Mark Johnson
- * A. Prof. Mark Dras
- A. Prof. Steve Cassidy
- * Dr. Diego Molla-Aliod
- Dr. Rolf Schwitter
About Us: Research Group on Natural Language Processing of Medical Texts

http://web.science.mq.edu.au/~diego/medicalnlp/

Active Members

Diego Mollá  Senior lecturer at Macquarie University.
Abeed Sarker  PhD student at Macquarie University.
Sara Faisal Shash  Masters student.

Past Members

María Elena Santiago-Martínez  Research programmer.
Patrick Davis-Desmond  Masters student.
Andreea Tutos  Masters student.
About Me: Diego Mollá-Aliod

Some Highlights

- ExtrAns and WebExtrAns projects at University of Zurich.
- AnswerFinder project and Medical NLP research at Macquarie University.

Research interests

- Question Answering.
- Summarisation.
- Information Extraction.
Contents

Evidence Based Medicine
  What is Evidence Based Medicine?
  EBM and NLP
  A Corpus for Summarisation

Text Summarisation
  Sentence Extraction
  Cohesion Check
  Balance and Coverage

Proposals for Text Summarisation
  Single-document Summarisation
  Optimisation and Summarisation
Contents

Evidence Based Medicine
What is Evidence Based Medicine?
EBM and NLP
A Corpus for Summarisation

Text Summarisation
Sentence Extraction
Cohesion Check
Balance and Coverage

Proposals for Text Summarisation
Single-document Summarisation
Optimisation and Summarisation
Evidence Based Medicine

Suggested Steps in EBM

### PICO for Asking the Right Question

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient or Problem</strong></td>
<td><strong>Intervention</strong> (a cause, prognostic factor, treatment, etc.)</td>
<td><strong>Comparison Intervention</strong> (if necessary)</td>
<td><strong>Outcomes</strong></td>
</tr>
<tr>
<td><strong>Tips for Building</strong></td>
<td><strong>Ask “Which main intervention am I considering?”</strong></td>
<td><strong>Ask “What is the main alternative to compare with the intervention?”</strong></td>
<td><strong>Ask “What can I hope to accomplish?”</strong>, or <strong>“What could this exposure really affect?”</strong></td>
</tr>
<tr>
<td><strong>Example</strong></td>
<td>“Starting with your patient, ask “How would I describe a group of patients similar to mine?” Balance precision with brevity.”</td>
<td>Again, be specific</td>
<td>Again, be specific</td>
</tr>
<tr>
<td></td>
<td>“In patients with heart failure from dilated cardiomyopathy who are in sinus rhythm...”</td>
<td>“...would adding anticoagulation with warfarin to standard heart failure therapy...”</td>
<td>“...lead to lower mortality or morbidity from thromboembolism. Is this enough to be worth the increased risk of bleeding?”</td>
</tr>
</tbody>
</table>
Where to search for external evidence?

1. Evidence-based Summaries (Systematic Reviews):
   - The Cochrane Library (http://www.thecochranelibrary.com/).
   - EBM Online (http://ebm.bmj.com).
   - UptoDate (http://www.uptodate.com).
   - ...

EBM Summarisation

Diego Mollá
Where to search for external evidence?

1. Evidence-based Summaries (Systematic Reviews):
   - The Cochrane Library (http://www.thecochranelibrary.com/).
   - EBM Online (http://ebm.bmj.com).
   - UptoDate (http://www.uptodate.com).
   - . . .

2. Search the Medical Literature:
Searching Cochrane

Use the tabs below to filter your search results:

**Search**

**Cochrane Reviews**  Colloquia abstracts  Methodology  Newsletters  Policy Manual  Registered titles  Whole site

Enter your keywords:
sleep apnea

Search

Showing results 1 to 10 of 12

*Anti-inflammatory drugs for the treatment of obstructive sleep* ...

Skip to main content Cochrane Summaries ... Anti-inflammatory drugs for the treatment of obstructive sleep apnea in children. ...

summaries.cochrane.org/CD00707...uctive-sleep-apnea-in-children

*Pharmacotherapy for hypertension in adults with obstructive* ...

Skip to main content Cochrane Summaries Cochrane Summaries beta. ... Pharmacotherapy for hypertension in adults with obstructive sleep apnea. ...

summaries.cochrane.org/CD00766...s-with-obstructive-sleep-apnea

*Adenotonsillectomy for obstructive sleep apnoea in children* | ...

