Semantic Kernels for Semantic Parsing

Iman Saleh -- Cairo University
Alessandro Moschitti, Preslav Nakov, Lluís Màrquez, Shafiq Joty -- Qatar Computing Research Institute

1. Motivation

How do we convert a spoken request like “cheap Italian restaurants in Doha with take out” into a database query?

Processing Steps

Text/Speech semantic parser

Semantic representation:

- "Price": "low"
- "City": "Doha"
- "Cuisine": "Italian"

DB query:

{"amenity": "restaurant", "cuisine": "italian", "name": "...", "price": "...", "type": "...

2. State-of-the-art Semantic Parser

Semi-Markov CRFs (Sarawagi & Cohen 04)
- Joint sequential segmentation/classification
- Discriminative probabilistic sequential model
- Undirected graphical model

3. Our Approach: Reranking with Semantic Kernels

- Semi-CRF generates n-best hypotheses (H_

- Pairwise reranking function trained with SVMs with the preference reranking kernel
- Hypotheses are represented with tree-based structures (tree kernels are applied)
- Focus on kernels using semantics: LSA-based smoothing lexical similarity and Brown clusters

Partial Tree Kernels (PTK)

- Soft matching of tree fragments via similarity
- Node similarity derived using LSA

Smoothing PTK or Semantic Kernel (SK)

- SK Structures
- SK + Brown Clusters (SK+BC)

Semantic Structures for CSL Reranking

- Tree Kernels capture structural similarities between trees
- Brown cluster IDs in the structures capture semantic dependencies between labels
- LSA-smoothed SKs enable to measure the similarity between tree fragments (based on LSA word similarity), thus powerful semantic patterns are generated

Extra features in a flat vector (+all)

- From the semi-CRF: probability of label sequences, label sequence n-grams, DB-based, etc.

4. Experiments

- Dataset: human annotations (McGraw et al. 12)
- Oracle: shows large room for improvement
- Brown clusters: from Yelp restaurants reviews
- Lexical similarity for SK: LSA on TripAdvisor reviews

Main results

5. Conclusions

- The LSA-smoothed semantic kernel (SK) improves significantly over the semi-CRF ("baseline") and over our previous state-of-the-art reranker, which uses shallow syntactic patterns and PTK (Saleh et al., COLING-2014; equivalent to PTK+all)
- ~10% relative improvement
- BCs do not significantly improve any model
- PTK+all is better than PTK but its accuracy is lower than any SK
- +all helps SK only with small sizes of the training set

Note: the state of the art on the task is very hard to beat

6. Future Work

- Explore semantic similarity from distributional and other sources, e.g., Wikipedia, Wiktionary, WordNet, FrameNet, BabelNet, and LSA for different domains.

Acknowledgements: This research is developed by the Arabic Language Technologies (ALT) group at Qatar Computing Research Institute (QCRI) within the Qatar Foundation in collaboration with MIT. It is part of the Interactive Systems for Answer Search (iSearch) project.