Harnessing Semantics for Answer Sentence Retrieval

Ruey-Cheng Chen, Damiano Spina, W. Bruce Croft[†],
Mark Sanderson, and Falk Scholer

RMIT University
[†] University of Massachusetts Amherst

Oct 23, 2015 (ESAIR '15)

Outline

Task

Answer sentence retrieval: What’s the challenge?

Methodology

Lexical and semantic matching

Learning-to-rank approach

Result

A ranking experiment

Answer Sentence Retrieval

Some questions solicit multiple-sentences answers

- Non-factoid;
- Verbose;
- Not necessarily posed as effective queries.

Research questions:

i. *What best characterizes "answer finding" as a specialized retrieval task?*
ii. *Will incorporating <u>semantics</u> result in better performance?*

"What was the role of Portugal in World War II?"

Google During World War II, the Portuguese Republic was an authoritarian political regime under António de Oliveira Salazar and the Estado Novo, often regarded as pro-fascist. Although Portugal was officially a neutral country, it exported goods to the Allies as well as Germany and other neutral countries.

Annotated answers from GOV2

<u>Perfect</u> Portugal supplied a variety of vital mineral resources for the Third Reich's war machine, including the ore for tungsten, a key additive used in the production of weapon-grade steel.

<u>Excel</u> The Portuguese Government allowed Jewish organizations to relocate from Occupied Europe to Lisbon during the war.

<u>Good</u> Portugal only provided some $4 million of the some $51 million the Allies initially sought with negotiations dragging on throughout the 1950s.

<u>None</u> Angola Civil war has been the norm in Angola since independence from Portugal in 1975.

Matching Questions to Answers

Lexical

Synonymy

Contextual

Structural

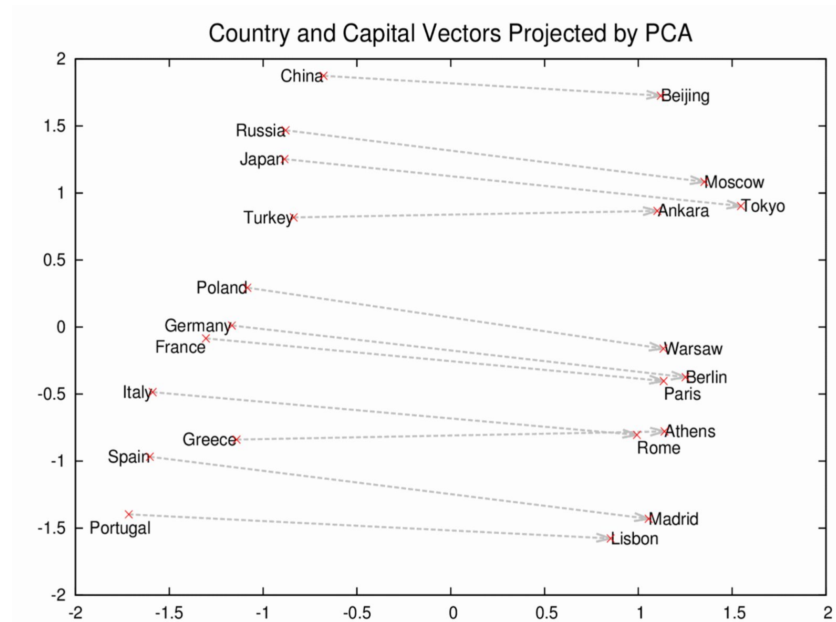
Factual

⇐ What we’re looking for

Semantic Representations

Explicit Semantic Analysis

Word2Vec



Gabrilovich and Markovitch. 2007. "Computing Semantic Relatedness Using Wikipedia-based Explicit Semantic Analysis." In *IJCAI*, vol. 7, pp. 1606-1611.

Mikolov et al. 2013. "Distributed representations of words and phrases and their compositionality." In *Advances in neural information processing systems*, pp. 3111-3119.

