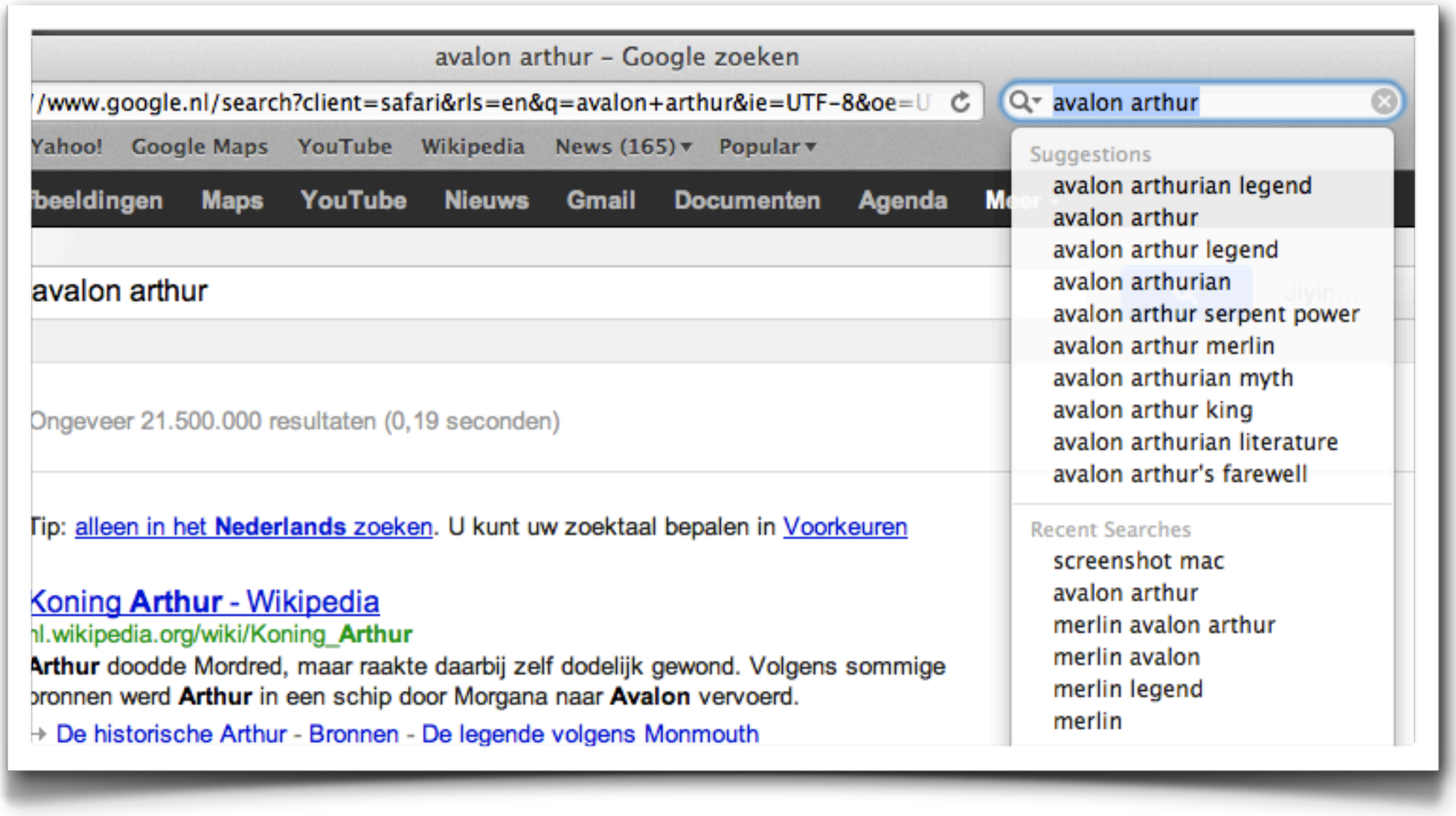
Explaining Query Modifications

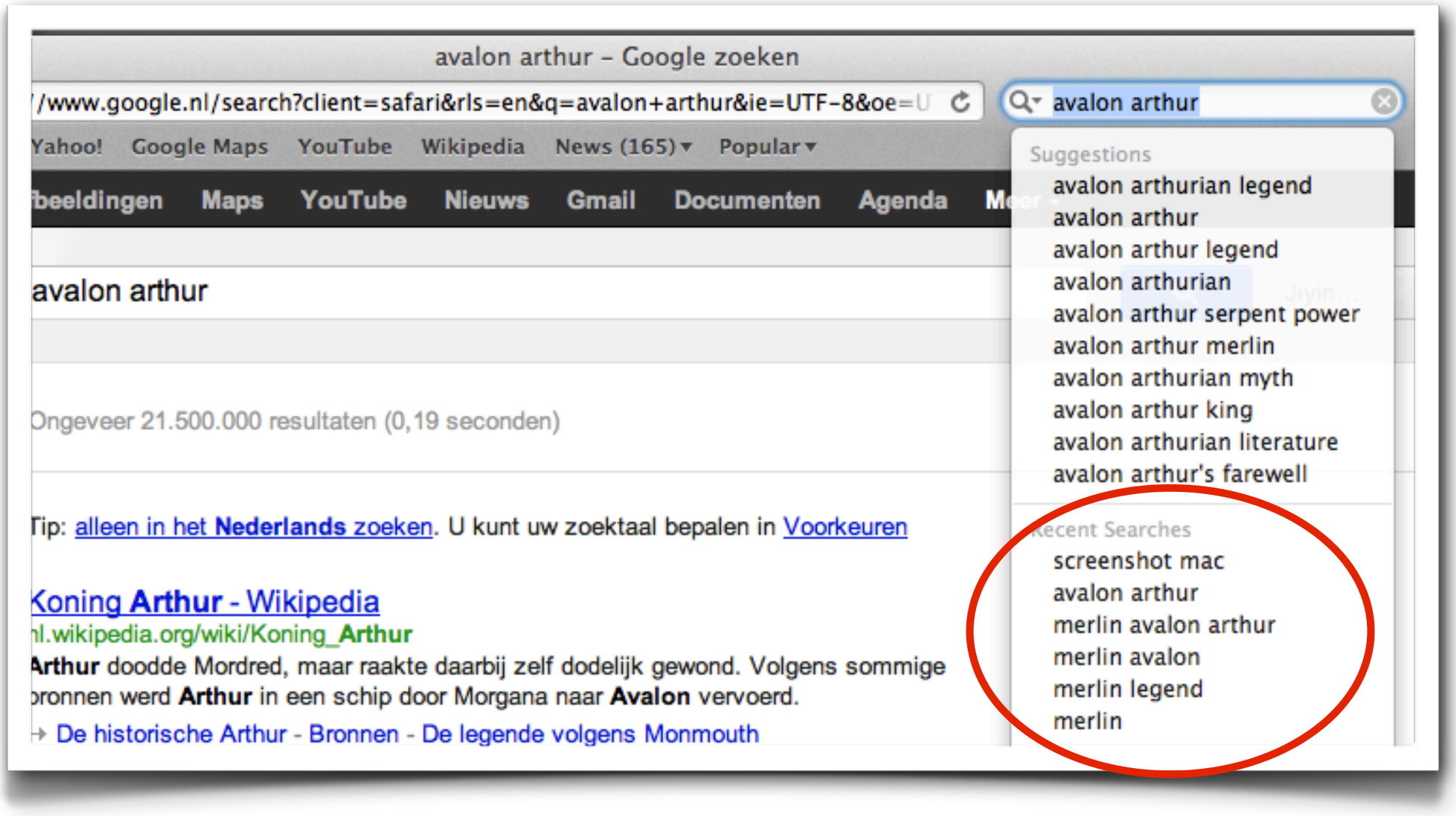
An alternative interpretation of term addition and removal

Vera Hollink, Jiyin He, Arjen de Vries
CWI, the Netherlands

Query modifications



Query modifications



Query modifications

merlin	
merlin legend	term addition