... The Cochrane Library ... Adenotonsillectomy for obstructive sleep apnoea in children. Have your say! Your ...
Searching PubMed

- Importance and management of chronic sleep apnoea in cardiology.
  - Jaffe LM, Kjekshus J, Gottlieb SS.
  - Eur Heart J. 2012 Mar 16. [Epub ahead of print]
  - PMID: 22427382 [PubMed - as supplied by publisher]
  - Related citations

- Association of Inflammatory biomarkers with sleep disorders in hemodialysis patients.
  - Razeghi E, Sahraian MA, Heidari R, Bagherzadeh M.
  - PMID: 22427289 [PubMed - in process]
  - Related citations

- Exploration of the relationship between sleep position and isolated tongue base or multilevel surgery in obstructive sleep apnea.
  - van Maanen JP, Ravesloot MJ, Witte BI, Grijseels M, de Vries N.
  - Eur Arch Otorhinolaryngol. 2012 Mar 20. [Epub ahead of print]
  - PMID: 22427104 [PubMed - as supplied by publisher]
  - Related citations

- Attention in children with obstructive sleep apnoea: An event-related potentials study.
  - Barnes ME, Gozal D, Molfese DL.
Searching the Trip Database

- **Continuous positive airway pressure delivery interfaces for obstructive sleep apnoea**
  - Cochrane Database of Systematic Reviews 2011
  - CPD/CME: Developing World
  - Related
  - Conclusion
  - Preview
  - DOI

- **Treatment of obstructive sleep apnoea for chronic cough in children**
  - Cochrane Database of Systematic Reviews 2011
  - CPD/CME: Developing World
  - Related
  - Conclusion
  - Preview
  - DOI

- **Anti-inflammatory medications for obstructive sleep apnoea in children**
  - Cochrane Database of Systematic Reviews 2011
  - CPD/CME: Developing World
  - Related
  - Conclusion
  - Preview
  - DOI

- **An integrated health-economic analysis of diagnostic and therapeutic strategies in the treatment of moderate-to-severe obstructive sleep apnoea**
  - NHS Economic Evaluation Database 2012
  - CPD/CME: Developing World
  - Related
  - Conclusion
  - Preview
  - DOI
## Appraising the Evidence

### The SORT Taxonomy

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level A</td>
<td>Consistent and good-quality patient-oriented evidence.</td>
</tr>
<tr>
<td>Level B</td>
<td>Inconsistent or limited-quality patient-oriented evidence.</td>
</tr>
<tr>
<td>Level C</td>
<td>Consensus, usual practise, opinion, disease-oriented evidence, or case series for studies of diagnosis, treatment, prevention, or screening.</td>
</tr>
</tbody>
</table>
Contents

Evidence Based Medicine
  What is Evidence Based Medicine?
  EBM and NLP
  A Corpus for Summarisation

Text Summarisation
  Sentence Extraction
  Cohesion Check
  Balance and Coverage

Proposals for Text Summarisation
  Single-document Summarisation
  Optimisation and Summarisation
Where can NLP Help?

- **Questions:**
  - Help formulate answerable questions.
  - From natural question to PICO frames?
  - Question analysis and classification.
Where can NLP Help?

Questions:
- Help formulate answerable questions.
- From natural question to PICO frames?
- Question analysis and classification.

Search:
- Retrieve and rank relevant literature.
- Extract the evidence-based information.
- Summarise the results.
Where can NLP Help? (II)

- **Appraisal**: Classify the evidence.
Contents

Evidence Based Medicine
  What is Evidence Based Medicine?
  EBM and NLP
  A Corpus for Summarisation

Text Summarisation
  Sentence Extraction
  Cohesion Check
  Balance and Coverage

Proposals for Text Summarisation
  Single-document Summarisation
  Optimisation and Summarisation
Where's the Corpus for Summarisation?

<table>
<thead>
<tr>
<th>Summarisation Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>▶ CENTRIFUSER/PERSIVAL: Developed and tested using user feedback (iterative design).</td>
</tr>
<tr>
<td>▶ Demner-Fushman &amp; Lin: ROUGE on original paper abstracts.</td>
</tr>
<tr>
<td>▶ Fiszman: Factoid-based evaluation.</td>
</tr>
</tbody>
</table>
Where’s the Corpus for Summarisation?

