Enriched syntax-based meaning representation for answer extraction

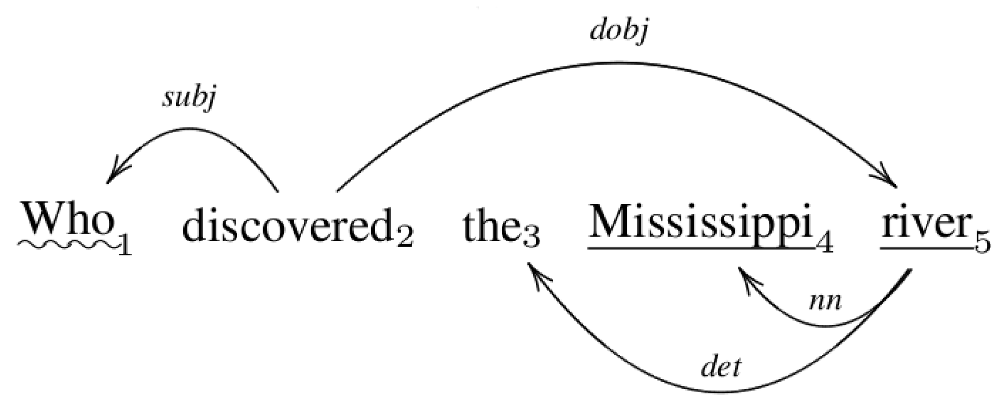
Grzegorz Chrupała, Georgiana Dinu and Benjamin Roth
Spoken Language Systems and Computational Linguistics
Saarland University, Germany

Introduction

- Challenge in QA: similar meaning expressed with different surface realizations in questions and in answer sentences.
- Propose an enriched syntax-based representation to provide a degree of generalization.
- Encode uncertainty about syntax of the question by using multiple alternative dependency parse trees.
- Augment question meaning representation by including multiple DIRT-like paraphrases of each dependency path.

Answer extraction and ranking

- Question parse tree



- Example question path

$$\text{who}_1 \xleftarrow{subj} \text{discover} \xrightarrow{dobj} \text{river} \xrightarrow{nn} \text{Mississippi}_4$$

- Example answer path

$$\text{Soto}_9 \xleftarrow{pobj} \text{by} \xleftarrow{prep} \text{discover} \xrightarrow{dobj} \text{river} \xrightarrow{nn} \text{Mississippi}_2$$

- Anchor: Mississippi
- Answer candidate: Soto

Rank an answer candidate according to the perplexity of the bigram LM trained on the set of matching paths in question. Perplexity is discounted to give some preference to shorter paths:

$$PP_Q(\pi) = 2^{-\frac{1}{N+d}\sum_{i=1}^{N} \log_2 P_Q(\pi_i|\pi_{i-1})}$$

where π is the sentence path of length N, and $P_Q(\pi_i|\pi_{i-1})$ is the probability of the i^{th} path element according to LM.

N-best parsing and paraphrasing

- Represent a question by the set of dependency paths extracted from n-best parses.
- Each question path is additionally paraphrased in multiple ways.
- Acquire DIRT-style paraphrases from 100-million-word of Gigaword using (Lin and Pantel 2001):

$$\mathrm{sim}_{\mathrm{Lin}}(\mathbf{w}, \mathbf{v}) = \frac{\sum_{i \in I(\mathbf{w}) \cap I(\mathbf{v})} (w_i + v_i)}{\sum_{i \in I(\mathbf{w})} w_i + \sum_{i \in I(\mathbf{v})} v_i}$$

- The most similar paraphrases for *X discover Y*:

$$\mathbf{X} \xleftarrow{subj} \textbf{discover} \xrightarrow{dobj} \mathbf{Y}$$

$$\text{X} \xleftarrow{pobj} \text{by} \xleftarrow{prep} \text{discover} \xrightarrow{subjpass} \text{Y}$$

$$\text{X} \xleftarrow{subj} \text{find} \xrightarrow{dobj} \text{Y}$$

$$\text{X} \xleftarrow{pobj} \text{by} \xleftarrow{prep} \text{find} \xrightarrow{subjpass} \text{Y}$$