merlin avalon	term substitution
merlin avalon arthur	term addition
avalon arthur	term removal
screenshot mac	different

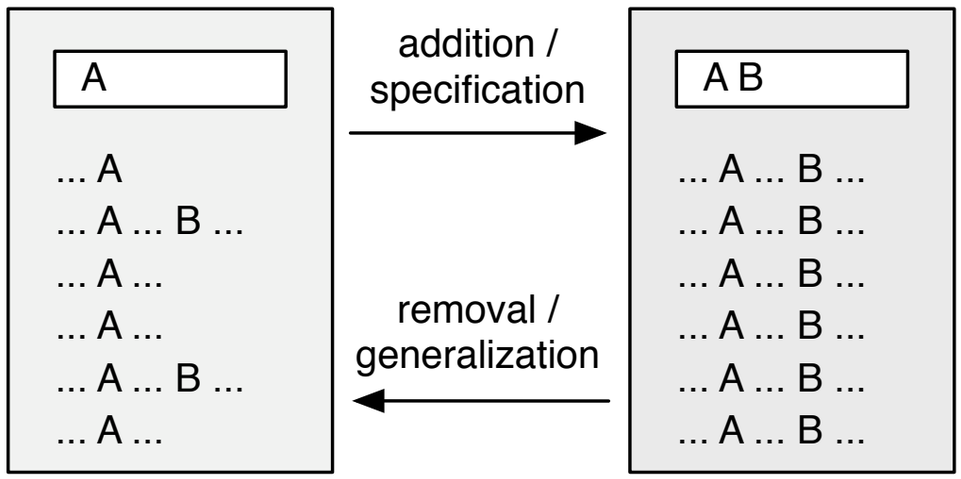
Query modifications

merlin	
merlin legend	term addition
merlin avalon	term substitution
merlin avalon arthur	term addition
avalon arthur	term removal
screenshot mac	different

