Mars: Modeling Context & State Representations with Contrastive Learning for End-to-End Task-Oriented Dialog

Haipeng Sun, Junwei Bao*, Youzheng Wu, Xiaodong He
JD AI Research, Beijing, China
{sunhaipeng6, baojunwei, wuyouzheng1, hexiaodong}@jd.com

Abstract

Traditional end-to-end task-oriented dialog systems first convert dialog context into belief state and action state before generating the system response. The system response performance is significantly affected by the quality of the belief state and action state. We first explore what dialog context representation is beneficial to improving the quality of the belief state and action state, which further enhances the generated response quality. To tackle our exploration, we propose **Mars**, an end-to-end task-oriented dialog system with two contrastive learning strategies to model the relationship between dialog context and belief/action state representations. Empirical results show dialog context representations, which are more different from semantic state representations, are more conducive to multi-turn task-oriented dialog. Moreover, our proposed Mars achieves state-of-the-art performance on the MultiWOZ 2.0, CamRest676, and CrossWOZ[1].

1 Introduction

Task-oriented dialog system (Zhang et al., 2020c) aims to assist users in completing some specific tasks such as table reservations, hotel reservations, ticket booking, and online shopping. Traditional task-oriented dialog system has been built through dialog state tracking (Lee et al., 2019; Wu et al., 2019), dialog policy (Schulman et al., 2017; Takanobu et al., 2019) and natural language generation (Wen et al., 2015) tasks. dialog state tracking transfers dialog context to belief state, which is the structured semantic state capturing the whole dialog context information. The belief state is used for the dialog system to query the database to obtain matched entities. Dialog policy selects an action state, a semantic state guiding the dialog system to generate a system response based on the current dialog context and database information. System response is generated through a natural language generation task.

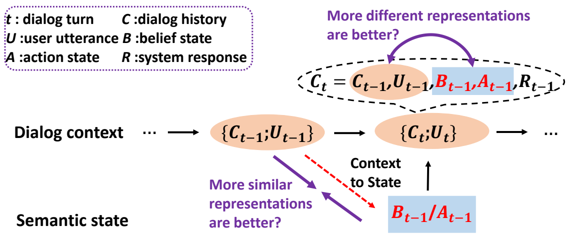


Figure 1: Illustration of the dialog context composition. Context to state represents dialog state tracking and dialog policy tasks. The previous dialog context $\{C_{t-1}, U_{t-1}\}$ is included in the dialog context $\{C_t, U_t\}$ of turn t. The database state is omitted for clarity.

With the widespread application of large-scale pre-training models (Devlin et al., 2019; Radford et al., 2019; Raffel et al., 2020), researchers gradually focus on the end-to-end task-oriented dialog system (Lin et al., 2020; Hosseini-Asl et al., 2020; Yang et al., 2021), which converts the whole dialog context into system response through multi-task training. Generally, an end-to-end task-oriented dialog modeling task is formulated as a cascaded generation problem (Su et al., 2021). Before generating a system response, the end-to-end task-oriented dialog system must first transfer dialog context into belief and action states, respectively. The quality of the belief state and action state greatly influence on the end-to-end task-oriented dialog performance[2].

In this paper, we explore what dialog context representation is beneficial to improving the quality of the belief/action state, which further enhances the generated response quality. As illustrated in Figure 1, dialog context is recursively hybrid of previous dialog context and semantic states[3],

*Corresponding author: baojunwei001@gmail.com

[1]The code is available at https://github.com/hpsun1109/Mars.

[2]The detailed analysis is provided in Appendix C.

[3]These previous semantic states are helpful references for

Findings of the Association for Computational Linguistics: ACL 2023, pages 11139–11160
July 9-14, 2023 ©2023 Association for Computational Linguistics

i.e., belief and action states, for multi-turn dialog. Intuitively, the representation of dialog context $\{C_{t-1}, U_{t-1}\}$, which is more similar with that of semantic states B_{t-1}/A_{t-1}, is beneficial to generate semantic states of the turn $t-1$. However, if their representations are too similar, there may be information redundancy in the representation of dialog context $\{C_t, U_t\}$ in turn t, as shown in Figure 1. Thus we raise another conjecture: whether representations of dialog context, which are more different from that of semantic states, are more conducive to multi-turn task-oriented dialog?

To tackle our conjectures, we propose **Mars**, an end-to-end task-oriented dialog system with two contrastive learning strategies, i.e., pair-aware context&state and group-aware context&state contrastive learning, to model the relationship between dialog context and semantic states from two different levels. Specifically, (1) the pair-aware context&state contrastive learning strategy focuses more on narrowing the gap in the continuous representation space between dialog context and corresponding semantic states for the same dialog turn. This strategy aims to obtain a continuous representation of the dialog context that is semantically more consistent with that of its semantic states. (2) Group-aware context&state contrastive learning strategy enlarges the overall continuous representation margin between dialog context and semantic states. The meaning behind this is to make representations between dialog context and semantic states more different. Extensive experiments and analysis on the response generation and dialog state tracking tasks verify our raised conjectures and the effectiveness of Mars. Mars achieves state-of-the-art performance on the MultiWOZ 2.0, CamRest676, and CrossWOZ. Moreover, Mars achieves remarkable performance in the low-resource scenario. Finally, we perform detailed error analysis and visualization to better apply our proposed Mars to real-world scenarios.

This paper primarily makes the following contributions: (1) We explore what dialog context representation is beneficial to improving task-oriented dialog performance. (2) We propose two contrastive learning strategies to model the relationship between dialog context and semantic state representations. (3) Empirical results show Mars achieves state-of-the-art performance on the MultiWOZ 2.0, CamRest676, and CrossWOZ.

the generation of the current turn (Yang et al., 2021).

2 Related Work

End-to-end task-oriented dialog systems (Lei et al., 2018; Zhang et al., 2020a,b) are established via copy-augmented seq2seq learning (Gu et al., 2016). Zhang et al. (2020b) proposes a multi-action data augmentation method to improve the diversity of generated system responses. Large-scale pre-trained language models, including BERT (Devlin et al., 2019), GPT-2 (Radford et al., 2019), T5 (Raffel et al., 2020), and UniLM (Dong et al., 2019), have been demonstrated effective for improving the performance of task-oriented dialog systems (Hosseini-Asl et al., 2020; Peng et al., 2021; Lin et al., 2020; Yang et al., 2021; Jeon and Lee, 2021; He et al., 2022) on MultiWOZ 2.0 (Budzianowski et al., 2018), a large-scale English multi-domain task-oriented dialog dataset. Recently, auxiliary tasks and auxiliary dialog corpora have been introduced to further improve dialog modeling ability. MTTOD (Lee, 2021) introduces a span prediction task to enhance the natural language understanding performance. BORT (Sun et al., 2022) proposes reconstruction strategies to alleviate the error propagation problem. PPTOD (Su et al., 2021) proposes a dialog multi-task pre-training strategy to model task completion from auxiliary heterogeneous dialog corpora. GALAXY (He et al., 2022) introduces a dialog act prediction task to explicitly learn dialog policy from auxiliary dialog corpora.

Recently, contrastive Learning (He et al., 2020; Chen et al., 2020; Grill et al., 2020; Chen and He, 2021) has attracted much attention in the computer vision community and has been applied to natural language processing to enhance sentence representation learning (Fang and Xie, 2020; Wu et al., 2020; Yan et al., 2021; Gao et al., 2021; Giorgi et al., 2021). In contrast, we propose contrastive learning strategies to model the relationship between dialog context and semantic state representations for task-oriented dialog. In addition, we don't introduce data augmentation methods, which are used in most contrastive learning works.

3 Task-Oriented Dialog Framework

Generally, an end-to-end task-oriented dialog modeling task is formulated as a cascaded generation problem (Su et al., 2021). Before generating a system response, the end-to-end task-oriented dialog system would transfer dialog

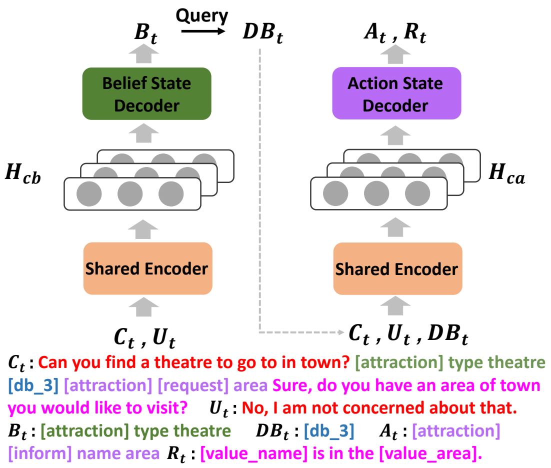


Figure 2: Illustration of a general task-oriented dialog system. For clarity, we take dialog turn $t = 1$ as an example. C_t is formulated as $\{U_0, B_0, DB_0, A_0, R_0\}$. [db_3] denotes the amount of matched entities.

context into belief state and action state, respectively. Belief state is a semantic state of dialog context, including dialog domain, slot name, and slot value. Action state is a semantic state of system response, including dialog domain, dialog act, and slot name. For example, the belief state is '*[attraction] type theatre*', and the action state is '*[attraction] [inform] name area*'.

We construct an end-to-end task-oriented dialog system via the seq2seq framework, including one shared encoder and two different decoders, as illustrated in Figure 2. One shared encoder encodes dialog context, one decoder $decoder_b(\cdot)$ decodes belief state, and another decoder $decoder_a(\cdot)$ decodes action state and system response. Consider a dialog in turn t, dialog history C_t, which contains dialog information for all previous turns, is formulated as $\{C_{t-1}, U_{t-1}, B_{t-1}, DB_{t-1}, A_{t-1}, R_{t-1}\}$, where U represents the user utterance, B represents the belief state, DB represents the database state, A represents the action state, and R represents the system response.

For end-to-end dialog modeling, a belief state is first generated. The dialog history C_t and the current user utterance U_t are firstly encoded into hidden representation H_{cb} through the shared encoder, and the belief state B_t is generated through the belief state decoder:

$$\begin{aligned} H_{cb} &= encoder(C_t, U_t), \\ B_t &= decoder_b(H_{cb}). \end{aligned} \tag{1}$$

The dialog state tracking process is optimized by minimizing the following objective function:

$$\mathcal{L}_B = -logP(B_t|C_t, U_t). \tag{2}$$

We use the generated belief state B_t to query the specific database to achieve the database state DB_t, which means the amount of matched entities.