## Summarisation Systems

- CENTRIFUSER/PERSIVAL: Developed and tested using user feedback (iterative design).
- Demner-Fushman & Lin: ROUGE on original paper abstracts.
- Fiszman: Factoid-based evaluation.

## Corpora

- Several corpora of questions/answers available.
- Answers lack explicit pointers to primary literature.
- Medical doctors want to know the primary sources.
Which treatments work best for hemorrhoids?

Evidence-based answer

Excision is the most effective treatment for thrombosed external hemorrhoids [strength of recommendation (SOR): A, systematic review]. Of nonoperative techniques, rubber band ligation produces the lowest rate of recurrence [SOR: A, systematic review].

Evidence summary

External hemorrhoids originate below the dentate line and become acutely painful with thrombosis. They can cause perianal pruritus and excoriation because of interference with perianal hygiene. Internal hemorrhoids become symptomatic when they bleed or prolapse (Table).

For thrombosed external hemorrhoids, surgery works best

Few studies have evaluated the best treatment for thrombosed external hemorrhoids. A retrospective study of 231 patients treated conservatively or surgically found that the 48.5% of patients treated surgically had a lower recurrence rate than the conservative group (number needed to treat [NNT]=2 for recurrence at mean follow-up of 7.6 months) and earlier resolution of symptoms (average 3.9 days compared with 24 days for conservative treatment).

Another retrospective analysis of 340 patients who underwent outpatient excision of thrombosed external hemorrhoids under local anesthesia reported a low recurrence rate of 6.5% at a mean follow-up of 17.3 months.

A prospective, randomized controlled trial (RCT) of 98 patients treated nonsurgically found improved pain relief with a combination of topical nifedipine 0.3% and lidocaine 1.5% compared with lidocaine alone. The NNT for complete pain relief at 7 days was 3.

Conventional hemorrhoidectomy beats stapling

Many studies have evaluated the best treatment for prolapsed hemorrhoids. A Cochrane systematic review of 12 RCTs that compared conventional hemorrhoidectomy with stapled hemorrhoidectomy in patients with grade I to III hemorrhoids found a lower rate of recurrence (follow-up ranged from 6 to 39 months) in patients who had conventional hemorrhoidectomy [NNT=14]. Conventional hemorrhoidectomy showed a nonsignificant trend in decreased bleeding and decreased incontinence.

A second systematic review of 25 studies, including some that were of lower quality, showed a higher recurrence rate at 1 year with stapled hemorrhoidectomy than with conventional surgery.

Nonoperative techniques?

Consider rubber band ligation

A systematic review of 3 poor-quality trials comparing rubber band ligation with excisional hemorrhoidectomy in patients with grade III hemorrhoids found that excisional hemorrhoidectomy produced better long-term symptom control but more immediate postoperative complications of anal stenosis and hemorrhage. Rubber band ligation had the lowest recurrence rate at 12 months compared with the other nonoperative techniques of sclerotherapy and infrared coagulation.

Fiber supplements help relieve symptoms

A Cochrane systematic review of 7 RCTs enrolling a total of 378 patients with grade I to III hemorrhoids evaluated the effect of fiber supplements on pain, itching, and bleeding. Persistent hemorrhoid symptoms decreased by 53% in the group receiving fiber.

When surgical hemorrhoidectomy is recommended

The American Society of Colon and Rectal Surgeons recommends adequate fluid and fiber intake for all patients with symptomatic hemorrhoids. For grade I to III hemorrhoids, the society states that banding is usually most effective. When office treatments fail, the society recommends surgical hemorrhoidectomy (SOR: B).

The society recommends excision of thrombosed hemorrhoids less than 72 hours old and expectant treatment with hemorrhoids that present early. Surgical hemorrhoidectomy should be reserved for when conservative treatment fails and for patients with symptomatic grade III and IV hemorrhoids.

