Entropy and Graph Based Modelling of Document Coherence using Discourse Entities: An Application to IR

Casper Petersen[1] Christina Lioma[1] Jakob Grue Simonsen[1] Birger Larsen[2]

[1]Department of Computer Science
University of Copenhagen, Denmark
{cazz,c.lioma,simonsen}@di.ku.dk

[2]Department of Communication
University of Aalborg, Denmark
birger@hum.aau.dk

ACM SIGIR International Conference on the Theory of Information Retrieval 2015
27/09 – 30/09, Northampton, Massachusetts, United States

Motivation

1. Text coherence (TC) is...
   - "*The extent that a reader can understand the relations between ideas in a text*" [McNamara and Kintsch, 1996]
   - A property of well-written texts
   - Major factor for comprehension
2. Recent effective TC models have not been used in IR tasks
3. Use TC models to improve select IR tasks
4. TC models based on *entity grids* will improve performance of:
   - Sentence ordering task (automatic summarisation)
   - Reranking results (ad hoc retrieval)

   over competitive baselines

Motivation

1. Text coherence (TC) is...
   - "*The extent that a reader can understand the relations between ideas in a text*" [McNamara and Kintsch, 1996]
   - A property of well-written texts
   - Major factor for comprehension

2. Recent effective TC models have not been used in IR tasks

3. Use TC models to improve select IR tasks

4. TC models based on *entity grids* will improve performance of:
   - Sentence ordering task (automatic summarisation)
   - Reranking results (ad hoc retrieval)

   over competitive baselines

Motivation

1. Text coherence (TC) is...
   - "*The extent that a reader can understand the relations between ideas in a text*" [McNamara and Kintsch, 1996]
   - A property of well-written texts
   - Major factor for comprehension

2. Recent effective TC models have not been used in IR tasks

3. Use TC models to improve select IR tasks

4. TC models based on *entity grids* will improve performance of:
   - Sentence ordering task (automatic summarisation)
   - Reranking results (ad hoc retrieval)

   over competitive baselines

Motivation

1. Text coherence (TC) is...
   - "*The extent that a reader can understand the relations between ideas in a text*" [McNamara and Kintsch, 1996]
   - A property of well-written texts
   - Major factor for comprehension

2. Recent effective TC models have not been used in IR tasks

3. Use TC models to improve select IR tasks

4. TC models based on *entity grids* will improve performance of:
   - Sentence ordering task (automatic summarisation)
   - Reranking results (ad hoc retrieval)

   over competitive baselines

Motivation

1. Text coherence (TC) is...
   - "*The extent that a reader can understand the relations between ideas in a text*" [McNamara and Kintsch, 1996]
   - A property of well-written texts
   - Major factor for comprehension
2. Recent effective TC models have not been used in IR tasks
3. Use TC models to improve select IR tasks
4. TC models based on *entity grids* will improve performance of:
   - Sentence ordering task (automatic summarisation)
   - Reranking results (ad hoc retrieval)

   over competitive baselines

Motivation

1. Text coherence (TC) is...
   - "*The extent that a reader can understand the relations between ideas in a text*" [McNamara and Kintsch, 1996]
   - A property of well-written texts
   - Major factor for comprehension
2. Recent effective TC models have not been used in IR tasks
3. Use TC models to improve select IR tasks
4. TC models based on *entity grids* will improve performance of:
   - Sentence ordering task (automatic summarisation)
   - Reranking results (ad hoc retrieval)

   over competitive baselines

Motivation

1. Text coherence (TC) is...
    - "*The extent that a reader can understand the relations between ideas in a text*" [McNamara and Kintsch, 1996]
    - A property of well-written texts
    - Major factor for comprehension
2. Recent effective TC models have not been used in IR tasks
3. Use TC models to improve select IR tasks
4. TC models based on *entity grids* will improve performance of:
    - Sentence ordering task (automatic summarisation)
    - Reranking results (ad hoc retrieval)

    over competitive baselines

Entity Grid Model

	MAN	HOPE	CONFIDENCE	BOY	YOU	THEM	I	THESE
S_1	s	s	s	—	—	—	—	—
S_2	—	—	—	s	—	—	—	—
S_3	s	—	—	—	s	o	—	—
S_4	—	—	—	s	—	—	s	o
S_5	s	—	—	—	o	—	—	—

ENTITY GRID

[Barzilay and Lapata, 2008]

1 - "One", the old man said; his hope and his confidence had never gone.