As described by MTTOD (Lee, 2021), the second decoder would be used to generate action state and system response simultaneously. The combination of the dialog history C_t, the current user utterance U_t, and the database state DB_t are encoded into hidden representation H_{ca} through the shared encoder. The action state A_t and system response R_t are generated in turn through the action state decoder:

$$\begin{aligned} H_{ca} &= encoder(C_t, U_t, DB_t), \\ A_t, R_t &= decoder_a(H_{ca}). \end{aligned} \tag{3}$$

Therefore, the action state and response generation process is optimized by minimizing the following objective function:

$$\mathcal{L}_{AR} = -logP(A_t, R_t|C_t, U_t, DB_t). \tag{4}$$

In summary, the entire end-to-end task-oriented dialog system can be optimized by minimizing:

$$\mathcal{L}_{all} = \mathcal{L}_B + \mathcal{L}_{AR}. \tag{5}$$

4 Methodology

To tackle our conjectures and enhance the relationship modeling between dialog context and corresponding semantic state representations of the task-oriented dialog system described in Section 3, we propose two contrastive learning methods: pair-aware context&state and group-aware context&state contrastive learning. Figure 3 illustrates the architecture of a task-oriented dialog system with our proposed methods. Generally, for any contrastive learning method, contrastive learning objective functions $\mathcal{L}_{bscl}$ and $\mathcal{L}_{ascl}$ are added for dialog state tracking and response generation tasks, respectively, to enhancing the relationship modeling between dialog context and semantic state representations during end-to-end dialog training. The general objective function can be reformulated as follows:

$$\begin{aligned} \mathcal{L}_{all} &= \mathcal{L}_{B'} + \mathcal{L}_{AR'}, \\ \mathcal{L}_{B'} &= \mathcal{L}_B + \lambda_1 \mathcal{L}_{bscl}, \\ \mathcal{L}_{AR'} &= \mathcal{L}_{AR} + \lambda_2 \mathcal{L}_{ascl}, \end{aligned} \tag{6}$$

where λ_1 and λ_2 are hyper-parameters that adjust the weight of the objective functions.

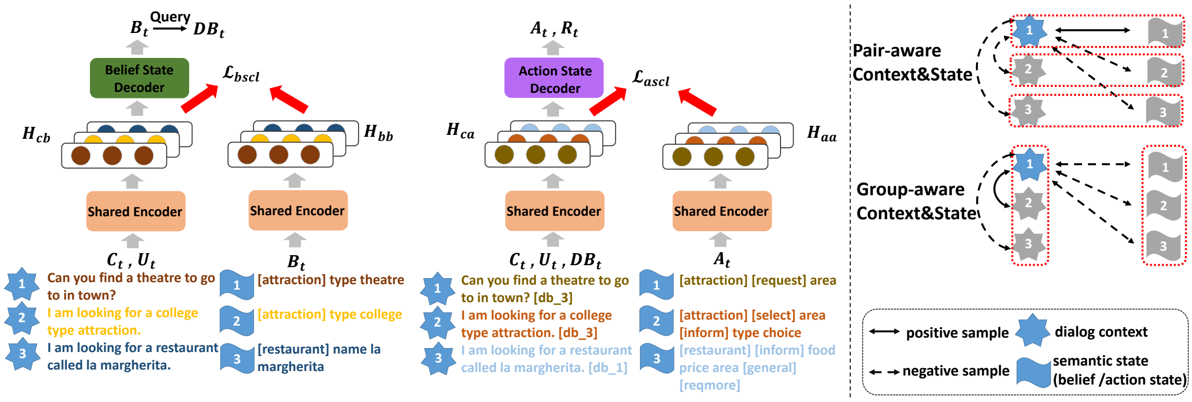


Figure 3: Illustration of the task-oriented dialog system with contrastive learning strategies. We take dialog turn $t = 0$ and batch size $N = 3$ as an example.

4.1 Pair-aware Context&State Contrastive Learning

To achieve dialog context representation, which is semantically more consistent with its semantic state representation, we propose a pair-aware context&state contrastive learning strategy (Mars-P) to close the continuous representation gap between dialog context $\{C_t, U_t\}$ and corresponding semantic states, including belief state B_t and action state A_t, for the same dialog turn.

We consider the dialog context $\{C_t, U_t\}$ and the belief state B_t from the same dialog to be as consistent as possible in the representation space, while the dialog context is as far away from other belief states as possible. As illustrated in Figure 3, the source continuous representation of the dialog context '*can you find a theater to go to in town?*' should be similar to that of the belief state '*[attraction] type theatre*' rather than other belief states '*[attraction] type college*' and '*[restaurant] name la margherita*'.

Specifically, the belief state B_t would be encoded into a hidden representation H_{bb} through the shared encoder:

$$H_{bb} = encoder(B_t). \tag{7}$$

For every dialog context input in a batch, we treat the corresponding belief state from the same dialog as a positive sample and other belief states and dialog contexts in the same batch as negative samples. Therefore, this dialog model is optimized by minimizing the objective function:

$$\mathcal{L}_{bscl} \triangleq \mathcal{L}_{bscl\_P} = -log \frac{e^{cos(H^i_{cb}, H^i_{bb})/T}}{\sum_{\substack{k=1\\k\neq i}}^{N} e^{cos(H^i_{cb}, H^k_{cb})/T} + \sum_{k=1}^{N} e^{cos(H^i_{cb}, H^k_{bb})/T}}, \tag{8}$$

where $cos(\cdot)$ denotes the cosine similarity function. T is a temperature hyperparameter. N is the batch size. In a batch, H^i_{cb} denotes the ith dialog context hidden representation after average pooling, and H^k_{bb} denotes the kth belief state hidden representation after average pooling.

During response generation, we would close the continuous representation gap of dialog context $\{C_t, U_t, DB_t\}$ and action state A_t. As illustrated in Figure 3, the source continuous representation of the user utterance '*i am looking for a restaurant called la margherita.*' and database information '*[db$_1$]*' should be similar to that of the action state '*[restaurant] [inform] food price area [general] [reqmore]*' rather than other action states '*[attraction] [request] area*' and '*[attraction] [select] area [inform] type choice*'. Specifically, the action state A_t would be encoded into a hidden representation H_{aa} through the shared encoder:

$$H_{aa} = encoder(A_t). \tag{9}$$

For every dialog context input in a batch, we treat the corresponding action state from the same dialog as a positive sample and other action states and dialog contexts in the same batch as negative samples. Therefore, this dialog model is optimized by minimizing the objective function:

$$\mathcal{L}_{ascl} \triangleq \mathcal{L}_{ascl\_P} = -log \frac{e^{cos(H^i_{ca}, H^i_{aa})/T}}{\sum\limits_{\substack{k=1 \\ k \neq i}}^{N} e^{cos(H^i_{ca}, H^k_{ca})/T} + \sum\limits_{k=1}^{N} e^{cos(H^i_{ca}, H^k_{aa})/T}}, \tag{10}$$

where H^i_{ca} denotes the ith dialog context hidden representation after average pooling, and H^k_{aa} denotes the kth action state hidden representation after average pooling.

4.2 Group-aware Context&State Contrastive Learning

To explore whether representations of dialog context, which are more different from that of semantic states, are more conducive to multi-turn task-oriented dialog, we propose a group-aware context&state contrastive learning strategy (Mars-G). Takes turn t as an example, Mars-G enlarges the overall continuous representation margin between dialog context and semantic states, regardless of the pairing relationship between specific dialog context, e.g., $\{C_i, U_i\}$, and semantic states, e.g.B_i/A_i (turn $i = 0, ..., t$). The meaning behind is to make representations between dialog context and semantic states more different, which makes it easy to distinguish dialog context $\{C_i, U_i\}$ and the corresponding semantic states B_i/A_i (turn $i = 0, ..., t$) inside the entire dialog context $\{C_{t+1}, U_{t+1}\}$ and achieve much richer dialog context representations.

Specifically, for every dialog context input, we treat all semantic states in the same batch as negative samples and any one dialog context in the same batch as a positive sample. Besides, considering that every dialog input contains a unique context, narrowing the in-batch context distance makes it hard to distinguish different contexts, which may be counterproductive to deriving the context representation. To resolve such an issue, we also select the rest in-batch dialog context inputs except the positive one as negative samples for every dialog context input. Therefore, the contrastive learning objective function can be reformulated as:

$$\mathcal{L}_{bscl} \triangleq \mathcal{L}_{bscl\_G} = -log \frac{e^{cos(H^i_{cb}, H^j_{cb})/T}}{\sum\limits_{\substack{k=1 \\ k \neq i}}^{N} e^{cos(H^i_{cb}, H^k_{cb})/T} + \sum\limits_{k=1}^{N} e^{cos(H^i_{cb}, H^k_{bb})/T}}, \tag{11}$$

$$\mathcal{L}_{ascl} \triangleq \mathcal{L}_{ascl\_G} = -log \frac{e^{cos(H^i_{ca}, H^j_{ca})/T}}{\sum\limits_{\substack{k=1 \\ k \neq i}}^{N} e^{cos(H^i_{ca}, H^k_{ca})/T} + \sum\limits_{k=1}^{N} e^{cos(H^i_{ca}, H^k_{aa})/T}}, \tag{12}$$

where H^j_{cb} and H^j_{ca} denote the jth ($j \neq i$) dialog context hidden representations after average pooling.

5 Experiments

5.1 Datasets and Evaluation Metrics

We conduct experiments on three task-oriented dialog datasets: MultiWOZ 2.0 (Budzianowski et al., 2018), CamRest676 (Wen et al., 2017), and CrossWOZ (Zhu et al., 2020). MultiWOZ 2.0 (Budzianowski et al., 2018) and CamRest676 (Wen et al., 2017) are English task-oriented dialog datasets. CrossWOZ (Zhu et al., 2020) is a Chinese multi-domain task-oriented dialog dataset. A detailed description of the datasets is provided in Appendix A.

