Dialogue Management based on Sentence Clustering

Wendong Ge
Institute of Automation,
Chinese Academy of Sciences,
Beijing, China
wendong.ge@ia.ac.cn

Bo Xu
Institute of Automation,
Chinese Academy of Sciences,
Beijing, China
xubo@ia.ac.cn

Abstract

Dialogue Management (DM) is a key issue in Spoken Dialogue System (SDS). Most of the existing studies on DM use Dialogue Act (DA) to represent semantic information of sentence, which might not represent the nuanced meaning sometimes. In this paper, we model DM based on sentence clusters which have more powerful semantic representation ability than DAs. Firstly, sentences are clustered not only based on the internal information such as words and sentence structures, but also based on the external information such as context in dialogue via Recurrent Neural Networks. Additionally, the DM problem is modeled as a Partially Observable Markov Decision Processes (POMDP) with sentence clusters. Finally, experimental results illustrate that the proposed DM scheme is superior to the existing one.

1 Introduction

Dialogue Management (DM) is an important issue in Spoken Dialogue Systems (SDS). (Paek et al., 2008) Most of the existing studies on DM use the abstract semantic representation such as Dialogue Act (DA) to represent the sentence intention. In (Bohus et al., 2009), authors propose a plan-based, task-independent DM framework, called RavenClaw, which isolates the domain-specific aspects of the dialogue control logic from domain-independent conversational skills. (Daubigney et al., 2010) proposes a Kalman Temporal Differences based algorithm to learn efficiently in an off-policy manner a strategy for a large scale dialogue system. In (Emmanuel et al., 2013), authors propose a scheme to utilize a socially-based reward function for reinforcement learning and use it to fit the user adaptation issue for DM. (Young et al., 2013) provides an overview of the current state of the art in the development of POMDP-based spoken dialog systems. (Hao et al., 2014) presents a dialog manager based on a log-linear probabilistic model and uses context-free grammars to impart hierarchical structure to variables and features.

As we know, sentences in human-human dialogues are extremely complicated. The sentences labeled with the same DA might contain different extra meanings. Thus, it is difficult for DA to represent the nuanced meaning of sentence in dialogue. In this paper, we propose a novel DM scheme based on sentence clustering. The contributions of this work are as follows.

- Semantic representation of sentence in dialogue is defined as sentence cluster which could represent more nuanced semantic information than DA. Sentence similarity for clustering is calculated via internal information such as words and sentence structures and external information such as the distributed representation of sentence (vector) from Recurrent Neural Networks (RNN).

- The DM problem is modeled as a POMDP, where state is defined as sequence of sentence clusters, reward is defined as slot-filling efficiency and sentence popularity, and state transition probability is calculated by the prediction model based on RNN, considering historical dialogue information sufficiently.

The rest of this paper is organized as follows. In Section 2, system model is introduced. Section 3 describes sentence clustering and prediction model based on RNN, and Section 4 models the DM problem as a POMDP. Extensive experimental results are provided in Section 5 to illustrate the performance comparison, and Section 6 concludes this study.

Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Short Papers), pages 800–805,
Beijing, China, July 26-31, 2015.

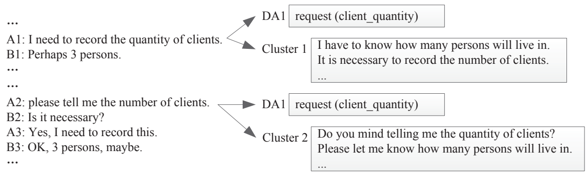

Figure 1: sentence cluster vs. DA

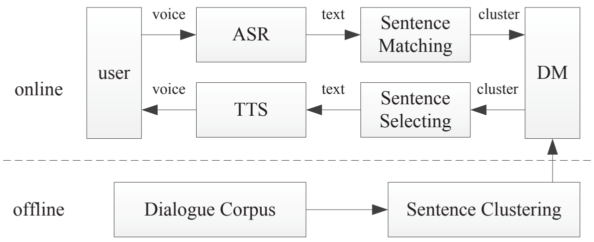

Figure 2: system model

2 System Model

In this paper, we establish a SDS via human-human dialogue corpus, where sentence cluster rather than DA is utilized to represent sentence intention due to its ability of catching finer-grained semantic information. For example, Fig. 1 shows some dialogue segments in hotel reservation. Both A1 and A2 could be labeled with "request (client_quantity)", because the aims of them are requesting the quantity of clients. However, A1 has an extra meaning that it is a necessity for the reception to record the quantity of clients, while A2 not, which might lead to different evolutions of dialogues. Probably, we could add this necessity to the DA corresponding to A1 manually, but it is infeasible for all the sentences to distinguish the fine-grained semantic information by adding abstract symbol to DA. Thus, in this paper, we automatically cluster all the sentences in dialogues, and utilize sentence clusters to represent sentence intentions, which has more powerful capability to capture semantic information.