References

The XML Contents

<record id="7843">
<url>http://www.jfponline.com/Pages.asp?AID=7843&amp;issue=September_2009&amp;UID=</url>
<question>Which treatments work best for hemorrhoids?</question>
<answer>
    <snip id="1">
        <snip text>Excision is the most effective treatment for thrombosed external hemorrhoids.</snip>
        <sor type="B">retrospective studies</sor>
    </snip>
    <long id="1_1">
        <longtext>A retrospective study of 231 patients treated conservatively or surgically found that the 48.5% of patients treated surgically had a lower recurrence rate than the conservative group (number needed to treat [NNT]=2 for recurrence at mean follow-up of 7.6 months) and earlier resolution of symptoms (average 3.9 days compared with 24 days for conservative treatment).</longtext>
    </long>
</record>
<long id="1_2">
    <longtext>A retrospective analysis of 340 patients who underwent outpatient excision of thrombosed external hemorrhoids under local anesthesia reported a low recurrence rate of 6.5% at a
The XML Contents II

mean follow-up of 17.3 months.<\longtext>

A prospective, randomized controlled trial (RCT) of 98 patients treated nonsurgically found improved pain relief with a combination of topical nifedipine 0.3% and lidocaine 1.5% compared with lidocaine alone. The NNT for complete pain relief at 7 days was 3.<\longtext>
Components of the Corpus

**Question**  Direct extract from the source.

**Answer**  Split from the source and manually checked.

**Evidence**  Extracted from the source.

**Additional text**  Manually extracted from the source and massaged.

**References**  PMID looked up in PubMed (automatic and manual procedure).
Corpus Statistics

Size

- 456 questions ("records").
- 1,396 answer parts ("snips").
- 3,036 answer justifications ("longs").
- 3,705 references:
  - 2,908 unique references.
  - 2,657 XML abstracts from PubMed.
Answer parts per Question

Snips per Record

<table>
<thead>
<tr>
<th>Number</th>
<th>Snips</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

Avg=3.06
Answer justifications per answer part

Avg = 2.17
References per answer justification

Avg = 1.22
References per question

Avg = 6.57

References per Record

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
0 17.5 35.0 52.5 70.0
Evidence Grade

![SOR Types Chart]

- A
- B
- C
- D
- No type

number of SORs
References

![Diagram showing the number of references in different categories: unique, repeated, XML, no XML, no abstract text. The y-axis represents the categories, and the x-axis represents the number of references, ranging from 0 to 3000.]
Contents

Evidence Based Medicine
  What is Evidence Based Medicine?
  EBM and NLP
  A Corpus for Summarisation

Text Summarisation
  Sentence Extraction
  Cohesion Check
  Balance and Coverage

Proposals for Text Summarisation
  Single-document Summarisation
  Optimisation and Summarisation
What is Document Summarisation

Summarisation (or automatic abstracting)

A summary is a text that is produced from one or more texts, that contains a significant portion of the information of the original text(s), and that is no longer than half of the original text(s).

(Hovy, 2003)
What is Document Summarisation Good For?

What for?

- For busy people to read the summary instead of the full text.
What is Document Summarisation Good For?

What for?

▶ For busy people to read the summary instead of the full text. → informative summary
What is Document Summarisation Good For?

What for?

- For busy people to read the summary instead of the full text. → informative summary
- For researchers, web surfers, ... to read the summary to decide if it is worth to read the original text.
What is Document Summarisation Good For?

What for?

- For busy people to read the summary instead of the full text. → informative summary
- For researchers, web surfers, ... to read the summary to decide if it is worth to read the original text. → indicative summary
What is Document Summarisation Good For?

What for?

- For busy people to read the summary instead of the full text. → informative summary
- For researchers, web surfers, ... to read the summary to decide if it is worth to read the original text. → indicative summary
- To avoid having to type out an abstract for a technical report when the publisher requests it.
What is Document Summarisation Good For?

What for?

- For busy people to read the summary instead of the full text. → informative summary

- For researchers, web surfers, ... to read the summary to decide if it is worth to read the original text. → indicative summary