We test our proposed Mars on two benchmark task-oriented dialog tasks: end-to-end dialog modeling response generation and dialog state tracking. We evaluate the performance of response generation on MultiWOZ 2.0 and CamRest676. Inconsistencies exist between previous task-oriented dialog works in data preprocessing and evaluation metrics on MultiWOZ 2.0 (Nekvinda and Dušek, 2021). To fairly compare our experiments with previous work, we use the pre-processing strategy (Zhang et al., 2020b) and the standalone standardized evaluation script released by Nekvinda and Dušek (2021). We follow the automatic evaluation metrics to evaluate the response quality for task-oriented dialog system on MultiWOZ 2.0. **Inform rate** measures whether a dialog system has provided an accurate entity; **Success rate** measures whether a dialog system has provided an accurate entity and answered all requested information; **BLEU score** (Papineni et al., 2002), which is computed with references, which have been obtained from the delexicalized MultiWOZ 2.2 span annotations, measures the fluency of the generated response; **Combined score**, which is calculated by $(Inform + Success) \times 0.5 + BLEU$, measures the overall quality of the dialog system. Moreover, we use the **Act F1** to measure the accuracy of generated action states. To make our experiments comparable with previous work (Zhang et al., 2020a; He et al., 2022)

Model	Pre-trained	Extra corpora	Dialog state tracking Joint Accuracy	Act F1	Response Generation Inform	Success	BLEU	Combined
DAMD (Zhang et al., 2020b)	-	no	-	-	57.9	47.6	16.4	69.2
LABES (Zhang et al., 2020a)	-	no	-	-	68.5	58.1	18.9	82.2
AuGPT (Kulhánek et al., 2021)	GPT-2	yes	-	-	76.6	60.5	16.8	85.4
MinTL (Lin et al., 2020)	T5-small	no	51.2	-	73.7	65.4	19.4	89.0
SOLOIST (Peng et al., 2021)	GPT-2	yes	53.2	-	82.3	72.4	13.6	91.0
DoTS (Jeon and Lee, 2021)	BERT-base	no	-	-	80.4	68.7	16.8	91.4
UBAR (Yang et al., 2021)	DistilGPT2	no	52.6	-	83.4	70.3	17.6	94.5
PPTOD (Su et al., 2021)	T5-base	yes	53.4	-	83.1	72.7	18.2	96.1
BORT (Sun et al., 2022)	T5-small	no	54.0	-	85.5	77.4	17.9	99.4
MTTOD (Lee, 2021)	T5-base	no	53.6	-	85.9	76.5	19.0	100.2
GALAXY (He et al., 2022)	UniLM-base	yes	-	-	85.4	75.7	19.6	100.2
Baseline	T5-small	no	53.8	53.0	83.2	70.3	19.4	96.2
Mars-P	T5-small	no	54.4	**53.9**	86.6	75.5	19.6	100.7
Mars-G	T5-small	no	**55.1**	53.7	**88.9**	**78.0**	**19.9**	**103.4**

Table 1: Comparison of end-to-end models evaluated on MultiWOZ 2.0. The results of previous work are reported on the official leaderboard of MultiWOZ (`https://github.com/budzianowski/multiwoz`).

on CamRest676, we use the same pre-processing strategy and use **Inform rate**, **Success F1**, **BLEU score**, and **Combined score**, which is computed by $(Inform + SuccessF1) \times 0.5 + BLEU$, to evaluate the response quality for the task-oriented dialog system. The success rate whether if the system answered all requested information to assess recall, while Success F1 balances recall and precision.

We evaluate the performance of dialog state tracking on MultiWOZ 2.0 and CrossWOZ. We use the **joint goal accuracy** to measure the accuracy of generated belief states.

5.2 Settings

We use a pre-trained T5 language model (Raffel et al., 2020) to initialize the dialog system based on the HuggingFace Transformers library (Wolf et al., 2020) and follow the settings of Lee (2021). We select T5-small (Raffel et al., 2020) for MultiWOZ 2.0 and CamRest676 and T5-base-Chinese (Raffel et al., 2020; Zhao et al., 2019) for CrossWOZ. The batch size is 8. The AdamW optimizer (Loshchilov and Hutter, 2019) optimizes the model parameters with linear learning rate decay. The initial learning rate is 0.0005, and the ratio of warm up is 0.2. The hyper-parameters λ_1 and λ_2 are set to 1 and 0.1, respectively. T is set to 0.1 for Mars-P, and T is set to 0.5 for Mars-G. The hyper-parameter analysis is provided in Appendix E. We train all dialog systems on one NVIDIA A100 GPU for 10 epochs and select the checkpoint model with the best performance on the validation dataset. One model is trained for approximately five hours. In addition, the model is trained for 20 epochs for the low resource scenarios. The description of baseline systems is provided in Appendix B.

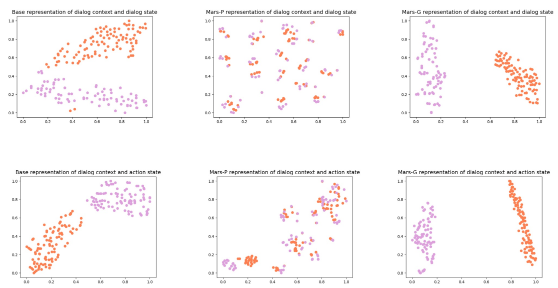


Figure 4: Visualization of dialog context and semantic state representations using t-sne. The three sub-figures on the first row show baseline/Mars-P/Mars-G representations of dialog context and dialog state. The three sub-figures on the second row show baseline/Mars-P/Mars-G representations of dialog context and action state. The coral dots denote dialog context representations. The plum dots denote semantic state representations. We plot 100 dialog examples.

Another baseline is the general architecture of a task-oriented dialog system, as illustrated in Figure 2.

5.3 Main Results

The detailed inform rates, success rates, BLEU scores, combined scores, act F1 scores, and joint goal accuracies for end-to-end task-oriented dialog models on the MultiWOZ 2.0 benchmark are presented in Table 1. Our re-implemented baseline system performs comparable with PPTOD (Su et al., 2021), and our proposed Mars-P and Mars-G outperform our re-implemented baseline system by 4.5 and 7.2 combined scores. Moreover, Mars-G, which doesn't use auxiliary corpora, substantially outperforms the previous state-of-the-art GALAXY (He et al., 2022) and MTTOD (Lee, 2021) by 3.2 combined scores, achieving the state-

Model	Match	Success F1	BLEU	Combined
Sequicity (Lei et al., 2018)	92.7	85.4	25.3	114.4
LABES (Zhang et al., 2020a)	96.4	82.3	25.6	115.0
SOLOIST (Peng et al., 2021)	94.7	87.1	25.5	116.4
GALAXY (He et al., 2022)	**98.5**	87.7	24.2	117.3
Mars-P	97.0	87.2	25.9	118.0
Mars-G	96.2	**89.6**	**26.1**	**119.0**

Table 2: Comparison of end-to-end task-oriented dialog systems on CamRest676.

Model	Joint Accuracy
TRADE (Wu et al., 2019)	36.1
BART-CSP (Moradshahi et al., 2021)	53.6
GEEX (Li et al., 2021)	54.7
Mars-P	59.3
Mars-G	**59.8**

Table 3: Comparison of dialog state tracking performance on CrossWOZ.

of-the-art performance in terms of inform rate, success rate, BLEU score, and combined score. In addition, Mar-G achieves the highest joint goal accuracy among the end-to-end task-oriented dialog systems, outperforming BORT (Sun et al., 2022) by 1.1 points. Compared with the baseline system, Mars-P and Mars-G achieve a better act F1 score. This demonstrates our proposed contrastive learning could effectively improve the quality of the belief state and action state, which further improves the generated response quality. Regarding the two proposed methods, Mars-G performs better than Mars-P. Figure 4 displays the visualization of dialog context and semantic state representations using t-sne. Compared with the baseline system, Mars-P could achieve dialog context representation that is semantically more consistent with its semantic state representation while Mars-G could make representations between dialog context and semantic states more different. These verify dialog context representations, which are more different from semantic state representations, are more beneficial to achieving task completion of task-oriented dialog. Further dialog context representation analysis is provided in Appendix D.

Table 2 presents the performance of task-oriented dialog systems on the CamRest676. Mars-G outperforms the previous state-of-the-art GALAXY (He et al., 2022) by 1.7 combined scores, achieving the state-of-the-art performance in terms of success F1, BLEU score, and combined score.

Table 3 reports the dialog state tracking performance on CrossWOZ. Mars-P and Mars-G substantially outperform the previous state-of-the-art GEEX (Li et al., 2021) by 4.6 and 5.1 points, achieving 59.3 and 59.8 joint goal accuracy. This further indicates that our proposed contrastive learning strategies could improve belief state learning ability, and Mars has good generalization ability. In addition, we provide an example to visualize our proposed Mars-G's dialog state tracking process in Appendix F.

Model	Inform	Success	BLEU	Combined
Mars-G	88.9	78.0	19.9	103.4
w/o BSC	88.3	76.6	19.5	102.0
w/o ASC	86.3	75.1	19.7	100.4
Mars-P	86.6	75.5	19.6	100.7
w/o BSC	85.4	75.0	19.7	99.9
w/o ASC	83.7	73.0	19.8	98.2
Baseline	83.2	70.3	19.4	96.2

Table 4: The performance of the different components of our proposed methods on MultiWOZ 2.0. BSC represents the belief state module of contrastive learning, and ASC represents the action state module.

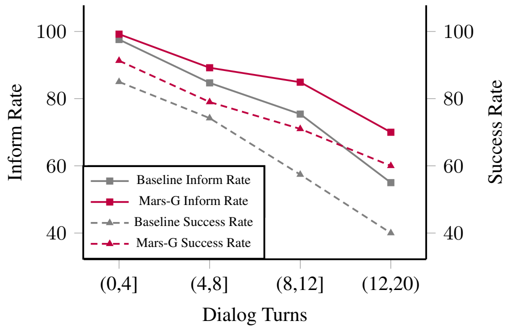


Figure 5: Performance of dialog systems on the test set with respect to different dialog turns.

5.4 Ablation Study

Table 4 shows the performance of the different components of Mars-P and Mars-G. Both state modules of Mars-P and Mars-G could improve the performance of the dialog system. Regarding two modules of contrastive learning strategies Mars-P and Mars-G, the action state module performs better than the belief state module by 1.7 and 1.6 combined scores, respectively, because the quality of the action state has a more direct impact on the response generation quality and action state module could improve action state learning ability. Moreover, the combination of both modules can complement each other to further improve end-to-end dialog modeling performance. The further ablation analysis is provided in Appendix G.

Model	5%				10%				20%				50%			
	Inform	Success	BLEU	Combined	Inform	Success	BLEU	Combined	Inform	Success	BLEU	Combined	Inform	Success	BLEU	Combined
DAMD	36.8	17.3	11.2	38.3	40.9	23.0	12.2	44.2	48.3	30.3	14.2	53.5	58.8	44.3	15.7	67.3
MinTL	52.5	38.1	13.9	59.2	55.5	44.9	15.6	65.8	64.3	54.9	16.2	75.8	70.3	62.2	18.0	84.3
UBAR	37.4	22.1	11.3	41.1	50.3	34.2	13.5	55.8	65.5	48.7	14.5	71.6	77.6	63.3	16.3	86.8
MTTOD	54.3	37.4	11.3	57.2	66.9	55.2	13.8	74.9	75.0	**63.3**	14.3	83.5	78.5	67.5	15.2	88.2
PPTOD	**65.5**	**48.3**	**14.3**	**71.2**	68.3	53.7	**15.7**	76.7	72.7	59.2	16.3	82.3	74.8	62.4	17.0	85.6
Mars-G	57.6	43.4	13.9	64.4	**69.4**	**55.3**	15.6	**78.0**	**76.7**	62.9	**17.2**	**87.0**	**82.2**	**71.2**	**18.6**	**95.3**

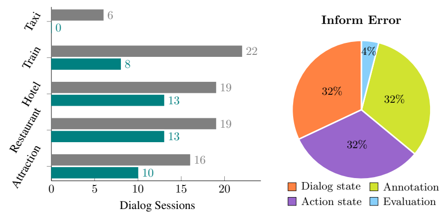
Table 5: Comparison of task-oriented dialog systems in the low resource scenarios on MultiWOZ 2.0. 5% (400 dialogs), 10% (800 dialogs), 20% (1600 dialogs), 50% (4000 dialogs) of training data is used to train each model.