The SDS based on sentence clustering could be divided into offline stage and online stage, illustrated in Fig. 2.

In offline stage:

Sentence Clustering: The sentence similarity is calculated based on not only internal information such as words and sentence structure, but also external information such as the distributed representation from RNN. And then the sentences in dialogue corpus are clustered into different tiny groups, which will be discussed in section 3.

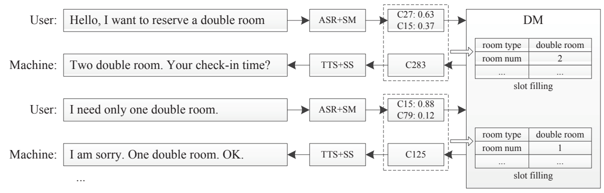

Figure 3: an online example

Dialogue Policy Training: We label the dialogues in corpus with the sentence clusters generated in the previous process. Thus, these labeled dialogues could be utilized to train the optimal dialogue policy with Reinforcement Learning, which will be introduced in section 4.

In online stage:

Automatic Speech Recognition (ASR): When receiving user voice, ASR module transforms it into text (Vinyals et al., 2012). As there might be ambiguity and errors in ASR, it is difficult to obtain the exact text corresponding to the input voice. Thus, the distribution over possible texts is used to represent the result of ASR.

Sentence Matching (SM): the function of SM is to establish a mapping from the distribution over possible texts to the distribution over possible sentence clusters.

DM: Based on the distribution of clusters, D-M model updates the belief state in POMDP and selects the optimal action, namely the optimal machine sentence cluster, according to the dialogue policy. The relevant slots are also filled based on the user and machine sentence clusters.

Sentence Selection: This module selects the most appropriate sentence from the output machine sentence cluster according to the user profile such as personality character (Ball et al., 2000).

Text To Speech (TTS): This model transforms the selected sentence text into the output voice as a response (Zen et al., 2007).

Fig. 3 is a human-machine dialogue example in online stage.

3 Sentence Clustering based on RNN

In this section, we cluster the sentences for DM modeling, which might be different from general sentence clustering. Sentence similarity for clustering are calculated from two aspects. Firstly, it is calculated traditionally based on internal information such as words and sentence structures, which is widely researched in (Li et al., 2006) (Achananuparp et al., 2008). (Word embedding

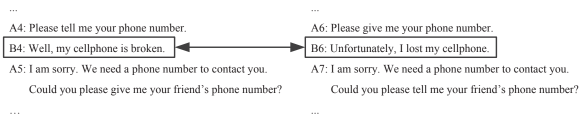

Figure 4: an example of sentence similarity

and sentence parsing might be used for this calculation.) Additionally, for DM-based sentence clustering, the sentences that we intend to put into the same cluster are not only the sentences with similar surface meaning, but also the sentences with similar intention (Semantics or Pragmatics), even if they might be different in surface meaning sometimes. For example, illustrated in Fig. 4, B4 and B6 are different in surface meaning, but they have similar intention, namely he or she might not provide his or her phone number right now. Thus, in the sentence clustering for DM modeling, they should be clustered into the same group. It is difficult to give a high similarity score between B4 and B6 only according to the internal information, but we could observe that the sentences around them in the context are similar. Thus, external information is also important to the sentence clustering for DM. In the following, we will discuss the clustering process.

We denote the sentence cluster set as $\mathscr{C}^k = \left\{c_1^k, c_2^k, \cdots, c_{N_C^k}^k\right\}$, and the dialogue set as $\mathscr{D}^k = \left\{d_1^k, d_2^k, \cdots, d_{N_D^k}^k\right\}$ in the k-th iteration. Thus, the steps of sentence clustering are:

Step 1: Initially, we only utilize the internal information to cluster the sentences via Affinity Propagation (AP) algorithm (Brendan et al., 2007) and denote the clustering result as $\mathscr{C}^0$. If $\mathscr{C}^0$ is used to label the sentences in dialogues, the j-th dialogue could be denoted as a sequence of clusters, namely $d_j^0 = \left\{c_1^0, c_2^0, \cdots, c_{N_j^d}^0\right\}$.