We study: term additions and removals between consecutive query pairs

A commonly accepted interpretation

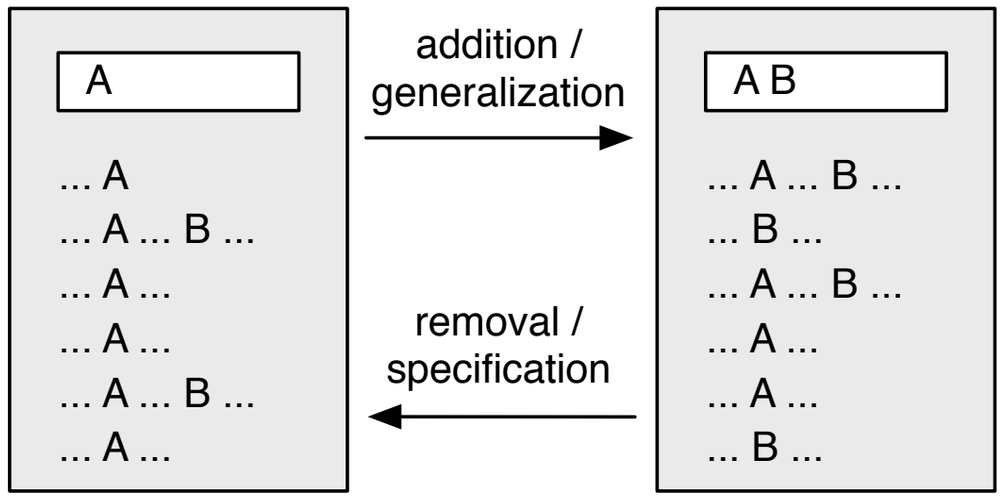
- An intersection-based interpretation



- It is valid if the retrieval system employs strict boolean operations, e.g., returned documents always contain *all* query terms
- Implicitly used in many studies (e.g., Boldi et al., 2010, Bruza and Dennis, 1997, Costa and Seco, 2008, He et al., 2002, Jansen et al., 2009)

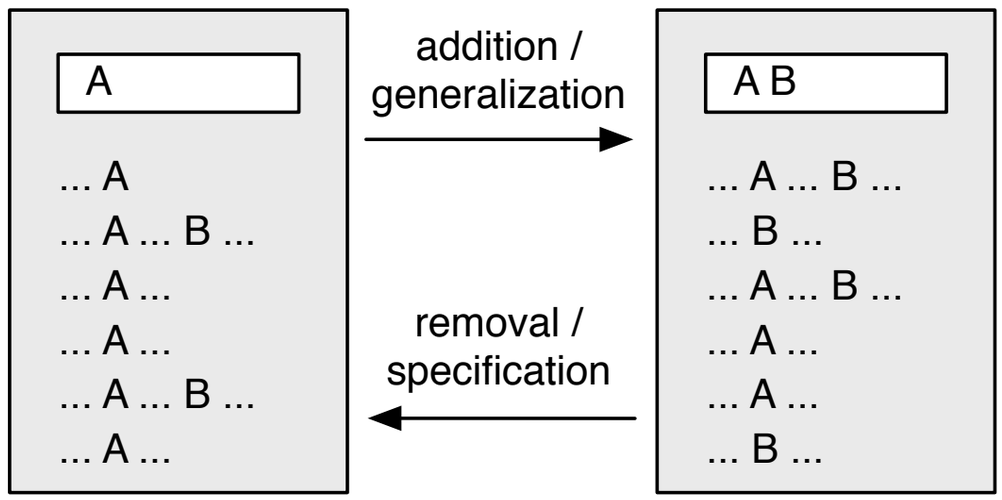
An alternative interpretation

- Modern search engines often return documents contain *some* of the query terms, i.e., non-boolean operations



An alternative interpretation

- An union-based interpretation
  - Removal may be used to get rid of non-relevant documents that contain all query terms
  - Addition may be used to include documents about the added term



Union-based interpretation: an example

shoulder pain and numbness in index finger

About 220,000 results (0.20 seconds)

Neck/**shoulder**/arm **pain**. Thumb, **index** and middle **finger**
www.medhelp.org/posts/...**shoulder**...**pain**...**index**...**finger**.../11901
5 Dec 2006 – is a lil crooket and the C shape in the back of my neck is
so he feels he can correct the problem. It doesn't seem to be working ...