5.5 Dialog Turn Analysis

To better assess the effectiveness of our proposed contrastive learning strategies, we investigate the performance (inform rate and success rate) of Mars-G and the baseline system on the test set with respect to different dialog turns. Specifically, we divide each test set into four groups according to the dialog turn. As shown in Figure 5, Mars-G is superior to the baseline system in every dialog turn group. This indicates our proposed contrastive learning strategies are beneficial to task-oriented dialog modeling. Especially, as dialog turn increases, the performance of the baseline system decreases rapidly, and the performance gap between the baseline system and our proposed Mars-G is increased. Because the baseline system is hard to model long-range semantic dependencies to generate inaccurate semantic states and system responses. In contrast, Mars-G enhances the relationship modeling between dialog context and semantic state representations and achieves better dialog context representations to capture long-range semantic dependencies in the long dialog turns.

5.6 Low Resource Scenario Analysis

To investigate the performance of task-oriented dialog systems in the low resource scenario, we choose 5%, 10%, 20%, and 50% of training dialog sessions to do stimulated experiments on the MultiWOZ 2.0. Considering the inconsistency of data distribution with different random seeds in the stimulated low resource scenario, we re-implement all baseline systems with the same random seed to ensure the consistency of data distribution. In addition, we train all dialog systems five times with different random seeds and report the average scores in Table 5. The detailed results of five runs are provided in Appendix H. As shown in Table 5, PPTOD achieves the best performance in the extreme low-resource scenario (5% training data) because auxiliary corpora used in PPTOD have many similar dialog sessions with MultiWOZ 2.0 and this benefits PPTOD in the stimulated low-resource scenario. In contrast, Mars-G doesn't use auxiliary corpora to improve performance in the low-resource scenario. Apart from this, Mars-G substantially outperforms all baseline systems in other low-resource scenarios. Moreover, Mars-G trained on the 50% training data performs better than some baseline systems such as MinTL and UBAR trained on all training data, as shown in Table 1. These further demonstrate that Mars-G is robust, achieving comparable performance in the low-resource scenario.



(a) (b)

Figure 6: The domain distribution (a) and primary reason distribution (b) of inaccurate dialog sessions according to the inform rate metric. The gray bars denote the total number of dialog sessions that contain the corresponding domain; the teal bars denote the number of dialog sessions with errors in the corresponding domain.

5.7 Error Analysis

To better apply our proposed Mars-G to real-world scenarios, we perform error analysis based on inform rate (informable slot) and success rate (requestable slot). In detail, we randomly extract 40 inaccurate dialog sessions from the MultiWOZ 2.0 testing set, respectively. The detailed domain distribution and primary reason distribution of informable slot errors are presented as shown in Figure 6. Given that there is no database in the taxi domain, the informable slots are consistently judged to be correct. The error rate of the dialogs in the hotel and restaurant domains is very high because some informable slots in these two domains are often mispredicted, such as '*type*' in the hotel domain. As illustrated in Figure 6(b), 64 percent of dialog informable slot errors are caused by the inaccurate belief states and action states, and the noisy dialog annotations generate 32 percent. 4 percent of that are caused by automatic evaluation scripts and are judged accurately by human evaluation. The detailed requestable slot error analysis and more examples are provided in Appendixes I and J, respectively. In the future, we will focus on solving errors caused by the inaccurate dialog/ action states to better apply Mars-G to real-world scenarios.

6 Conclusion

This study explores what dialog context representation is beneficial to improving task-oriented dialog performance. Specifically, we propose two contrastive learning strategies to explicitly model the relationship between dialog context and semantic state representations, achieving better task completion of a task-oriented dialog system. Extensive experiments and analysis demonstrate that dialog context representations that are more different from semantic state representations are more beneficial to multi-turn task-oriented dialog. Moreover, our proposed Mars achieves state-of-the-art performance on three datasets.

Limitations

The training process of Mars needs to rely on manually annotated belief states and action states as semantic states to explicitly model the relationship between dialog context and semantic state representations through contrastive learning methods. We propose Mars in the research community and hope it can be better applied to real-world scenarios in the industry. However, the annotated data is expensive, which makes our methods have some limitations in the landing process of real scenarios. In the future, to better apply our proposed Mars to real-world scenarios, we will introduce semi-supervised methods to reduce the dependence on annotated dialog corpus.

Acknowledgments

This work was supported by the National Key R&D Program of China under Grant No. 2020AAA0108600.

References

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. Association for Computational Linguistics.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey E. Hinton. 2020. A simple framework for contrastive learning of visual representations. In *Proceedings of the 37th International Conference on Machine Learning*, pages 1597–1607. PMLR.

Xinlei Chen and Kaiming He. 2021. Exploring simple siamese representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15750–15758. Computer Vision Foundation / IEEE.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186. Association for Computational Linguistics.

Li Dong, Nan Yang, Wenhui Wang, Furu Wei, Xiaodong Liu, Yu Wang, Jianfeng Gao, Ming Zhou, and Hsiao-Wuen Hon. 2019. Unified language model pre-training for natural language understanding and generation. In *Advances in Neural Information Processing Systems 32*, pages 13042–13054. Curran Associates, Inc.

Hongchao Fang and Pengtao Xie. 2020. CERT: contrastive self-supervised learning for language understanding. *CoRR*, abs/2005.12766.

Tianyu Gao, Xingcheng Yao, and Danqi Chen. 2021. SimCSE: Simple contrastive learning of sentence embeddings. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 6894–6910. Association for Computational Linguistics.

John Giorgi, Osvald Nitski, Bo Wang, and Gary Bader. 2021. DeCLUTR: Deep contrastive learning for unsupervised textual representations. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 879–895. Association for Computational Linguistics.

Jean-Bastien Grill, Florian Strub, Florent Altché, Corentin Tallec, Pierre H. Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Ávila Pires, Zhaohan Guo, Mohammad Gheshlaghi Azar, Bilal

Piot, Koray Kavukcuoglu, Rémi Munos, and Michal Valko. 2020. Bootstrap your own latent - A new approach to self-supervised learning. In *Advances in Neural Information Processing Systems 33*. Curran Associates, Inc.

Jiatao Gu, Zhengdong Lu, Hang Li, and Victor O.K. Li. 2016. Incorporating copying mechanism in sequence-to-sequence learning. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1631–1640. Association for Computational Linguistics.

Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross B. Girshick. 2020. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9726–9735. Computer Vision Foundation / IEEE.

Wanwei He, Yinpei Dai, Yinhe Zheng, Yuchuan Wu, Zheng Cao, Dermot Liu, Peng Jiang, Min Yang, Fei Huang, Luo Si, et al. 2022. Galaxy: A generative pre-trained model for task-oriented dialog with semi-supervised learning and explicit policy injection. *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence*.

Ehsan Hosseini-Asl, Bryan McCann, Chien-Sheng Wu, Semih Yavuz, and Richard Socher. 2020. A simple language model for task-oriented dialogue. In *Advances in Neural Information Processing Systems 33*, pages 20179–20191. Curran Associates, Inc.

Hyunmin Jeon and Gary Geunbae Lee. 2021. Domain state tracking for a simplified dialogue system. *CoRR*, abs/2103.06648.

Jonás Kulhánek, Vojtech Hudecek, Tomás Nekvinda, and Ondrej Dusek. 2021. Augpt: Dialogue with pre-trained language models and data augmentation. *CoRR*, abs/2102.05126.

Hwaran Lee, Jinsik Lee, and Tae-Yoon Kim. 2019. SUMBT: Slot-utterance matching for universal and scalable belief tracking. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5478–5483. Association for Computational Linguistics.

Yohan Lee. 2021. Improving end-to-end task-oriented dialog system with a simple auxiliary task. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 1296–1303. Association for Computational Linguistics.

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. Association for Computational Linguistics.

Xinmeng Li, Qian Li, Wansen Wu, and Quanjun Yin. 2021. Generation and extraction combined dialogue state tracking with hierarchical ontology integration. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 2241–2249. Association for Computational Linguistics.

Zhaojiang Lin, Andrea Madotto, Genta Indra Winata, and Pascale Fung. 2020. MinTL: Minimalist transfer learning for task-oriented dialogue systems. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, pages 3391–3405. Association for Computational Linguistics.

Ilya Loshchilov and Frank Hutter. 2019. Decoupled weight decay regularization. In *Proceedings of the 7th International Conference on Learning Representations*. OpenReview.net.

Mehrad Moradshahi, Victoria Tsai, Giovanni Campagna, and Monica S. Lam. 2021. Contextual semantic parsing for multilingual task-oriented dialogues. *CoRR*, abs/2111.02574.

Tomáš Nekvinda and Ondřej Dušek. 2021. Shades of BLEU, flavours of success: The case of MultiWOZ. In *Proceedings of the 1st Workshop on Natural Language Generation, Evaluation, and Metrics*, pages 34–46. Association for Computational Linguistics.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318. Association for Computational Linguistics.

Baolin Peng, Chunyuan Li, Jinchao Li, Shahin Shayandeh, Lars Liden, and Jianfeng Gao. 2021. Soloist: Building task bots at scale with transfer learning and machine teaching. *Transactions of the Association for Computational Linguistics*, 9:807–824.

Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140):1–67.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *CoRR*, abs/1707.06347.

Yixuan Su, Lei Shu, Elman Mansimov, Arshit Gupta, Deng Cai, Yi-An Lai, and Yi Zhang. 2021. Multi-task pre-training for plug-and-play task-oriented dialogue system. *CoRR*, abs/2109.14739.

Haipeng Sun, Junwei Bao, Youzheng Wu, and Xiaodong He. 2022. BORT: Back and denoising reconstruction for end-to-end task-oriented dialog. In *Findings of the Association for Computational Linguistics: NAACL2022*. Association for Computational Linguistics.

Ryuichi Takanobu, Hanlin Zhu, and Minlie Huang. 2019. Guided dialog policy learning: Reward estimation for multi-domain task-oriented dialog. 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 100–110. Association for Computational Linguistics.

