DERIVING LOCAL RELATIONAL SURFACE FORMS FROM DEPENDENCY-BASED ENTITY EMBDIEDEMENTS FOR UNSUPERVISED SPOKEN LANGUAGE UNDERSTANDING

Yun-Nung (Vivian) Chen, Dilek Hakkani-Tür, and Gokhan Tur

1. Introduction

Main Idea
- Spoken Language Understanding (SLU): convert ASR outputs into pre-defined semantic format
- Relation: semantic interpretation of input utterances
- Unsupervised SLU: utilize external knowledge to help relation detection without labelled data

Semantic Interpretation via Relations

User Utterance: find movies produced by james cameron

Knowledge Graph: graph with
- strongly typed and uniquely identified entities (nodes)
- facts/literals connected by relations (edges)

2. Framework

“find me some films directed by james cameron”

Input Utterance

Proven increase the relation detection accuracy.

3. Relation Inference from Gazetteers (Entity List)

If the relation is not included by the entity, how to detect such local relational expressions/surface forms?

4. Local Surface Form Derivation

Given the Knowledge Graph resource and collection of unlabelled raw data, can we use the information to automatically detect relations of entities to perform SLU in an unsupervised way?

5. Probabilistic Enrichment & Bootstrapping

- Integrated Relations for Words by keeping the most possible relations
- Bootstrapping: train a multi-label multi-class classifier estimating relations given an utterance

6. Experiments

Knowledge Base: Freebase (movie domain)
- 670K entities
- 78 entity types

- Relation Detection Data
  - Crowd-sourced utterances
  - Manually annotated with SPARQL queries → relations

- The best result is the combination of all approaches, since prob. from different resources can complement each other.

- Only adding entity surface forms performs similarly, showing that the major improvement comes from entity surface forms.

Boosting significantly improves performance.


7. Conclusion

- We propose an unsupervised approach to capture the relational surface forms including entity surface forms and entity contexts based on dependency-based entity embeddings.
- The detected relations viewed as local observations can be integrated with background knowledge by probabilistic enrichment methods.
- Experiments show that involving derived relational surface forms as local cues together with prior knowledge can significantly improve the relation detection task and help open domain SLU.

yvchen@cs.cmu.edu, vivian.ychen