Generalized Zero-shot Intent Detection via Commonsense Knowledge

A.B. Siddique, Fuad Jamour, Luxun Xu, and Vagelis Hristidis
{msidd005,fuadj,lxu051}@ucr.edu, vagelis@cs.ucr.edu

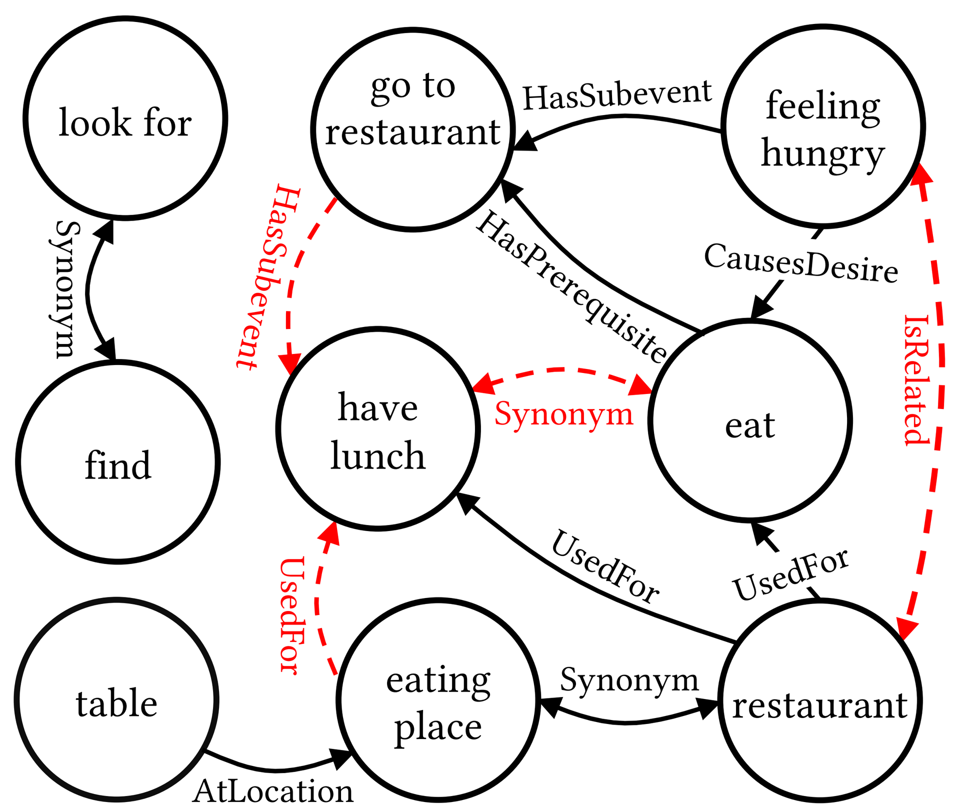
UC RIVERSIDE

sigir21

Motivation

- Identifying user intents from natural language utterances is crucial in conversational systems
- It has been extensively studied as a supervised classification problem
- However, in practice, **unseen intents** emerge after deploying the model and they do not have any training data
- We propose **RIDE**: a generalized zero-shot intent detection model that seamlessly adapts and classifies natural language utterances with **both seen and unseen** intents
- RIDE computes robust and generalizable relationship meta-features that capture deep semantic relationships between utterances and intent labels
- These **meta-features** are computed by considering how the concepts in an utterance are linked to those in an intent label via commonsense knowledge

Computation of Relationship Meta-features



(a) Automatic Commonsense KG Completion

User Utterance:
Look for something nearby, I am feeling hungry

Intent Label:
FindRestaurant

Relationship Meta-features:

⟨look for, *Synonym*, find⟩ 0.99

⋮

⟨feeling hungry, *IsRelated*, restaurant⟩ 0.91

⋮

(b) Relationship Meta-feature Generation

Relationship Meta-features Generator (RMG)

