Unsupervised Learning and Modeling of Knowledge and Intent for Spoken Dialogue Systems

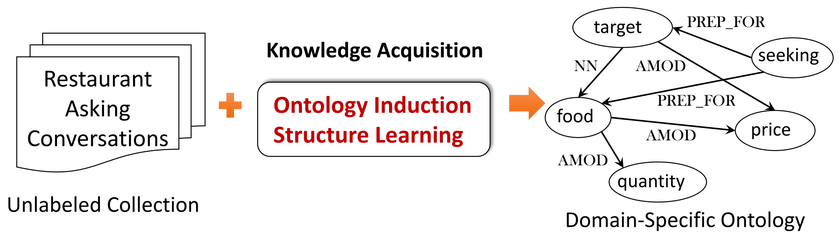
Carnegie Mellon University

Yun-Nung (Vivian) Chen

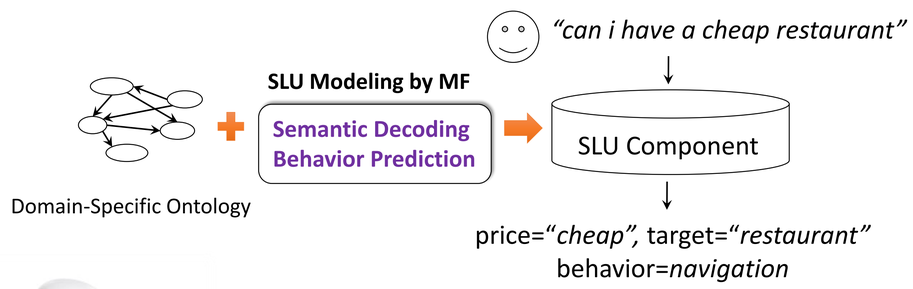
yvchen@cs.cmu.edu

Goal

