

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

Yun-Nung Chen

School of Computer Science, Carnegie Mellon University

5000 Forbes Avenue, Pittsburgh, PA 15213-3891, USA

yvchen@cs.cmu.edu

## Abstract

Spoken dialogue systems (SDS) are rapidly appearing in various smart devices (smartphone, smart-TV, in-car navigating system, etc). The key role in a successful SDS is a spoken language understanding (SLU) component, which parses user utterances into semantic concepts in order to understand users' intentions. However, such semantic concepts and their structure are manually created by experts, and the annotation process results in extremely high cost and poor scalability in system development. Therefore, the dissertation focuses on improving SDS generalization and scalability by automatically inferring domain knowledge and learning structures from unlabeled conversations through a matrix factorization (MF) technique. With the automatically acquired semantic concepts and structures, we further investigate whether such information can be utilized to effectively understand user utterances and then show the feasibility of reducing human effort during SDS development.

## 1 Introduction

Various smart devices (e.g. smartphone, smart-TV, in-car navigating system) are incorporating spoken language interfaces, a.k.a. spoken dialogue systems (SDS), in order to help users finish tasks more efficiently. The key role in a successful SDS is a spoken language understanding (SLU) component; in order to capture the language variation from dialogue participants, the SLU component must create a mapping between the natural language inputs and semantic representations that correspond to users' intentions.

The semantic representation must include "concepts" and a "structure": concepts are the domain-

specific topics, and the structure describes the relations between concepts and conveys intentions. However, most prior work focused on learning the mapping between utterances and semantic representations, where such knowledge still remains predefined. The need of annotations results in extremely high cost and poor scalability in system development. Therefore, current technology usually limits conversational interactions to a few narrow predefined domains/topics. With the increasing conversational interactions, this dissertation focuses on improving *generalization* and *scalability* of building SDSs with little human effort.

In order to achieve the goal, two questions need to be addressed: 1) Given unlabelled conversations, how can a system automatically induce and organize the domain-specific concepts? 2) With the automatically acquired knowledge, how can a system understand user utterances and intents? To tackle the above problems, we propose to acquire the domain knowledge that captures human's semantics, intents, and behaviors. Then based on the acquired knowledge, we build an SLU component to understand users and to offer better interactions in dialogues.

The dissertation shows the feasibility of building a dialogue learning system that is able to understand how particular domains work based on unlabeled conversations. As a result, an initial SDS can be automatically built according to the learned knowledge, and its performance can be quickly improved by interacting with users for practical usage, presenting the potential of reducing human effort for SDS development.

## 2 Related Work

**Unsupervised SLU** Tur et al. (2011; 2012) were among the first to consider unsupervised approaches for SLU, where they exploited query logs for slot-filling. In a subsequent study, Heck and Hakkani-Tür (2012) studied the Semantic Web for