Features

Lexical Features (Metzler & Kanungo, 2008)

- SentenceLength
- SentenceLocation
- ExactMatch
- TermOverlap
- SynonymOverlap
- LanguageModel

Semantic Features

- ESACosineSimilarity
- Word2Vec

1. Build index over a recent dump of English Wikipedia
2. Retrieve top-100 concepts for each sentence/query
3. Compute distance by cosine similarity

1. Use the pretrained 100B word model on Google News
2. Compute distance by

$$\frac{1}{|Q||S|}\sum_{\vec{u}\in Q}\sum_{\vec{v}\in S}\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}$$

Learning to Rank

Collection	WebAP Dataset (82 topics, 991K sentences)
Feature Extraction	• Metzler-Kanungo: 6 features • Semantic: 2 features (ESA and word2vec)
Rankers	• Coordinate Ascent • MART

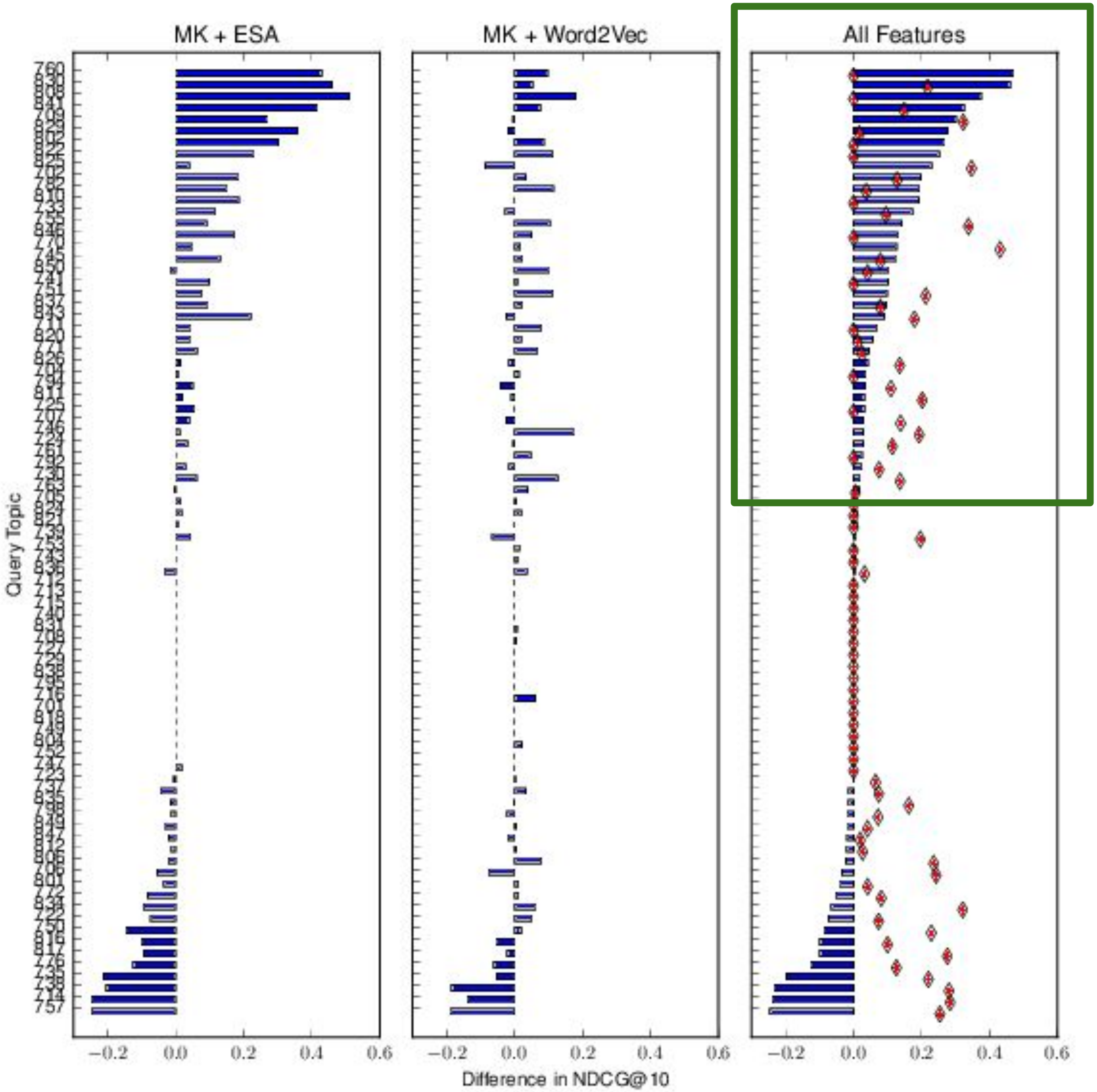
WebAP Dataset https://ciir.cs.umass.edu/downloads/WebAP/
RankLib http://sourceforge.net/p/lemur/wiki/RankLib/