Input: $\mathcal{R} = \{r_1, \cdots, r_t\}$: relations in KG

$\mathcal{G}_i = \{g_1, \cdots, g_q\}$: utterance n-grams

$\mathcal{I}_j = \{\mathcal{A}, \mathcal{O}\}$: intent's Action and Object

Output: $\mathbf{e}_{relationship}$: $\mathcal{X}_i$-$\mathcal{I}_j$ relationship meta-features

Let $\mathbf{e}^{\overrightarrow{\mathcal{A}}}_{\mathcal{X}_i} = \mathtt{RM}\ (\mathcal{A}, \mathcal{G}_i, \rightarrow)$ // Action to utterance

Let $\mathbf{e}^{\overrightarrow{\mathcal{O}}}_{\mathcal{X}_i} = \mathtt{RM}\ (\mathcal{O}, \mathcal{G}_i, \rightarrow)$ // Object to utterance

Let $\mathbf{e}^{\overleftarrow{\mathcal{A}}}_{\mathcal{X}_i} = \mathtt{RM}\ (\mathcal{A}, \mathcal{G}_i, \leftarrow)$ // utterance to Action

Let $\mathbf{e}^{\overleftarrow{\mathcal{O}}}_{\mathcal{X}_i} = \mathtt{RM}\ (\mathcal{O}, \mathcal{G}_i, \leftarrow)$ // utterance to Object

Let $\mathbf{e}_{relationship} = [\mathbf{e}^{\overrightarrow{\mathcal{A}}}_{\mathcal{X}_i}, \mathbf{e}^{\overrightarrow{\mathcal{O}}}_{\mathcal{X}_i}, \mathbf{e}^{\overleftarrow{\mathcal{A}}}_{\mathcal{X}_i}, \mathbf{e}^{\overleftarrow{\mathcal{O}}}_{\mathcal{X}_i}]$

return $\mathbf{e}_{relationship}$

Function RM(*concept*, *phrases*, *direction*):

  Let $\mathbf{e} = []$

  foreach $r \in \mathcal{R}$ **do**

    if *direction* $= \rightarrow$ **then**

      Let $p = \mathtt{Max}\ (LP(concept, r, g))$ for $g \in phrases$

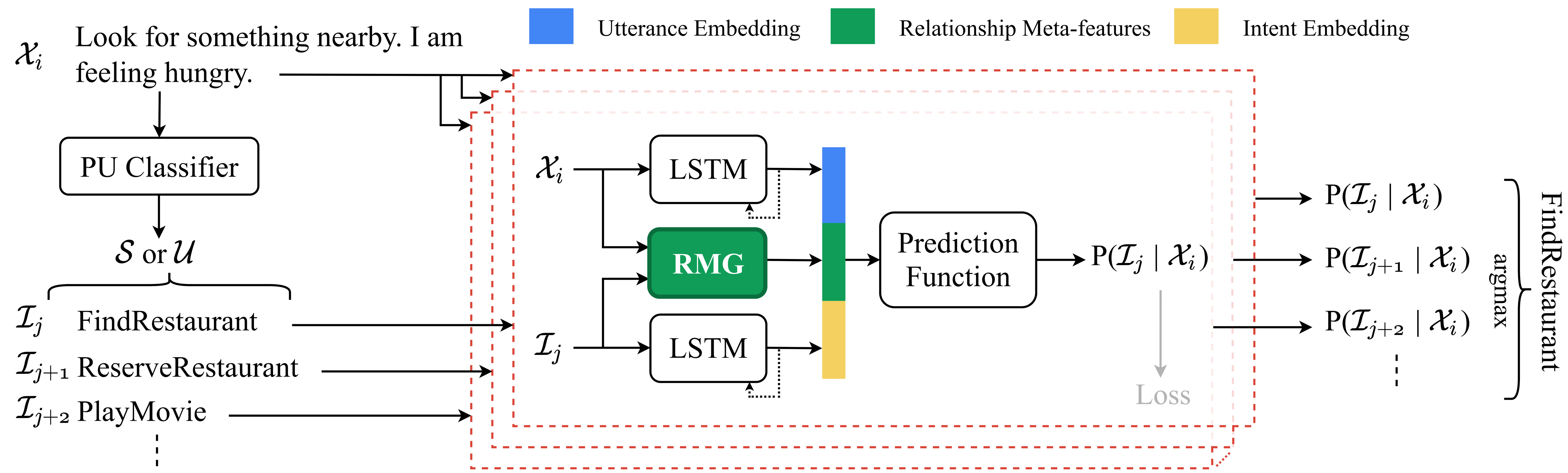
    if *direction* $= \leftarrow$ **then**