Shoulder Blade **pain and numb index finger** - Neurology
www.medhelp.org/posts/.../**Shoulder**...**pain**-and...**index-finger**/.../123...
30 Apr 2010 – Went back to clinic for second round of **pain** meds. This
give me naflex, naproxen and steroids. Also Vicadin pills so I can sleep

nerve **pain shoulder numbness finger**
www.arthritis-treatment-and-relief.com/nerve-**pain-shoulder-numbne**...
nerve **pain shoulder numbness** finger. ... Enter your First Name (optio
Weakness will be felt in the biceps, with **numbness** along the thumb an

shoulder pain numbness in pointer **finger**
www.arthritis-treatment-and-relief.com/**shoulder-pain-numbness**-in-p...
Pain, **numbness**, or tingling in the hand, wrist, palm, thumb, or **first** thr
May complain of diffuse **numbness** in the forearm, upper arm, and **sho**

Numbness in index finger-under nail - Carpal Tunnel Syn
www.healthboards.com/.../547916-**numbness-index-finger**-under-nai...
28 Oct 2007 – **Shoulder** and Arm **Pain** with left pointer finger **numbnes**

numbness in index finger

About 299,000 results (0.11 seconds)

Paresthesia (**Finger Numbness**): Check Your Symptoms
www.medicinenet.com › home › arthritis center › arthritis az list
15+ items – **Numbness Fingers** Forum.
Arthritis — Arthritis is infla
Carpal Tunnel Syndrome And Tarsal Tunnel Syndrome — Carpal tunnel ..

Numbness Tingling
www.handhealthresources.com/.../**numbness**_tingling.htm
If you experience **numbness** and tingling in the thumb, **index finger**, n
and/or part of the ring finger, you may have a compression or irritation

Index Finger Numbness
www.buzzle.com/articles/**index-finger-numbness**.html
12 Oct 2011 – **Index finger numbness** can lead to tingling and pain in
fingers. Let's read more about symptoms, causes and treatment of this

Numbness in index finger left hand, please help ? - Yaho
answers.yahoo.com › ... › Pain & Pain Management
4 Feb 2012 – This happen out of no where. Its the same color it...

- Why is my right **index finger** numb? - 14 Feb 2011
- The tip of my middle **finger** is numb.? - 11 Jan 2010
- **Numbness** in my right middle **finger**? - 7 Jul 2007
- What causes **numbness** on my right **fingers** , thumb, pointer and ...

CWI

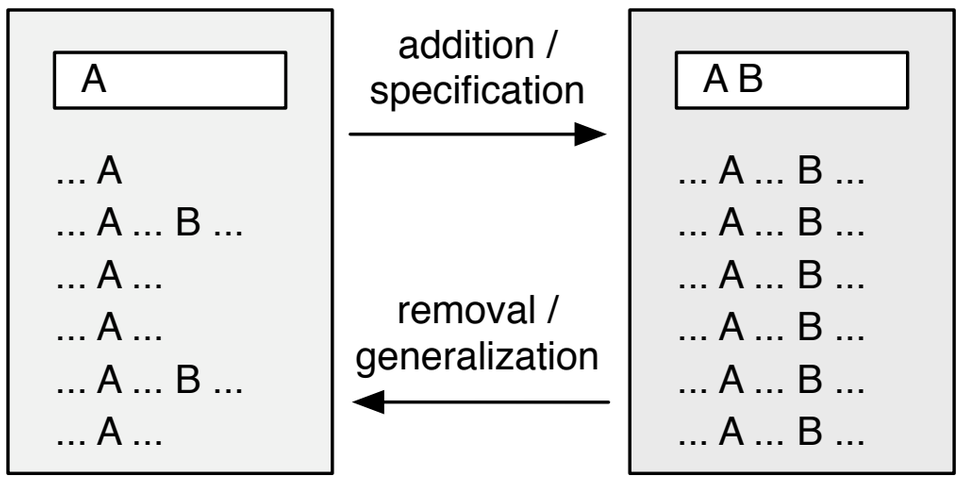
A research question

- How well can each of the two interpretations of term additions and removals explain the query modification behaviors of the searchers?