- To avoid having to type out an abstract for a technical report when the publisher requests it. → not realistic
Abraham Lincoln was born in Kentucky on February 12th, 1809. His family moved to Indiana when he was eight years old. His mother died when he was ten. Abraham Lincoln loved to read. He only went to school for a few months. He had to study and learn on his own. From his childhood, Abraham Lincoln was known as a hard worker. He worked on a farm and in a store and on a boat. He studied law and practiced for many years in Illinois. He married Mary Todd Lincoln. Together they had four sons. Abraham Lincoln was elected as the 16th President of the United States in 1860. He did many things as President. Many people think he was the best American President of all time. He is most remembered for freeing the slaves. He was President of the United States during the time the Civil War was fought. The Civil War was fought between the Northern and Southern states. He was known as a great speaker. Some of his most famous speeches include the Emancipation Proclamation - in which he freed the slaves - and the Gettysburg Address, which starts ”Four score and seven years ago ...” While attending a play on April 14th, 1865, Abraham Lincoln was shot by John Wilkes Booth. Lincoln died the next morning. Lincoln’s birthday is celebrated in February of each year.
Examples II

Word’s Autosummarize

Abraham Lincoln was born in Kentucky on February 12th, 1809.
Abraham Lincoln loved to read.
Abraham Lincoln was elected as the 16th President of the United States in 1860.
The Civil War was fought between the Northern and Southern states.
While attending a play on April 14th, 1865, Abraham Lincoln was shot by John Wilkes Booth.
Lincoln died the next morning.

MEAD

Abraham Lincoln was born in Kentucky on February 12th 1809.
His family moved to Indiana when he was eight years old.
From his childhood Abraham Lincoln was known as a hard worker.
Abraham Lincoln was elected as the 16th President of the United States in 1860.
While attending a play on April 14th 1865 Abraham Lincoln was shot by John Wilkes Booth.
An Ideal Document Summarisation System

Understanding Stage

Document(s) $\rightarrow$ Knowledge base
An Ideal Document Summarisation System

Understanding Stage

Document(s) → Knowledge base

Generation Stage

Knowledge base → Summary
A Compromise Solution

**Sentence Extraction**

Document $\rightarrow$ Sentence candidates
A Compromise Solution

Sentence Extraction
Document → Sentence candidates

Cohesion Check
Sentence candidates → Coherent text
A Compromise Solution

Sentence Extraction

Document → Sentence candidates

Cohesion Check

Sentence candidates → Coherent text

Balance and Coverage

Coherent text → Summary
A Compromise Solution

Sentence Extraction
Document \(\rightarrow\) Sentence candidates
This is what most commercial and free summarisers do

Cohesion Check
Sentence candidates \(\rightarrow\) Coherent text

Balance and Coverage
Coherent text \(\rightarrow\) Summary
Contents

Evidence Based Medicine
  What is Evidence Based Medicine?
  EBM and NLP
  A Corpus for Summarisation

Text Summarisation
  Sentence Extraction
  Cohesion Check
  Balance and Coverage

Proposals for Text Summarisation
  Single-document Summarisation
  Optimisation and Summarisation
General Approach

For each sentence . . .

1. Look for clues to its importance.
2. Compute a score for the sentence based on the clues found.
3. Select all sentences whose scores exceed some threshold.
   ▶ Or select the highest scoring sentences up to a certain total.
The Frequency-keyword Approach

1. Compute the keywords of the document:
   - Ignore the function words by using a stop word list.
   - Sort all remaining words according to frequency or measures such as \( tf.idf \) (next slide).
   - Select the top words (say, the top 5%).
The Frequency-keyword Approach

1. Compute the **keywords** of the document:
   - Ignore the function words by using a stop word list.
   - Sort all remaining words according to frequency or measures such as $tf.idf$ (next slide).
   - Select the top words (say, the top 5%).

2. Score the document sentences according to the presence of keywords:
   - Simple keyword count.
   - Weighted keyword count (keyword weights for each sentence).
   - Looking for keyword clusters in the sentence.
Finding Most Informative Sentences

\(tf \cdot idf\) to find keywords

- **Term Frequency (tf):** Words that are very frequent in a document are more “important”.

  \[ tf(w) = \# \text{ times word } w \text{ is in document} \]

- **Inverse Document Frequency (idf):** Words that appear in many documents are less “important”.

  \[ idf(w) = \log \frac{\# \text{ documents}}{\# \text{ documents that contain word } w} \]
The Biased Keyword Approach

Title and headings biased

Compute a list of keywords on the basis of document structure:

- select candidates from titles and headings only, or
- candidates from titles and headings have more importance:
  - e.g. they are counted as being more frequent.
The Biased Keyword Approach

Title and headings biased

Compute a list of keywords on the basis of document structure:

▶ select candidates from titles and headings only, or
▶ candidates from titles and headings have more importance:
  ▶ e.g. they are counted as being more frequent.