      Let $p = \mathtt{Max}\ (LP(g, r, concept))$ for $g \in phrases$

    $\mathbf{e}$.append(p)

  return $\mathbf{e}$

RIDE Overview

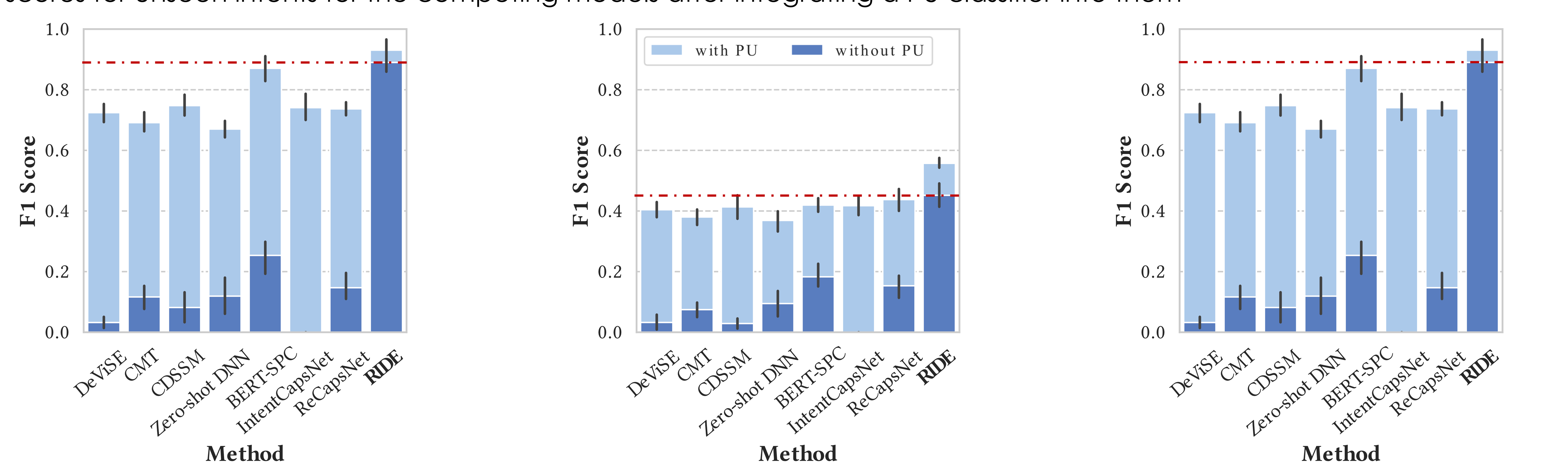


Evaluation

- Main result: F1 scores for competing models in the generalized zero-shot setting

Method	SNIPS		SGD		MultiWOZ	
	Unseen	Seen	Unseen	Seen	Unseen	Seen
BERT-SPC	0.2761	0.7152	0.1872	0.6401	0.1932	0.6413
IntentCapsNet	0.0000	0.6532	0.0000	0.5508	0.0000	0.6038
ReCapsNet	0.1601	0.6783	0.1331	0.5751	0.1467	0.6170
SEG	0.6991	0.8651	0.4032	0.6356	0.4143	0.6456
RIDE w/o PU	0.9103	0.8799	0.4634	0.8295	0.4645	0.8816
RIDE /w PU	**0.9254**	**0.9080**	**0.5734**	**0.8298**	**0.5206**	**0.8847**

- F1 scores for unseen intents for the competing models after integrating a PU classifier into them



(a) SNIPS dataset (b) SGD dataset (c) MultiWOZ dataset