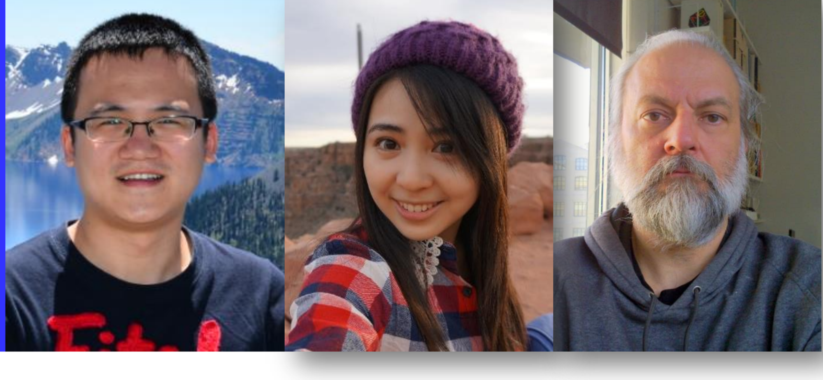
Learning OOV through Semantic Relatedness in Spoken Dialog Systems

Ming Sun, Yun-Nung (Vivian) Chen, and Alexander I. Rudnicky

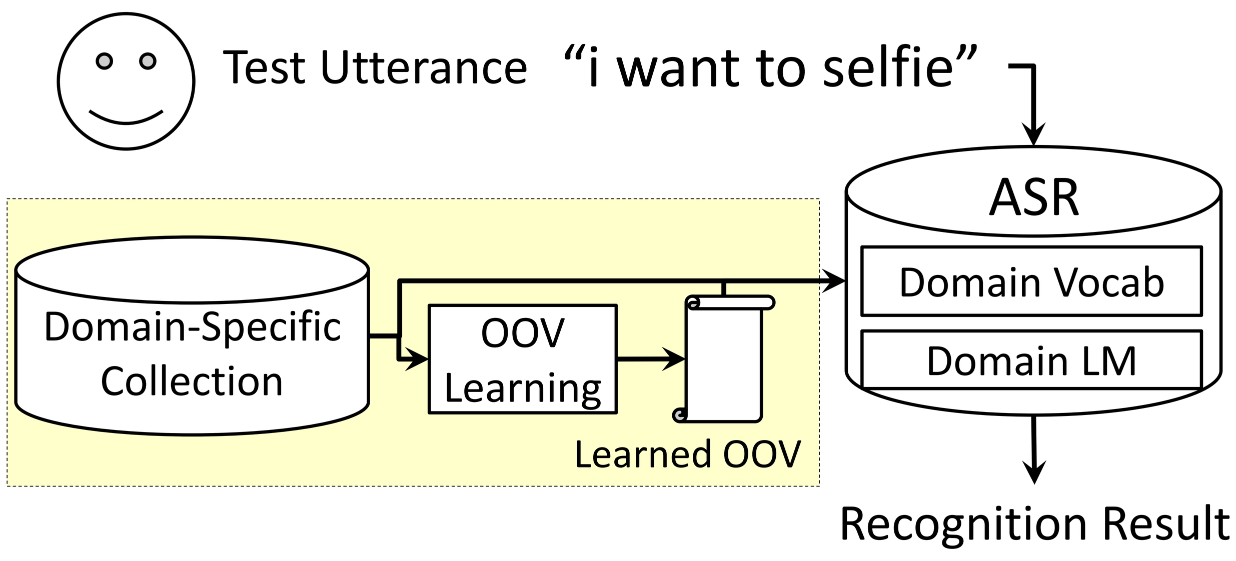


Summary

- Motivation
  - Domain language may drift over time so that ensuring language coverage in dialog systems can be a challenge (Furnas et al., 1987).
  - The mismatch between training data and current input increases recognition errors and misunderstanding.
  - *Detect-and-Learn* strategy requires human effort and takes more time to adapt the vocabulary and LM.
- Approach: ***Expect-and-Learn***
  - Automatically acquiring potential out-of-vocabulary (OOV) words by leveraging different types of words relatedness.
- Result
  - Both recognition and semantic parsing accuracy can be improved after acquiring potential OOVs.

OOV Learning Method

- Detect-and-Learn (Qin et al., 2011; 2012):
  - Discover OOV words during the conversation
  - Example:
    S: "I heard something like SELF, can you repeat it?"
    U: "It's SELFIE."
  - Drawbacks
    - Limited number of new words
    - Required human efforts to correct spellings and pronunciations
- **Expect-and-Learn (proposed):**
  - Use semantic relatedness to automatically enrich the vocabulary and language model beforehand



  - Advantages
    - Large amount of potentially useful new words can be learned
    - No human involved