$$\text{X} \xleftarrow{subj} \text{unearth} \xrightarrow{dobj} \text{Y}$$

$$\text{X} \xleftarrow{subj} \text{uncover} \xrightarrow{dobj} \text{Y}$$

Evaluation

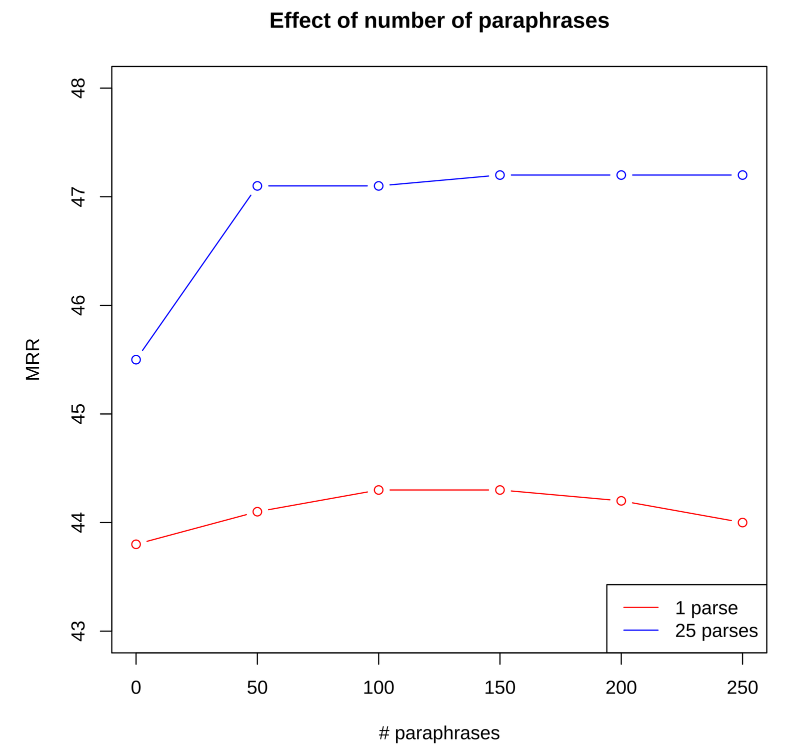
- Data set: QASP (Kaisser and Lowe 2008). For subset of TREC 02-06 questions provides answer sentences and answer strings.
- **Sentence level** instead of **question level** task
- Results on development set (2002). Exp(n,m) – experiment using n parse trees and m paraphrases.

Configuration	MRR
Exp(1,0)	43.7
Exp(1,100)	44.3
Exp(25,0)	45.5
Exp(25,100)	47.2

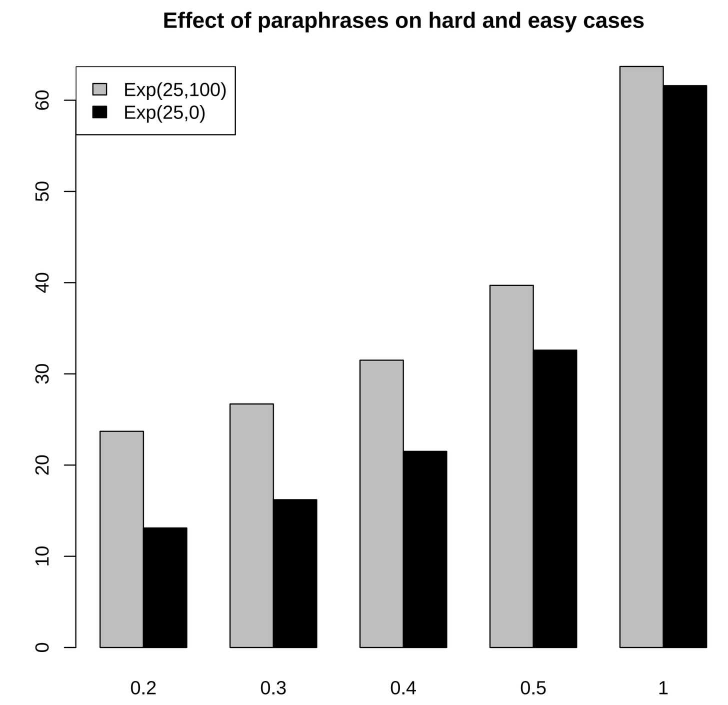
Analysis

- Varying the number of parses (no paraphrasing)

# parses	1	5	10	15	20	25	30
MRR	43.8	45.3	45.3	45.3	45.4	45.5	45.4



Effect of varying the number of paraphrases with 1-best parse and 25-best parses settings.



Effect of paraphrasing on sentences of different difficulty. Each bin corresponds to the sentences where no-para configuration had RR $\leq x$

- Cross-validation on 2002-2006, tuning number of parse trees and paraphrases on 1 year in turn.

Model	MRR
Baseline	39.03
Enriched	40.64

Conclusion

- Paraphrasing together with alternative question parses allows us to enhance question meaning representation.
- This brings improvements in answer extraction MRR scores, particularly for sentences which are difficult for the baseline syntactic method.

Acknowledgements

Grzegorz Chrupała was funded by the BMBF project NL-Search under contract number 01IS08020B. Georgiana Dinu was funded by the IRTG PhD program.