Query biased (customised summaries)

Use the user’s query to determine the keyword’s weights:

▶ the user’s query determines all the keywords, or
▶ the user’s query introduces additional keywords or updates the weights of existing keywords.
The Location Method

Observation

First and last sentence of a paragraph are usually most central to the theme of a text.

Increase the score of a sentence according to its position in the paragraph:

- Beginning of paragraph.
- End of paragraph.
Cues, Indicator Phrases I

Cues

▶ Certain words (not necessarily keywords) provide an indication of the importance of the sentence.
▶ Use these words to determine the sentence score:
  ▶ **bonus words** increase the sentence score:
    ▶ “greatest”, “significant”
  ▶ **stigma words** decrease the sentence score:
    ▶ “hardly”, “impossible”, “now”
Cues, Indicator Phrases II

Indicator Phrases

Indicator phrases are specific phrases or patterns of phrases that can be used to determine the sentence importance:

- “The main aim of the present paper is . . .”
- “The purpose of this article is . . .”
- “In this report, we outline . . .”
- “Our investigation has shown that . . .”
Relational Criteria

1. Build a semantic structure for the document:
   - sentences are vertices
   - inter-sentence links are edges
     - Rhetorical links (elaboration, sequence, etc)
     - Cooccurrence of keywords
     - ...

2. Use the link structure to determine the most important sentences
   - Degree of the vertex
   - Eigenvalues (PageRank style)
Contents

Evidence Based Medicine
What is Evidence Based Medicine?
EBM and NLP
A Corpus for Summarisation

Text Summarisation
Sentence Extraction
Cohesion Check
Balance and Coverage

Proposals for Text Summarisation
Single-document Summarisation
Optimisation and Summarisation
Textual Cohesion

- Lack of cohesion results in “odd” extracts.
- Sentences include references to other sentences:
  - Anaphoric reference:
    - “John saw Mary. She was talking over the phone”
  - Rhetorical connectives:
    - “So, the following example . . .”
  - Lexical or definite reference:
    - “I saw a man with a book. The book was . . .”
- Possible solutions:
  - Aggregation: Add preceding sentences until there are no external references.
  - Deletion: Remove the difficult sentences.
  - Modification: Alter the sentences to eliminate or disguise the problem.
Contents

Evidence Based Medicine
  What is Evidence Based Medicine?
  EBM and NLP
  A Corpus for Summarisation

Text Summarisation
  Sentence Extraction
  Cohesion Check
  Balance and Coverage

Proposals for Text Summarisation
  Single-document Summarisation
  Optimisation and Summarisation
Balance and Coverage

- We need to process the selected sentences in order to produce a real abstract:
  - Delete redundant sentences.
  - Harmonise tense and voice of verbs.
  - Ensure balance and proper coverage.

- Combination of information extraction and text generation.

- Need to consider text structure:
  - Each sentence plays a role in the text and in relation with the other sentences.

- Problem to address:
  - Lack of balance and coverage:
    - Missing important information.
    - Too much emphasis on less important information.
Contents

Evidence Based Medicine
  What is Evidence Based Medicine?
  EBM and NLP
  A Corpus for Summarisation

Text Summarisation
  Sentence Extraction
  Cohesion Check
  Balance and Coverage

Proposals for Text Summarisation
  Single-document Summarisation
  Optimisation and Summarisation
Contents

Evidence Based Medicine
  What is Evidence Based Medicine?
  EBM and NLP
  A Corpus for Summarisation

Text Summarisation
  Sentence Extraction
  Cohesion Check
  Balance and Coverage

Proposals for Text Summarisation
  Single-document Summarisation
  Optimisation and Summarisation
Single-document Summarisation

**Input**

- Question.
- Document Abstract.

