Statistical NLP
Spring 2010

Lecture 22: Summarization

Dan Klein – UC Berkeley

Includes slides from Aria Haghighi, Dan Gillick

Sentence Extraction

Selection

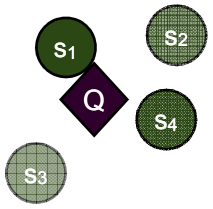
mid-'90s

present

- Maximum Marginal Relevance

  [Carbonell and Goldstein, 1998]

Greedy search over sentences



Minimize redundancy

Maximize similarity to the query

$$MMR \quad \underset{D_i \in R}{\operatorname{argmax}} \left(\lambda [Sim_1(D_i, Q)] \quad (1 \quad \lambda) \max_{\ldots} [Sim_2(D_i, D_j)] \right)$$

Selection

mid-'90s

present

- Maximum Marginal Relevance
- Graph algorithms [Mihalcea 05++]

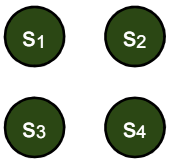
Selection

mid-'90s

present

- Maximum Marginal Relevance
- Graph algorithms

Nodes are sentences



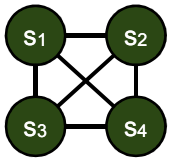
Selection

mid-'90s

present

- Maximum Marginal Relevance
- Graph algorithms

Nodes are sentences



Edges are similarities

Selection

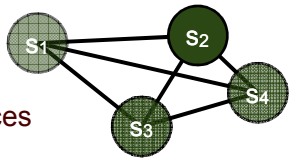
mid-'90s

present

- Maximum Marginal Relevance
- Graph algorithms

Stationary distribution
represents node centrality

Nodes are sentences



Edges are similarities

Selection

mid-'90s

present

- Maximum Marginal Relevance
- Graph algorithms
- Word distribution models

w	$P_D(w)$
Obama	0.017
speech	0.024
health	0.009
Montana	0.002

Input document distribution

~

w	$P_A(w)$
Obama	?
speech	?
health	?
Montana	?

Summary distribution

Selection

mid-'90s

present

- Maximum Marginal Relevance
- Graph algorithms
- Word distribution models

SumBasic [Nenkova and Vanderwende, 2005]

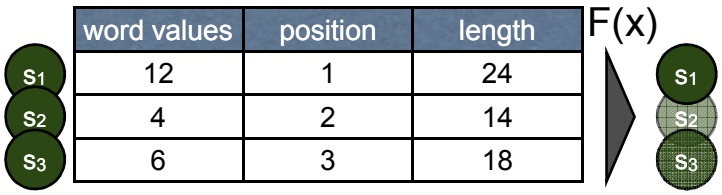
```
Value(wi) = PD(wi)
Value(si) = sum of its word values
Choose si with largest value
Adjust PD(w)
Repeat until length constraint
```

Selection

mid-'90s

present

- Maximum Marginal Relevance
- Graph algorithms
- Word distribution models
- Regression models



frequency is just one of many features

Selection

mid-'90s

present

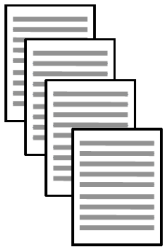
- Maximum Marginal Relevance
- Graph algorithms
- Word distribution models
- Regression models
- Topic model-based

  [Haghighi and Vanderwende, 2009]

Summarization Criterion

$P_{\mathcal{C}}(\cdot)$

Barack Obama: 0.15
Serve America Act: 0.13
signed: 0.12



$P_{\mathbf{S}}(\cdot)$

Barack Obama: 0.18
Serve America Act: 0.16
signed: 0.10

Summarization Criterion

$$\mathbf{S}^* = \min_{\mathbf{S}:words(\mathbf{S})\leq L} KL(P_{\mathcal{C}} \| P_{\mathbf{S}})$$

$P_{\mathcal{C}}(\cdot)$

Barack Obama: 0.15
Serve America Act: 0.13
signed: 0.12

$P_{\mathbf{S}}(\cdot)$

Barack Obama: 0.18
Serve America Act: 0.16
signed: 0.10

[Haghighi & Vanderwende, NAACL '09]

Raw Count Content Model

$P_C(\cdot)$

Barack Obama: 0.15
Serve America Act: 0.13
signed: 0.12

President Barack Obama received the Serve America Act after congress' vote. The ailing senator was instrumental in its passage.

Document Structure

General

President Barack Obama received the Serve America Act after congress' vote...