Result: Single Features

Feature	NDCG@10	P@10	MRR
SentenceLength	0.0036	0.0037	0.0174
SentenceLocation	0.0000	0.0000	0.0056
ExactMatch	0.0194	0.0220	0.0529
TermOverlap	0.0618	0.0622	0.1978
SynonymOverlap	0.0272	0.0293	0.1058
LanguageModel	0.0721	0.0866	0.1980
ESACosineSimilarity	**0.1053**	**0.1171**	**0.2690**
Word2Vec	0.0634	0.0720	0.1924

Result: Learning to Rank

Feature Set	Algorithm	NDCG@10	P@10	MRR
MK	Coordinate Ascent	0.0667	0.0788	0.1954
MK + ESA		0.1080*	0.1221*	0.2694
MK + Word2Vec		0.0810	0.0936	0.2278
All Features		**0.1114****	**0.1240***	**0.2778**
MK	MART	0.0603	0.0699	0.1754
MK + ESA		0.0994**	0.1119*	0.2404
MK + Word2Vec		0.0699	0.0769	0.1985
All Features		0.0953*	0.1088**	0.2363


MK + ESA
MK + Word2Vec
All Features
Query Topic
Difference in NDCG@10
−0.2
0.0
0.2
0.4
0.6

Run	Rel	Top-1 Sentence
711: *What security measures have been employed at train stations due to heightened security concerns?*		
MK	0	The biggest concern in the minds of security personnel is the possibility of a person boarding a bus or train with a gun or other weapon.
MK+ESA	0	Two major cooperatives in the fertilizer industry, Farmland and CF Industries,have always been aware of potential security concerns, but both have increased their guard as security threats have become a heightened concern in a post-Sept. 11 world.
MK+Word2Vec	3	"Amtrak responded admirably to the crisis, quickly training personnel on heightened security and safety procedures, assigning more security officers to stations and trains, and requiring passengers to bring photo identifications for security checks," Schumer wrote.
770: *What is the state of Kyrgyzstan-United States relations?*		
MK	0	(3) Kyrgyzstan concluded a bilateral investment treaty with the United States in 1994.
MK+ESA	4	The extension of unconditional normal trade relations treatment to the products of Kyrgyzstan will enable the United States to avail itself of all rights under the World Trade Organization with respect to Kyrgyzstan.
MK+Word2Vec	0	(begin text) U.S. DEPARTMENT OF STATE Office of the Spokesman January 15, 2002 Media Note RELIGIOUS LEADERS FROM KYRGYZSTAN EXAMINE ISLAM IN THE UNITED STATES