**Output**

- Extractive summary that answers the question.
- Target summary is the annotated answer justification ("long").
- Evaluated using ROUGE-L with Stemming.
General Approach (Sarker et al., CBMS 2012)

**In a Nutshell**

1. Gather statistics from the best 3-sentence extracts.
   - Exhaustive search to find these best extracts.

2. Build **three classifiers**, one per sentence in the final extract.
   - Classifier 1 based on statistics from best 1st sentence.
   - Classifier 2 based on statistics from best 2nd sentence.
   - Classifier 3 based on statistics from best 3rd sentence.
The Statistics Gathered

1. Source sentence position.
2. Sentence length.
3. Sentence similarity.
4. Sentence type.
1. Source Sentence Position

- Compute relative positions (0 \ldots 1).
- Create normalised frequency histograms $f_1, f_2, \ldots, f_{10}$.
- Score every relative position in bin $i$ with its bin frequency: $S_{pos}(i) = f_{bin(i)}$. 
2. Sentence Length

Reward larger sentences and penalise shorter sentences:

Normalised sentence length

\[ S_{len}(i) = \frac{l_s - l_{avg}}{l_d} \]

- \( l_s \): sentence length
- \( l_{avg} \): average sentence length in the corpus
- \( l_d \): document length
3. Sentence Similarity

**Sentence Similarity**

- Lowercase, stem, remove stop words.
- Build vector of \( tf.idf \) with remaining words and UMLS semantic types.
- \( \text{CosSim}(X, Y) = \frac{X \cdot Y}{||X|| \cdot ||Y||} \)

**Maximal Marginal Relevance (Carbonell & Goldstein, 1998)**

Reward sentences similar to the query and penalise those similar to other summary sentences.

\[
\text{MMR} = \lambda(\text{CosSim}(S_i, Q)) - (1 - \lambda) \max_{S_j \in S}(\text{CosSim}(S_i, S_j))
\]
4. PIBOSO (Kim et al. 2011) I

1. Classify all sentences into PIBOSO types (a variant of PICO).
2. Generate normalised frequency histograms of resulting PIBOSO types.
4. PIBOSO (Kim et al. 2011) II

Position independent

$$S_{PIPS}(i) = \frac{P_{best}}{P_{all}}$$

**Position dependent**

$$S_{PDPS}(i) = \frac{P_{pos}}{P_{best}}$$

- **$P_{best}$**: proportion of this PIBOSO type among all best summary sentences.
- **$P_{all}$**: proportion of this PIBOSO type among all sentences.
- **$P_{pos}$**: proportion of this PIBOSO type among all best summary sentences at this position.
Classification

Edmunsonian Formula

\[ S_{Si} = \alpha S_{rpos_i} + \beta S_{len_i} + \gamma S_{PIPS_i} \\
+ \delta S_{PDPS_i} + \epsilon S_{MMR_i} \]

- MMR is replaced with cosine similarity for first sentence.
- In case of ties, the sentence with greatest length is chosen.
- Parameters are fine-tuned through exhaustive search (grid search) using training set.

\[ \alpha = 1.0, \beta = 0.8, \gamma = 0.1, \delta = 0.8, \epsilon = 0.1, \lambda = 0.1. \]
Percentile-based Evaluation (Ceylan et al. 2010) I

We compare against all possible 3-sentence extracts in the test set.

1. Bin all possible three-sentence combinations of each abstract.
   ▶ 1,000 bins.

2. Normalise the resulting histograms.

3. Combine all histograms.
   ▶ convolution.

4. The result approximates the probability density distribution of all three-sentence summaries in all abstracts.
Percentile-based Evaluation (Ceylan et al. 2010) II
Systems

L3  Last three sentences.
O3  Last three PIBOSO outcome sentences.
   Random.
O   All outcome sentences.
PI  Sentence position independent.

PD  Sentence position dependent (our proposal).
## Results

<table>
<thead>
<tr>
<th>System</th>
<th>F-Score</th>
<th>95% CI</th>
<th>Percentile (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L3</td>
<td>0.159</td>
<td>0.155–0.163</td>
<td>60.3</td>
</tr>
<tr>
<td>O3</td>
<td>0.161</td>
<td>0.158–0.165</td>
<td>77.5</td>
</tr>
<tr>
<td>R</td>
<td>0.158</td>
<td>0.154–0.161</td>
<td>50.3</td>
</tr>
<tr>
<td>O</td>
<td>0.159</td>
<td>0.155–0.164</td>
<td>60.3</td>
</tr>
<tr>
<td>PI</td>
<td>0.160</td>
<td>0.157–0.164</td>
<td>69.4</td>
</tr>
<tr>
<td>PD</td>
<td>0.166</td>
<td>0.162–0.170</td>
<td>97.3</td>
</tr>
</tbody>
</table>
Towards Multi-document Summarisation

- Evidence suggests that a two-step process is promising (Sarker et al., ALTA 2012).
  1. Single document summarisation.
  2. Multi-document summarisation from the single-document summaries.