Step 2: In the k-th iteration, we use cluster set $\mathscr{C}^k$ to label dialogue set $\mathscr{D}^k$.

Step 3: We utilize RNN to obtain the distributed representation of sentence, illustrated in Fig. 5. The input of RNN is sentence cluster in each turn, namely c_t^k. The input layer $\mathbf{I}(t)$ is the one-hot representation of c_t^k. (Turian et al., 2010) (The size of $\mathbf{I}(t)$ is equivalent to $\left|\mathscr{C}^k\right|$. There is only one 1 in $\mathbf{I}(t)$ corresponding to the c_t^k position, and other elements are zeros.) $\mathbf{H}(t)$ is defined as the hidden layer. The output layer $\mathbf{O}(t)$ is the distribution over possible c_{t+1}^k, which could be calculated as

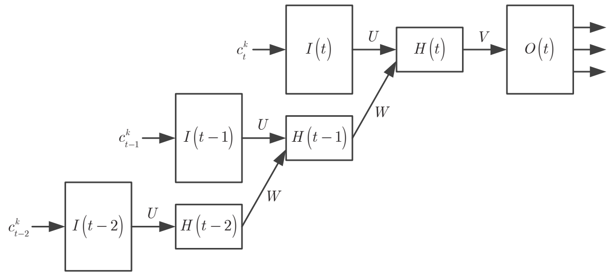

Figure 5: RNN for sentence clustering

follow. (Mikolov et al., 2010)

$$\begin{cases} \mathbf{H}(t) = f(\mathbf{UI}(t) + \mathbf{WH}(t-1)) \\ \mathbf{O}(t) = g(\mathbf{VH}(t)) \end{cases} \tag{1}$$

where $f(x) = 1/(1+e^{-x})$ and $g(x_i) = e^{x_i} \Big/ \sum_{i=1}^{N_e} e^{x_i}$. The parameters of this RNN could be trained by the Back Propagation Through Time (BPTT) algorithm. (Mikolov, 2012) From RNN, we could obtain two significant results: one is the distributed representation (vectors) of the sentence clusters ($\mathbf{U}$), which is used for sentence clustering; the other is the prediction model for sentence clusters, which is used for DM.

Step 4: we calculate the sentence similarity based on vectors obtained in **Step 3**, and combine it with the sentence similarity from internal information (weighted mean), in order to cluster the set $\mathscr{C}^k$ via AP algorithm, which is denoted as $\mathscr{C}^{k+1}$.

Step 5: $\bar{N}_C = \sum_{i=k-k_{th}+2}^{k+1} N_C^i$ is defined as the average number of clusters in the last k_{th} iteration. If $\sum_{i=k-k_{th}+2}^{k+1} \left|N_C^i - \bar{N}_C\right| < N_{th}$, stop the iteration of clustering, or go to **Step 2**, where N_{th} is the variation threshold of quantity of clusters.

Thus, in the last iteration, we get the cluster set $\mathscr{C}^{\bar{k}} = \left\{c_1^{\bar{k}}, c_2^{\bar{k}}, \cdots, c_{N_C^{\bar{k}}}^{\bar{k}}\right\}$ and prediction model for these sentence clusters. We divide all the sentences in dialogue corpus into the sentence set spoken by customers and the sentence set spoken by customer service representatives, and then utilize $\mathscr{C}^{\bar{k}}$ to label them respectively, which is denoted as $\mathscr{C}^u = \left\{c_1^u, c_2^u, \cdots, c_{N_u}^u\right\}$, namely the clusters of user sentences, and $\mathscr{C}^m = \left\{c_1^m, c_2^m, \cdots, c_{N_m}^m\right\}$, namely the clusters of machine sentences.