2 - "Two", the boy said.

3 - "Two", the old man agreed; "you didn't steal them?"

4 - "I would", the boy said, "but I bought these".

5 - "Thank you", the old man said.

SAMPLE TEXT (*"The Old Man and the Sea"*)

From Entity Grid Model to Entropy Models

EXAMPLE BIGRAMS

(s,s) (BOY,MAN)
(s,o) (I,THESE)

	MAN	HOPE	CONFIDENCE	BOY	YOU	THEM	I	THESE
S_1	s	s	s	—	—	—	—	—
S_2	—	—	—	s	—	—	—	—
S_3	s	—	—	—	s	o	—	—
S_4	—	—	—	s	—	—	s	o
S_5	s	—	—	—	o	—	—	—

ENTITY GRID

[Barzilay and Lapata, 2008]

1. Extract n-grams in row-wise fashion
2. Calculate n-gram probabilities using MLE
   $p(e_i) = \frac{f(e_i)}{|E|}, \qquad p(e_i|e_{i-1}) = \frac{f(e_{i-1}, e_i)}{f(e_i)}$
3. Entropy score:
   $H_{k=0}(E) = -\sum_{e_i \in E} p(e_i) \log_2 p(e_i)$
4. Coherence score:
   $C = \frac{1}{H_k(E)}$

From Entity Grid Model to Entropy Models

EXAMPLE BIGRAMS

(s,s) (BOY,MAN)
(s,o) (I,THESE)

	MAN	HOPE	CONFIDENCE	BOY	YOU	THEM	I	THESE
S_1	s	s	s	—	—	—	—	—
S_2	—	—	—	s	—	—	—	—
S_3	s	—	—	—	s	o	—	—
S_4	—	—	—	s	—	—	s	o
S_5	s	—	—	—	o	—	—	—

ENTITY GRID

[Barzilay and Lapata, 2008]

1. Extract *n*-grams in row-wise fashion
2. Calculate *n*-gram probabilities using MLE
   $p(e_i) = \frac{f(e_i)}{|E|}, \qquad p(e_i|e_{i-1}) = \frac{f(e_{i-1}, e_i)}{f(e_i)}$
3. Entropy score:
   $H_{k=0}(E) = -\sum_{e_i \in E} p(e_i) \log_2 p(e_i)$
4. Coherence score:
   $C = \frac{1}{H_k(E)}$

From Entity Grid Model to Entropy Models

EXAMPLE BIGRAMS

(s,s) (BOY,MAN)
(s,o) (I,THESE)

	MAN	HOPE	CONFIDENCE	BOY	YOU	THEM	I	THESE
S_1	s	s	s	–	–	–	–	–
S_2	–	–	–	s	–	–	–	–
S_3	s	–	–	–	s	o	–	–
S_4	–	–	–	s	–	–	s	o
S_5	s	–	–	–	o	–	–	–

ENTITY GRID

[Barzilay and Lapata, 2008]

1. Extract *n*-grams in row-wise fashion
2. Calculate *n*-gram probabilities using MLE

$$p(e_i) = \frac{f(e_i)}{|E|}, \qquad p(e_i|e_{i-1}) = \frac{f(e_{i-1}, e_i)}{f(e_i)}$$

3. Entropy score:

$$H_{k=0}(E) = -\sum_{e_i \in E} p(e_i) \log_2 p(e_i)$$

4. Coherence score:

$$C = \frac{1}{H_k(E)}$$

From Entity Grid Model to Entropy Models

EXAMPLE BIGRAMS

(s,s) (BOY,MAN)
(s,o) (I,THESE)