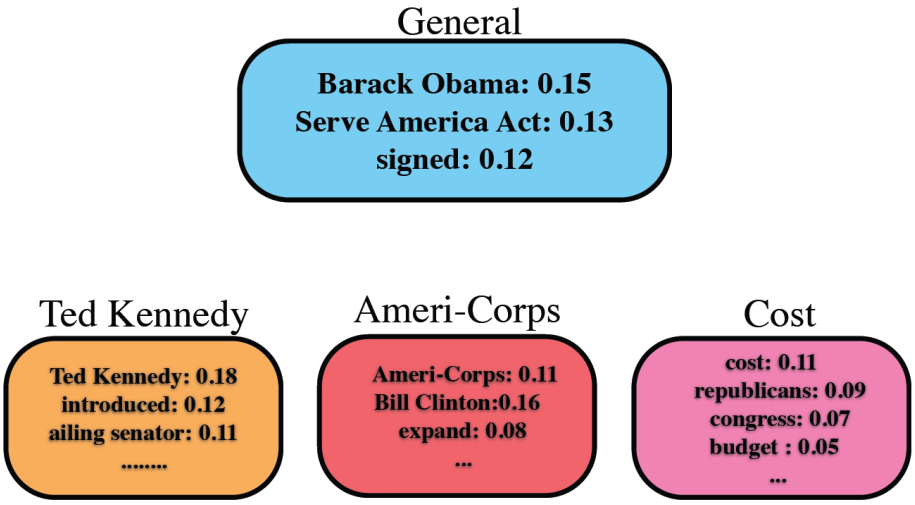
Ted Kennedy

The bill is named after Massachusetts Senator Ted Kennedy who was present at its signing. The ailing senator was instrumental...

Ameri Corps

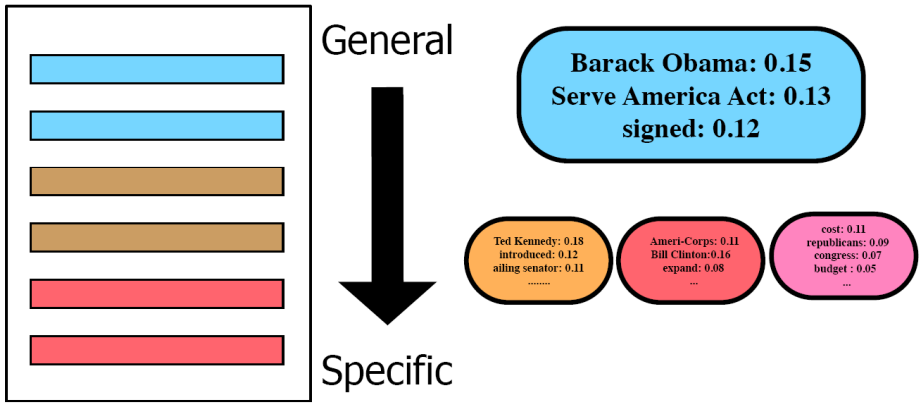
The legislation would greatly expand the ranks of Ameri-Corps, which was created by President Bill Clinton in 1993...

Structured Content Models

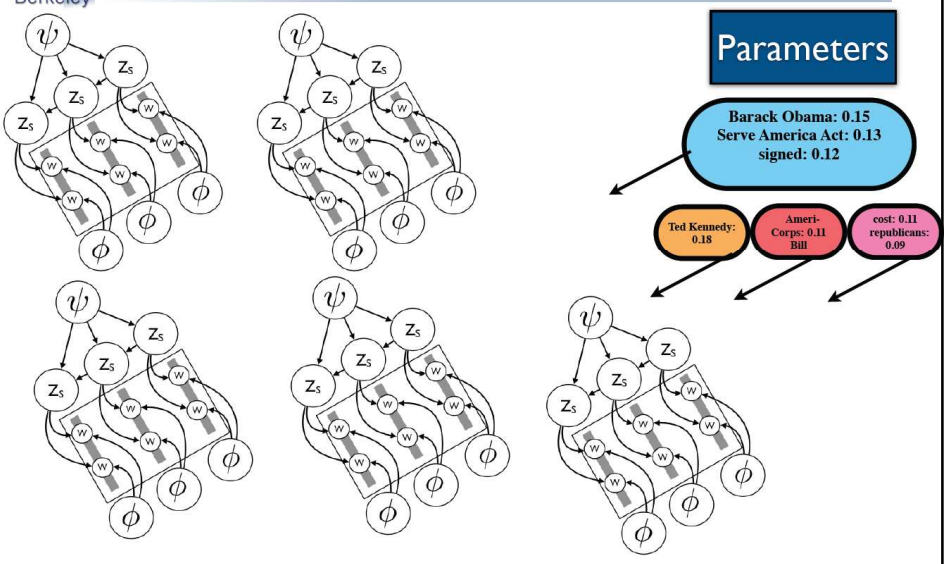


[Haghighi & Vanderwende, NAACL '09]

