

Multijugate Dual Learning for Low-Resource Task-Oriented Dialogue System

Shimin Li¹, Xiaotian Zhang¹, Yanjun Zheng¹, Linyang Li¹, Xipeng Qiu^{1,2*},

¹ School of Computer Science, Fudan University

² Shanghai Key Laboratory of Intelligent Information Processing, Fudan University
sml20@fudan.edu.cn, {xiaotianzhang21, yanjunzheng21}@m.fudan.edu.cn,
{linyangli19, xpqiu}@fudan.edu.cn

Abstract

Dialogue data in real scenarios tend to be sparsely available, rendering data-starved end-to-end dialogue systems trained inadequately. We discover that data utilization efficiency in low-resource scenarios can be enhanced by mining alignment information uncertain utterance and deterministic dialogue state. Therefore, we innovatively implement dual learning in task-oriented dialogues to exploit the correlation of heterogeneous data. In addition, the one-to-one duality is converted into a multijugate duality to reduce the influence of spurious correlations in dual training for generalization. Without introducing additional parameters, our method could be implemented in arbitrary networks. Extensive empirical analyses demonstrate that our proposed method improves the effectiveness of end-to-end task-oriented dialogue systems under multiple benchmarks and obtains state-of-the-art results in low-resource scenarios.

1 Introduction

With the emergence of dialogue data (Zhang et al., 2020b), and the evolution of pre-trained language models (Qiu et al., 2020), end-to-end task-oriented dialogue (TOD) systems (Su et al., 2022; Lee, 2021; Tian et al., 2022) gradually replaced the previous modular cascading dialogue systems (Gao et al., 2018). The end-to-end TOD system adopts a uniform training objective, preventing the error propagation problem in pipelined dialogue systems (Gao et al., 2018). Nonetheless, the end-to-end paradigm requires more training data to perform better (Su et al., 2022). Meanwhile, TOD data is enormously expensive to annotate (Budzianowski et al., 2018) as it simultaneously contains dialogue state tracking, dialogue action prediction, and response generation. It is also expensive to annotate large amounts of complicated dialogue data for

*Corresponding Author.

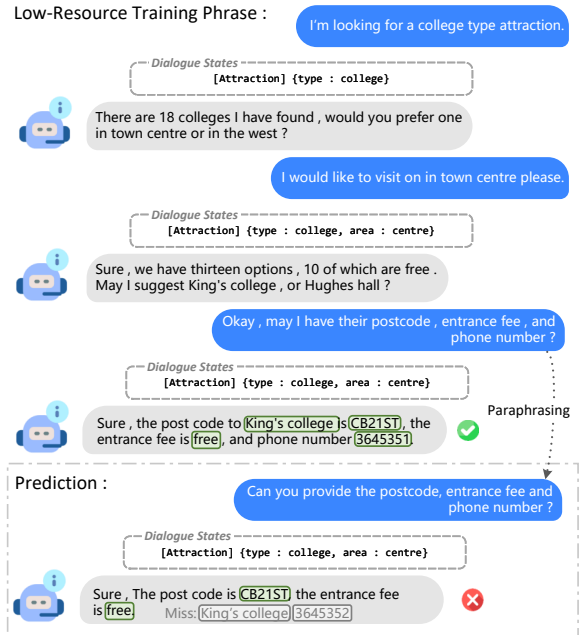


Figure 1: The TOD training and prediction procedure in the low-resource scenario. When the user utterance is rephrased, the predictions miss some entities.

each emerging domain (Mi et al., 2022). Therefore, improving data utilization efficiency in low-resource scenarios becomes critical for end-to-end TOD.

Previous approaches (Zhang et al., 2020b; Su et al., 2022) improve the transferability of models on downstream tasks and capacity to handle small samples by conducting self-supervised or semi-supervised further-pretraining (He et al., 2022) of models on data from additional dialogue domains. However, these further pre-trains on million-level datasets may require hundreds of GPU hours and are resource-constrained. Then on specific downstream dialogue tasks, a unified multi-task generative paradigm (Lee, 2021; Su et al., 2022) was applied to end-to-end dialogue tasks. Although this generative approach demonstrates better generalization and outcomes, we argue that heterogeneity

and duality between data are ignored. Here, heterogeneity refers to the formative discrepancy between uncertain, unstructured discourse (e.g., user utterances and system responses) and deterministic, structured dialogue states. Accordingly, the underlying alignment information and knowledge contained within the heterogeneous data is not fully exploited in the above approach.

