

# Dialogue State Tracking with Explicit Slot Connection Modeling

Yawen Ouyang<sup>1\*</sup>, Moxin Chen<sup>1\*</sup>, Xinyu Dai<sup>1†</sup>, Yinggong Zhao<sup>2</sup>  
Shujiang Huang<sup>1</sup> and Jiajun Chen<sup>1</sup>

<sup>1</sup>National Key Laboratory for Novel Software Technology, Nanjing University, China

<sup>2</sup>Leyan Technologies, China

{ouyangyw, chenmx}@smail.nju.edu.cn  
{daixinyu, huangsj, chenjj}@nju.edu.cn  
yinggong.zhao@gmail.com

## Abstract

Recent proposed approaches have made promising progress in dialogue state tracking (DST). However, in multi-domain scenarios, ellipsis and reference are frequently adopted by users to express values that have been mentioned by slots from other domains. To handle these phenomena, we propose a Dialogue State Tracking with Slot Connections (DST-SC) model to explicitly consider slot correlations across different domains. Given a target slot, the slot connecting mechanism in DST-SC can infer its source slot and copy the source slot value directly, thus significantly reducing the difficulty of learning and reasoning. Experimental results verify the benefits of explicit slot connection modeling, and our model achieves state-of-the-art performance on MultiWOZ 2.0 and MultiWOZ 2.1 datasets.

## 1 Introduction

Task-oriented dialogue systems assist users to achieve their certain goals, such as making a restaurant reservation or booking a taxi. To fulfill users' goals, dialogue state tracking (DST) is employed to estimate dialogue states at each turn. Dialogue states consist of constraints and requests conveyed by user utterances, typically are represented by a set of predefined slots and their corresponding values. For instance, the user utterance “*I am looking for a Korean restaurant in the centre*” mentions two slots, *food* and *area*, whose values are *Korean* and *centre* respectively.

Numerous methods are proposed to tackle the challenge of DST recently, and these methods can be mainly categorized into two types: fixed vocabulary and open vocabulary (Eric et al., 2019). Fixed vocabulary models are designed in the paradigm of multi-class classification, relying on a predefined

\* Equal contributions.

† Corresponding author.

Turns	Target Slot	Source Slot
U <sub>0</sub> : <i>I am looking for a Korean restaurant in the centre.</i>	restaurant-food	–
S <sub>0</sub> : <i>I have 1 restaurant name Little Seoul in the expensive price range.</i>	restaurant-area	–
	⋮	⋮
U <sub>2</sub> : <i>Are there any places to go in the same area as the restaurant?</i>	attraction-area	restaurant-area
S <sub>2</sub> : <i>There are dozens of places to go in city centre. What type of attraction are you interested in today?</i>	restaurant-area	–
	⋮	⋮
U <sub>5</sub> : <i>I also need a taxi to commute and need it to arrive at the restaurant.</i>	taxi-departure	attraction-name
	taxi-destination	restaurant-name
S <sub>5</sub> : <i>I have booked a cab to take you to the restaurant when you leave All Saint's church. The booked car type is a yellow volkswagen.</i>	attraction-area	restaurant-area
	restaurant-food	–
	restaurant-area	–
	⋮	⋮

Table 1: An example of multi-domain dialogue with slot connections expressed by ellipsis and reference. (We omit some turns and slots for simplicity.)

ontology (Henderson et al., 2014a; Mrkšić et al., 2017; Zhong et al., 2018). Open vocabulary approaches (Xu and Hu, 2018; Wu et al., 2019; Gao et al., 2019; Ren et al., 2019) break the assumption of predefined ontologies, turning to generate values only given target slots. Wu et al. (2019) propose a copy-augmented encoder-decoder model to track dialogue states, which outperforms fixed vocabulary models and achieves the state-of-the-art result in multi-domain DST.

Despite significant improvements achieved by those open vocabulary models, they always suffer from understanding enormous ellipsis and reference expressions in multi-domain scenarios. As shown in Table 1, there are several slot connections across multiple domains and turns. For example, at the second turn, the value of the target slot *attraction-area* is informed by a referring expression “*in the same area as the restaurant*”. Thus, the system needs to retrieve the value of its source slot *restaurant-area*. The last turn shows an obscurer utterance with multiple slot connections, in which target slots *taxi-departure* and *taxi-destination* are implicitly connected to their source slots *attraction-name* and *restaurant-name* respectively. For those

slots that need connections, existing methods attempt to find their values out from the lengthy dialogue history, which usually fail because of high learning complexity.