1) Given unlabeled conversations, how can a system automatically induce and organize domain-specific concepts?



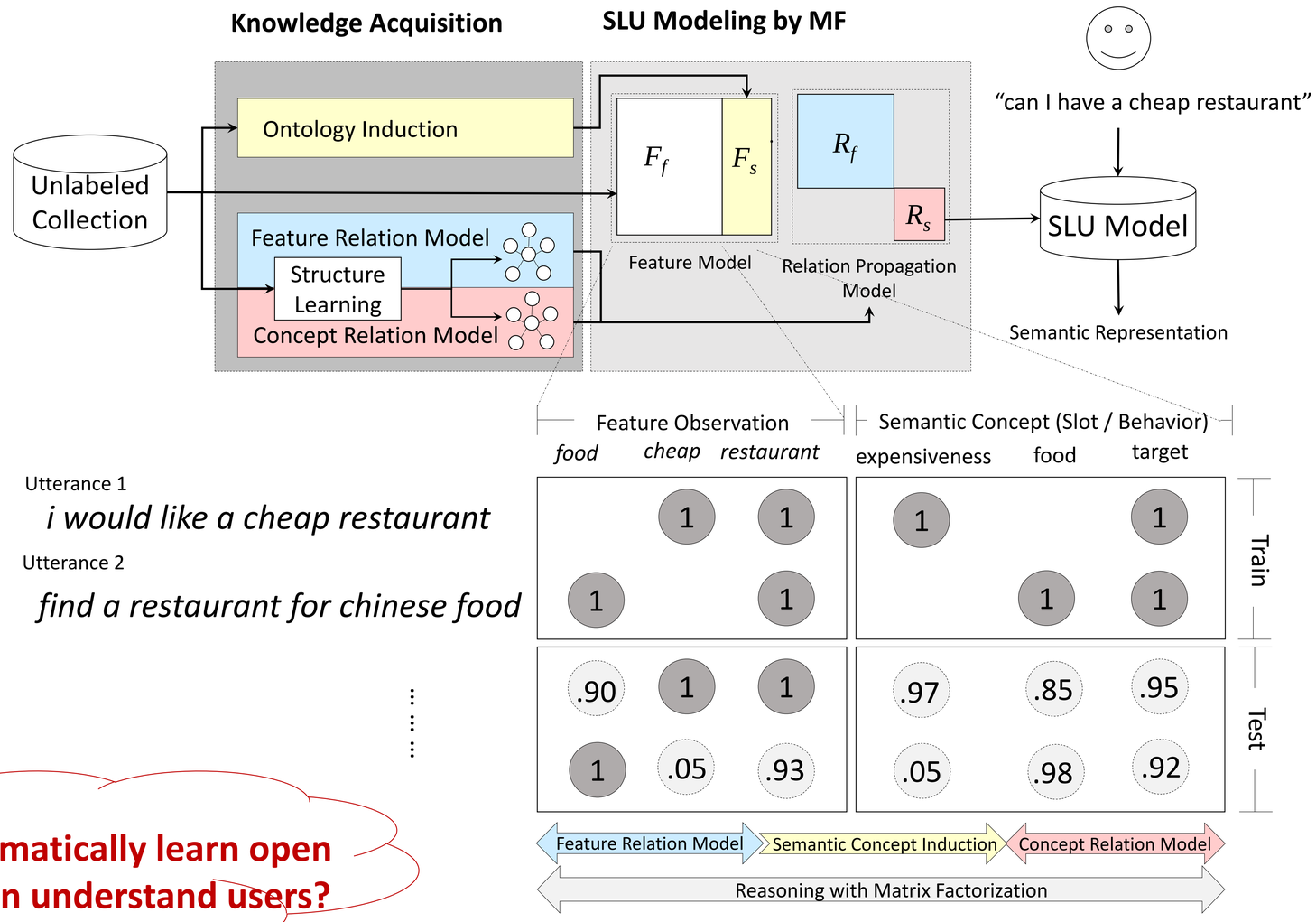
2) With the automatically acquired knowledge, how can a system understand utterances?