To address the above challenges, we propose an innovative multijugate dual learning framework in TOD (MDTOD). Contrary to previous work on reconstructing user discourse based on belief states (Sun et al., 2022; Chen et al., 2020), we observed that modeling the duality between user utterance and system responses can further uncover alignment information of entities between user utterance, system responses and dialogue states. Specifically, the model is required to reconstruct the user discourse based on the dialogue state, and also to deduce the user utterance backwards based on the system response. Consequently, the model can further learn the mapping relationship between the heterogeneous information, and improve the performance of the end-to-end TOD system in low-resource scenarios.

However, proper dual training increases the likelihood of the model learning spurious data correlations. It is evidenced by the fact that comparable model performance can be attained using only high-frequency phrases as the training set (Yang et al., 2022). As a result, the model does not generalize well to test samples with significant expression variations or domain differences, as illustrated in Figure 1. To accomplish this, we expand the one-to-one dual learning paradigm to multijugate dual learning by capitalizing on the property of semantic representation variety. Given a deterministic dialog state as a constraint (Hokamp and Liu, 2017), a specific user utterance (system response) is rewritten into multiple utterances (responses) with the same semantics but various expressions utilizing decoding methods such as beam search or random sampling. Consequently, the richer representation of information permits the spurious correlation of shallow statistical patterns acquired by the model to be effectively mitigated, thereby enhancing the model’s generalization (Cui et al., 2019).

Our proposed method exploits the entity alignment information among heterogeneous data by designing a dual learning task; it also mitigates the phenomenon of false correlations and increases

the generalization capacity of models via rephrase-enhanced multijugate dual learning. As a result, the method does not introduce any additional trainable model parameters. It can be directly integrated into end-to-end TOD systems in arbitrary low-resource scenarios as a training approach to increase data utilization efficiency. We show the effectiveness of our method in several task-oriented datasets, including MultiWOZ2.0 (Budzianowski et al., 2018), MultiWOZ2.1 (Eric et al., 2020), and KVRET (Eric et al., 2017). We also demonstrate the advantages of our approach in low-resource scenarios. All code and parameters will be made public.

Our primary contributions are summarized below:

- A novel, model-independent, dual learning technique intended for low-resource end-to-end TOD systems is presented that can be incorporated directly into the training of any TOD system.
- To address the issue of spurious correlations impacting the generalization of models, a paradigm of paraphrase-enhanced multijugate dual learning is presented.
- We empirically evaluate the technique on several datasets, achieving competitive results without introducing extra model parameters or further pre-training and state-of-the-art results in low-resource circumstances.

2 Related Work

2.1 Task-Oriented Dialogue Systems

TOD aims to complete user-specific goals via multiple turns of dialogue. Prior work focused mainly on TOD subtasks based on the pipeline paradigm (Gao et al., 2018), but it was prone to error propagation between modules. Therefore, recent research has attempted to model dialogue tasks from an end-to-end generation approach. DAMD (Zhang et al., 2020a) generates the different outputs of a conversation process via multiple decoders and expands multiple dialogue actions dependent on the dialogue state. A portion of the study (Hosseini-Asl et al., 2020; Yang et al., 2020; Peng et al., 2021) models the individual dialogue tasks in the TOD as cascading generation tasks using GPT2 (Radford et al., 2019) of the decoder architecture as the backbone network. Multi-task approaches (Lin et al., 2020; Su et al., 2022; Lee, 2021) utilizing