In this paper, we formally consider the above challenge as *related-slot problem* and propose a novel model DST-SC (Dialogue State Tracking with Slot Connections) to address it. We follow previous work to build a copy-augmented encoder-decoder model. Specially, DST-SC is designed with a slot connecting mechanism to establish the connection between the target slot and its source slot explicitly. Thus it can take advantage of the source slot value directly instead of reasoning from preceding turns. The contributions of this work are two-fold:

- To the best of our knowledge, this work is the first one to discuss the related-slot problem in multi-domain DST and address it by explicitly modeling slot connections across domains.
- We demonstrate that DST-SC is more effective for handling the related-slot problem and outperforms state-of-the-art baselines.

## 2 Model

In this section, we will describe DST-SC model in detail. DST-SC is an open vocabulary model based on the encoder-decoder architecture. As shown in Figure 1, there are three components that contribute to obtain the target slot value: (1) word generation from the vocabulary; (2) word copying from the dialogue history; (3) value copying from the source slot. To reduce the burden on the decoder, DST-SC also equips with a slot gate (Wu et al., 2019) to predict for slot values of *none* and *dontcare*.

### 2.1 Encoder

Our model uses a bi-directional GRU (Cho et al., 2014) to encode the dialogue history  $\mathbf{x} = \{w_1, w_2, \dots, w_m\}$ , where  $m$  is the number of tokens in the dialogue history. Each input token is first embedded using a word embedding function  $\phi^{emb}$  and then encoded into a fix-length vector  $\mathbf{h}_i$ .

$$\mathbf{h}_i = \text{GRU}(\phi^{emb}(w_i)). \quad (1)$$

### 2.2 Word Generation

We employ another GRU to decode slot values. Each slot is comprised of a domain name and a slot name, e.g., *hotel-area*. While decoding the

$j$ -th slot  $s_j$ , its summed embedding is fed as the first input. The last hidden state of the encoder initializes the decoder hidden state. At decoding step  $t$ , the hidden state is represented as  $\tilde{\mathbf{h}}_t^j$ . (The superscript  $j$  will be omitted for simplicity.)

Following the vanilla attention-based decoder architecture (Bahdanau et al., 2014),  $\tilde{\mathbf{h}}_t$  is used to apply attention over encoder outputs and aggregate them to get the context vector  $\mathbf{c}_t$ .

$$a_i^t = \text{softmax}(f_{mlp}([\tilde{\mathbf{h}}_t, \mathbf{h}_i])), \quad (2)$$

$$\mathbf{c}_t = \sum_{i=1}^m a_i^t \mathbf{h}_i. \quad (3)$$

The distribution of generating token  $y_t$  is given by:

$$P_{gen}(y_t) = \text{softmax}(W_{gen}[\tilde{\mathbf{h}}_t, \mathbf{c}_t]). \quad (4)$$

### 2.3 Word Copying

The copy mechanism is shown to be effective in DST (Lei et al., 2018; Xu and Hu, 2018; Wu et al., 2019). Here, we follow Wu et al. (2019) to augment the vanilla attention-based decoder with pointer-generator copying, enabling it to capture slot values that explicitly occur in the dialogue history.

$$P_{wc}(y_t = w) = \sum_{i:w_i=w} a_i^t. \quad (5)$$

A soft gate  $g_1$  is used to combine word copying distribution and generative distribution.

$$g_1 = \text{sigmoid}(W_{g_1}[\tilde{\mathbf{h}}_t, \mathbf{c}_t, \phi^{emb}(y_{t-1})]), \quad (6)$$

$$P_{orig}(y_t) = g_1 P_{gen}(y_t) + (1 - g_1) P_{wc}(y_t). \quad (7)$$

### 2.4 Slot Connecting Mechanism

As claimed in Section 1, connecting the target slot with its source slot helps to decrease the reasoning difficulty. Therefore, we enhance the copy-augmented encoder-decoder model with a slot connecting mechanism to model slot correlations directly. When decoding the target slot  $s_j$ , DST-SC infers its source slot from last dialogue states, then copies its value for the final distribution.