Document Structure

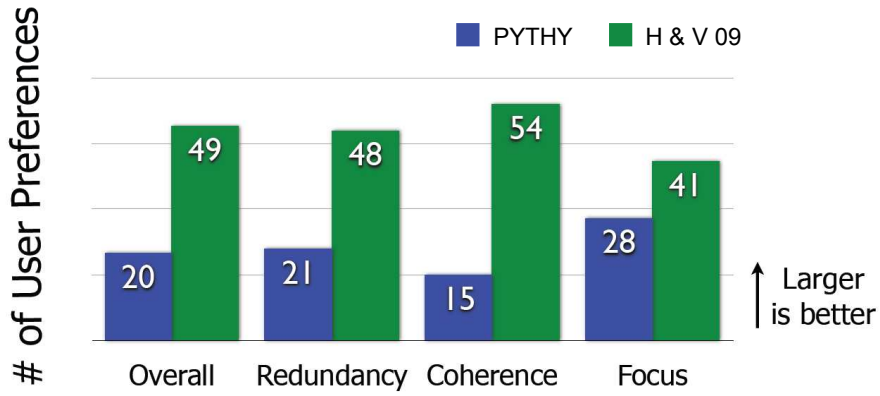


Generative Model



Evaluation: User Study

Pairwise Comparison



Example General Summary

Former House Speaker Newt Gingrich is asking a judge to force his estranged wife to turn over money he says she is hoarding.

On Thursday, accusations of wrongdoing and the mining of dirt in the former U.S. House speaker's divorce case gave way to a secret settlement between Gingrich and his wife of 18 years, Marianne Gingrich.

Gingrich filed for divorce July 29 amid allegations he is having an affair with 33-year-old congressional aide Callista Bisek.

Example Topical Summary

Gingrich Bio / Post-Speaker Life

Gingrich is best known leading the Republican Party's takeover of the House in 1994. During that so-called Republican Revolution, Gingrich emphasized that "family values" should be a core pillar in American society.

Since resigning as speaker and from the congressional seat he held for 20 years, Gingrich has been making a living giving speeches, sitting on corporate boards, consulting and appearing as a political analyst on Fox News.

U.S. Rep. J.D. Hayworth (R-Ariz.) argued that Gingrich's new job as a political commentator for Fox News makes it inappropriate to include him in political gatherings. "Time marches on. He's gone on to other pursuits," Hayworth said.

Selection

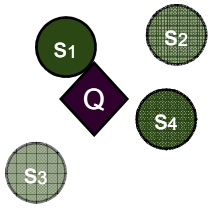
mid-'90s

present

- Maximum Marginal Relevance
- Graph algorithms
- Word distribution models
- Regression models
- Topic models
- Globally optimal search

[McDonald, 2007]

Optimal search using MMR



Integer Linear Program

Maximize: $\sum_i Rel_i s_i - \sum_{ij} Red_{ij} s_{ij}$

Subject to: $\sum_j l_j s_j \leq L$

$s_{ij} \leq s_i \quad s_{ij} \leq s_j \quad \forall i, j$

$s_i + s_j - s_{ij} \leq 1 \quad \forall i, j$

$s_i \in \{0, 1\} \quad \forall i$

$s_{ij} \in \{0, 1\} \quad \forall i, j$

Selection

[Gillick and Favre, 2008]

S1 The health care bill is a major test for the Obama administration.

S2 Universal health care is a divisive issue.

S3 President Obama remained calm.

S4 Obama addressed the House on Tuesday.

concept	value

Selection

[Gillick and Favre, 2008]

S1 The health care bill is a major test for the Obama administration.

S2 Universal health care is a divisive issue.

S3 President Obama remained calm.

S4 Obama addressed the House on Tuesday.

concept	value
obama	3

Selection

[Gillick and Favre, 2008]

S1 The health care bill is a major test for the Obama administration.

S2 Universal health care is a divisive issue.

S3 President Obama remained calm.