	MAN	HOPE	CONFIDENCE	BOY	YOU	THEM	I	THESE
S_1	s	s	s	—	—	—	—	—
S_2	—	—	—	s	—	—	—	—
S_3	s	—	—	—	s	o	—	—
S_4	—	—	—	s	—	—	s	o
S_5	s	—	—	—	o	—	—	—

ENTITY GRID

[Barzilay and Lapata, 2008]

1. Extract *n*-grams in row-wise fashion
2. Calculate *n*-gram probabilities using MLE

$$p(e_i) = \frac{f(e_i)}{|E|}, \qquad p(e_i|e_{i-1}) = \frac{f(e_{i-1}, e_i)}{f(e_i)}$$

3. Entropy score:

$$H_{k=0}(E) = -\sum_{e_i \in E} p(e_i) \log_2 p(e_i)$$

4. Coherence score:

$$C = \frac{1}{H_k(E)}$$

From Entity Grid Model to Entropy Models

EXAMPLE BIGRAMS

(s,s) (BOY,MAN)
(s,o) (I,THESE)

	MAN	HOPE	CONFIDENCE	BOY	YOU	THEM	I	THESE
S_1	s	s	s	–	–	–	–	–
S_2	–	–	–	s	–	–	–	–
S_3	s	–	–	–	s	o	–	–
S_4	–	–	–	s	–	–	s	o
S_5	s	–	–	–	o	–	–	–

ENTITY GRID

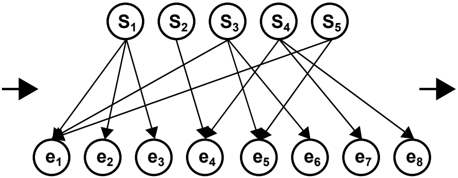
[Barzilay and Lapata, 2008]

1. Extract *n*-grams in row-wise fashion
2. Calculate *n*-gram probabilities using MLE
$$p(e_i) = \frac{f(e_i)}{|E|}, \qquad p(e_i|e_{i-1}) = \frac{f(e_{i-1}, e_i)}{f(e_i)}$$
3. Entropy score:
$$H_{k=0}(E) = -\sum_{e_i \in E} p(e_i) \log_2 p(e_i)$$
4. Coherence score:
$$C = \frac{1}{H_k(E)}$$

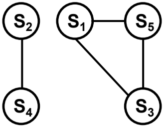
From Entity Grid Model to Graph Models (1/2)

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
S_1	s	s	s	—	—	—	—	—
S_2	—	—	—	s	—	—	—	—
S_3	s	—	—	—	s	o	—	—
S_4	—	—	—	s	—	—	s	o
S_5	s	—	—	—	o	—	—	—

ENTITY GRID



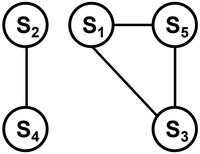
BIPARTITE GRAPH



SIMPLE GRAPH

Approach by [Guinaudeau and Strube, 2013]

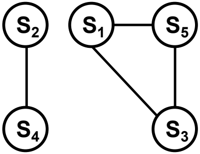
From Entity Grid Model to Graph Models (2/2)



SIMPLE GRAPH

1. *Assumption*:
   Graph topology reflects text coherence
2. Graph topology captured using graph metrics
   - PageRank
   - Clustering coefficient
   - Betweenness
   - Entity distance
   - Adjacent Topic Flow
   - Adjacent Weighted Topic Flow
   - Non adjacent Topic Flow
   - Non adjacent Weighted Topic Flow
3. Captures either local or global coherence

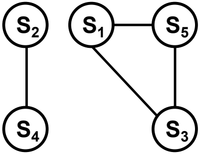
From Entity Grid Model to Graph Models (2/2)