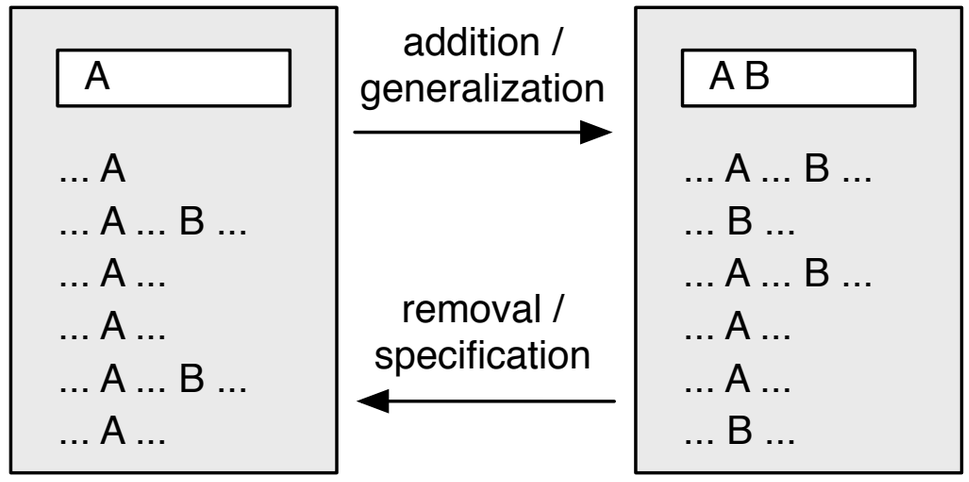
Method

- Assumptions

Intersection-based



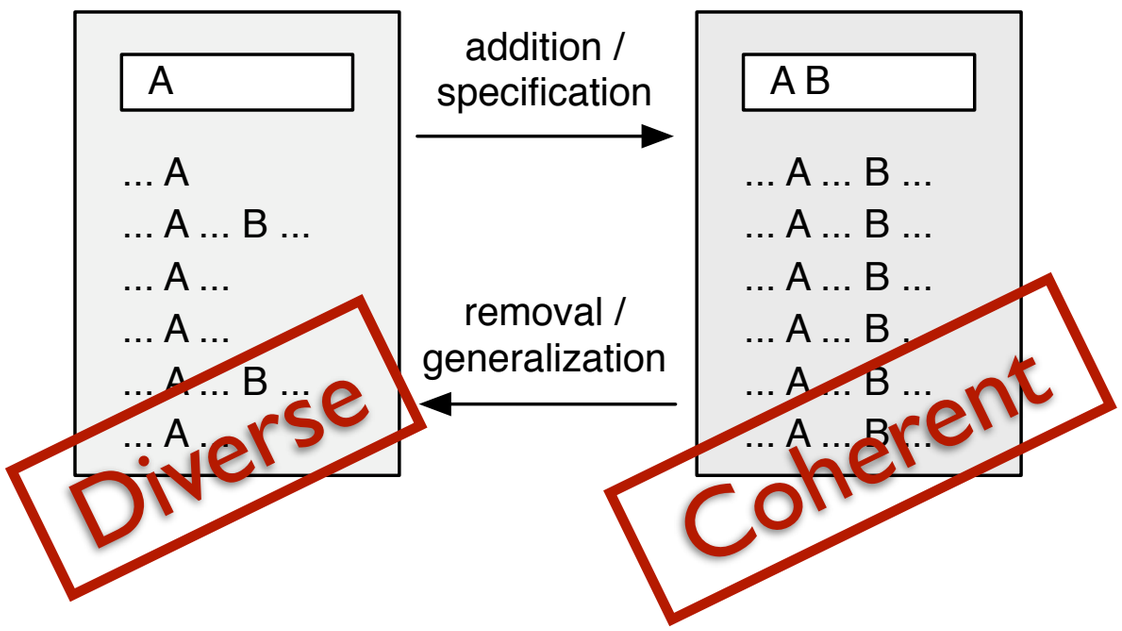
Union-based



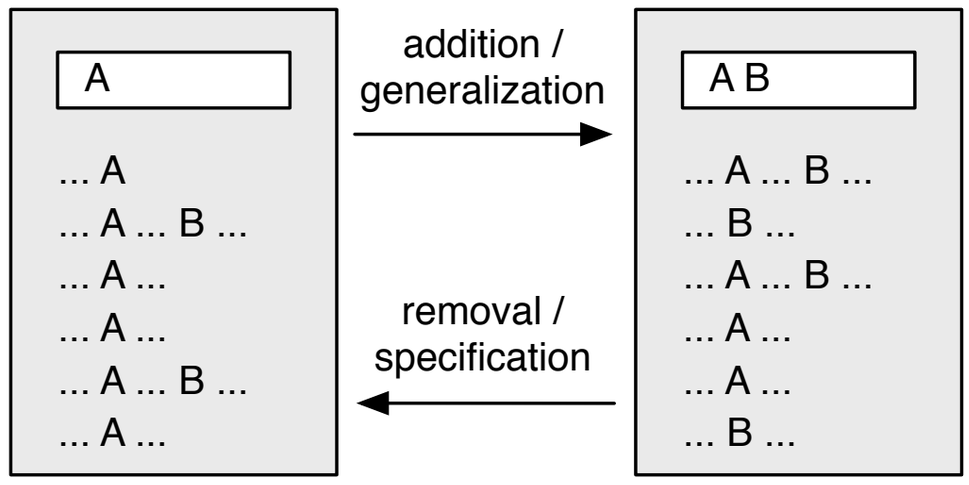
Method

- Assumptions

Intersection-based

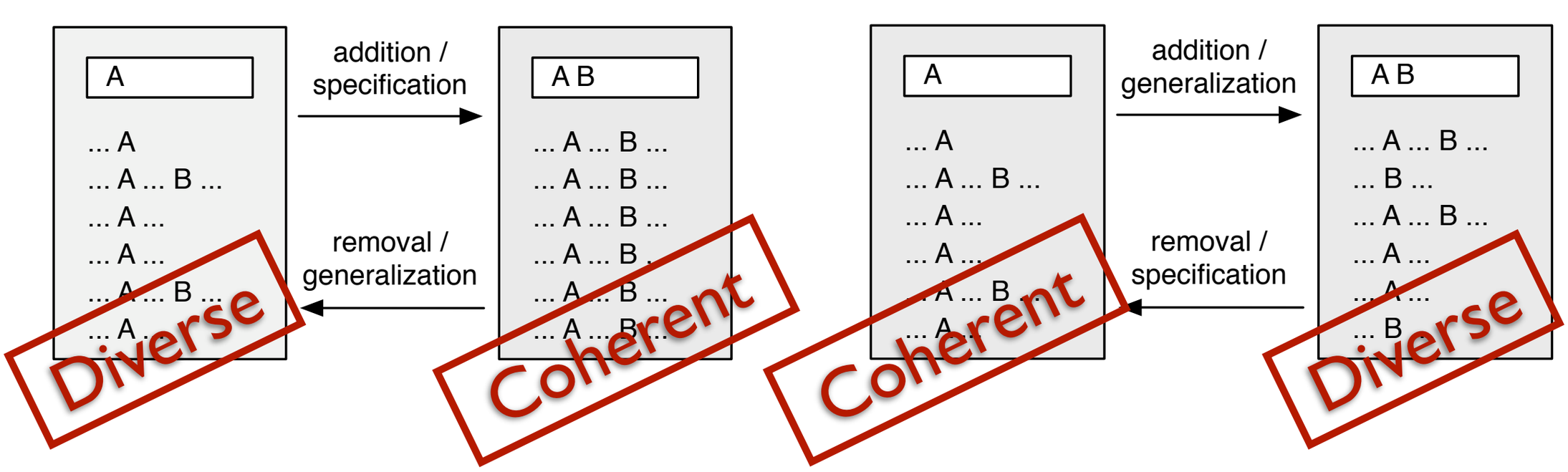


Union-based



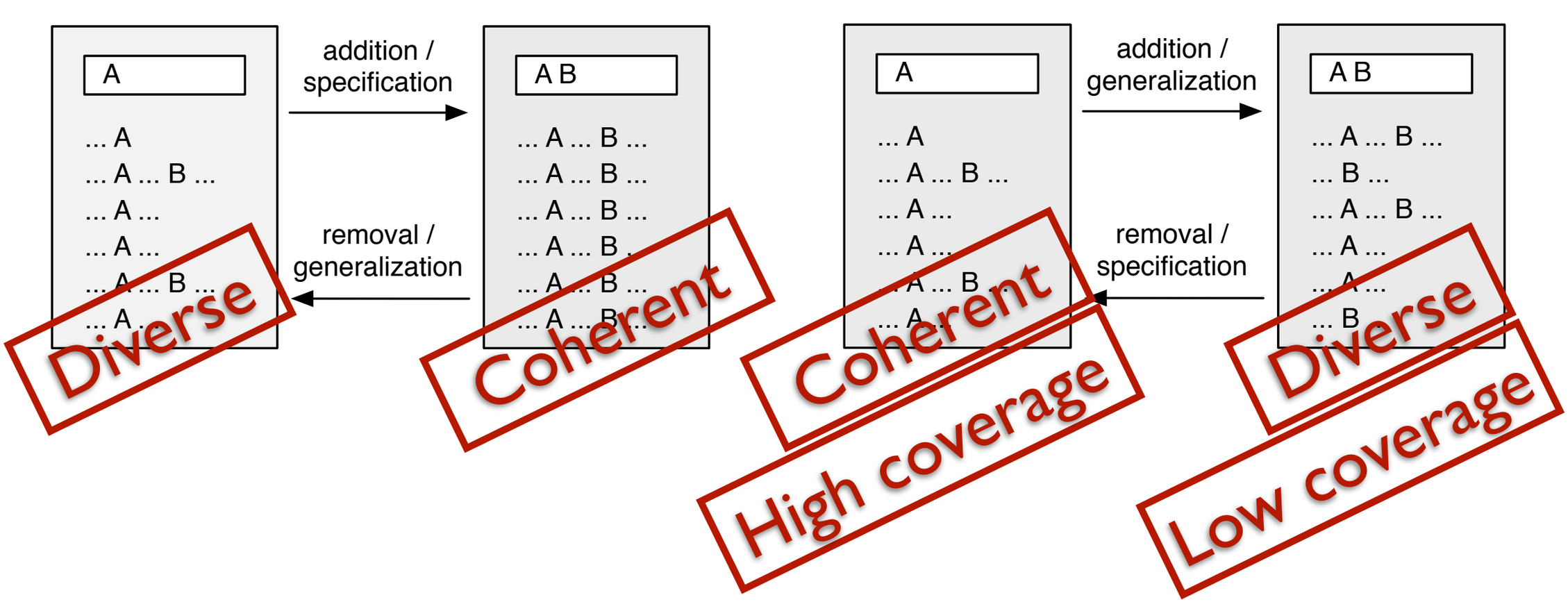
Method

- Assumptions



Method

- Assumptions



Method

- Empirical validate:
  - Do more coherent or less coherent result sets more often lead to term removals and term additions?
  - Do term removals and term additions increase or decrease the coherence of the result sets?
  - Do term removals often occur when many of the original result sets do not contain all query terms and term additions occur when all results do contain all terms?