Tsung-Hsien Wen, Milica Gašić, Nikola Mrkšić, Pei-Hao Su, David Vandyke, and Steve Young. 2015. Semantically conditioned LSTM-based natural language generation for spoken dialogue systems. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1711–1721. Association for Computational Linguistics.

Tsung-Hsien Wen, David Vandyke, Nikola Mrkšić, Milica Gašić, Lina M. Rojas-Barahona, Pei-Hao Su, Stefan Ultes, and Steve Young. 2017. A network-based end-to-end trainable task-oriented dialogue system. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 438–449. Association for Computational Linguistics.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander Rush. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45. 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. Association for Computational Linguistics.

Zhuofeng Wu, Sinong Wang, Jiatao Gu, Madian Khabsa, Fei Sun, and Hao Ma. 2020. CLEAR: contrastive learning for sentence representation. *CoRR*, abs/2012.15466.

Yuanmeng Yan, Rumei Li, Sirui Wang, Fuzheng Zhang, Wei Wu, and Weiran Xu. 2021. ConSERT: A contrastive framework for self-supervised sentence representation transfer. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 5065–5075. Association for Computational Linguistics.

Yunyi Yang, Yunhao Li, and Xiaojun Quan. 2021. UBAR: towards fully end-to-end task-oriented dialog system with GPT-2. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence*, pages 14230–14238. AAAI Press.

Yichi Zhang, Zhijian Ou, Min Hu, and Junlan Feng. 2020a. A probabilistic end-to-end task-oriented dialog model with latent belief states towards semi-supervised learning. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, pages 9207–9219. Association for Computational Linguistics.

Yichi Zhang, Zhijian Ou, and Zhou Yu. 2020b. Task-oriented dialog systems that consider multiple appropriate responses under the same context. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence*, pages 9604–9611. AAAI Press.

Zheng Zhang, Ryuichi Takanobu, Minlie Huang, and Xiaoyan Zhu. 2020c. Recent advances and challenges in task-oriented dialog system. *CoRR*, abs/2003.07490.

Zhe Zhao, Hui Chen, Jinbin Zhang, Xin Zhao, Tao Liu, Wei Lu, Xi Chen, Haotang Deng, Qi Ju, and Xiaoyong Du. 2019. UER: An open-source toolkit for pre-training models. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing: System Demonstrations*, pages 241–246. Association for Computational Linguistics.

Qi Zhu, Kaili Huang, Zheng Zhang, Xiaoyan Zhu, and Minlie Huang. 2020. CrossWOZ: A large-scale Chinese cross-domain task-oriented dialogue dataset. *Transactions of the Association for Computational Linguistics*, 8:281–295.

A Datasets

MultiWOZ 2.0 (Budzianowski et al., 2018) is a large-scale English multi-domain task-oriented dialog dataset containing 8438, 1000, and 1000 dialog sessions for training, validation, and testing datasets. It consists of seven domains: attraction, hotel, restaurant, taxi, train, hospital, and police. CamRest676 (Wen et al., 2017) is a small-scale English restaurant-domain dataset, which is split 3/ 1/ 1 for training, validation, and testing datasets. CrossWOZ (Zhu et al., 2020) is a large-scale Chinese multi-domain task-oriented dialog dataset containing 5012, 500, and 500 dialog sessions

Model	Inform	Success	BLEU
End-to-end model	83.2	70.3	19.4
w/ oracle state	90.8	87.4	30.6
Reference Corpus	93.7	90.9	100.0

Table 6: Comparison of task-oriented dialog models evaluated on MultiWOZ 2.0. w/ oracle state denotes the system using ground truth belief state and action state for the response generation. Reference results are reported on the official leaderboard of MultiWOZ.

for training, validation, and testing datasets. It comprises five domains: attraction, restaurant, hotel, taxi, and metro.

B Baselines

Sequicity (Lei et al., 2018), DAMD (Zhang et al., 2020b), and LABES (Zhang et al., 2020a) are copy-augmented GRU-based end-to-end task-oriented dialog systems. Bidirectional auto-encoding language model BERT (Devlin et al., 2019) is used for the context encoder in DoTS (Jeon and Lee, 2021). Unidirectional auto-regressive language model GPT-2 (Radford et al., 2019) is used in AuGPT (Kulhánek et al., 2021), SOLOIST (Peng et al., 2021), and UBAR (Yang et al., 2021). Seq2seq language model T5 (Raffel et al., 2020) is used in MinTL (Lin et al., 2020), PPTOD (Su et al., 2021), and MTTOD (Lee, 2021). The unified language model UniLM (Dong et al., 2019) is used in GALAXY (He et al., 2022). In addition, auxiliary task-oriented dialog corpora are used to pre-train in AuGPT (Kulhánek et al., 2021), SOLOIST (Peng et al., 2021), PPTOD (Su et al., 2021), and GALAXY (He et al., 2022). TRADE (Wu et al., 2019), BART-CSP (Moradshahi et al., 2021), and GEEX (Li et al., 2021) are some additional dialog state tracking models.

C States Analysis

To investigate the impact of belief state and action state on the performance of end-to-end task-oriented dialog, we empirically conduct preliminary experiments on MultiWOZ 2.0 (Budzianowski et al., 2018). As shown in Table 6, the system using ground truth belief state and action state substantially outperforms the traditional end-to-end task-oriented dialog systems, achieving performance comparable to reference in terms of task completion. This demonstrates that the quality of belief state and action state greatly

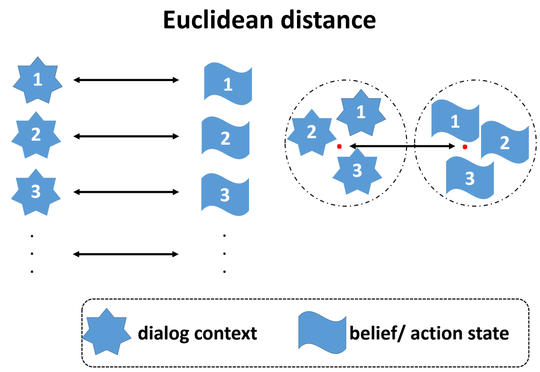


Figure 7: The calculation of Euclidean distance between dialog context and semantic state representations.

influence on the end-to-end task-oriented dialog performance.

D Dialog Context Representation Analysis

To further analyze dialog context and semantic state representations of Mars-P and Mars-G, we would measure the similarity of continuous encoder representation between dialog context and corresponding belief/action state on the MultiWOZ 2.0 test set, as illustrated in Figure 7. Table 7 shows the average L2-normalized Euclidean distance between dialog context and corresponding belief/action state representations. Table 8 shows the Euclidean distance between the centroids of these two L2-normalized representation spaces. The centroid is the average of all the points in the representation space. T5 denotes the result before training on the MultiWOZ 2.0. We find the distance between dialog context and corresponding semantic state representations changes a little before and after training. Mars-P achieves a smaller distance, thus obtaining a continuous representation of the dialogue context that is semantically more consistent with its semantic state representation. The distance of Mars-G is enormous, demonstrating Mars-G achieves more diverse dialog context representations, different from semantic state representations.

E Hyper-parameter Analysis

We empirically investigate how the hyper-parameters λ and T for both modules of Mars-G affect the performance of task-oriented dialog on the MultiWOZ 2.0, respectively. The selection of λ influences the role of the contrastive learning objective function across the entire task-oriented

Model	Context&Belief State	Context&Action State
T5	0.797	1.018
Baseline	0.844	1.017
Mars-P	0.340	0.542
Mars-G	1.996	1.993

Table 7: The distance between dialog context and corresponding semantic state representations on MultiWOZ 2.0.

Model	Context&Belief State	Context&Action State
T5	0.555	0.807
Baseline	0.598	0.699
Mars-P	0.042	0.046
Mars-G	1.993	1.987

Table 8: The distance between the centroids of these two representation spaces on MultiWOZ 2.0.

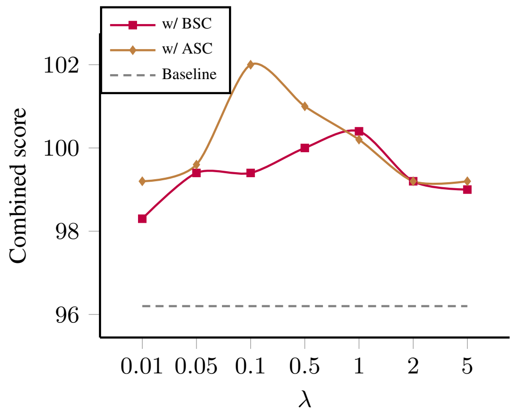


Figure 8: The Mars-G performance with different levels of hyper-parameter λ on the MultiWOZ 2.0. w/ BSC denotes belief state module, w/ ASC denotes action state module. T is set to 0.5.

dialog training process. As Figure 8 shows, λ ranging from 0.01 to 5 nearly all improve task-oriented dialog performance. This indicates our proposed Mars-G is robust and effective. When $\lambda = 0.1$, w/ ASC achieves the best performance. When $\lambda = 1$, w/ BSC achieves the best performance. The selection of T affects the differentiation of hard negative samples. The smaller the value of T is, the more attention is paid to distinguishing complex negative samples. As shown in Figure 9, combined scores increase for almost all T values ranging from 0.01 to 10, and the best performance is achieved when $T = 0.5$ for both modules of Mars-G.

F Visualization

We provide an example to visualize the dialog state tracking process of our proposed Mars-G and baseline system. The cross-attention weights

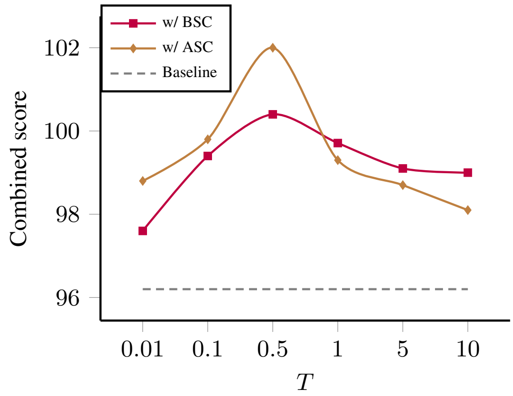


Figure 9: The Mars-G performance with different levels of hyper-parameter T on the MultiWOZ 2.0. λ_1 is set to 1 for w/ BSC, and λ_2 is set to 0.1 for w/ ASC.