Last dialogue states are represented by (slot, value) tuples:  $\{(s_1, v_1), (s_2, v_2), \dots, (s_n, v_n)\}$ . We use  $\mathbf{h}_0$  as the query to attend the potential source slot.

$$a_k = \text{softmax}(f_{mlp}([\tilde{\mathbf{h}}_0, \mathbf{s}_k])), \quad (8)$$

where  $\mathbf{s}_k$  is the summed slot embedding,  $k \in \{1, 2, \dots, n\} \setminus \{j\}$ . Attention score  $a_k$  measures

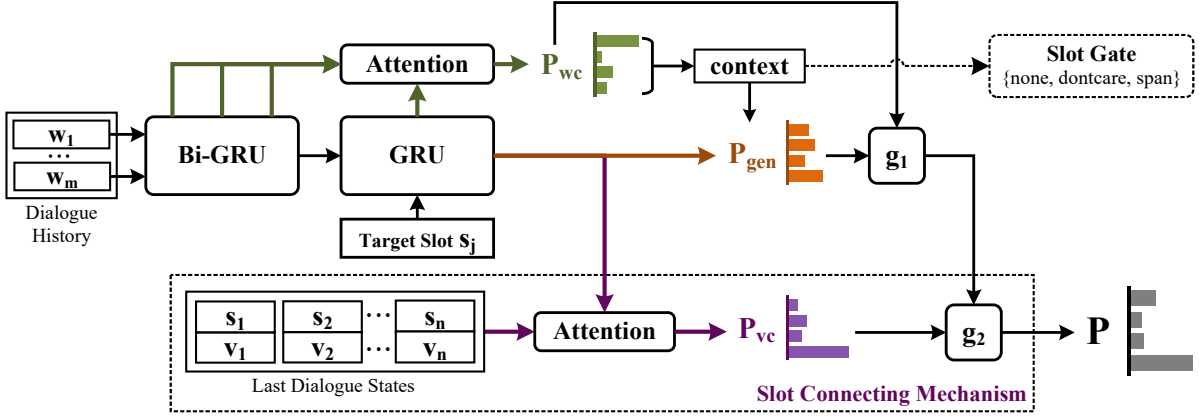


Figure 1: DST-SC model architecture (best viewed in color). Three processing flows leading to  $P_{gen}$ ,  $P_{wc}$ ,  $P_{vc}$  are respectively generation (brown), copying from dialogue history (green), copying from last dialogue states (purple).

how related  $s_k$  is to the target slot  $s_j$ . It is computed only once at the first decoding step and maintained consistency to subsequent tokens in the value  $v_k$ . At the  $t$ -th decoding step, the  $t$ -th token  $v_{kt}$  contributes to form value copying distribution  $P_{vc}(y_t)$ .

$$P_{vc}(y_t = w) = \sum_{k: v_{kt}=w} a_k. \quad (9)$$

Similar to the copy-augmented decoder, we combine value copying distribution and original distributions using a soft gate  $g_2$  to get final output distribution.

$$g_2 = \text{sigmoid}(W_{g_2} \mathbf{c}_0), \quad (10)$$

$$P(y_t) = g_2 P_{vc}(y_t) + (1 - g_2) P_{orig}(y_t). \quad (11)$$

### 3 Experimental Setup

#### 3.1 Datasets

To evaluate the effectiveness of DST-SC, we conducted experiments on MultiWOZ 2.0 (Budzianowski et al., 2018) and MultiWOZ 2.1 datasets (Eric et al., 2019). MultiWOZ 2.0 is a multi-domain dialogues corpus, and some annotation errors are corrected in MultiWOZ 2.1.

#### 3.2 Baselines

We compare DST-SC with several baseline methods. FJST and HJST (Eric et al., 2019) apply a separate feed-forward network to classify for every single state slot. HyST (Goel et al., 2019) is a hybrid approach, which combines the joint tracking fixed vocabulary approach and open vocabulary approach. COMER (Ren et al., 2019) adopts three hierarchically stacked decoders to generate

dialogue states. TRADE (Wu et al., 2019) generates dialogue states from the dialogue history using a copy-augmented decoder.

#### 3.3 Implementation Details

In our experiments, we used Glove (Pennington et al., 2014) and character embeddings (Hashimoto et al., 2017) to initialize word embeddings, each word is represented by a 400-dimensional vector. The hidden sizes of all GRU layers are set to 400. In the training phase, we used ground truth prior-turn dialogue states in the slot connecting mechanism. Adam optimizer (Kingma and Ba, 2015) is applied with 0.001 learning rate initially. The learning rate then reduced by a factor of 0.2, and the training stopped early when the performance in validation set was not improved for 6 consecutive epochs. We used a batch size of 32 and dropout rate of 0.2. Greedy search strategy is used for decoding, with maximum 10 decoded tokens and 50% probability of teacher forcing. Also, we followed previous work to utilize our model with the word dropout (Wu et al., 2019) by masking input tokens with a 20% probability. All experiments are averaged across 3 seeds.