Method

- Measuring coherence
  - Average similarity scores

$$AvgSim(D_q) = \frac{\sum_{i<j \in \{1,\ldots,N\}} Sim(d_i, d_j)}{\frac{1}{2}N(N-1)}$$

  - Coherence score (He et al. 2008)

$$Coherence(D_q) = \frac{\sum_{i<j \in \{1,\ldots,N\}} \sigma(d_i, d_j)}{\frac{1}{2}N(N-1)} \quad , \text{ where } \quad \sigma(d_i, d_j) = \begin{cases} 1 \text{ if } Sim(d_i, d_j) \geq \theta \\ 0 \text{ otherwise} \end{cases}$$

- Measuring query term coverage

$$Coverage(D_q) = \frac{|\{d | d \in D_q, T_q \subseteq T_d\}|}{|D_q|}$$

Experiments

- Data sets: query pairs (additions/removals) from 3 query logs

Logs	News	iCLEF 08/09	Web
All	556,007	49,174	20, 000
2 terms	282,039	15,713	4,842
>=2 terms	355,660	44,132	17,659

- Retrieval systems : top 16 documents are used as result set
  - News: lemur toolkit
  - iCLEF: FlickLing - a Flickr API
  - Web : Bing API
- A user study verifies that the coherence score agrees with human judgements in determining the coherency of a result set (Cohen's kappa = 0.70)

Validation of the two interpretations

- Do more coherent or less coherent result sets more often lead to term removals and term additions?

		Coherence		Avg Sim		Coverage	
Data		A	R	A	R	A	R
News	all	0.65	>> 0.56	0.56	>> 0.52	0.90	>> 0.29
	2 terms	0.66	>> 0.57	0.56	>> 0.52	0.78	>> 0.40
	>=2 terms	0.66	>> 0.56	0.56	>> 0.52	0.73	>> 0.29
iCLEF	all	0.94	>> 0.71	0.32	>> 0.29	0.80	>> 0.39
	2 terms	0.94	>> 0.73	0.34	>> 0.27	0.81	>> 0.51
	>=2 terms	0.94	>> 0.71	0.35	>> 0.29	0.75	>> 0.39
Web	all	0.68	>> 0.64	0.28	>> 0.27	0.69	>> 0.35
	2 terms	0.70	>> 0.58	0.29	>> 0.25	0.80	>> 0.61
	>=2 terms	0.73	>> 0.64	0.30	>> 0.27	0.64	>> 0.35

≫/≪ indicates significantly larger/smaller with p-value <0.01 using the Wilcoxon rank sum test.

Validation of the two interpretations

- Do term removals and term additions increase or decrease the coherence of the result sets?

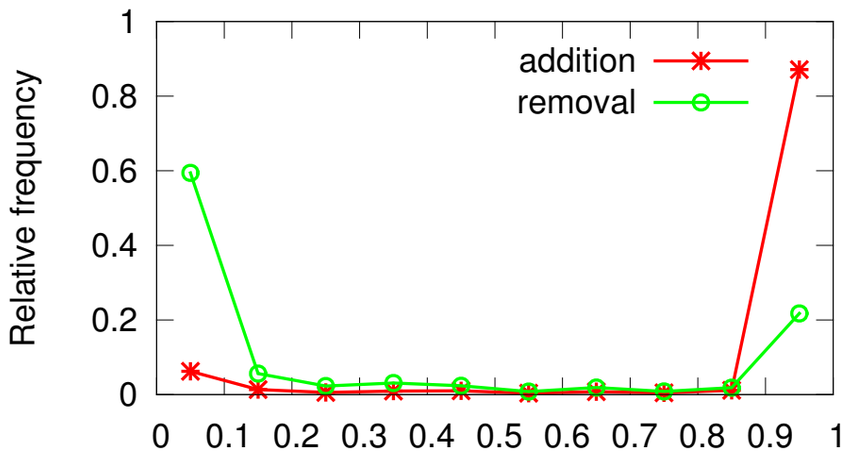
		Coherence			Avg Sim			Coverage		
Data		A		R	A		R	A		R
News	all	-0.035	<<	0.072	-0.016	<<	0.034	-0.449	>>	0.554
	2 terms	-0.031	<<	0.078	-0.012	<<	0.034	-0.455	>>	0.601
	>=2 terms	-0.031	<<	0.072	-0.013	<<	0.034	-0.424	<<	0.554
iCLEF	all	-0.138	<<	0.186	-0.012	<<	0.025	-0.282	<<	0.323
	2 terms	-0.151	<<	0.190	-0.029	<<	-0.015	-0.296	<<	0.406
	>=2 terms	-0.148	<<	0.186	-0.033	<<	0.025	-0.278	<<	0.323
Web	all	-0.013	<<	0.039	0.002	<<	0.010	-0.320	<<	0.337
	2 terms	-0.024	>>	-0.08	-0.000	>>	-0.042	-0.384	<<	0.256
	>=2 terms	-0.054	<<	0.039	-0.014	<<	0.010	-0.321	<<	0.338

≫/≪ indicates significantly larger/smaller with p-value <0.01 using the Wilcoxon rank sum test.