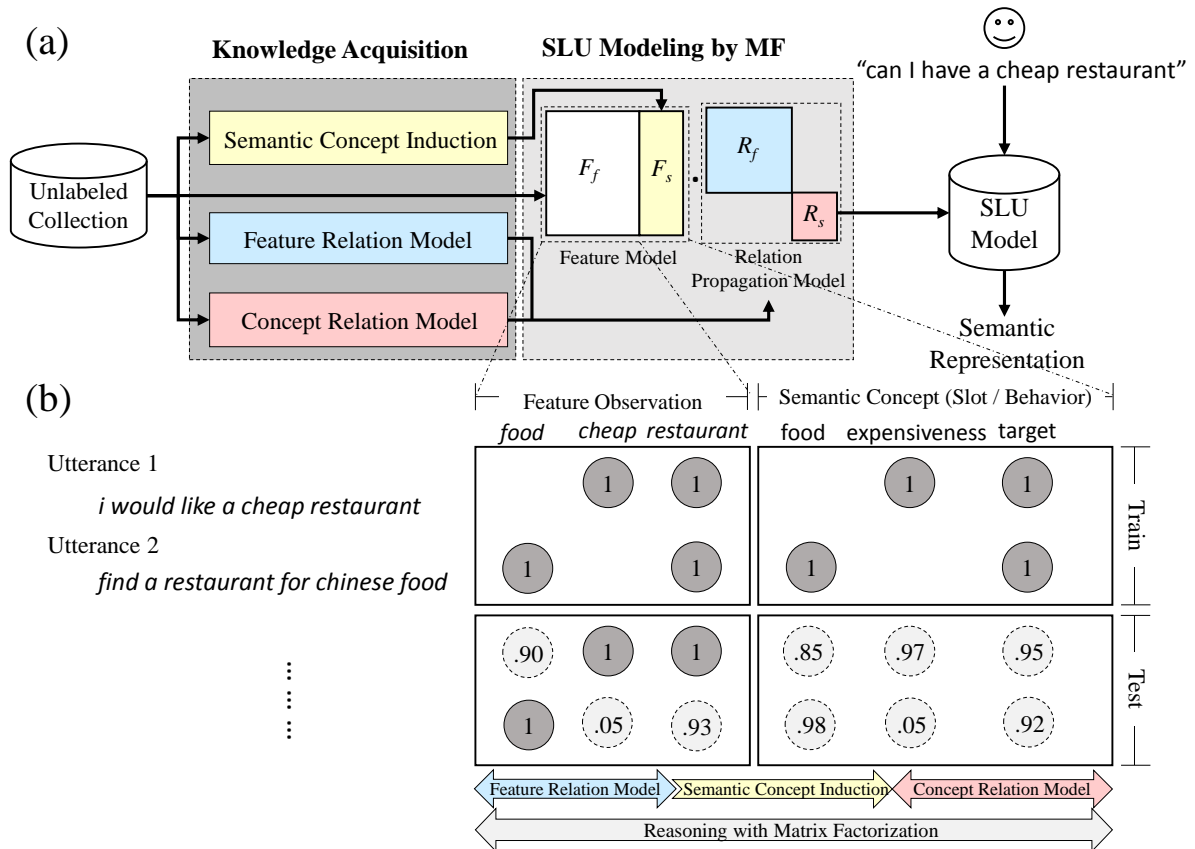


Figure 1: (a): The proposed framework. (b): Our MF method completes a partially-missing matrix for semantic decoding/behavior prediction. Dark circles are observed facts, shaded circles are inferred facts. The ontology induction maps observed feature patterns to semantic concepts. The feature relation model constructs correlations between observed feature patterns. The concept relation model learns the high-level semantic correlations for inferring hidden semantic slots or predicting subsequent behaviors. Reasoning with matrix factorization incorporates these models jointly, and produces a coherent and domain-specific SLU model.

the intent detection problem in SLU, showing that results obtained from the unsupervised training process align well with the performance of traditional supervised learning. Following their success of unsupervised SLU, recent studies have also obtained interesting results on the tasks of relation detection (Hakkani-Tür et al., 2013; Chen et al., 2014a), entity extraction (Wang et al., 2014), and extending domain coverage (El-Kahky et al., 2014; Chen and Rudnicky, 2014). However, most studies above do not explicitly learn latent factor representations from the data—while we hypothesize that the better robustness can be achieved by explicitly modeling the measurement errors (usually produced by automatic speech recognizers (ASR)) using latent variable models and taking additional local and global semantic constraints into account.

**Latent Variable Modeling in SLU** Early studies on latent variable modeling in speech included

the classic hidden Markov model for statistical speech recognition (Jelinek, 1997). Recently, Celikyilmaz et al. (2011) were the first to study the intent detection problem using query logs and a discrete Bayesian latent variable model. In the field of dialogue modeling, the partially observable Markov decision process (POMDP) (Young et al., 2013) model is a popular technique for dialogue management, reducing the cost of hand-crafted dialogue managers while producing robustness against speech recognition errors. More recently, Tur et al. (2013) used a semi-supervised LDA model to show improvement on the slot filling task. Also, Zhai and Williams (2014) proposed an unsupervised model for connecting words with latent states in HMMs using topic models, obtaining interesting qualitative and quantitative results. However, for unsupervised SLU, it is not obvious how to incorporate additional information in the HMMs. With increasing works about learn-

ing the feature matrices for language representations (Mikolov et al., 2013), matrix factorization (MF) has become very popular for both implicit and explicit feedback (Rendle et al., 2009; Chen et al., 2015a).