### 4 Results and Analysis

#### 4.1 Experimental Results

We follow previous work to compare the performance of joint goal accuracy. We get the joint goal correct if the predicted state exactly matches the ground truth state for every slot. As shown in Table 2, open vocabulary approaches achieve higher accuracy than fixed vocabulary approaches. DST-SC achieves state-of-the-art performance on

Model	MultiWOZ 2.0	MultiWOZ 2.1
FJST <sup>†</sup>	40.20%	38.00%
HJST <sup>†</sup>	38.40%	35.55%
HyST <sup>†</sup>	42.33%	38.10%
COMER <sup>†</sup>	48.79%	–
TRADE <sup>1</sup>	50.83%	48.29%
DST-SC	<b>52.24%</b>	<b>49.58%</b>

Table 2: Joint goal accuracy on MultiWOZ 2.0 and MultiWOZ 2.1. Results marked with † are from original papers.

MultiWOZ 2.0 and MultiWOZ 2.1, with the joint goal accuracy of 52.24% and 49.58%.

## 4.2 Related-slot Tests

We conducted further related-slot tests to verify the effectiveness of DST-SC in solving the related-slot problem. The dataset for related-slot tests is constructed by manually extracting dialogues with the related-slot problem from MultiWOZ 2.1 test set. We made an observation that slot connections are common at target slots such as *attraction-area*, *hotel-area*, *hotel-book day* and so on. We only need to focus on target slot accuracy of turns with slot connections. However, some target slots occur infrequently in the extracted dataset. Considering that target slots from different domains with the same slot type always correspond to similar slot connection expressions, we can neglect their domains and calculate the accuracy of each slot type instead. For example, we can calculate the accuracy of slot type price instead of calculating the accuracy of *hotel-price range* and *restaurant-price range* separately. Table 3 lists slot types and their corresponding target slots.

To make more convincing tests, we performed data augmentations to get more samples for each slot type. We used two heuristic rules to augment the extracted data and obtained 100 dialogues for each slot type. **(1) Paraphrasing:** we rewrote some utterances to get multiple phrases with the same intent. For example, the phrase “*in the same area as the restaurant*” can be rewritten as “*close to the restaurant*”. **(2) Replacing values:** we replaced some slot values to exclude the influence of overfitting. For example, the phrase “*stay in the east*” can be replaced as “*stay in the west*”.

<sup>1</sup>We re-implemented TRADE as described in section 2.2 and section 2.3 and got a stronger baseline.

Slot Type	Target Slots
area	attraction-area, hotel-area, restaurant-area
day	hotel-book day, train-day, restaurant-book day
people	hotel-book people, restaurant-book people, train-book people
departure	taxi-departure
destination	taxi-destination
price	hotel-price range, restaurant-price range
time	restaurant-book time, taxi-arrive by, taxi-leave at, train-arrive by, train-leave at

Table 3: Slot types and corresponding target slots involved in related-slot tests.

As shown in Table 4, DST-SC outperforms TRADE by a large margin at most slot types. Case 1 in Table 5 illustrates the advantage of DST-SC explicitly. We find that both generation and word copying miss the correct token. However, the slot connecting mechanism in DST-SC helps to find out the correct source slot and merges its value into  $P$  under the control of gate  $g_2$ .

Note that there are no obvious improvements on slot types departure and destination. We suspect that this is caused by lots of missing annotations for *attraction-name*, *hotel-name* and *restaurant-name*, which usually act as source slots for departure and destination. The absence of these critical information makes DST-SC pay less attention to values from source slots. As shown in case 2 in Table 5, even if the slot connection mechanism has inferred the correct source slot, the unconfidence of  $g_2$  leads to the final incorrect output.