SIMPLE GRAPH

1. *Assumption*:
   Graph topology reflects text coherence
2. Graph topology captured using graph metrics
   - PageRank
   - Clustering coefficient
   - Betweenness
   - Entity distance
   - Adjacent Topic Flow
   - Adjacent Weighted Topic Flow
   - Non adjacent Topic Flow
   - Non adjacent Weighted Topic Flow
3. Captures either local or global coherence

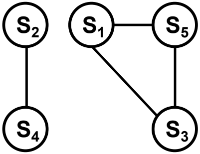
From Entity Grid Model to Graph Models (2/2)



SIMPLE GRAPH

1. *Assumption*:
   Graph topology reflects text coherence
2. Graph topology captured using graph metrics
   - PageRank
   - Clustering coefficient
   - Betweenness
   - Entity distance
   - Adjacent Topic Flow
   - Adjacent Weighted Topic Flow
   - Non adjacent Topic Flow
   - Non adjacent Weighted Topic Flow
3. Captures either local or global coherence

From Entity Grid Model to Graph Models (2/2)



SIMPLE GRAPH

1. *Assumption*:
   Graph topology reflects text coherence
2. Graph topology captured using graph metrics
   - PageRank
   - Clustering coefficient
   - Betweenness
   - Entity distance
   - Adjacent Topic Flow
   - Adjacent Weighted Topic Flow
   - Non adjacent Topic Flow
   - Non adjacent Weighted Topic Flow
3. Captures either local or global coherence

Experiment 1: Setup

- Evaluate using *sentence reordering task*
- Datasets:
  - Earthquake (100 documents, articles, curated, 257.3 avg.)
  - Accidents (100 documents, narratives, curated, 223.5 avg.)
- Performance measure: accuracy
- Baselines:
  - Entity Grid Model [Barzilay and Lapata, 2008]
  - Hidden Markov Model [Barzilay and Lee, 2004]
- *Tuned* baselines against *untuned* coherence models

Experiment 1: Setup

- Evaluate using *sentence reordering task*
- Datasets:
  - Earthquake (100 documents, articles, curated, 257.3 avg.)
  - Accidents (100 documents, narratives, curated, 223.5 avg.)
- Performance measure: accuracy
- Baselines:
  - Entity Grid Model [Barzilay and Lapata, 2008]
  - Hidden Markov Model [Barzilay and Lee, 2004]
- *Tuned* baselines against *untuned* coherence models

Experiment 1: Setup

- Evaluate using *sentence reordering task*
- Datasets:
  - Earthquake (100 documents, articles, curated, 257.3 avg.)
  - Accidents (100 documents, narratives, curated, 223.5 avg.)
- Performance measure: accuracy
- Baselines:
  - Entity Grid Model [Barzilay and Lapata, 2008]
  - Hidden Markov Model [Barzilay and Lee, 2004]
- *Tuned* baselines against *untuned* coherence models

Experiment 1: Setup

- Evaluate using *sentence reordering task*
- Datasets:
  - Earthquake (100 documents, articles, curated, 257.3 avg.)
  - Accidents (100 documents, narratives, curated, 223.5 avg.)
- Performance measure: accuracy
- Baselines:
  - Entity Grid Model [Barzilay and Lapata, 2008]
  - Hidden Markov Model [Barzilay and Lee, 2004]
- *Tuned* baselines against *untuned* coherence models

Experiment 1: Setup

- Evaluate using *sentence reordering task*
- Datasets:
  - Earthquake (100 documents, articles, curated, 257.3 avg.)
  - Accidents (100 documents, narratives, curated, 223.5 avg.)
- Performance measure: accuracy
- Baselines:
  - Entity Grid Model [Barzilay and Lapata, 2008]
  - Hidden Markov Model [Barzilay and Lee, 2004]
- *Tuned* baselines against *untuned* coherence models