- Traditional clustering techniques seem to produce good clustering of references (Shash & Molla, unpublished).

- We are still looking at means to obtain the answer parts.
  - Topics as cluster centroids.
  - Overlap with the question.
Contents

Evidence Based Medicine
  What is Evidence Based Medicine?
  EBM and NLP
  A Corpus for Summarisation

Text Summarisation
  Sentence Extraction
  Cohesion Check
  Balance and Coverage

Proposals for Text Summarisation
  Single-document Summarisation
  Optimisation and Summarisation
Many NLP tasks are based on optimisation

- **Text classification**: minimise the classification error.
- **Part of speech tagging**: Find the optimal sequence of labels.
- **Parsing**: Find the most likely parse.
- **Machine translation**: Dual optimisation.
  - The target sentence must keep the most meaning.
  - The target sentence must try to follow the language model of the target language.
Many ML tasks are about optimising parameters

\[
\arg\min_{\theta} J(h_\theta(X), Y)
\]

- \(J\) Cost (error) function.
- \(\theta\) Machine Learning parameters.
- \(h_\theta(X)\) Hypothesis function.
- \(X\) Inputs.
- \(Y\) Observed results.
Optimisation in Classification Tasks

Hypothesis function in Logistic Regression

Sigmoid (logistic) function

\[ h_{\theta}(X) = \frac{1}{1 + e^{-\sum_{i} \theta_i x_i}} = \frac{1}{1 + e^{-\theta^T X}} \]

Cost function in Logistic Regression

Cross entropy

\[ J(h_{\theta}(X), Y) = -\frac{1}{m} \left[ \sum_{i=1}^{m} y^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)})) \right] \]
Query Based Summarisation I

A summary sentence $S_i$ must maximise its similarity with the question

$$\arg \max_{S} \sum_{i} \cos \text{Sim}(S_i, Q)$$

A summary sentence $S_i$ must minimise its similarity with other summary sentences

$$\arg \min_{S} \sum_{i,j} \cos \text{Sim}(S_i, S_j)$$
Query Based Summarisation II

Maximal Marginal Relevance (Carbonell & Goldstein, 1998)

Greedy approach

- Each iteration, select the sentence $S_i$ with highest MMR score.

$$\text{MMR} = \lambda \left( \text{CosSim}(S_i, Q) \right) - (1 - \lambda) \max_{S_j \in S} \left( \text{CosSim}(S_i, S_j) \right)$$
Single-document Summarisation

Information contents in summary sentences must be maximal

- The sum of all weighted concepts in a summary must be maximal.

\[ \arg \max_S \sum_{c \in C_S} w(c) \]

- This is the knapsack problem (NP-hard).

Readability in summary sentences must be maximal

- We’ve seen some aspects of readability above . . .
- . . . now we need to express them as a problem of optimisation.
Query-based Multi-document Abstractive Summarisation

The summary must fit the question topics best

1. Fitting to cluster centroids.
2. Topic modelling (LDA) variants.

The summary must be most readable

Find the word extracts that are most likely produced by language models:

1. Word sequences (e.g. 2-grams).
2. Best likely parse (e.g. the $k$-minimum spanning tree in a graph).
Latent Dirichlet Allocation

- $\Pi_d$: topic probability distribution for document $d$.
- $Z_{dn}$: actual topic selected for word $n$ in document $d$.
- $\theta_k$: word probability distribution for topic $k$.
- $\alpha, \beta$: hyperparameters.
Possible LDA Variant for Query-focused Summarisation
Evidence Based Medicine (EBM) is an important problem that medical doctors face.

EBM can benefit from Natural Language Processing (NLP) in general, and text summarisation in particular.

Text summarisation, like many NLP tasks, relies on optimisation.

We need expertise on optimisation techniques!
Evidence Based Medicine (EBM) is an important problem that medical doctors face.

EBM can benefit from Natural Language Processing (NLP) in general, and text summarisation in particular.

Text summarisation, like many NLP tasks, relies on optimisation.

We need expertise on optimisation techniques!

Questions?

Further information about our research:
http://web.science.mq.edu.au/~diego/medicalnlp/