This thesis proposal is the first to propose a framework about unsupervised SLU modeling, which is able to simultaneously consider various local and global knowledge automatically learned from unlabelled data using a matrix factorization (MF) technique.

### 3 The Proposed Work

The proposed framework is shown in Figure 1(a), where there are two main parts, one is *knowledge acquisition* and another is *SLU modeling by MF*. The first part is to acquire the domain knowledge that is useful for building the domain-specific dialogue systems, which addresses the question about how to induce and organize the semantic concepts (the first problem). Here we propose ontology induction and structure learning procedures. The ontology induction refers to the semantic concept induction (yellow block) and the structure learning refers to relation models (blue and pink blocks) in Figure 1(a). The details are described in Section 4. The second part is to self-train an SLU component using the acquired knowledge for the domain-specific SDS, and this part answers to the question about how to utilize the obtained information in SDSs to understand user utterances and intents. There are two aspects regarding to SLU modeling, semantic decoding and behavior prediction. The semantic decoding is to parse the input utterances into semantic forms for better understanding, and the behavior prediction is to predict the subsequent user behaviors for providing better system interactions. This dissertation plans to apply MF techniques to unsupervised SLU modeling, including both semantic decoding and behavior prediction.

In the proposed model, we first build a feature matrix to represent training utterances, where each row refers to an utterance and each column refers to an observed feature pattern or a learned semantic concept (either a slot or a behavior). Figure 1(b) illustrates an example of the matrix. Then given a testing utterance, we can convert it into a vector based on the observed patterns, and fill in the missing values of the semantic concepts. In the first example utterance of the figure, although semantic slot `FOOD` is not observed, the ut-

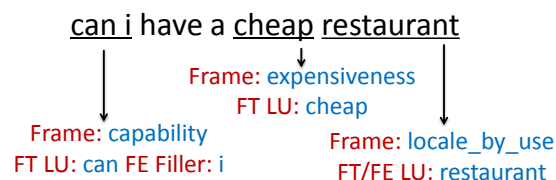


Figure 2: An example of probabilistic frame-semantic parsing on ASR output. FT: frame target. FE: frame element. LU: lexical unit.

terance implies the meaning facet `food`. The MF approach is able to learn the latent feature vectors for utterances and semantic concepts, inferring implicit semantics to improve the decoding process—namely, by filling the matrix with probabilities (lower part of the matrix in Figure 1(b)).

The feature model is built on the observed feature patterns and the learned concepts, where the concepts are obtained from the knowledge acquisition process (Chen et al., 2013; Chen et al., 2015b). Section 5.1 explains the detail of the feature model. In order to consider the additional structure information, we propose a relation propagation model based on the learned structure, which includes a feature relation model (blue block) and a concept relation model (pink block) described in Section 5.2.

Finally we train an SLU model by learning latent feature vectors for utterances and slots/behaviors through MF techniques. Combining with a relation propagation model, the trained SLU model is able to estimate the probability that each concept occurs in the testing utterance, and how likely each concept is domain-specific simultaneously. In other words, the SLU model is able to transform testing utterances into domain-specific semantic representations or predicted behaviors without human involvement.

## 4 Knowledge Acquisition

Given unlabeled conversations and available knowledge resources, we plan to extract organized knowledge that can be used for domain-specific SDSs. The ontology induction and structure learning are proposed to automate an ontology building process.

### 4.1 Ontology Induction

Chen et al. (2013; 2014b) proposed to automatically induce semantic slots for SDSs by frame-semantic parsing, where all ASR-decoded utter-

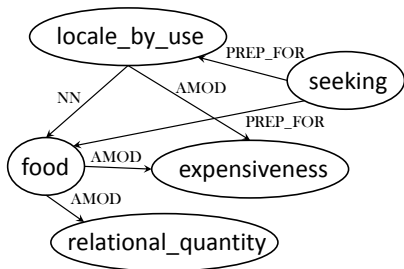


Figure 3: A simplified example of the automatically derived knowledge graph.

ances are parsed using SEMAFOR<sup>1</sup>, a state-of-the-art frame-semantic parser (Das et al., 2010; Das et al., 2013), and then all frames from parsed results are extracted as slot candidates (Dinarelli et al., 2009). For example, Figure 2 shows an example of an ASR-decoded text output parsed by SEMAFOR. There are three frames (*capability*, *expensiveness*, and *locale\_by\_use*) in the utterance, which we consider as slot candidates.