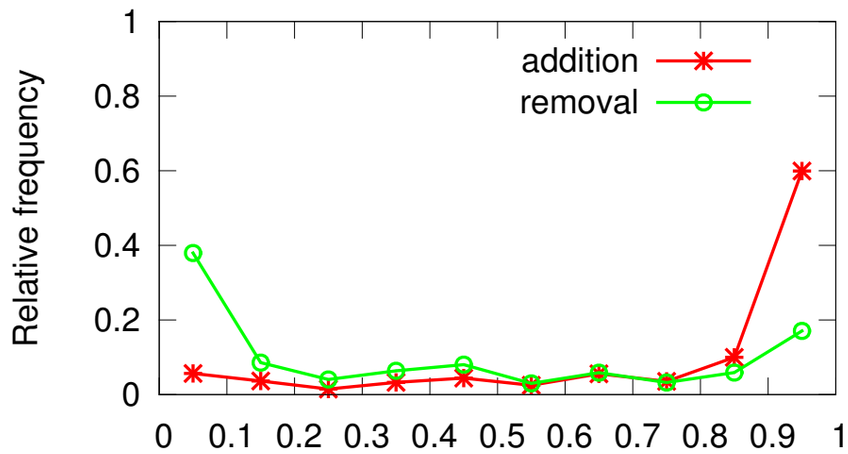
Validation of the two interpretations

- Query term coverage

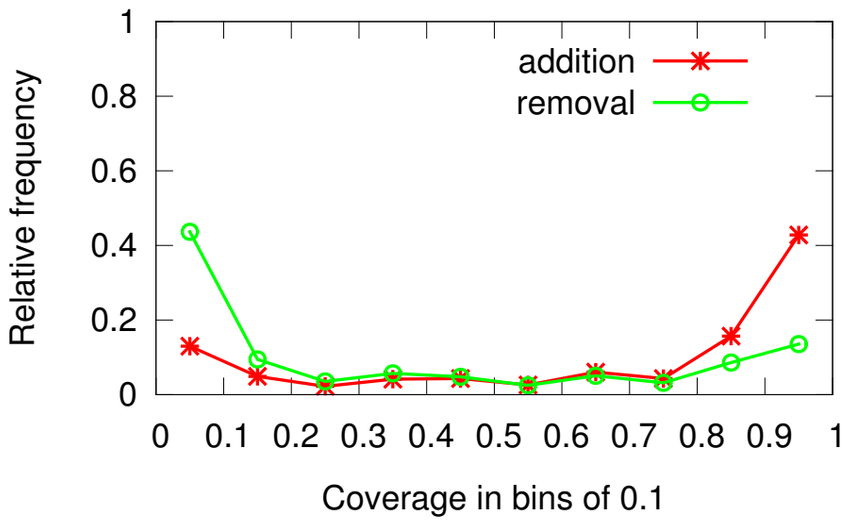
News



iCLEF



Web



Conclusion

- We presented a method to study the relation between query modification and result set coherence
- The widely accepted intersection-based interpretation is not always valid
- An union-based interpretation provides alternative explanation to query modifications
- Implication: log analysis based purely on intersection-based interpretation may lead to biased view on the intentions behind query modifications

References

1. Boldi, P., Bonchi, F., Castillo, C., Vigna, S.: Query reformulation mining: models, patterns, and applications. Information Retrieval 14(3), 257–289 (2010)
2. Bruza, P., Dennis, S.: Query reformulation on the internet: empirical data and the hyperindex search engine. In: RIAO'97. pp. 488–499 (1997)
3. Costa, R.P., Seco, N.: Hyponymy extraction and web search behavior analysis based on query reformulation. In: IBERAMIA'08 (2008)
4. He, D., Goker, A., Harper, D.J.: Combining evidence for automatic web session identification. Information Processing and Management 38(5), 727–742 (2002)
5. Jansen, B.J., Booth, D.L., Spink, A.: Patterns of query reformulation during web searching. JASIST 60(7), 1358–1371 (2009)

CWI

Reliability of the coherence score (skip)

- User study with 3 subjects
- 120 modification pairs
  - for each pair we compute the coherence score of the original and the modified queries and the difference of the two
    - 40 pairs with strong positive difference
    - 40 pairs with strong negative difference
    - 40 paris with roughly equal coherence
- User task
  - Classify the result sets of the 120 pairs into 3 categories:
    - “less specific”, “more specific”, “equally specific”, “incomparable”

Reliability of the coherence score

- Majority vote is used to determine the label of the query pairs
  - 10 cases where majority label is not available
  - 24 cases where majority label is “incomparable”
- Results
  - 80% agreement between coherence score and user judgements (Cohen’s kappa = 0.70)

	Coherence scores		
Majority	Strong negative	Roughly equal	Strong positive
More specific	**23**	3	0
Equal specific	3	**26**	6
Less specific	0	5	**20**