Can a dialogue system automatically learn open domain knowledge and then understand users?

- Domain: restaurant recommendation in an in-car setting (WER = 37%)
  - Dialogue slots: addr, area, food, phone, postcode, pricerange, task, type

Framework



Knowledge Acquisition

➢ **Ontology Induction** (Chen et al., 2013 & 2014)

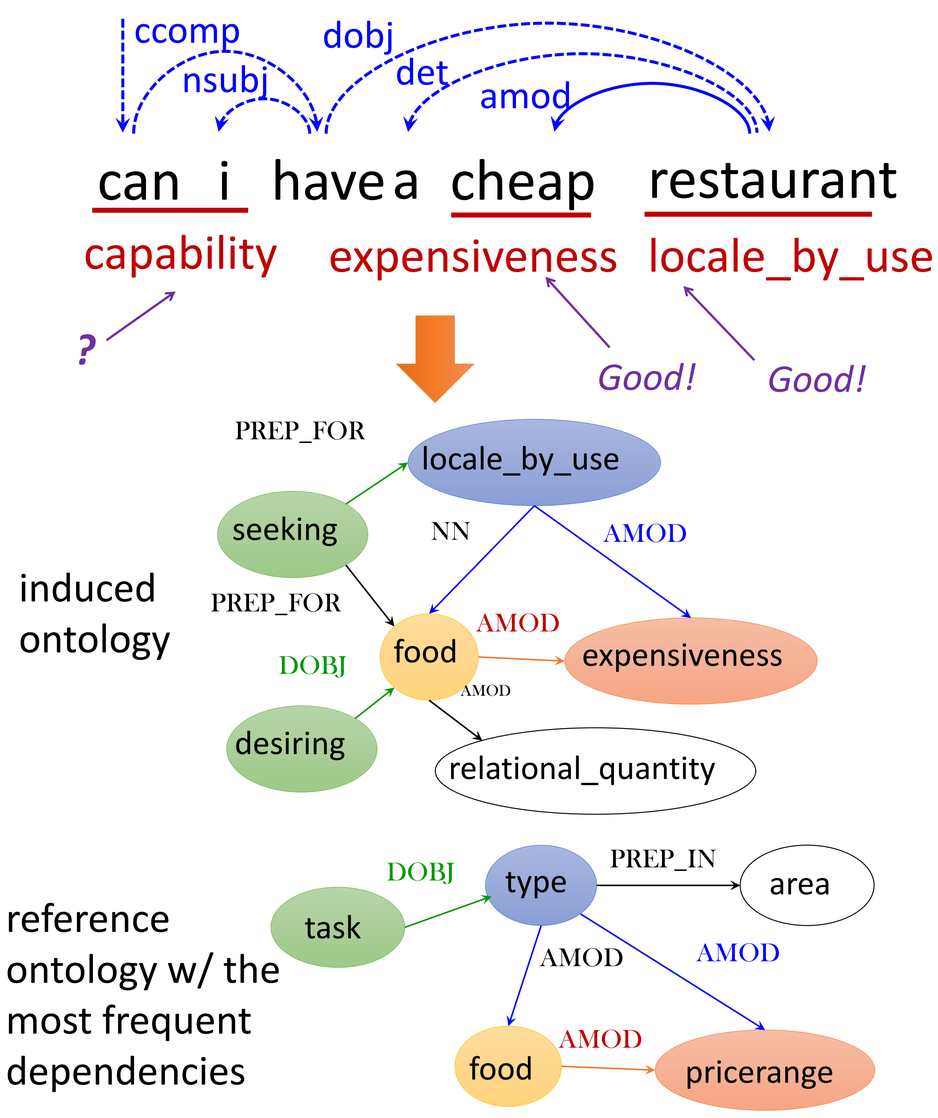
Frame-semantic parsing on ASR results (Das et al., 2013)

- frame → slot candidate
- lexical unit → slot filler

➢ **Structure Learning** (Chen et al., 2015a)

Typed syntactic dependencies on ASR

https://github.com/yvchen/MRRW/



SLU Modeling by Matrix Factorization

➢ **Semantic Decoding** (Chen et al., 2015b)

- concept → semantic slot

➢ **Behavior Prediction**

- concept → user behavior

❖ **1st Issue:** How to induce domain-specific concepts?

- Relation Propagation Model
  - Feature Knowledge Graph
  - Concept Knowledge Graph

✓ Assumption: The domain-specific features/concepts have more dependency to each other.

Relation matrices allow each node to propagate scores to its neighbors in the knowledge graph, so that domain-specific features/concepts have higher scores during training.