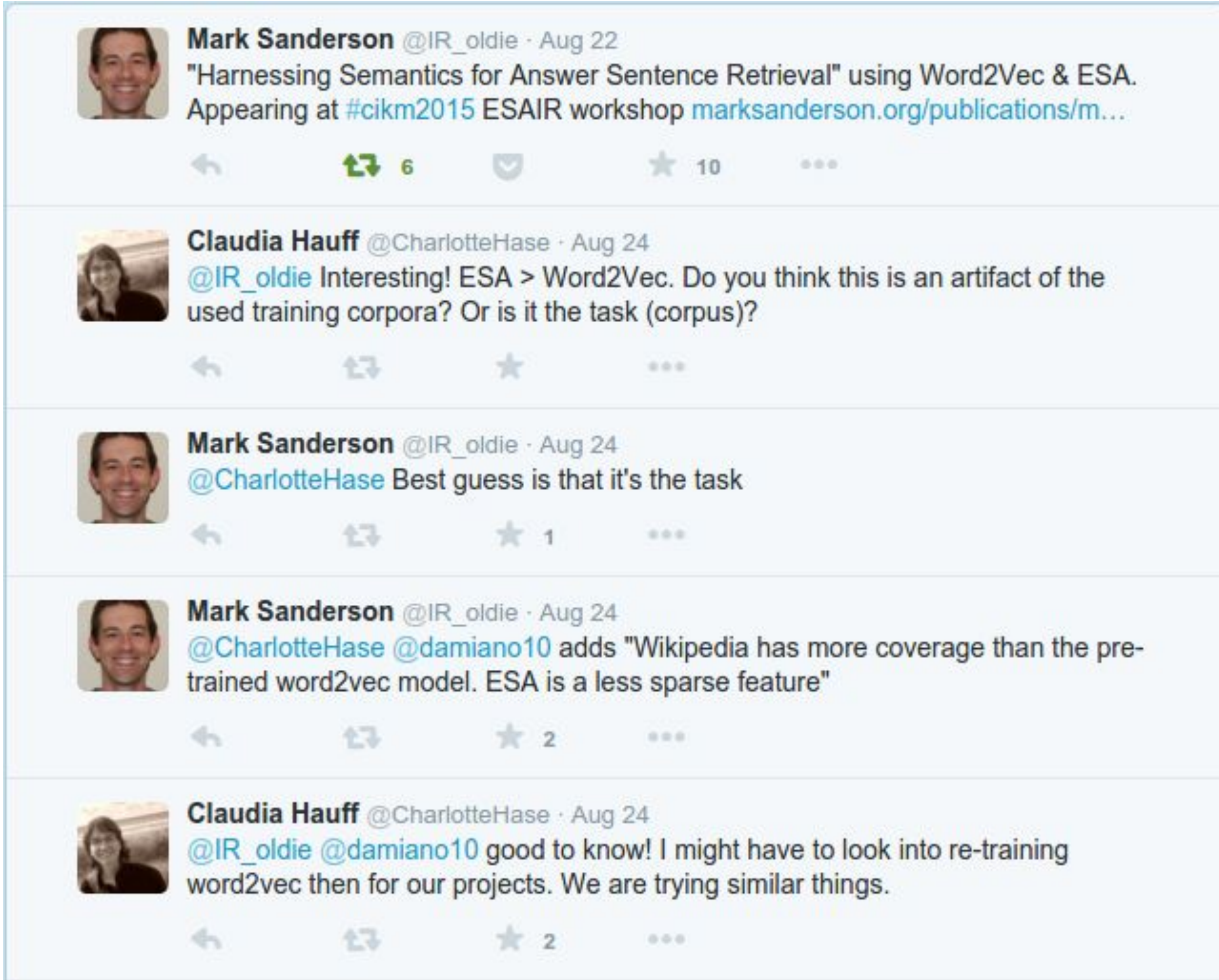
Feedbacks

“It is somewhat disappointing that the improvements over ESACosineSimilarity are rather limited.”