Since SEMAFOR was trained on FrameNet annotation, which has a more generic frame-semantic context, not all the frames from the parsing results can be used as the actual slots in the domain-specific dialogue systems. For instance, in Figure 2, “*expensiveness*” and “*locale\_by\_use*” frames are essentially the key slots for the purpose of understanding in the restaurant query domain, whereas the “*capability*” frame does not convey particularly valuable information for the domain-specific SDS. In order to fix this issue, Chen et al. (2014b) proved that integrating continuous-valued word embeddings with a probabilistic frame-semantic parser is able to identify key semantic slots in an unsupervised fashion, reducing the cost of designing task-oriented SDSs.

## 4.2 Structure Learning

A key challenge of designing a coherent semantic ontology for SLU is to consider the structure and relations between semantic concepts. In practice, however, it is difficult for domain experts and professional annotators to define a coherent slot set, while considering various lexical, syntactic, and semantic dependencies. The previous work exploited the typed syntactic dependency theory for unsupervised induction and organization of semantic slots in SDSs (Chen et al., 2015b). More specifically, two knowledge

graphs, a slot-based semantic knowledge graph and a word-based lexical knowledge graph, are automatically constructed. To jointly consider the word-to-word, word-to-slot, and slot-to-slot relations, we use a random walk inference algorithm to combine these two knowledge graphs, guided by dependency grammars. Figure 3 is a simplified example of the automatically built semantic knowledge graph corresponding to the restaurant domain. The experiments showed that considering inter-slot relations is crucial for generating a more coherent and complete slot set, resulting in a better SLU model, while enhancing the interpretability of semantic slots.

## 5 SLU Modeling by Matrix Factorization

For two aspects of SLU modeling: semantic decoding and behavior prediction, we plan to apply MF to both tasks by treating learned concepts as semantic slots and human behaviors respectively.

Considering the benefits brought by MF techniques, including 1) modeling the noisy data, 2) modeling hidden information, and 3) modeling the dependency between observations, the dissertation applies an MF approach to SLU modeling for SDSs. In our model, we use  $U$  to denote the set of input utterances,  $F$  as the set of observed feature patterns, and  $S$  as the set of semantic concepts we would like to predict (slots or human behaviors). The pair of an utterance  $u \in U$  and a feature/concept  $x \in \{F + S\}$ ,  $\langle u, x \rangle$ , is a *fact*. The input to our model is a set of observed facts  $\mathcal{O}$ , and the observed facts for a given utterance is denoted by  $\{\langle u, x \rangle \in \mathcal{O}\}$ . The goal of our model is to estimate, for a given utterance  $u$  and a given feature pattern/concept  $x$ , the probability,  $p(M_{u,x} = 1)$ , where  $M_{u,x}$  is a binary random variable that is true if and only if  $x$  is the feature pattern/domain-specific concept in the utterance  $u$ . We introduce a series of exponential family models that estimate the probability using a natural parameter  $\theta_{u,x}$  and the logistic sigmoid function:

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

We construct a matrix  $M_{|U| \times (|F| + |S|)}$  as observed facts for MF by integrating a feature model and a relation propagation model below.

<sup>1</sup><http://www.ark.cs.cmu.edu/SEMAFOR/>

## 5.1 Feature Model

First, we build a binary feature pattern matrix  $F_f$  based on the observations, where each row refers to an utterance and each column refers to a feature pattern (a word or a phrase). In other words,  $F_f$  carries the basic word/phrase vector for each utterance, which is illustrated as the left part of the matrix in Figure 1(b). Then we build a binary matrix  $F_s$  based on the induced semantic concepts from Section 4.1, which also denotes the slot/behavior features for all utterances (right part of the matrix in Figure 1(b)).

For building the feature model  $M_F$ , we concatenate two matrices and obtain  $M_F = [F_f \ F_s]$ , which refers to the upper part of the matrix in Figure 1(b) for training utterances.