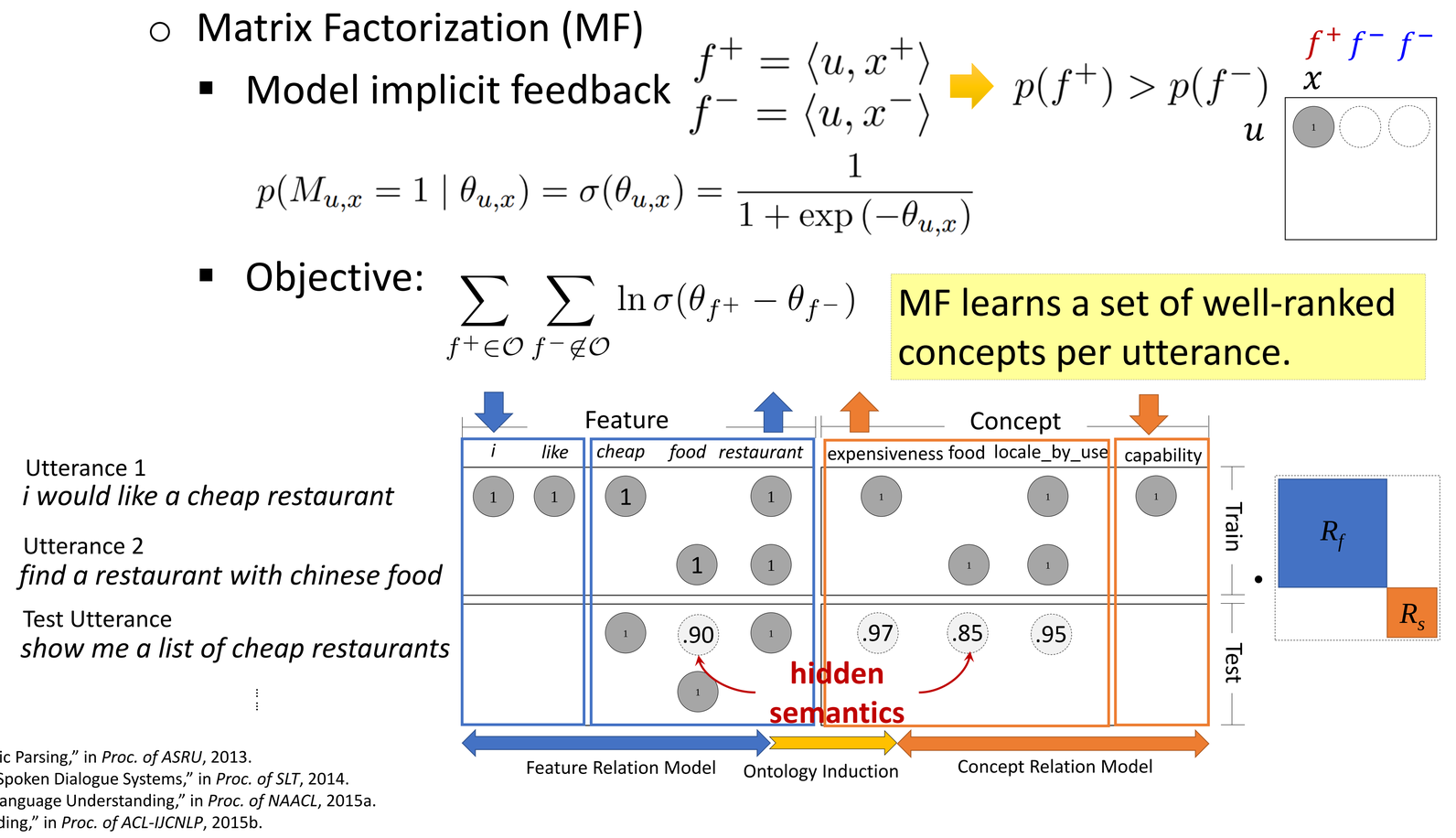
❖ **2nd Issue:** Hidden semantics cannot be observed but may benefit understanding performance.

- Matrix Factorization (MF)
  - Model implicit feedback $f^+ = \langle u, x^+ \rangle$, $f^- = \langle u, x^- \rangle$ ➔ $p(f^+) > p(f^-)$

$$p(M_{u,x} = 1 \mid \theta_{u,x}) = \sigma(\theta_{u,x}) = \frac{1}{1 + \exp(-\theta_{u,x})}$$

  - Objective: $\sum_{f^+ \in \mathcal{O}} \sum_{f^- \notin \mathcal{O}} \ln \sigma(\theta_{f^+} - \theta_{f^-})$

MF learns a set of well-ranked concepts per utterance.



Chen et al., "Unsupervised Induction and Filling of Semantic Slots for Spoken Dialogue Systems Using Frame-Semantic Parsing," in *Proc. of ASRU*, 2013.
Chen et al., "Leveraging Frame Semantics and Distributional Semantics for Unsupervised Semantic Slot Induction in Spoken Dialogue Systems," in *Proc. of SLT*, 2014.
Chen et al., "Jointly Modeling Inter-Slot Relations by Random Walk on Knowledge Graphs for Unsupervised Spoken Language Understanding," in *Proc. of NAACL*, 2015a.
Chen et al., "Matrix Factorization with Knowledge Graph Propagation for Unsupervised Spoken Language Understanding," in *Proc. of ACL-IJCNLP*, 2015b.