Model	Inform	Success	BLEU	Combined
Baseline	83.2	70.3	19.4	96.2
Mars-variant	85.7	74.8	19.6	99.9
Mars-P	86.6	75.5	19.6	100.7

Table 9: The performance of the different methods on MultiWOZ 2.0. Mars-variant denotes similarity strategy.

between dialog context and generated belief states from the last layer of the transformer decoder stack are shown in Figures 10 and 11. Compared with the baseline system, Mars-G could achieve more accurate attention weights. The slot '*arrive 09:00*' assigns high attention weights for the user utterance '*09:00*' and previous belief state '*arrive 09:00*'. Similarly, the slots '*destination mumford theatre*' and '*departure wagamama*' accurately give high attention weights for the corresponding user utterance. The visualization further demonstrates that Mars-G could achieve more reasonable dialog context representation to generate accurate belief states.

G Further Ablation Analysis

To get a more complete picture of the effectiveness of Mars-P, we introduce a similarity strategy (Mars-variant). We use the cosine similarity function to narrow the distance between the continuous representation of dialog contexts and semantic states for the same dialog session to model the relationship between dialog context and corresponding semantic state representations. We don't distinguish the continuous representation of dialog context and states for different dialog sessions. As shown in Table 9, Mars-variant outperforms the baseline system by 3.7

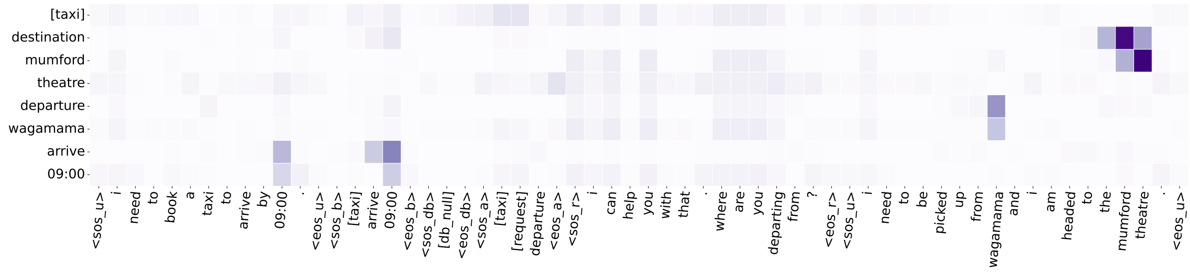


Figure 10: Visualization of the cross-attention weights between dialog context and generated belief states for our proposed Mars-G. The horizontal axis is the dialog context, and the vertical axis is the generated belief state.

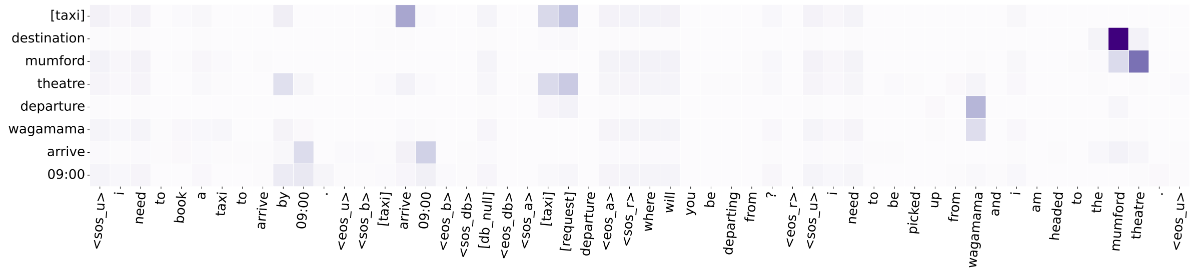
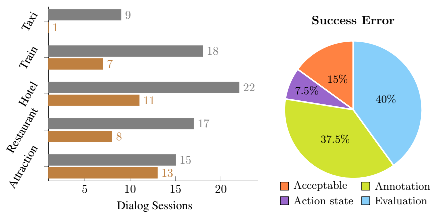


Figure 11: Visualization of the cross-attention weights between dialog context and generated belief states for the baseline system.



(a) (b)

Figure 12: The domain distribution (a) and primary reason distribution (b) of inaccurate dialog sessions according to success rate metric. The gray bars denote the total number of dialog sessions that contain the corresponding domain; the brown bars denote the number of dialog sessions with errors in the corresponding domain.

combined scores, indicating the effectiveness of the relationship modeling between dialog context and corresponding semantic representations. In addition, Mars-variant underperforms Mars-P by 0.8 combined scores. This demonstrates that distinguishing the continuous representation of dialog context and states for different dialog sessions is beneficial for dialog modeling.

H Low Resource Scenario Results

We train all dialog systems five times with different random seeds in the low resource scenario. The detailed results of 5 runs are provided in Table 10.

I Requestable Slot Error Analysis

Considering the inclusion relationship of the two metrics described in Section 5.1, we select dialog sessions with the wrong success rate and accurate inform rate for success rate error analysis. The detailed domain distribution and primary reason distribution of requestable slot errors are presented as shown in Figure 12. The error rate of the dialogs in the taxi and train domains is very low because requestable slots in these two domains are few and simple. For example, the requestable slot in the taxi domain only has '*phone*'. The error rate of the dialogs in the attraction domain is very high. As illustrated in Figure 12(b), 77.5 percent of dialog requestable slot errors are caused by the noisy dialog annotations and automatic evaluation scripts. 15 percent of generated system responses are acceptable. When users request some information about something and do not ask for a specific requestable slot, Mars-G generates system responses that lack some requestable slots such as '*postcode*' and '*address*'. In addition, Mars-G requests users some other useful information instead of providing booked reference directly. We think system responses generated by Mars-G in both cases are reasonable. Inaccurate action states cause 7.5 percent of dialog requestable slot errors.

Model	5%				10%				20%				50%			
	Inform	Success	BLEU	Combined	Inform	Success	BLEU	Combined	Inform	Success	BLEU	Combined	Inform	Success	BLEU	Combined
DAMD																
run 1	35.4	17.2	10.9	37.2	41.3	23.7	12.4	44.9	51.8	31.9	14.1	56.0	60.1	44.2	15.6	67.8
run 2	40.8	20.9	12.0	42.9	41.5	25.3	11.2	44.6	50.4	32.4	13.8	55.2	54.7	39.5	14.8	61.9
run 3	38.5	14.5	10.6	37.1	40.0	23.9	12.3	44.3	42.5	26.8	15.4	50.1	59.1	45.7	15.1	67.5
run 4	35.4	16.5	10.2	36.2	42.3	20.2	12.0	43.3	46.1	29.2	14.0	51.7	57.2	43.0	16.1	66.2
run 5	34.0	17.6	12.4	38.2	39.5	21.9	13.0	43.7	50.8	31.3	13.6	54.7	63.1	49.3	17.1	73.3
Average	36.8	17.3	11.2	38.3	40.9	23.0	12.2	44.2	48.3	30.3	14.2	53.5	58.8	44.3	15.7	67.3
MinTL																
run 1	54.4	41.1	14.2	62.0	55.8	44.0	15.3	65.2	62.5	54.2	17.3	75.7	71.7	62.7	16.9	84.1
run 2	54.8	36.8	13.6	59.4	51.6	42.0	15.7	62.5	65.8	56.3	15.5	76.6	67.4	59.7	18.7	82.3
run 3	53.3	39.3	14.2	60.5	55.1	44.7	16.1	66.0	68.0	59.0	16.6	80.1	70.6	62.6	17.5	84.1
run 4	52.4	37.1	13.8	58.6	58.4	47.3	15.2	68.1	58.3	48.4	14.4	67.8	68.9	61.3	18.2	83.3
run 5	47.5	36.3	13.8	55.7	56.8	46.4	15.9	67.5	66.9	56.8	17.0	78.9	73.1	64.5	18.6	87.4
Average	52.5	38.1	13.9	59.2	55.5	44.9	15.6	65.8	64.3	54.9	16.2	75.8	70.3	62.2	18.0	84.3
UBAR																
run 1	37.4	23.0	11.6	41.8	52.3	34.8	13.0	56.6	61.7	45.7	15.9	69.6	77.2	61.5	15.5	84.9
run 2	33.3	20.6	11.2	38.2	48.5	35.9	14.5	56.7	63.4	47.8	15.5	71.1	78.0	63.8	16.9	87.8
run 3	40.0	23.1	11.7	43.3	50.3	33.2	13.6	55.4	67.8	50.0	13.1	72.0	77.4	64.6	16.2	87.2
run 4	38.2	22.4	10.7	41.0	52.5	34.6	12.5	56.1	68.3	51.7	14.4	74.4	78.5	64.1	16.8	88.1
run 5	38.0	21.3	11.3	41.0	47.8	32.3	13.7	53.8	66.2	48.3	13.8	71.1	76.8	62.4	16.2	85.8
Average	37.4	22.1	11.3	41.1	50.3	34.2	13.5	55.8	65.5	48.7	14.5	71.6	77.6	63.3	16.3	86.8
MTTOD																
run 1	51.4	37.5	12.0	56.5	70.9	58.0	13.8	78.3	71.1	59.0	14.2	79.3	74.7	64.4	15.2	84.8
run 2	53.8	41.7	11.3	59.1	64.1	53.7	13.8	72.7	69.5	60.7	14.0	79.1	79.3	67.7	15.0	88.5
run 3	55.7	31.1	11.5	54.9	61.0	50.8	13.7	69.6	78.4	65.1	14.7	86.5	82.3	71.1	15.5	92.2
run 4	52.4	33.3	10.6	53.5	73.0	59.3	14.0	80.2	80.2	67.4	14.5	88.3	76.6	65.6	15.3	86.4
run 5	58.0	43.2	11.3	61.9	65.4	54.2	13.7	73.5	75.9	64.3	14.1	84.2	79.8	68.7	15.1	89.4
Average	54.3	37.4	11.3	57.2	66.9	55.2	13.8	74.9	75.0	63.3	14.3	83.5	78.5	67.5	15.2	88.2
PPTOD																
run 1	70.7	46.8	13.7	72.5	65.2	50.6	14.2	72.1	72.3	55.0	14.9	78.6	74.8	60.4	15.8	83.4
run 2	64.6	45.8	13.8	69.0	69.3	52.9	15.3	76.4	70.5	57.7	17.7	81.8	74.1	64.2	16.4	85.6
run 3	64.4	51.1	15.1	72.9	65.7	53.6	15.8	75.5	74.8	64.6	16.9	86.6	74.3	61.8	17.2	85.3
run 4	63.9	47.0	14.7	70.2	70.1	55.4	17.8	80.6	71.8	57.3	16.0	80.6	76.4	63.7	18.0	88.1
run 5	63.7	50.7	14.4	71.6	71.2	55.8	15.6	79.1	74.1	61.6	15.8	83.7	74.4	61.9	17.5	85.7
Average	65.5	48.3	14.3	71.2	68.3	53.7	15.7	76.7	72.7	59.2	16.3	82.3	74.8	62.4	17.0	85.6
Mars-G																
run 1	55.8	41.1	14.0	62.5	68.7	55.0	16.7	78.6	72.4	60.2	18.1	84.4	82.6	70.2	18.8	95.2
run 2	57.0	43.2	12.9	63.0	68.4	55.9	15.2	77.4	76.0	61.4	17.1	85.8	78.4	66.9	18.7	91.4
run 3	61.4	46.7	14.5	68.6	68.9	53.8	14.0	75.4	76.6	63.8	17.0	87.2	82.8	73.6	17.9	96.1
run 4	56.1	42.4	14.1	63.4	73.1	60.3	16.6	83.3	80.6	63.9	17.1	89.4	82.5	71.3	19.0	95.9
run 5	57.8	43.5	13.8	64.5	67.7	51.5	15.7	75.3	77.7	65.0	16.8	88.2	84.6	74.2	18.7	98.1
Average	57.6	43.4	13.9	64.4	69.4	55.3	15.6	78.0	76.7	62.9	17.2	87.0	82.2	71.2	18.6	95.3

Table 10: Comparison of task-oriented dialog systems on the MultiWOZ 2.0 in the low resource scenarios.