Experiment 1: Results

	Method	Earthquakes Acc.	Earthquakes ±%	Accidents Acc.	Accidents ±%
BASELINES	Entity grid model	69.7*	–	67.0*	–
	HMM-based model	60.3*	–	31.7*	–
ENTROPY	Entropy-0 order	**75.0**	+7.6%	**73.0***	+9.0%
	Entropy-1 order	64.0	−8.2%	**70.0***	+4.5%
	Entropy-2 order	64.0	−8.2%	**70.0***	+4.5%
GRAPH	PageRank	**75.0**	+7.6%	**73.0***	+9.0%
	Clustering Coef.	67.0	−3.9%	66.0*	−1.5%
	Betweenness	**73.0***	+4.7%	‡**77.0***	+14.9%
	Entity Distance	‡**76.0**	+9.0%	**75.0***	+11.9%
	Adj. Topic Flow	**70.0***	+0.4%	**74.0***	+10.4%
	Adj W. Topic Flow	61.0*	−12.5%	66.0*	−1.5%
	nAdj. Topic Flow	**70.0**	+0.4%	**70.0**	+4.5%
	nAdj. W. Topic Flow	**70.0**	+0.4%	**70.0***	+4.5%

- Large length sentences
- Spatial proximity != semantic relatedness

Experiment 1: Results

	Method	Earthquakes Acc.	Earthquakes ±%	Accidents Acc.	Accidents ±%
BASELINES	Entity grid model	69.7*	–	67.0*	–
	HMM-based model	60.3*	–	31.7*	–
ENTROPY	Entropy-0 order	**75.0**	+7.6%	**73.0***	+9.0%
	Entropy-1 order	64.0	−8.2%	**70.0***	+4.5%
	Entropy-2 order	64.0	−8.2%	**70.0***	+4.5%
GRAPH	PageRank	**75.0**	+7.6%	**73.0***	+9.0%
	Clustering Coef.	67.0	−3.9%	66.0*	−1.5%
	Betweenness	**73.0***	+4.7%	‡**77.0***	+14.9%
	Entity Distance	‡**76.0**	+9.0%	**75.0***	+11.9%
	Adj. Topic Flow	**70.0***	+0.4%	**74.0***	+10.4%
	Adj W. Topic Flow	61.0*	−12.5%	66.0*	−1.5%
	nAdj. Topic Flow	**70.0**	+0.4%	**70.0**	+4.5%
	nAdj. W. Topic Flow	**70.0**	+0.4%	**70.0***	+4.5%

- Large length sentences
- Spatial proximity != semantic relatedness

Experiment 1: Results

	Method	Earthquakes Acc.	Earthquakes ±%	Accidents Acc.	Accidents ±%
BASELINES	Entity grid model	69.7*	–	67.0*	–
	HMM-based model	60.3*	–	31.7*	–
ENTROPY	Entropy-0 order	**75.0**	+7.6%	**73.0***	+9.0%
	Entropy-1 order	64.0	−8.2%	**70.0***	+4.5%
	Entropy-2 order	64.0	−8.2%	**70.0***	+4.5%
GRAPH	PageRank	**75.0**	+7.6%	**73.0***	+9.0%
	Clustering Coef.	67.0	−3.9%	66.0*	−1.5%
	Betweenness	**73.0***	+4.7%	‡**77.0***	+14.9%
	Entity Distance	‡**76.0**	+9.0%	**75.0***	+11.9%
	Adj. Topic Flow	**70.0***	+0.4%	**74.0***	+10.4%
	Adj W. Topic Flow	61.0*	−12.5%	66.0*	−1.5%
	nAdj. Topic Flow	**70.0**	+0.4%	**70.0**	+4.5%
	nAdj. W. Topic Flow	**70.0**	+0.4%	**70.0***	+4.5%