## 5 Related Work

Traditional approaches for dialogue state tracking (Henderson et al., 2014b; Sun et al., 2014; Zilka and Jurcicek, 2015; Mrkšić et al., 2015) rely on manually constructed semantic dictionaries to extract features from input text, known as delexicalisation. These methods are vulnerable to linguistic variations and difficult to scale. To overcome these problems, Mrkšić et al. (2017) propose the first data-driven model for DST, the employed deep learning approaches provide stronger representa-

Model	area	day	departure	destination	people	price	time
TRADE	49.33%	16.00%	<b>49.66%</b>	48.33%	12.00%	26.33%	86.66%
DST-SC	<b>86.33%</b>	<b>92.00%</b>	46.66%	<b>48.66%</b>	<b>87.00%</b>	<b>53.33%</b>	<b>87.33%</b>

Table 4: Slot type accuracy of related-slot tests.

Case 1: dialogue_idx=PMUL0129 (success)							Case 2: dialogue_idx=MUL1228 (failure)						
U <sub>1</sub> : I want to book a table for 4 people ...							S <sub>3</sub> : I have 1 hotel in the moderate range, cityroomz. Would you like ...						
...							U <sub>4</sub> : Yes, please. Can you book a room for Friday for 1 person, 3 nights?						
S <sub>3</sub> : The Bridge guest house is available. Would you like ...							...						
U <sub>4</sub> : Yes, please. For the same number of people, 2 nights ...							U <sub>6</sub> : ... I need the taxi to take me to the hotel.						
Target slot: hotel-book people=4							Target slot: taxi-destination=cityroomz						
Source slot: restaurant-book people=4							Source slot: hotel-name=cityroomz						
Model	$P_{gen}$	$P_{wc}$	$P_{vc}$	$g_1$	$g_2$	$P$	Model	$P_{gen}$	$P_{wc}$	$P_{vc}$	$g_1$	$g_2$	$P$
TRADE	“3”	“people”	–	0.999	–	“3”	TRADE	“none”	“peking”	–	0.148	–	“peking”
DST-SC	“1”	“the”	“4”	0.999	<b>0.991</b>	“4”	DST-SC	“lensfield”	“hotel”	“cityroomz”	0.942	<b>0.078</b>	“lensfield”

Table 5: Case study. We only list tokens with the highest output probability in  $P_{gen}$ ,  $P_{wc}$ ,  $P_{vc}$  and  $P$ .

tion learning ability. By sharing parameters among slots (Ren et al., 2018; Zhong et al., 2018; Nouri and Hosseini-Asl, 2018), the model is further improved to track rare slot values. These approaches are all designed in the paradigm of multi-class classification over predefined slot value candidates and usually referred to as fixed vocabulary approaches.

Fixed vocabulary approaches always require a predefined ontology, which is usually impractical. Their applications are usually limited in a single domain. Therefore, several open vocabulary approaches in generative fashion (Xu and Hu, 2018; Wu et al., 2019; Gao et al., 2019; Ren et al., 2019) are proposed to handle unlimited slot values in more complicated dialogues. Open vocabulary models show the promising performance in multi-domain DST. However, ellipsis and reference phenomena among multi-domain slots are still less explored in existing literature.

## 6 Conclusion

In this paper, we highlight a regularly appeared yet rarely discussed problem in multi-domain DST, namely the related-slot problem. We propose a novel dialogue state tracking model DST-SC, which equips with the slot connecting mechanism to build slot connections across domains. Our model achieves significant improvements on two public datasets and shows effectiveness on related-slot problem tests. Annotations complement for MultiWOZ dataset in the future might enable DST-SC to handle the related-slot problem more effectively and further improve the joint accuracy.

## Acknowledgments

We would like to thank the anonymous reviewers for their constructive comments. This work is supported by the National Natural Science Foundation of China (Nos. 61976114 and 61936012), the National Key R&D Program of China (No. 2018YFB1005102).

## References

- Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.
- Paweł Budzianowski, Tsung-Hsien Wen, Bo-Hsiang Tseng, Iñigo Casanueva, Stefan Ultes, Osman Ramadan, and Milica Gašić. 2018. **MultiWOZ - a large-scale multi-domain wizard-of-Oz dataset for task-oriented dialogue modelling**. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 5016–5026, Brussels, Belgium. Association for Computational Linguistics.
- Kyunghyun Cho, Bart van Merriënboer, Dzmitry Bahdanau, and Yoshua Bengio. 2014. **On the properties of neural machine translation: Encoder-decoder approaches**. In *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*, pages 103–111, Doha, Qatar. Association for Computational Linguistics.
- Mihail Eric, Rahul Goel, Shachi Paul, Abhishek Sethi, Sanchit Agarwal, Shuyag Gao, and Dilek Hakkani-Tur. 2019. **Multiwoz 2.1: Multi-domain dialogue state corrections and state tracking baselines**. In *arXiv preprint arXiv:1907.01669*.