J Examples for Error Analysis

Tables 11 - 19 show several examples generated by Mars-G for detailed error analysis. As shown in Table 11, Mars-G generates the inaccurate belief state ‘*food jamaican*’ rather than ‘*food italian*’, leading to the informable slot error. Table 12 shows that Mars-G generates the inadequate action state, not including the slot name ‘*name*’, leading to the informable slot error. Table 13 shows that the informable slot error is caused by automatic evaluation. Mars-G provides the accurate response in turn 7. However, the automatic evaluation script estimates the wrong active domain ‘*[taxi]*’ rather than ‘*[attraction]*’ from the belief state. The informable slot error in Table 14 is caused by noisy dialog annotations. The informable slot ‘pricerange moderate’ does not appear in the conversation.

As shown in Table 15, Mars-G generates the inaccurate action state ‘*[request] people*’ provided in the belief state ‘*people 1*’, leading to the requestable slot error. Table 16 shows that the requestable slot error is caused by automatic evaluation. Mars-G provides the accurate response in turn 2, while the automatic evaluation script only determines if it offers a reference in turn 3. We think system responses generated by Mars-G in Tables 17 and 18 are acceptable. When users request some information about something and do not ask for a specific requestable slot, Mars-G generates system responses, lack of the requestable slot ‘*postcode*’, as shown in Table 17. In addition, Mars-G requests users whether to make a reservation instead of providing booked references directly, as shown in Table 18. Noisy dialog annotations cause the requestable slot error in Table 19. The requestable slot ‘postcode’ does not appear in the conversation.

Goal:	[restaurant]: {informable: food italian area east} {requestable: reference}
User (Turn 0):	i am looking for a jamaican restaurant in the east.
Golden R:	there are no [value_food] restaurants in the [value_area]. would you like to try another food type or area?
Golden BS:	[restaurant] food jamaican area east
Golden AS:	[restaurant] [nooffer] food area [request] area food
Mars-G R:	i am sorry, there are no [value_food] restaurants in the [value_area]. would you like to try another area or another type of food?
Mars-G BS:	[restaurant] food jamaican area east
Mars-G AS:	[restaurant] [nooffer] food area [request] area food
User (Turn 1):	i am looking for a place that serves jamaican food in the east. if not, italian will do.
Golden R:	there is [value_choice] [value_food] place in the [value_area], [value_name].
Golden BS:	[restaurant] food italian area east
Golden AS:	[restaurant] [inform] choice name food area
Mars-G R:	i am sorry, there are no [value_food] restaurants in the [value_area]. would you like to try another type of food?
Mars-G BS:	[restaurant] food jamaican area east
Mars-G AS:	[restaurant] [nooffer] food area [request] food
User (Turn 2):	ok, i would like to make a reservation there.
Golden R:	i can do that for you, i just need details such as day, time, and the size of your party.
Golden BS:	[restaurant] food italian area east
Golden AS:	[restaurant] [offerbook] [request] time day people
Mars-G R:	what day and time would you like to dine? and how many people will be in your party?
Mars-G BS:	[restaurant] food jamaican area east
Mars-G AS:	[restaurant] [request] day time people
User (Turn 3):	i need a table for 7 at 14:45 on friday.
Golden R:	i booked you a table for [value_people] on [value_day] at [value_time] at [value_name], reference: [value_reference]. can i assist you further today?
Golden BS:	[restaurant] food italian area east time 14:45 day friday people 7
Golden AS:	[restaurant] [offerbooked] time reference name day people [general] [reqmore]
Mars-G R:	booking was successful. the table will be reserved for 15 minutes. reference number is: [value_reference].
Mars-G BS:	[restaurant] food jamaican area east time 14:45 day friday people 7
Mars-G AS:	[restaurant] [offerbooked] reference

Table 11: An example in dialog session pmul0367. The informable slot error is caused by the inaccurate belief state. R denotes response, BS denotes belief state, and AS denotes action state.

Goal:	[attraction]: {informable: area centre}
User (Turn 3):	i am also looking for places to go in town. perhaps an attraction in the city centre.
Golden R:	i have the [value_name] located at [value_address] and it's [value_price] entrance.
Golden BS:	[attraction] area centre
Golden AS:	[attraction] [recommend] address price name
Mars-G R:	there are [value_choice] attractions in the city centre. do you have a preference for type?
Mars-G BS:	[attraction] area centre
Mars-G AS:	[attraction] [inform] choice [request] type
User (Turn 4):	that sounds perfect. thanks!
Golden R:	can i help you with anything else?
Golden BS:	[attraction] area centre
Golden AS:	[general] [reqmore]
Mars-G R:	is there anything else i can help you with?
Mars-G BS:	[attraction] area centre
Mars-G AS:	[general] [reqmore]

Table 12: An example in dialog session pmul2563. The informable slot error is caused by the inaccurate action state.

Goal:	[attraction]: {informable: area south} {requestable: postcode}
User (Turn 4):	i would also like to go to some attractions in the same area as the restaurant.
Golden R:	there are [value_choice] attractions in the area. there's a [value_type], [value_type], [value_type], [value_type], [value_type]... do any of these appeal to you in particular?
Golden BS:	[attraction] area south [restaurant] food indian pricerange expensive area south
Golden AS:	[attraction] [select] type [inform] choice
Mars-G R:	there are [value_choice] attractions in the area. do you have a preference for type?
Mars-G BS:	[attraction] area south [restaurant] food indian pricerange expensive area south
Mars-G AS:	[attraction] [inform] choice [request] type
User (Turn 5):	restaurant is what i am looking for, entrance fee and postcode please. get me a taxi from the attraction by 10.15, get me its contact number and car type.
Golden R:	just to clarify you are looking for another restaurant and not an attraction?
Golden BS:	[attraction] area south [restaurant] food indian pricerange expensive area south
Golden AS:	[attraction] [select] [restaurant] [select]
Mars-G R:	[value_name] is located in postcode [value_postcode]. the phone number is [value_phone]. the entrance fee is [value_price]. can i help you with anything else?
Mars-G BS:	[taxi] leave 10.15 destination tandoori palace departure the nirala [attraction] area south [restaurant] food indian pricerange expensive area south
Mars-G AS:	[attraction] [inform] price phone name postcode [general] [reqmore]
User (Turn 6):	oh, yeah, i do want information on attractions! in the same area. after that, i will need a taxi.
Golden R:	if you like parks the [value_name] has [value_price] admission and is in the [value_postcode] postcode. or i can tell you about the other attractions?
Golden BS:	[attraction] area south [restaurant] food indian pricerange expensive area south
Golden AS:	[attraction] [recommend] name postcode price [general] [reqmore]
Mars-G R:	there are [value_choice] attractions in the [value_area]. [value_choice] are [value_type], [value_choice] are [value_type], and [value_choice] is [value_type]. do you have a preference?
Mars-G BS:	[taxi] leave 10.15 destination tandoori palace departure the nirala [attraction] area south [restaurant] food indian pricerange expensive area south
Mars-G AS:	[attraction] [inform] area type choice [request] type
User (Turn 7):	that sounds wonderful, is it also in the south part of town? if so, i would like a taxi from there to my hotel.
Golden R:	yes, [value_name] is in the [value_area]. in order to help you book a taxi between the park and your hotel, i need to know what hotel you're at.
Golden BS:	[taxi] departure wandlebury country park [attraction] area south [restaurant] food indian pricerange expensive area south
Golden AS:	[attraction] [inform] area name [taxi] [request] destination
Mars-G R:	yes, [value_name] is in the [value_area].
Mars-G BS:	[taxi] leave 10.15 destination tandoori palace departure the nirala [attraction] area south [restaurant] food indian pricerange expensive area south
Mars-G AS:	[attraction] [inform] area name

Table 13: An example in dialog session mul2491. The informable slot error is caused by automatic evaluation.

Goal:	[restaurant]: {informable: pricerange moderate area west }{requestable: reference}
User (Turn 0):	hi! i am planning a trip in cambridge and i am looking for a restaurant in the west part of town.
Golden R:	there are [value_choice] restaurants in the [value_area]. is there any kind of food you're looking for?
Golden BS:	[restaurant] area west
Golden AS:	[restaurant] [inform] choice area [request] food
Mars-G R:	there are [value_choice] restaurants in the [value_area] part of town. do you have a price range or type of food in mind?
Mars-G BS:	[restaurant] area west
Mars-G AS:	[restaurant] [inform] choice area [request] food price
User (Turn 1):	how about you suggest one a book a table for 4 on tuesday at 15:45 and i need the reference number too please
Golden R:	i was able to book [value_name], reference number is [value_reference]
Golden BS:	[restaurant] area west time 15:45 day tuesday people 4
Golden AS:	[restaurant] [offerbooked] name reference
Mars-G R:	i have booked you at [value_name]. your reference number is [value_reference]. is there anything else i can help you with?
Mars-G BS:	[restaurant] area west time 15:45 day tuesday people 4
Mars-G AS:	[restaurant] [offerbooked] reference name [general] [reqmore]

Table 14: An example in dialog session pmul3992. The informable slot error is caused by noisy dialog annotations.