- Large length sentences
- Spatial proximity != semantic relatedness

Experiment 2: Setup

- *Assumption*:
  More coherent documents are more relevant documents
- Rerank top-1000 retrieved documents:
  $\widehat{RSV}_d = RSV_d \times \alpha + (1 - \alpha)\, COH_d$
- Spam filtered ClueWeb09 cat. B. ($\approx$ 16M documents)
- Dirichlet-smoothed unigram query likelihood language model
- Queries 150–200 (TREC WebTrack 2012)
- Performance measures: MRR, P@10, MAP, ERR@20
- Tuned baseline and $\widehat{RSV}$. 5-fold cross-validation

Experiment 2: Setup

- *Assumption*:
  More coherent documents are more relevant documents
- Rerank top-1000 retrieved documents:
  $\widehat{RSV}_d = RSV_d \times \alpha + (1 - \alpha)\, COH_d$
- Spam filtered ClueWeb09 cat. B. ($\approx$ 16M documents)
- Dirichlet-smoothed unigram query likelihood language model
- Queries 150–200 (TREC WebTrack 2012)
- Performance measures: MRR, P@10, MAP, ERR@20
- Tuned baseline and $\widehat{RSV}$. 5-fold cross-validation

Experiment 2: Setup

- *Assumption*:
  More coherent documents are more relevant documents
- Rerank top-1000 retrieved documents:
  $\widehat{RSV}_d = RSV_d \times \alpha + (1 - \alpha)\, COH_d$
- Spam filtered ClueWeb09 cat. B. ($\approx$ 16M documents)
- Dirichlet-smoothed unigram query likelihood language model
- Queries 150–200 (TREC WebTrack 2012)
- Performance measures: MRR, P@10, MAP, ERR@20
- Tuned baseline and $\widehat{RSV}$. 5-fold cross-validation

Experiment 2: Setup

- *Assumption*:
  More coherent documents are more relevant documents
- Rerank top-1000 retrieved documents:
  $\widehat{RSV}_d = RSV_d \times \alpha + (1 - \alpha)\, COH_d$
- Spam filtered ClueWeb09 cat. B. ($\approx$ 16M documents)
- Dirichlet-smoothed unigram query likelihood language model
- Queries 150–200 (TREC WebTrack 2012)
- Performance measures: MRR, P@10, MAP, ERR@20
- Tuned baseline and $\widehat{RSV}$. 5-fold cross-validation

Experiment 2: Setup

- *Assumption*:
  More coherent documents are more relevant documents
- Rerank top-1000 retrieved documents:
  $\widehat{RSV}_d = RSV_d \times \alpha + (1 - \alpha)\, COH_d$
- Spam filtered ClueWeb09 cat. B. ($\approx$ 16M documents)
- Dirichlet-smoothed unigram query likelihood language model
- Queries 150–200 (TREC WebTrack 2012)
- Performance measures: MRR, P@10, MAP, ERR@20
- Tuned baseline and $\widehat{RSV}$. 5-fold cross-validation

Experiment 2: Setup

- *Assumption*:
  More coherent documents are more relevant documents
- Rerank top-1000 retrieved documents:
  $\widehat{RSV}_d = RSV_d \times \alpha + (1 - \alpha)\, COH_d$
- Spam filtered ClueWeb09 cat. B. ($\approx$ 16M documents)
- Dirichlet-smoothed unigram query likelihood language model
- Queries 150–200 (TREC WebTrack 2012)
- Performance measures: MRR, P@10, MAP, ERR@20
- Tuned baseline and $\widehat{RSV}$. 5-fold cross-validation

Experiment 2: Setup

- *Assumption*:
  More coherent documents are more relevant documents
- Rerank top-1000 retrieved documents:
  $\widehat{RSV}_d = RSV_d \times \alpha + (1 - \alpha)\, COH_d$
- Spam filtered ClueWeb09 cat. B. ($\approx$ 16M documents)
- Dirichlet-smoothed unigram query likelihood language model
- Queries 150–200 (TREC WebTrack 2012)
- Performance measures: MRR, P@10, MAP, ERR@20
- Tuned baseline and $\widehat{RSV}$. 5-fold cross-validation