## 5.2 Relation Propagation Model

It is shown that the structure of semantic concepts helps decide domain-specific slots and further improves the SLU performance (Chen et al., 2015b). With the learned structure from Section 4.2, we can model the relations between semantic concepts, such as inter-slot and inter-behavior relations. Also, the relations between feature patterns can be modeled in the similar way. We construct two knowledge graphs to model the structure:

- **Feature knowledge graph** is built as  $G_f = \langle V_f, E_{ff} \rangle$ , where  $V_f = \{f_i \in F\}$  and  $E_{ff} = \{e_{ij} \mid f_i, f_j \in V_f\}$ .
- **Semantic concept knowledge graph** is built as  $G_s = \langle V_s, E_{ss} \rangle$ , where  $V_s = \{s_i \in S\}$  and  $E_{ss} = \{e_{ij} \mid s_i, s_j \in V_s\}$ .

The structured graph can model the relation between the connected node pair  $(x_i, x_j)$  as  $r(x_i, x_j)$ . Here we compute two matrices  $R_s = [r(s_i, s_j)]_{|S| \times |S|}$  and  $R_f = [r(f_i, f_j)]_{|F| \times |F|}$  to represent concept relations and feature relations respectively. With the built relation models, we combine them as a relation propagation matrix  $M_R^2$ :

$$M_R = \begin{bmatrix} R_f & 0 \\ 0 & R_s \end{bmatrix}. \quad (2)$$

The goal of this matrix is to propagate scores between nodes according to different types of relations in the constructed knowledge graphs (Chen and Metze, 2012).

<sup>2</sup>The values in the diagonal of  $M_R$  are 0 to model the propagation from other entries.

## 5.3 Integrated Model

With a feature model  $M_F$  and a relation propagation model  $M_R$ , we integrate them into a single matrix.

$$\begin{aligned} M &= M_F \cdot (M_R + I) \\ &= \begin{bmatrix} F_f R_f + F_f & 0 \\ 0 & F_s R_s + F_s \end{bmatrix}, \end{aligned} \quad (3)$$

where  $M$  is final matrix and  $I$  is the identity matrix in order to remain the original values. The matrix  $M$  is similar to  $M_F$ , but some weights are enhanced through the relation propagation model. The feature relations are built by  $F_f R_f$ , which is the matrix with internal weight propagation on the feature knowledge graph (the blue arrow in Figure 1(b)). Similarly,  $F_s R_s$  models the semantic concept correlations, and can be treated as the matrix with internal weight propagation on the semantic concept knowledge graph (the pink arrow in Figure 1(b)). The propagation model can be treated as running a random walk algorithm on the graphs.

By integrating with the relation propagation model, the relations can be propagated via the knowledge graphs, and the hidden information may be modeled based on the assumption that mutual relations usually help inference (Chen et al., 2015b). Hence, the structure information can be automatically involved in the matrix. In conclusion, for each utterance, the integrated model not only predicts the probabilities that semantic concepts occur but also considers whether they are domain-specific.

## 5.4 Model Learning

The proposed model is parameterized through weights and latent component vectors, where the parameters are estimated by maximizing the log likelihood of observed data (Collins et al., 2001).

$$\begin{aligned} \theta^* &= \arg \max_{\theta} \prod_{u \in U} p(\theta \mid M_u) \\ &= \arg \max_{\theta} \prod_{u \in U} p(M_u \mid \theta) p(\theta) \\ &= \arg \max_{\theta} \sum_{u \in U} \ln p(M_u \mid \theta) - \lambda_{\theta}, \end{aligned} \quad (4)$$

where  $M_u$  is the vector corresponding to the utterance  $u$  from  $M_{u,x}$  in (1), because we assume that each utterance is independent of others.

To avoid treating unobserved facts as designed negative facts, we consider our positive-only data

as *implicit feedback*. Bayesian Personalized Ranking (BPR) is an optimization criterion that learns from implicit feedback for MF, which uses a variant of the ranking: giving observed true facts higher scores than unobserved (true or false) facts (Rendle et al., 2009). Riedel et al. (2013) also showed that BPR learns the implicit relations and improves a relation extraction task.