S4 Obama addressed the House on Tuesday.

concept	value
obama	3
health	2

Selection

[Gillick and Favre, 2008]

s_1	The health care bill is a major test for the Obama administration.
s_2	Universal health care is a divisive issue.
s_3	President Obama remained calm.
s_4	Obama addressed the House on Tuesday.

concept	value
obama	3
health	2
house	1

Selection

[Gillick and Favre, 2008]

s_1	The health care bill is a major test for the Obama administration.
s_2	Universal health care is a divisive issue.
s_3	President Obama remained calm.
s_4	Obama addressed the House on Tuesday.

concept	value
obama	3
health	2
house	1

Length limit: 18 words

summary	length	value	
$\{s_1, s_3\}$	17	5	← greedy
$\{s_2, s_3, s_4\}$	17	6	← optimal

Selection

Integer Linear Program for the maximum coverage model

[Gillick, Riedhammer, Favre, Hakkani-Tur, 2008]

$$\text{Maximize: } \sum_i w_i c_i$$ ← total concept value

$$\text{Subject to: } \sum_j l_j s_j \leq L$$ ← summary length limit

$$s_j Occ_{ij} \leq c_i, \quad \forall i, j$$

$$\sum_j s_j Occ_{ij} \geq c_i \quad \forall i$$

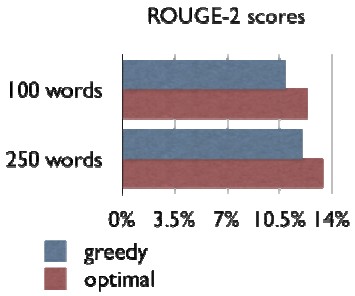
← maintain consistency between selected sentences and concepts

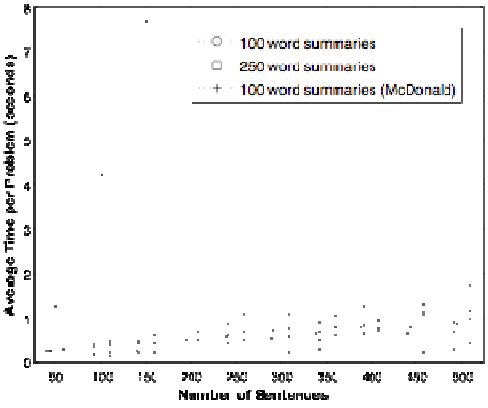
$$c_i \in \{0,1\} \quad \forall i$$

$$s_j \in \{0,1\} \quad \forall j$$

Selection

[Gillick and Favre, 2009]

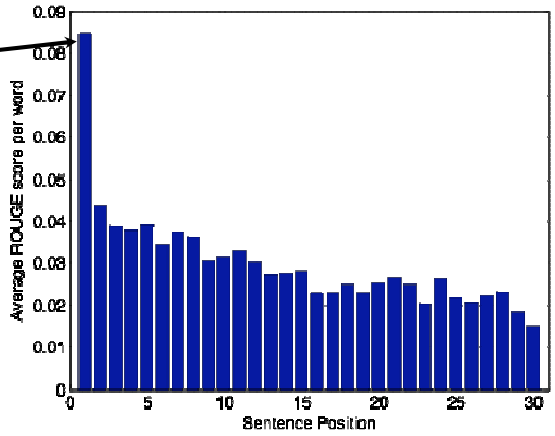




This ILP is tractable for reasonable problems

Selection

First sentences are unique



How to include sentence position?

Selection

Only allow first sentences in the summary	surprisingly strong baseline
Up-weight concepts appearing in first sentences	included in TAC 2009 system
Identify more sentences that look like first sentences	first sentence classifier is not reliable enough yet

How to include sentence position?

Selection

Some interesting work on sentence ordering

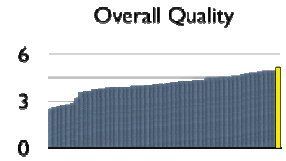
[Barzilay et. al., 1997; 2002]

But choosing independent sentences is easier

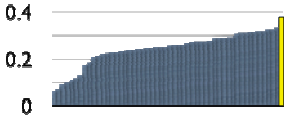
- First sentences usually stand alone well
- Sentences without unresolved pronouns
- Classifier trained on OntoNotes: <10% error rate

Baseline ordering module (chronological) is not obviously worse than anything fancier

Results [G & F, 2009]



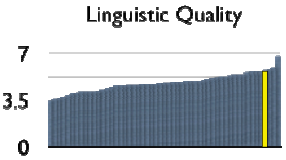
- Rating scale: 1-10
- Humans in [8.3, 9.3]



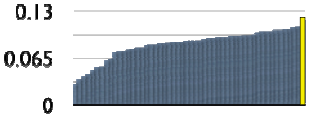
- Rating scale: 0-1
- Humans in [0.62, 0.77]