4 DM based on Sentence Clustering

The dialogue process mentioned in section 2 could be formulized as follows, illustrated in Fig. 6. It is defined $X = \{x_1, \cdots, x_T\}$ as inner (or exact) sentence cluster corresponding to the user input in each turn, which is unobservable and $x_t \in$

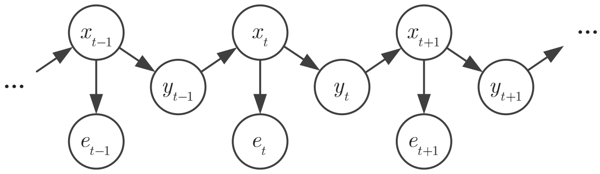

Figure 6: dialogue process

$\mathscr{C}^u$. $E = \{e_1, \cdots, e_T\}$ is defined as the input voice, which is observable to infer x_t in each turn. $Y = \{y_1, \cdots, y_T\}$ is defined as the output cluster of machine, where $y_t \in \mathscr{C}^m$. Thus, the DM problem is to find out the optimal y_t according to $\{e_1, y_1, \cdots, e_t\}$. In the following, the DM problem is modeled as a POMDP.

State in the t-th epoch is defined as the sequence of clusters, namely $s_t = \{x_{t-\tau}, y_{t-\tau}, \cdots, x_{t-1}, y_{t-1}, x_t\}$, where $s_t \in \mathscr{S}$. Action in the t-th epoch is defined as $a_t = y_t$, where $a_t \in \mathscr{A}$. The state transition probability $\Pr\{s_{t+1} | s_t, a_t\}$ could be shown as

$$\begin{aligned} &\Pr\{s_{t+1} | s_t, a_t\} \\ &= \Pr\{x_{t+1} | y_t, x_t, \cdots, y_{t-\tau}, x_{t-\tau}\} \end{aligned} \tag{2}$$

which is calculated by the prediction model based on RNN in section 3.

Observation is defined as $o_t = \{e_{t-\tau}, \cdots, e_t\}$, where $o_t \in \mathscr{O}$. As $\{x_{t-\tau}, \cdots, x_t\}$ in state s_t is unobservable, belief state is defined to represent the distribution over possible states, which is denoted as $b(t) \in \mathscr{B}$. According to (Kaelbling et al., 1998), the belief state updating could be represented as

$$b_{t+1}(s_{t+1}) = \frac{\Pr\{o_{t+1} | s_{t+1}, a_t\} p_{s_{t+1}}}{\Pr\{o_{t+1} | b_t, a_t\}} \tag{3}$$

where $p_{s_{t+1}} = \sum_{s_t \in \mathscr{S}} \Pr\{s_{t+1} | s_t, a_t\} b_t(s_t)$. According to Fig. 5, $\Pr\{o_{t+1} | s_{t+1}, a_t\}$ could be shown as

$$\begin{aligned} &\Pr\{o_{t+1} | s_{t+1}, a_t\} \\ &= \Pr\{o_{t+1} | s_{t+1}\} \\ &= \Pr\{e_{t-\tau+1}, \cdots, e_{t+1} | x_{t-\tau+1}, \cdots, y_t, x_{t+1}\} \\ &= \Pr\{e_{t-\tau+1}, \cdots, e_{t+1} | x_{t-\tau+1}, \cdots, x_{t+1}\} \\ &= \prod_{i=t-\tau+1}^{t+1} \Pr\{e_i | x_i\} \end{aligned} \tag{4}$$

However, it is difficult to obtain the probability $\Pr\{e_t | x_t\}$, as different people have different habits of expression and pronunciation. Fortunately, $\Pr\{x_t | e_t\}$ could be estimated based on ASR and SM. Thus, based on Bayes Rules, we have the following equation.

$$\Pr\{e_i | x_i\} = \frac{\Pr\{x_i | e_i\} \Pr\{e_i\}}{\Pr\{x_i\}} \tag{5}$$

where $\Pr\{x_t\}$ is the prior distribution of x_t and could be counted by corpus. With (4) and (5), (3) could be rewritten as

$$b_{t+1}(s_{t+1}) = \frac{\kappa \cdot p_{s_{t+1}} \cdot \prod\limits_{i=t-\tau+1}^{t+1} \Pr\{x_i | e_i\}}{\prod\limits_{i=t-\tau+1}^{t+1} \Pr\{x_i\}} \tag{6}$$

where

$$\kappa = \prod\nolimits_{i=t-\tau+1}^{t+1} \Pr\{e_i\} \Big/ \Pr\{o_{t+1} | b_t, a_t\} \tag{7}$$

is a normalization constant.