Experiment 2: Results

Method	MRR	±%	P@10	±%
Baseline	20.57	–	19.80	–
Entropy-0 order	**49.50**	+140.6%	**33.00**	+66.7%
PageRank	**49.85**	+142.3%	**34.40**	+73.7%
Clustering Coef.	**51.82**	+151.9%	**34.60**	+74.7%
Betweenness	**49.74**	+141.8%	**36.40**	+83.8%
Entity Distance	**34.18**	+66.2%	**22.40**	+13.1%
Adj. Topic Flow	**55.73**	+170.9%	**34.20**	+72.7%
Adj. W. Topic Flow	**51.60**	+150.8%	**34.20**	+72.7%
nAdj. Topic Flow	**50.62**	+146.1%	**34.40**	+73.7%
nAdj. W. Topic Flow	**50.79**	+146.9%	**34.60**	+74.7%

- Coherence improves early precision
- Entity distance is consistently the weakest
- Coherence a discriminative feature of relevance

Experiment 2: Results

Method	MRR	±%	P@10	±%
Baseline	20.57	–	19.80	–
Entropy-0 order	**49.50**	+140.6%	**33.00**	+66.7%
PageRank	**49.85**	+142.3%	**34.40**	+73.7%
Clustering Coef.	**51.82**	+151.9%	**34.60**	+74.7%
Betweenness	**49.74**	+141.8%	**36.40**	+83.8%
Entity Distance	**34.18**	+66.2%	**22.40**	+13.1%
Adj. Topic Flow	**55.73**	+170.9%	**34.20**	+72.7%
Adj. W. Topic Flow	**51.60**	+150.8%	**34.20**	+72.7%
nAdj. Topic Flow	**50.62**	+146.1%	**34.40**	+73.7%
nAdj. W. Topic Flow	**50.79**	+146.9%	**34.60**	+74.7%

- Coherence improves early precision
- Entity distance is consistently the weakest
- Coherence a discriminative feature of relevance

Experiment 2: Results

Method	MRR	±%	P@10	±%
Baseline	20.57	–	19.80	–
Entropy-0 order	**49.50**	+140.6%	**33.00**	+66.7%
PageRank	**49.85**	+142.3%	**34.40**	+73.7%
Clustering Coef.	**51.82**	+151.9%	**34.60**	+74.7%
Betweenness	**49.74**	+141.8%	**36.40**	+83.8%
Entity Distance	**34.18**	+66.2%	**22.40**	+13.1%
Adj. Topic Flow	**55.73**	+170.9%	**34.20**	+72.7%
Adj. W. Topic Flow	**51.60**	+150.8%	**34.20**	+72.7%
nAdj. Topic Flow	**50.62**	+146.1%	**34.40**	+73.7%
nAdj. W. Topic Flow	**50.79**	+146.9%	**34.60**	+74.7%

- Coherence improves early precision
- Entity distance is consistently the weakest
- Coherence a discriminative feature of relevance

Experiment 2: Results

Method	MRR	±%	P@10	±%
Baseline	20.57	–	19.80	–
Entropy-0 order	**49.50**	+140.6%	**33.00**	+66.7%
PageRank	**49.85**	+142.3%	**34.40**	+73.7%
Clustering Coef.	**51.82**	+151.9%	**34.60**	+74.7%
Betweenness	**49.74**	+141.8%	**36.40**	+83.8%
Entity Distance	**34.18**	+66.2%	**22.40**	+13.1%
Adj. Topic Flow	**55.73**	+170.9%	**34.20**	+72.7%
Adj. W. Topic Flow	**51.60**	+150.8%	**34.20**	+72.7%
nAdj. Topic Flow	**50.62**	+146.1%	**34.40**	+73.7%
nAdj. W. Topic Flow	**50.79**	+146.9%	**34.60**	+74.7%

- Coherence improves early precision
- Entity distance is consistently the weakest
- Coherence a discriminative feature of relevance

