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

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

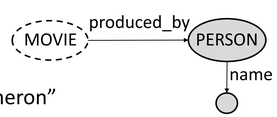
Carnegie Mellon

Microsoft Research

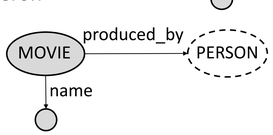
1. Introduction

- Main Idea
  - **Spoken Language Understanding (SLU)**: convert ASR outputs into pre-defined semantic format

    "when was james cameron's avatar released" → Intent: FIND_RELEASE_DATE; Slot-Val: MOVIE_NAME="avatar", DIRECTOR_NAME="james cameron"
  - **Relation**: semantic interpretation of input utterances
    movie.release_date, movie.name, movie.directed_by, director.name
  - **Unsupervised SLU:** utilize external knowledge to help relation detection without labelled data
- Knowledge Graph: graph with
  - strongly typed and uniquely identified entities (nodes)
  - facts/literals connected by relations (edges)
- Semantic Interpretation via Relations



User Utterance: find movies produced by james cameron
Logical Form: λx. ∃y. movie.produced_by(x, y) ∧ person.name(y, z) ∧ z="James Cameron"
Relation: movie.produced_by producer.name

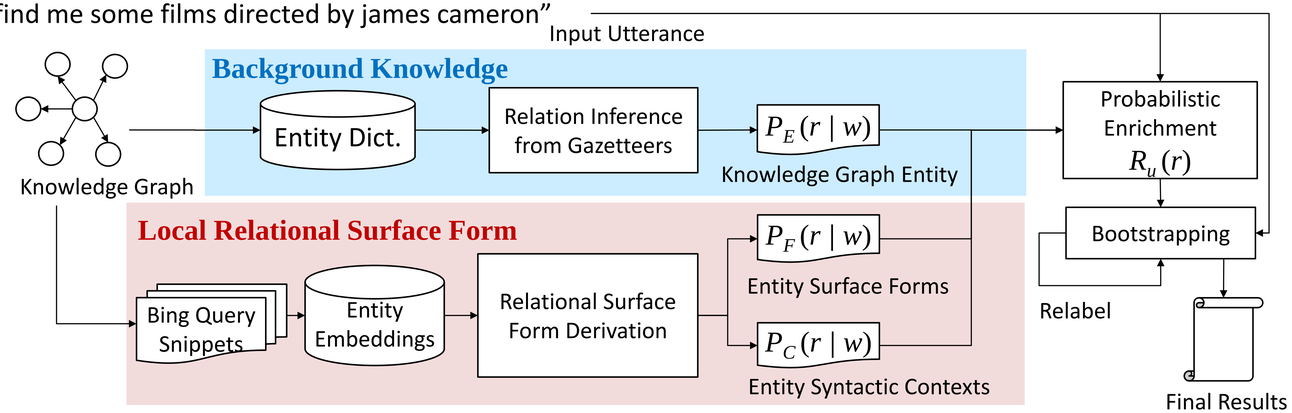


User Utterance: who produced avatar
Logical Form: λy. ∃x. movie.produced_by(x, y) ∧ movie.name(x, z) ∧ z="Avatar"
Relation: movie.name movie.produced_by

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

2. Framework

"find me some films directed by james cameron"



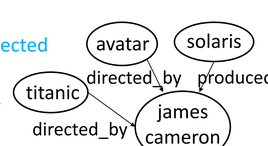
3. Relation Inference from Gazetteers (Entity List)

r: relation; w:observed word; t: entity type

"james cameron" #movies James Cameron directed

$$P_E(r_i \mid w) = \frac{C(w, t_i)}{\sum_{t_k \in T(w)} C(w, t_k)}$$

r_i: movie.directed_by, director.name; t_k: director

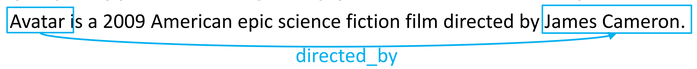


➢ If the relation is not included by the entity, how to detect such local relational expressions/surface forms?