Expect-and-Learn Procedure

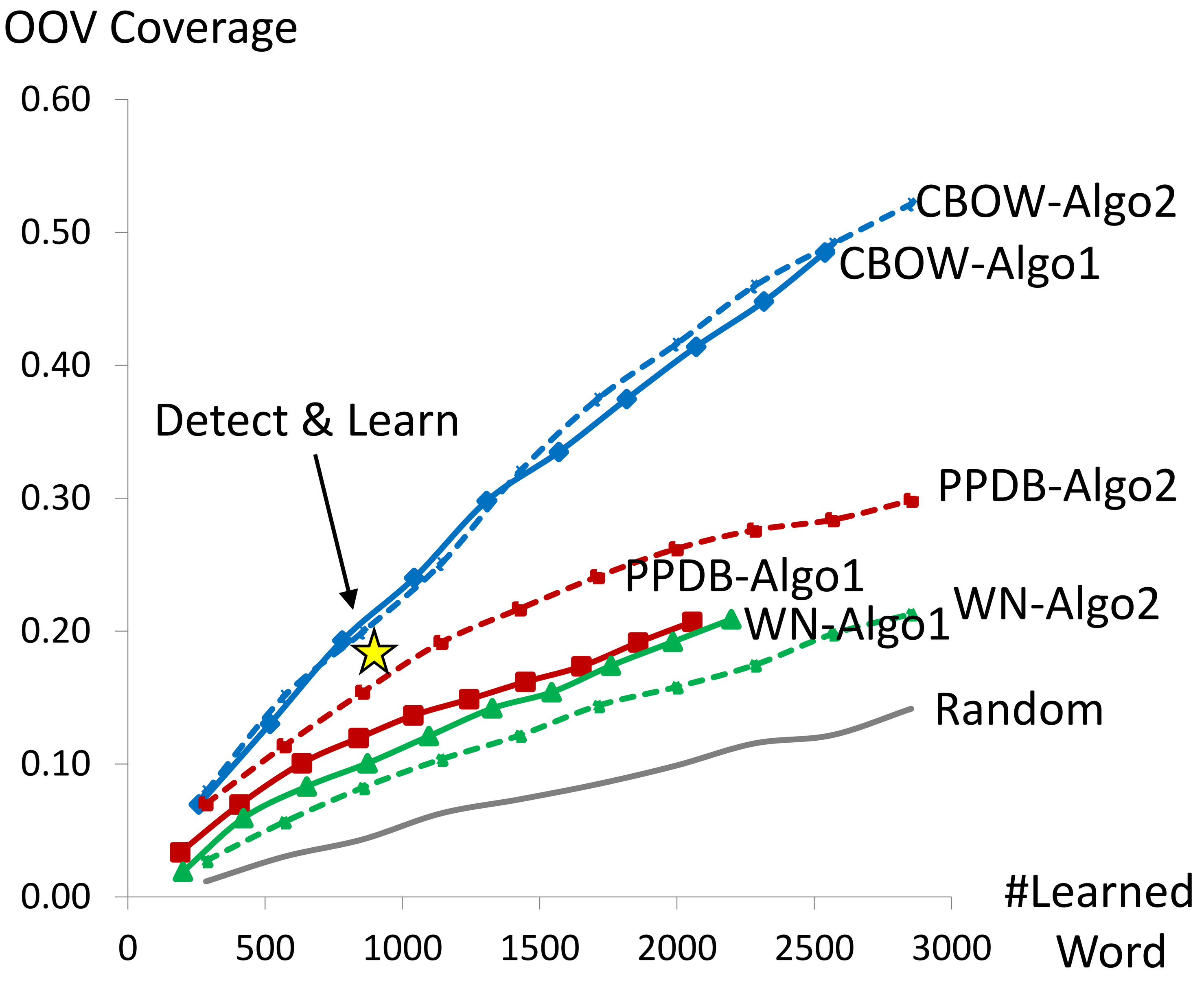
- **Vocabulary Expansion**
- **Idea: learn new words related to the current domain represented by in-vocabulary words (IVs)**
  1. From the IV with the highest frequency v*, one unseen word w* is extracted from the resource according to:
     - Local relatedness (Algo1): w* is mostly related to v*
     - Global relatedness (Algo2): w* is mostly related to the complete IV set
  2. Repeat until the size of vocabulary satisfies a threshold
- **Language Model Expansion**
  - Use Kneser-Ney smoothing to estimate the unigram for the newly learned OOVs.

Relatedness Resources

1. **Linguistically semantic relatedness**
   - Defined by linguistics, e.g., WordNet (WN), Paraphrase Database (PPDB) (Ganitkevitch et al., 2013)
2. **Data-driven semantic relatedness**
   - Distributional semantics, e.g., continuous bag-of-word embeddings (CBOW) (Mikolov et al., 2013)

Experimental Results

- **Dataset:** Wall Street Journal
  - Acoustic model: WSJ GMM-HMM semi continuous
  - Pronunciation: CMU Dictionary + Logios Lexicon Tool
- **OOV Coverage Evaluation**
  - How much OOV tokens in test set can be covered by using different relatedness resources.



Recognition and Understanding Performance

Only Domain Specific Models

Learning Strategy	Vocab Size	OOV Rate (%)	Recog. WER (%)	SLU F1 (%)
Baseline	2854	22.6	49.9	57.0
Algo1	5394	11.7	**41.6**	**65.4**
Algo2	5394	11.6	42.0	65.1
Oracle	4254	0.0	23.5	80.9

Domain + Generic Models

Learning Strategy	Vocab Size	OOV Rate (%)	Recog. WER (%)	SLU F1 (%)
Baseline	20175	3.6	21.7	82.2
Algo1	22599	3.0	**20.3**	**83.2**
Algo2	22599	3.0	20.4	**83.2**
Oracle	20431	0.0	15.1	87.1

Conclusion

- Speech recognition and language understanding performance can be improved through an OOV expect-and-learn procedure.
- A limited domain vocabulary can be utilized to effectively acquire OOVs by the **word relatedness theory** through web knowledge bases.
- With data-driven semantic relatedness, both the global and local learning procedures are able to successfully harvest more than 50% of OOVs, leading to better recognition and understanding performance.
- This work demonstrates that
  - OOV learning may benefit dialog system
  - the proposed ***expect-and-learn*** strategy outperforms the traditional *detect-and-learn* in both higher effectiveness and no human involvement.

Carnegie Mellon University