Conclusion

- Text coherence (TC) is crucial for conveying and acquiring information from documents
- TC models based on *entity grids* not used before for IR tasks
- We make two contributions:
  1. Propose two classes of TC models that may be useful for NLP
  2. We show that several of these TC models are useful for retrieval
- TC may be discriminative feature of relevance
- Complements findings by [Bendersky et al., 2011] and [Tan et al., 2012]

Conclusion

- Text coherence (TC) is crucial for conveying and acquiring information from documents
- TC models based on *entity grids* not used before for IR tasks
- We make two contributions:
  1. Propose two classes of TC models that may be useful for NLP
  2. We show that several of these TC models are useful for retrieval
- TC may be discriminative feature of relevance
- Complements findings by [Bendersky et al., 2011] and [Tan et al., 2012]

Conclusion

- Text coherence (TC) is crucial for conveying and acquiring information from documents
- TC models based on *entity grids* not used before for IR tasks
- We make two contributions:
  1. Propose two classes of TC models that may be useful for NLP
  2. We show that several of these TC models are useful for retrieval
- TC may be discriminative feature of relevance
- Complements findings by [Bendersky et al., 2011] and [Tan et al., 2012]

Conclusion

- Text coherence (TC) is crucial for conveying and acquiring information from documents
- TC models based on *entity grids* not used before for IR tasks
- We make two contributions:
  1. Propose two classes of TC models that may be useful for NLP
  2. We show that several of these TC models are useful for retrieval
- TC may be discriminative feature of relevance
- Complements findings by [Bendersky et al., 2011] and [Tan et al., 2012]

Conclusion

- Text coherence (TC) is crucial for conveying and acquiring information from documents
- TC models based on *entity grids* not used before for IR tasks
- We make two contributions:
  1. Propose two classes of TC models that may be useful for NLP
  2. We show that several of these TC models are useful for retrieval
- TC may be discriminative feature of relevance
- Complements findings by [Bendersky et al., 2011] and [Tan et al., 2012]

Conclusion

- Text coherence (TC) is crucial for conveying and acquiring information from documents
- TC models based on *entity grids* not used before for IR tasks
- We make two contributions:
  1. Propose two classes of TC models that may be useful for NLP
  2. We show that several of these TC models are useful for retrieval
- TC may be discriminative feature of relevance
- Complements findings by [Bendersky et al., 2011] and [Tan et al., 2012]

Conclusion

- Text coherence (TC) is crucial for conveying and acquiring information from documents
- TC models based on *entity grids* not used before for IR tasks
- We make two contributions:
    1. Propose two classes of TC models that may be useful for NLP
    2. We show that several of these TC models are useful for retrieval
- TC may be discriminative feature of relevance
- Complements findings by [Bendersky et al., 2011] and [Tan et al., 2012]

Thank you for your attention

cazz@di.ku.dk – @cpdiku

Bibliography

[Barzilay and Lapata, 2008] Barzilay, R. and Lapata, M. (2008).
Modeling local coherence: An entity-based approach.
ACL, pages 1–34.

[Barzilay and Lee, 2004] Barzilay, R. and Lee, L. (2004).
Catching the drift: Probabilistic content models with applications to generation and summarization.
HLT-NAACL, pages 113–120.

[Bendersky et al., 2011] Bendersky, M., Croft, W. B., and Diao, Y. (2011).
Quality-biased ranking of web documents.
In *WSDM*, pages 95–104. ACM.

[Guinaudeau and Strube, 2013] Guinaudeau, C. and Strube, M. (2013).
Graph-based local coherence modeling.
ACL, pages 93–103.

[McNamara and Kintsch, 1996] McNamara, D. S. and Kintsch, W. (1996).
Learning from texts: Effects of prior knowledge and text coherence.
Discourse processes, 22(3):247–288.

[Tan et al., 2012] Tan, C., Gabrilovich, E., and Pang, B. (2012).
To each his own: personalized content selection based on text comprehensibility.
In *WSDM*, pages 233–242. ACM.