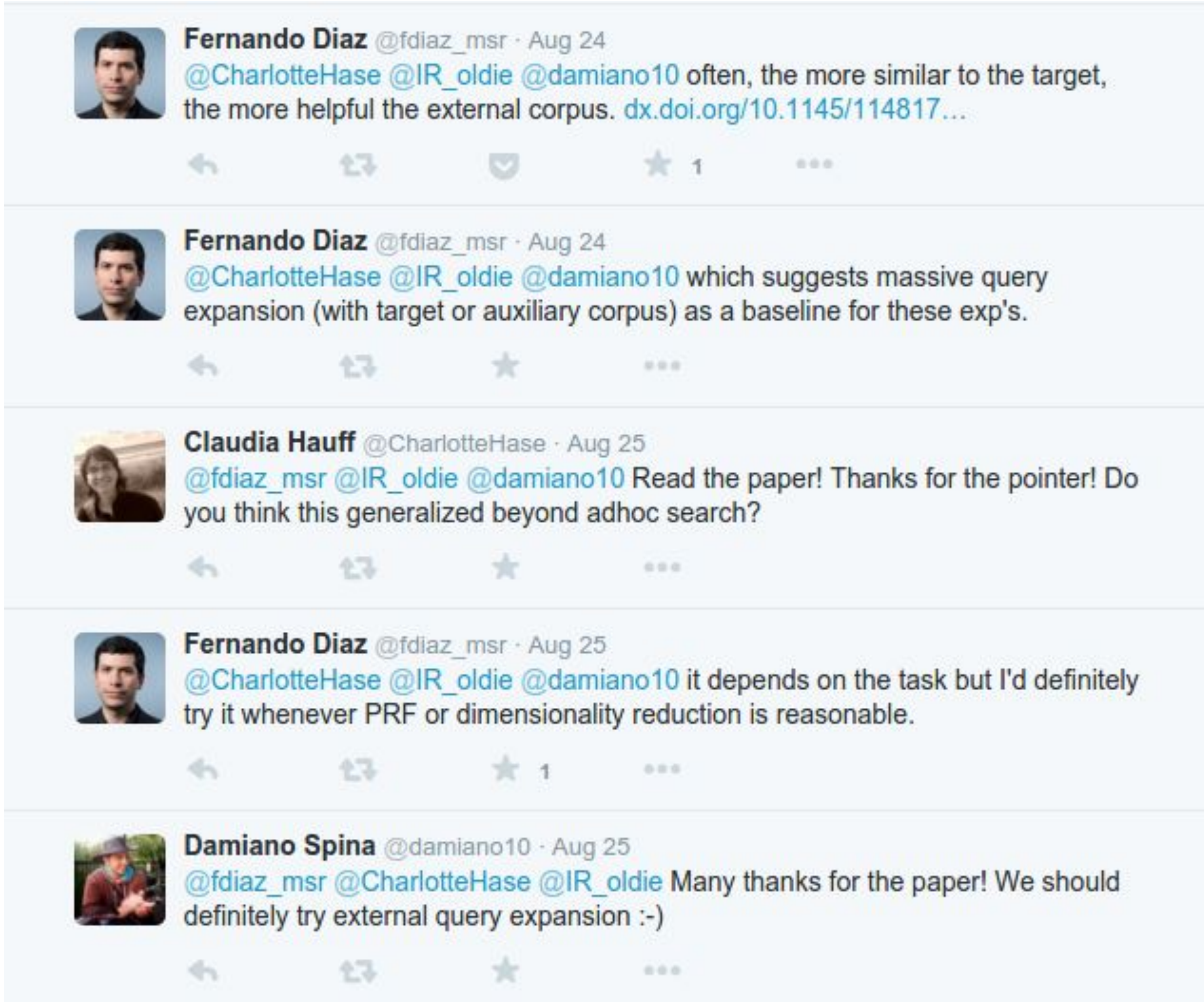
Author 1: I share your feeling

“Enhance the by-topic analysis and more detailed analysis…”

Author 2: Don’t wanna read too much out of the data.



ESA remains the single most useful feature as of today



Query expansion is your baseline.

Conclusions

Significant improvements via adding both features

- And there's still room for improvement

What to expect next?

Enhancements: Linked entities, substructures

Tasks: Snippet generation, summarization

Thank You!

Harnessing Semantics for Answer Sentence Retrieval

Ruey-Cheng Chen, Damiano Spina, W. Bruce Croft[†],
Mark Sanderson, and Falk Scholer

RMIT University
[†] University of Massachusetts Amherst