Goal:	[train]: {informable: leave 17:00 destination birmingham new street day monday departure cambridge} {requestable: reference}
User (Turn 0):	what train do i take to get to birmingham new street? i need to leave after 17:00.
Golden R:	there are trains departing hourly at 1 minute past the hour [value_leave]. all are [value_price], and take [value_time].
Golden BS:	[train] destination birmingham new street leave 17:00
Golden AS:	[train] [inform] price leave time
Mars-G R:	where will you be departing from and what day would you like to travel ?
Mars-G BS:	[train] destination birmingham new street leave 17:00
Mars-G AS:	[train] [request] departure day
User (Turn 1):	i am looking to travel on monday. the first train at 17:01 would be fine. can i get that train info?
Golden R:	sure the trainid is [value_id]. it leaves at [value_leave] and arrives at [value_arrive]. would you like to make a reservation?
Golden BS:	[train] destination birmingham new street day monday departure cambridge leave 17:00
Golden AS:	[train] [inform] arrive leave id [offerbook]
Mars-G R:	[value_id] leaves at [value_leave] and arrives at [value_arrive]. would you like me to book that for you?
Mars-G BS:	[train] destination birmingham new street day monday departure cambridge leave 17:00
Mars-G AS:	[train] [inform] arrive leave id [offerbook]
User (Turn 2):	yes please make a reservation and give me the reference number.
Golden R:	your reservation is book. the reference number is [value_reference]. can i assist you with anything else today?
Golden BS:	[train] destination birmingham new street day monday departure cambridge people 1 leave 17:00
Golden AS:	[train] [offerbooked] reference [general] [reqmore]
Mars-G R:	how many tickets do you need?
Mars-G BS:	[train] destination birmingham new street day monday departure cambridge people 1 leave 17:00
Mars-G AS:	[train] [request] people

Table 15: An example in dialog session mul1624. The inaccurate action state causes the requestable slot error.

Goal:	[restaurant]: {informable: food international pricerange moderate area centre} {requestable: reference}
User (Turn 0):	what restaurants in the centre serve international cuisine?
Golden R:	the [value_name] and the [value_name] serve [value_food] food and are in the [value_area] of town.
Golden BS:	[restaurant] food international area centre
Golden AS:	[restaurant] [inform] food area name
Mars-G R:	there are [value_choice] restaurants that meet your criteria. do you have a price range in mind?
Mars-G BS:	[restaurant] food international area centre
Mars-G AS:	[restaurant] [inform] choice [request] price
User (Turn 1):	how about a place in the moderate price range?
Golden R:	[value_choice] of the named restaurants are in the [value_pricerange] price range.
Golden BS:	[restaurant] food international pricerange moderate area centre
Golden AS:	[restaurant] [inform] choice price
Mars-G R:	there are [value_choice] restaurants that meet your criteria. [value_name] and [value_name]. would you like me to book one of these for you?
Mars-G BS:	[restaurant] food international pricerange moderate area centre
Mars-G AS:	[restaurant] [inform] choice name [offerbook]
User (Turn 2):	ok, can you book a table for 6 at 12:00 on tuesday at the varsity restaurant? i will need a reference number too, please.
Golden R:	certainly. i will have that reference number for you in just one second.
Golden BS:	[restaurant] food international pricerange moderate area centre
Golden AS:	[restaurant] [offerbook]
Mars-G R:	booking was successful. the table will be reserved for 15 minutes. reference number is: [value_reference]. is there anything else i can help you with?
Mars-G BS:	[restaurant] food international pricerange moderate area centre time 12:00 day tuesday people 6
Mars-G AS:	[restaurant] [offerbooked] reference [general] [reqmore]
User (Turn 3):	thanks. looking forward to it.
Golden R:	yes i can. i have booked you a table for [value_time] on [value_day] for [value_people] people. your reference number is [value_reference]. anything else i can help you with?
Golden BS:	[restaurant] food international pricerange moderate area centre time 12:00 day tuesday people 6
Golden AS:	[restaurant] [offerbooked] time reference people day
Mars-G R:	you're welcome. have a great day!
Mars-G BS:	[restaurant] food international pricerange moderate area centre time 12:00 day tuesday people 6
Mars-G AS:	[general] [welcome] [bye]

Table 16: An example in dialog session sng0636. The requestable slot error is caused by automatic evaluation.

Goal:	[attraction]: {informable: area north} {requestable: postcode phone }
User (Turn 0):	i am looking for something to do in town. are there any interesting attraction in the north part of town?
Golden R:	yes , there are [value_choice] places: i see a [value_type], a [value_type], and two [value_type]. would one of these work for you?
Golden BS:	[attraction] area north
Golden AS:	[attraction] [inform] type choice [general] [reqmore]
Mars-G R:	there are [value_choice] [value_type] and [value_type] in the [value_area]. do you have a preference?
Mars-G BS:	[attraction] area north
Mars-G AS:	[attraction] [inform] area type choice [request] type
User (Turn 1):	ooh, a pool would be fun. can you give me their contact info?
Golden R:	great! [value_address] is located at [value_address], kings hedges postcode is [value_postcode] and their phone is [value_phone]. is there anything else i can help you with?
Golden BS:	[attraction] area north
Golden AS:	[attraction] [inform] name postcode phone address [general] [reqmore]
Mars-G R:	[value_name] is located at [value_address]. their phone number is [value_phone].
Mars-G BS:	[attraction] area north type swimming pool
Mars-G AS:	[attraction] [inform] name phone address

Table 17: An example in dialog session mul1811. The requestable slot error is acceptable.

Goal:	[train]: {informable: destination stansted airport day friday arrive 15:30 departure cambridge} {requestable: reference}
User (Turn 3):	i will also need a train leaving cambridge.
Golden R:	sure, what is your arrival information?
Golden BS:	[train] departure cambridge
Golden AS:	[train] [request] destination
Mars-G R:	what day will you be traveling, and where will you be traveling to?
Mars-G BS:	[train] departure cambridge
Mars-G AS:	[train] [request] destination day
User (Turn 4):	i need to go to stansted airport on friday.
Golden R:	we have [value_choice] different departure times. any time table your looking at?
Golden BS:	[train] destination stansted airport departure cambridge
Golden AS:	[train] [inform] choice [request] leave
Mars-G R:	there are [value_choice] trains that meet your criteria. do you have a preferred departure or arrival time?
Mars-G BS:	[train] destination stansted airport day friday departure cambridge
Mars-G AS:	[train] [inform] choice [request] leave arrive
User (Turn 5):	i need to be in stansted airport by 15:30 please.
Golden R:	the soonest it will get you there is [value_arrive] your reference number is [value_reference], anything else?
Golden BS:	[train] destination stansted airport arrive 15:30 departure cambridge people 1
Golden AS:	[train] [offerbooked] arrive reference [general] [reqmore]
Mars-G R:	[value_id] will arrive at [value_arrive]. would you like me to book that for you?
Mars-G BS:	[train] destination stansted airport day friday arrive 15:30 departure cambridge
Mars-G AS:	[train] [inform] arrive id [offerbook]

Table 18: An example in dialog session pmul3247. The requestable slot error is acceptable.

Goal:	[attraction]: {informable: type architecture area centre} {requestable: address phone postcode}
User (Turn 0):	i want to find a place to go in the centre to enjoy some entertainment. can you help?
Golden R:	there are [value_choice] attraction types in the [value_area], did you have one in mind?
Golden BS:	[attraction] area centre
Golden AS:	[attraction] [inform] area choice [request] name
Mars-G R:	i am sorry, there are no [value_type] attractions in the [value_area]. would you like to try another area?
Mars-G BS:	[attraction] area centre type entertainment
Mars-G AS:	[attraction] [nooffer] area type [request] area
User (Turn 1):	yes, i am looking for something with the type of architecture. thanks!
Golden R:	i would recommend [value_name] as a great place to see interesting [value_type].
Golden BS:	[attraction] area centre type architecture
Golden AS:	[attraction] [recommend] name type
Mars-G R:	there are [value_choice] [value_type] attractions in the [value_area]. i recommend [value_name].
Mars-G BS:	[attraction] area centre type architecture
Mars-G AS:	[attraction] [recommend] name [inform] type choice area
User (Turn 2):	okay, could you give me the address and phone number?
Golden R:	sure. the address is [value_address]. the phone number is [value_phone]. is there anything else i can help you with today?
Golden BS:	[attraction] area centre type architecture
Golden AS:	[attraction] [inform] address phone [general] [reqmore]
Mars-G R:	the address is [value_address] and the phone number is [value_phone].
Mars-G BS:	[attraction] area centre type architecture
Mars-G AS:	[attraction] [inform] address phone

Table 19: An example in dialog session pmul1320. Noisy dialog annotations cause the requestable slot error.

ACL 2023 Responsible NLP Checklist

A For every submission:

☑ A1. Did you describe the limitations of your work?
Yes. In Section Limitations

☐ A2. Did you discuss any potential risks of your work?
Not applicable. Left blank.

☑ A3. Do the abstract and introduction summarize the paper's main claims?
Yes. In Section 1, we list all the contributions.

☒ A4. Have you used AI writing assistants when working on this paper?
Left blank.

B ☑ Did you use or create scientific artifacts?

Left blank.

☑ B1. Did you cite the creators of artifacts you used?
Yes. In Section 5.2

☒ B2. Did you discuss the license or terms for use and / or distribution of any artifacts?
No. The datasets are public data.

☒ B3. Did you discuss if your use of existing artifact(s) was consistent with their intended use, provided that it was specified? For the artifacts you create, do you specify intended use and whether that is compatible with the original access conditions (in particular, derivatives of data accessed for research purposes should not be used outside of research contexts)?
☒

☒ B4. Did you discuss the steps taken to check whether the data that was collected / used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect / anonymize it?
No. We only use public data.

☑ B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?
Yes. in Section 5.1.

☑ B6. Did you report relevant statistics like the number of examples, details of train / test / dev splits, etc. for the data that you used / created? Even for commonly-used benchmark datasets, include the number of examples in train / validation / test splits, as these provide necessary context for a reader to understand experimental results. For example, small differences in accuracy on large test sets may be significant, while on small test sets they may not be.
Yes. in Section 5.1.

C ☑ Did you run computational experiments?

Left blank.

☑ C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?
Yes. in Section 5.2.

The Responsible NLP Checklist used at ACL 2023 is adopted from NAACL 2022, with the addition of a question on AI writing assistance.

☑ C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
Yes. in Section 5.2 and Appendix E.

☑ C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?
Yes. in Section 5.2.

☑ C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation), did you report the implementation, model, and parameter settings used (e.g., NLTK, Spacy, ROUGE, etc.)?
Yes. in Section 5.2.

D ☒ Did you use human annotators (e.g., crowdworkers) or research with human participants?

Left blank.

☐ D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?
No response.

☐ D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?
No response.

☐ D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating? For example, if you collected data via crowdsourcing, did your instructions to crowdworkers explain how the data would be used?
No response.

☐ D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
No response.

☐ D5. Did you report the basic demographic and geographic characteristics of the annotator population that is the source of the data?
No response.