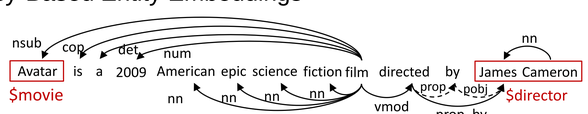
- Dilek Hakkani-Tur, Asli Celikyilmaz, Larry Heck, and Gokhan Tur, Probabilistic enrichment of knowledge graph entities for relation detection in conversational understanding, in *Interspeech*, 2014.

4. Local Surface Form Derivation

- Bing query snippets including entity pairs connected with specific relations in KG



- Dependency-Based Entity Embeddings



  - Each word w is associated with a vector v_w and each context c is represented as a vector v_c
  - Learn vector representations for both words and contexts such that the dot product $v_w \cdot v_c$ associated with good word-context pairs belonging to the training data D is maximized
- Entity Surface Forms
  - learn the surface forms corresponding to entities

$$S_i^F(w_j) = \frac{\text{sim}(w_j, e_i)}{\sum_{e_k \in E} \text{sim}(w_j, e_k)} \quad P^F(r_i \mid w_j)$$

based on word vector v_w

$char, $director, etc.

$char: "character", "role", "who"
$director: "director", "filmmaker"
$genre: "action", "fiction"
→ with similar contexts

  - learn the important contexts of entities

$$S_i^C(w_j) = \frac{\text{sim}(\hat{w_j}, e_i)}{\sum_{e_k \in E} \text{sim}(\hat{w_j}, e_k)} \quad P^C(r_i \mid w_j)$$

based on context vector v_c

$char: "played"
$director: "directed"
→ frequently occurring together

➢ The relational surface forms can be extracted by similar dependency contexts and dependency-based neighboring words.

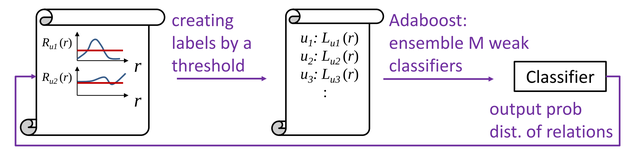
Word	Contexts
$movie	film/nsub^{-1}
is	film/cop^{-1}
a	film/det^{-1}
2009	film/num^{-1}
american, epic, science, fiction	film/nn^{-1}

Word	Contexts
film	film/nsub, is/cop, a/det, 2009/num, american/nn, epic/nn, science/nn, fiction/nn, directed/vmod
directed	$director/prep_by
$director	directed/prep_by^{-1}

5. Probabilistic Enrichment & Bootstrapping

- Integrated Relations for Words by keeping the most possible relations
- Integrated Relations for Utterances by $R_u(r_i) = \max_{w \in u} R_w(r_i)$
- Bootstrapping: train a multi-label multi-class classifier estimating relations given an utterance

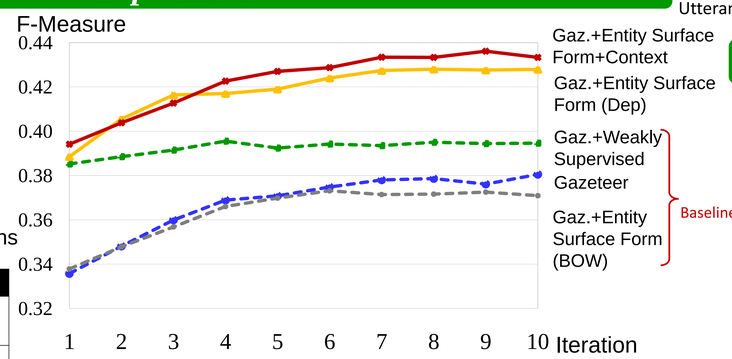
➢ Unsupervised self-training helps improve the results.



6. Experiments

- **Knowledge Base: Freebase (movie domain)**
  - 670K entities
  - 78 entity types
- **Relation Detection Data**
  - Crowd-sourced utterances
  - Manually annotated with SPARQL queries → relations

Entity Tag	Derived Word (Dep.)
$character	character, role, who, girl, she, he, officier
$director	director, dir, filmmaker
$genre	comedy, drama, fantasy, cartoon, horror, sci
$language	language, spanish, english, german
$producer	producer, filmmaker, screenwriter



➢ 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 most performance.

Dataset available: http://research.microsoft.com/en-us/projects/nl-sparql/default.aspx

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.ynchen