To estimate the parameters in (4), we create a dataset of *ranked pairs* from  $M$  in (3): for each utterance  $u$  and each observed fact  $f^+ = \langle u, x^+ \rangle$ , where  $M_{u,x} \geq \delta$ , we choose each semantic concept  $x^-$  such that  $f^- = \langle u, x^- \rangle$ , where  $M_{u,x} < \delta$ , which refers to the semantic concept we have not observed in utterance  $u$ . That is, we construct the observed data  $\mathcal{O}$  from  $M$ . Then for each pair of facts  $f^+$  and  $f^-$ , we want to model  $p(f^+) > p(f^-)$  and hence  $\theta_{f^+} > \theta_{f^-}$  according to (1). BPR maximizes the summation of each ranked pair, where the objective is

$$\sum_{u \in U} \ln p(M_u | \theta) = \sum_{f^+ \in \mathcal{O}} \sum_{f^- \notin \mathcal{O}} \ln \sigma(\theta_{f^+} - \theta_{f^-}). \quad (5)$$

The BPR objective is an approximation to the per utterance AUC (area under the ROC curve), which directly correlates to what we want to achieve – well-ranked semantic concepts per utterance, which denotes the better estimation of semantic slots or human behaviors.

To maximize the objective in (5), we employ a stochastic gradient descent (SGD) algorithm (Rendle et al., 2009). For each randomly sampled observed fact  $\langle u, x^+ \rangle$ , we sample an unobserved fact  $\langle u, x^- \rangle$ , which results in  $|\mathcal{O}|$  fact pairs  $\langle f^-, f^+ \rangle$ . For each pair, we perform an SGD update using the gradient of the corresponding objective function for matrix factorization (Gantner et al., 2011).

## 6 Conclusion and Future Work

This thesis proposal proposes an unsupervised SLU approach by automating the dialogue learning process on speech conversations. The preliminary results show that for the automatic speech recognition (ASR) transcripts (word error rate is about 37%), the acquired knowledge can be successfully applied to SLU modeling through MF techniques, guiding the direction of the methodology.

The main planned tasks include:

- Semantic concept identification
- Semantic concept annotation

- SLU modeling by matrix factorization

In this thesis proposal, ongoing work and future plans have been presented towards an automatically built domain-specific SDS. With increasing semantic resources, such as Google’s Knowledge Graph and Microsoft Satori, the dissertation shows the feasibility that utilizing available knowledge improves the generalization and the scalability of dialogue system development for practical usage.

## Acknowledgements

I thank my committee members, Prof. Alexander I. Rudnicky, Prof. Anatole Gershman, Prof. Alan W Black, and Dr. Dilek Hakkani-Tür for their advising and anonymous reviewers for their useful comments. I am also grateful to Prof. Mei Ling Meng for her helpful mentoring.

## References

- Asli Celikyilmaz, Dilek Hakkani-Tür, and Gokhan Tür. 2011. Leveraging web query logs to learn user intent via bayesian discrete latent variable model. In *Proceedings of ICML*.
- Yun-Nung Chen and Florian Metze. 2012. Two-layer mutually reinforced random walk for improved multi-party meeting summarization. In *Proceedings of The 4th IEEE Workshop on Spoken Language Technology*, pages 461–466.
- Yun-Nung Chen and Alexander I. Rudnicky. 2014. Dynamically supporting unexplored domains in conversational interactions by enriching semantics with neural word embeddings. In *Proceedings of 2014 IEEE Spoken Language Technology Workshop (SLT)*, pages 590–595. IEEE.
- Yun-Nung Chen, William Yang Wang, and Alexander I Rudnicky. 2013. Unsupervised induction and filling of semantic slots for spoken dialogue systems using frame-semantic parsing. In *Proceedings of 2013 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, pages 120–125. IEEE.
- Yun-Nung Chen, Dilek Hakkani-Tür, and Gokan Tur. 2014a. Deriving local relational surface forms from dependency-based entity embeddings for unsupervised spoken language understanding. In *Proceedings of 2014 IEEE Spoken Language Technology Workshop (SLT)*, pages 242–247. IEEE.
- Yun-Nung Chen, William Yang Wang, and Alexander I. Rudnicky. 2014b. Leveraging frame semantics and distributional semantics for unsupervised semantic slot induction in spoken dialogue systems. In *Proceedings of 2014 IEEE Spoken Language Technology Workshop (SLT)*, pages 584–589. IEEE.