- Shuyang Gao, Abhishek Sethi, Sanchit Agarwal, Tagyoung Chung, and Dilek Hakkani-Tur. 2019. [Dialog state tracking: A neural reading comprehension approach](#). In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*, pages 264–273, Stockholm, Sweden. Association for Computational Linguistics.
- Rahul Goel, Shachi Paul, and Dilek Hakkani-Tür. 2019. Hyst: A hybrid approach for flexible and accurate dialogue state tracking. *arXiv preprint arXiv:1907.00883*.
- Kazuma Hashimoto, Caiming Xiong, Yoshimasa Tsuruoka, and Richard Socher. 2017. [A joint many-task model: Growing a neural network for multiple NLP tasks](#). In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1923–1933, Copenhagen, Denmark. Association for Computational Linguistics.
- Matthew Henderson, Blaise Thomson, and Steve Young. 2014a. [Word-based dialog state tracking with recurrent neural networks](#). In *Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 292–299.
- Matthew Henderson, Blaise Thomson, and Steve Young. 2014b. [Word-based dialog state tracking with recurrent neural networks](#). In *Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 292–299, Philadelphia, PA, U.S.A. Association for Computational Linguistics.
- Diederik P. Kingma and Jimmy Ba. 2015. [Adam: A method for stochastic optimization](#). In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.
- Wenqiang Lei, Xisen Jin, Min-Yen Kan, Zhaochun Ren, Xiangnan He, and Dawei Yin. 2018. [Sequicity: Simplifying task-oriented dialogue systems with single sequence-to-sequence architectures](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1437–1447, Melbourne, Australia. Association for Computational Linguistics.
- Nikola Mrkšić, Diarmuid Ó Séaghdha, Blaise Thomson, Milica Gašić, Pei-Hao Su, David Vandyke, Tsung-Hsien Wen, and Steve Young. 2015. [Multi-domain dialog state tracking using recurrent neural networks](#). In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 794–799, Beijing, China. Association for Computational Linguistics.
- Nikola Mrkšić, Diarmuid Ó Séaghdha, Tsung-Hsien Wen, Blaise Thomson, and Steve Young. 2017. [Neural belief tracker: Data-driven dialogue state tracking](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1777–1788, Vancouver, Canada. Association for Computational Linguistics.
- Elnaz Nouri and Ehsan Hosseini-Asl. 2018. [Toward scalable neural dialogue state tracking model](#). *CoRR*, abs/1812.00899.
- Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. [Glove: Global vectors for word representation](#). In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar. Association for Computational Linguistics.
- Liliang Ren, Jianmo Ni, and Julian McAuley. 2019. [Scalable and accurate dialogue state tracking via hierarchical sequence generation](#). In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 1876–1885, Hong Kong, China. Association for Computational Linguistics.
- Liliang Ren, Kaige Xie, Lu Chen, and Kai Yu. 2018. [Towards universal dialogue state tracking](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2780–2786, Brussels, Belgium. Association for Computational Linguistics.
- Kai Sun, Lu Chen, Su Zhu, and Kai Yu. 2014. [The SJTU system for dialog state tracking challenge 2](#). In *Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 318–326, Philadelphia, PA, U.S.A. Association for Computational Linguistics.
- Chien-Sheng Wu, Andrea Madotto, Ehsan Hosseini-Asl, Caiming Xiong, Richard Socher, and Pascale Fung. 2019. [Transferable multi-domain state generator for task-oriented dialogue systems](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 808–819, Florence, Italy. Association for Computational Linguistics.
- Puyang Xu and Qi Hu. 2018. [An end-to-end approach for handling unknown slot values in dialogue state tracking](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1448–1457, Melbourne, Australia. Association for Computational Linguistics.
- Victor Zhong, Caiming Xiong, and Richard Socher. 2018. [Global-locally self-attentive encoder for dialogue state tracking](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1458–1467, Melbourne, Australia. Association for Computational Linguistics.
- Lukás Zilka and Filip Jurčiček. 2015. Incremental lstm-based dialog state tracker. *2015 IEEE Workshop on*

*Automatic Speech Recognition and Understanding*  
(ASRU), pages 757–762.