The reward function is defined as

$$r_t(s_t, a_t, s_{t+1}) = \lambda_f r^f_{(s_t, a_t, s_{t+1})} + \lambda_p r^p_{(s_t, a_t, s_{t+1})} \tag{8}$$

where $\lambda_f + \lambda_p = 1$ and $r_t(s_t, a_t, s_{t+1}) \in \mathscr{R}$. Firstly, $r^f_{(s_t, a_t, s_{t+1})}$ stands for the number of unfilled slots that are filled by the sequence of sentence clusters corresponding to (s_t, a_t, s_{t+1}). This slot-filling process could be achieved by a classifier trained by the dialogues labeled with sentence clusters and slot-filling information. (Inputs are cluster sequences, and outputs are filled slots.) Additionally, $r^p_{(s_t, a_t, s_{t+1})}$ is defined as the normalized quantity of s_{t+1} conditioned by s_t and a_t, which could be counted in corpus and stands for the popularity features of human-human dialogues. Thus, for the belief state, the reward function could be represented as

$$\begin{aligned} r_t(b_t, a_t) = &\sum_{s_{t+1} \in \mathscr{S}} \sum_{s_t \in \mathscr{S}} r_t(s_t, a_t, s_{t+1}) \\ &\cdot \Pr(s_{t+1} | s_t, a_t) b_t(s_t) \end{aligned} \tag{9}$$

Therefore, if we define the policy as a mapping from belief state to action, namely $\zeta \in \mathscr{Z} : \mathscr{B} \to \mathscr{A}$, the POMDP-based DM problem is shown as

$$\begin{aligned} &\max_{\zeta \in \mathscr{S}} E_\zeta \left[\sum_{t=1}^{T} \beta r_t(b_t, a_t) \right] \\ &s.t.\ b_{t+1}(s_{t+1}) = \frac{\kappa \prod\limits_{i=t-\tau+1}^{t+1} \Pr\{x_i | e_i\}}{\prod\limits_{i=t-\tau+1}^{t+1} \Pr\{x_i\}} \\ &\qquad \cdot \sum_{s_t \in \mathscr{S}} \Pr\{s_{t+1} | s_t, a_t\} b_t(s_t) \end{aligned} \tag{10}$$

where β is the time discount factor and $0 < \beta < 1$. This problem is a MDP problem with continuous states, which could be solved by the Natural Actor and Critic algorithm (Peters et al., 2008).

5 Experimental Results

In this section, we compare the performances of the proposed Sentence Clustering based Dialogue Management (SCDM) scheme and the existing DM scheme. The existing scheme is designed according to (Young et al., 2013), where DA is utilized to represent the semantic information of sentence and the dialogue policy is trained via Reinforcement Learning. It is also an extrinsic (or end-to-end) evaluation to compare the semantic representation ability between sentence cluster and DA.

In order to compare the performances of the DM schemes, we collect 171 human-human dialogues in hotel reservation and utilize 100 dialogues of them to establish a SDS. The residual 71 dialogues are used to establish a simulated user for testing (Schatzmann et al., 2006). We define the slots requested from machine to user as "room type", "room quantity", "checkin time", "checkout time", "client name" and "client phone". We also define the slots requested from users to machine as "hotel address = No.95 East St.", "room type set = single room, double room, and deluxe room", "single room price = \$80", "double room price = \$100", "deluxe room price = \$150". The hotel reservation task could be considered as a process of exchanging the slot information between machine and user to some extent.

Fig. 7 illustrates the dialogue turn in the DM schemes, using different training corpus. Here, we vary the size of training corpus from 10 dialogues to 100 dialogues and define average turn as the average dialogue turn cost to complete the task. From this picture, we find out that the SCDM scheme has lower average turn than the existing scheme, partly because the sentence are automatically clustered into many small groups that could represent more nuanced semantic information than DAs, partly because RNN could estimate next sentence cluster according to the vector in hidden layer that contains abundant historical dialogue information. As the number of sentence clusters is greater than number of DAs, RNN could also solve the scarcity problem and smoothing problem in the predicting process. Additionally, with the increment of training dialogue size, the average turn

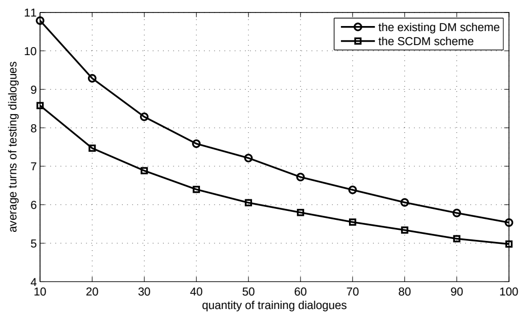

Figure 7: comparison of average turn

of dialogue decreases, which ought to be ascribed to the fact that more training data could let SDS reach more states with more times and increase the accuracy of the parameter estimation in RNN and POMDP. Furthermore, with the increment of training dialogue size, the dialogue turn improvement of the proposed scheme turns less obvious, because the number of new sentence pattern deceases with the training size increment.