- Yun-Nung Chen, William Yang Wang, Anatole Gershman, and Alexander I. Rudnicky. 2015a. Matrix factorization with knowledge graph propagation for unsupervised spoken language understanding. In *Proceedings of The 53rd Annual Meeting of the Association for Computational Linguistics and The 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing (ACL-IJCNLP 2015)*. ACL.
- Yun-Nung Chen, William Yang Wang, and Alexander I. Rudnicky. 2015b. Jointly modeling inter-slot relations by random walk on knowledge graphs for unsupervised spoken language understanding. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies*. ACL.
- Michael Collins, Sanjoy Dasgupta, and Robert E Schapire. 2001. A generalization of principal components analysis to the exponential family. In *Proceedings of Advances in Neural Information Processing Systems*, pages 617–624.
- Dipanjan Das, Nathan Schneider, Desai Chen, and Noah A Smith. 2010. Probabilistic frame-semantic parsing. In *Proceedings of The Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 948–956.
- Dipanjan Das, Desai Chen, André F. T. Martins, Nathan Schneider, and Noah A. Smith. 2013. Frame-semantic parsing. *Computational Linguistics*.
- Marco Dinarelli, Silvia Quarteroni, Sara Tonelli, Alessandro Moschitti, and Giuseppe Riccardi. 2009. Annotating spoken dialogs: from speech segments to dialog acts and frame semantics. In *Proceedings of the 2nd Workshop on Semantic Representation of Spoken Language*, pages 34–41. ACL.
- Ali El-Kahky, Derek Liu, Ruhi Sarikaya, Gökhan Tür, Dilek Hakkani-Tür, and Larry Heck. 2014. Extending domain coverage of language understanding systems via intent transfer between domains using knowledge graphs and search query click logs. In *Proceedings of ICASSP*.
- Zeno Gantner, Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2011. Mymedialite: A free recommender system library. In *Proceedings of the fifth ACM conference on Recommender systems*, pages 305–308. ACM.
- Dilek Hakkani-Tür, Larry Heck, and Gokhan Tur. 2013. Using a knowledge graph and query click logs for unsupervised learning of relation detection. In *Proceedings of ICASSP*, pages 8327–8331.
- Larry Heck and Dilek Hakkani-Tür. 2012. Exploiting the semantic web for unsupervised spoken language understanding. In *Proceedings of SLT*, pages 228–233.
- Frederick Jelinek. 1997. *Statistical methods for speech recognition*. MIT press.
- Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of Advances in Neural Information Processing Systems*, pages 3111–3119.
- Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, pages 452–461. AUAI Press.
- Sebastian Riedel, Limin Yao, Andrew McCallum, and Benjamin M Marlin. 2013. Relation extraction with matrix factorization and universal schemas. In *Proceedings of NAACL-HLT*, pages 74–84.
- Gokhan Tur, Dilek Z Hakkani-Tür, Dustin Hillard, and Asli Celikyilmaz. 2011. Towards unsupervised spoken language understanding: Exploiting query click logs for slot filling. In *Proceedings of INTERSPEECH*, pages 1293–1296.
- Gokhan Tur, Minwoo Jeong, Ye-Yi Wang, Dilek Hakkani-Tür, and Larry P Heck. 2012. Exploiting the semantic web for unsupervised natural language semantic parsing. In *Proceedings of INTERSPEECH*.
- Gokhan Tur, Asli Celikyilmaz, and Dilek Hakkani-Tür. 2013. Latent semantic modeling for slot filling in conversational understanding. In *Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8307–8311. IEEE.
- Lu Wang, Dilek Hakkani-Tür, and Larry Heck. 2014. Leveraging semantic web search and browse sessions for multi-turn spoken dialog systems. In *Proceedings of 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4082–4086. IEEE.
- Steve Young, Milica Gasic, Blaise Thomson, and Jason D Williams. 2013. POMDP-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, 101(5):1160–1179.
- Ke Zhai and Jason D Williams. 2014. Discovering latent structure in task-oriented dialogues. In *Proceedings of the Association for Computational Linguistics*.