- 52 submissions
- 27 teams
- 44 topics
  - 10 input docs
  - 100 word summaries

Gillick & Favre



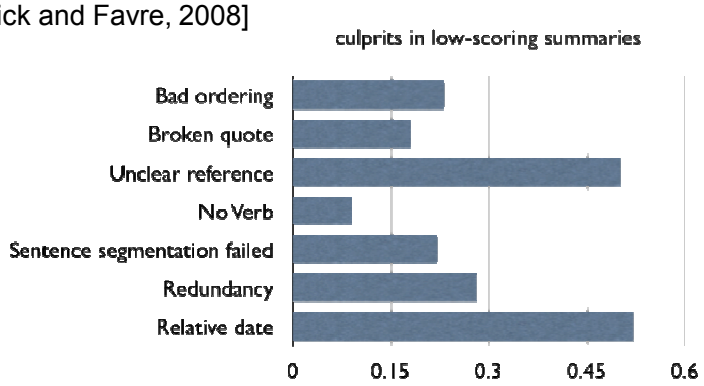
- Rating scale: 1-10
- Humans in [8.5, 9.3]



- Rating scale: 0-1
- Humans in [0.11, 0.15]

Error Breakdown?

[Gillick and Favre, 2008]

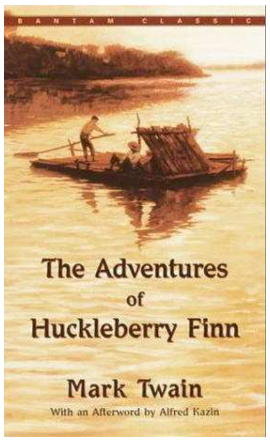


Beyond Extraction?

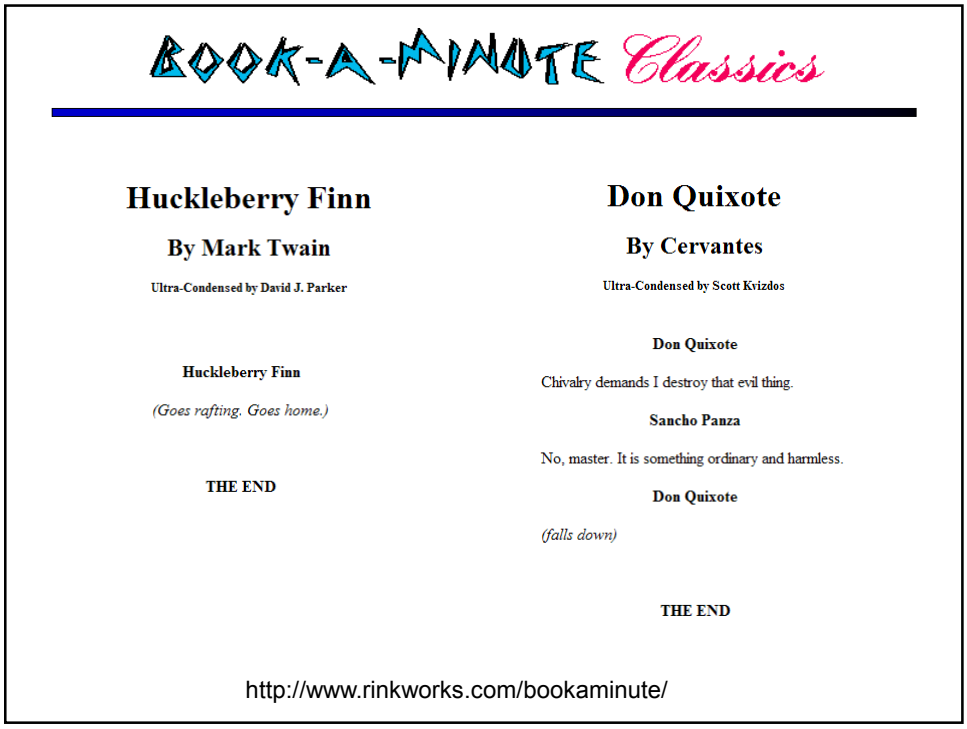
Sentence extraction is limiting

... and boring!

But abstractive summaries are much harder to generate…



in 25 words?


BOOK-A-MINUTE Classics
Huckleberry Finn
By Mark Twain
Ultra-Condensed by David J. Parker
Huckleberry Finn
(Goes rafting. Goes home.)
THE END
Don Quixote
By Cervantes
Ultra-Condensed by Scott Kvizdos
Don Quixote
Chivalry demands I destroy that evil thing.
Sancho Panza
No, master. It is something ordinary and harmless.
Don Quixote
(falls down)
THE END
http://www.rinkworks.com/bookaminute/