6 Conclusion

In this paper, we focused on the DM scheme based on sentence clustering. Firstly, sentence cluster is defined as the semantic representation of sentence in dialogue, which could describe more naunced sentence intention than DA. Secondly, RNN is established for sentence clustering, where sentence similarity is calculated not only based on the internal information such as words and sentence structure, but also based on the external information such as context in dialogue. Thirdly, the DM problem is modeled as a POMDP, where the state is defined as the sequence of sentence clusters and the state transition probability is estimated by RNN, considering the whole information of historical dialogue. Finally, the experimental results illustrated that the proposed DM scheme is superior to the existing one.

Acknowledgments

This work is supported by the National Program on Key Basic Research Project (973 Program), basic theories and methods of Chinese Language Processing and Deep Computing in Internet environment, multi-lingual Automatic Speech Recognition for complex environments. (No. 2013CB329302)

References

Dan Bohus, Alexander, I. Rudnicky. 2009. The RavenClaw dialog management framework: Architecture and systems *Computer Speech and Language*, vol. 23, pages: 332-361, 2009.

Emmanuel Ferreira, Fabrice Lefvre. 2013. Social signal and user adaptation in reinforcement learning-based dialogue management. *MLIS '13*, Aug, 2013.

Brendan J. Frey, Delbert Dueck. 2007 Clustering by Passing Messages Between Data Points. *Science*, 2007.

Ball, G., Breese, J. 2000. Emotion and personality in a conversational agent. *Embodied conversational agents*, pages: 189-219, 2000.

Zen, H., Nose, T., Yamagishi, J., Sako, S., Masuko, T., Black, A. W., Tokuda, K. 2007 The HMM-based speech synthesis system version 2.0. *In Proc. 6th ISCA Workshop on Speech Synthesis*, Aug, 2007.

Schatzmann, J., Weilhammer, K., Stuttle, M., Young, S. 2006 A survey of statistical user simulation techniques for reinforcement-learning of dialogue management strategies. *The knowledge engineering review*, 21(02), 97-126, 2006.

Turian, J., Ratinov, L., Bengio, Y. 2010 Word representations: a simple and general method for semi-supervised learning. *In Proceedings of the 48th annual meeting of the association for computational linguistics*, Jul, 2010.

Peters, J., Schaal, S. 2008 Natural actor-critic. *Neurocomputinge*, 71(7), 1180-1190, 2008.

Kaelbling, L., Littman, M., and Cassandr, A. 1998 Planning and acting in partially observable stochastic domains. *Artif. Intell.*, vol.101, pages: 99-134, 1998.

Daubigney, L., Geist, M., Pietquin, O. 2012. Off-policy learning in large-scale POMDP-based dialogue systems. *EEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).*, pages: 4989-499, Mar, 2012.

Vinyals, O., Ravuri, S. V., Povey, D. 2012. Revisiting Recurrent Neural Networks for robust ASR. *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012.

Achananuparp, P., Hu, X., Shen, X. 2008. The evaluation of sentence similarity measures. *In Data Warehousing and Knowledge Discovery*, pages: 305-316, 2008.

Young, S., Gasic, M., Thomson, B., Williams, J. D. (2013). 2013. Pomdp-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, 101(5), pages: 1160-1179, 2013.

Mikolov, T. 2012 Statistical language models based on neural networks. *Presentation at Google, Mountain View*, 2012.

Mikolov, T., Karafit, M., Burget, L., Cernocky, J., Khudanpur, S. 2010 Recurrent neural network based language model. *11th Annual Conference of the International Speech Communication Association*, Sep, 2010.

Paek, T., Pieraccini, R. 2008. Automating spoken dialogue management design using machine learning: An industry perspective. *Speech communication*, 50(8), 716-729, 2008.

Hao Tang, Watanabe, S., Marks, T. K., Hershey, J. R. 2014. Log-linear dialog manager. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May, 2014.

Li Y, McLean D, Bandar Z A, et al. 2006. Sentence similarity based on semantic nets and corpus statistics. *Knowledge and Data Engineering, IEEE Transactions on*, 18(8), 1138-1150, 2006.