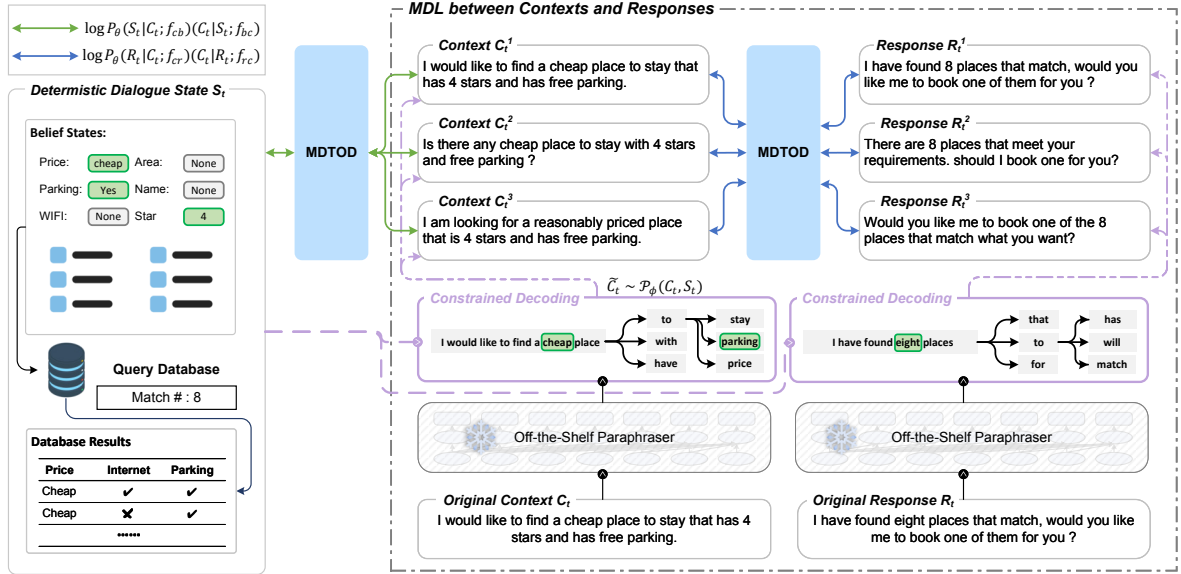


Figure 2: The overall structure of multijugate dual learning. To get paraphrase-enhanced multiple contexts \tilde{C}_t and responses \tilde{R}_t , the contexts and responses in each dialogue turn will be paraphrased based on deterministic dialogue states using an off-the-shelf paraphrase model. Then, the multijugate dual learning is performed between the paraphrase-enhanced contexts \tilde{C}_t and dialogue states and between the paraphrase-enhanced responses \tilde{R}_t and dialogue states, respectively.

encoder-decoder architectures such as T5 (Raffel et al., 2020) or BART (Lewis et al., 2020) exist for modeling dialogue sub-tasks as sequence-to-sequence generating tasks.

Although the methods mentioned above use a uniform end-to-end approach to model TOD, none performs well in low-resource scenarios. To this end, we devise a rephrase-enhanced multijugate dual learning to exploit the entity alignment information more adequately and to obtain more robust performance.

2.2 Dual Learning for Generation

Dual learning aims to utilize the paired structure of data to acquire effective feedback or regularization information, thus enhancing model training performance. Dual learning was initially introduced in unsupervised machine translation (He et al., 2016) and combined with reinforcement learning to optimize two agents iteratively. DSL (Xia et al., 2017) then extended dual learning to supervised settings to take advantage of pairwise relationships of parallel corpora. Similar work (Guo et al., 2020) employs cycle training to enable unsupervised mutual generation of structured graphs and text. MPDL (Li et al., 2021) expands the duality in dialogue tasks to stylized dialogue generation without the parallel corpus. A portion of the work (Sun et al.,

2022; Chen et al., 2020) integrates the idea of duality into the dialogue state tracking. Some of the work (Zhang et al., 2018; Yang et al., 2018; Cui et al., 2019) introduces dual learning in dialogue generation to enhance responses’ diversity, personality, or coherence. However, each method mentioned above requires multiple models or combines reinforcement learning and dual modeling, considerably increasing the task’s complexity and training difficulty.

In contrast to previous work, our proposed multijugate dual learning objectives share the same model parameters. It does not require modifications to the original training objectives of the maximum likelihood estimation, making training more straightforward and more readily applicable to other tasks.

3 Methodology

3.1 End-to-End Task-Oriented Dialogue System

Typically, end-to-end TOD systems consist of sub-tasks such as dialogue state prediction and response generation (Lee, 2021). End-to-end TOD systems typically model the several subtasks of the dialogue process as sequence generation tasks to facilitate the unification of model structure, and training objectives (Hosseini-Asl et al., 2020). Denote the

TOD dataset as $\mathcal{D}_{\text{TOD}} = \{Dial_i, DB\}_{i=1}^N$, where DB is the database. In a multi-turn dialogue $Dial_i$, where the user utterance in the t -th turn is U_t , and the system response is R_t , the dialogue history or dialogue context can be expressed as follows:

$$C_t = [U_0, R_0, \dots, U_{t-1}, R_{t-1}, U_t]. \quad (1)$$

After that, the model generates the dialogue state B_t based on the previous dialogue context C_t :

$$\mathcal{L}_B = \sum_{i=1}^N \sum_{t=1}^{T_i} -\log P_\theta(B_t|C_t), \quad (2)$$

where N represents the total number of sessions in the dataset, T_i symbolizes the total number of turns per session and θ denotes an arbitrary generation model. The system then searches the database with the criterion B_t and retrieves the database result D_t . Then, the TOD system generate the response R_t based on the context U_t , dialogue state B_t and database query result D_t for each round:

$$\mathcal{L}_R = \sum_{i=1}^N \sum_{t=1}^{T_i} -\log P_\theta(R_t|C_t, B_t, D_t). \quad (3)$$

Finally, a human-readable response text containing the entity is obtained by combining the belief state and the search results from the database.

3.2 Multijugate Dual Learning

This section describes how to design dual learning objectives in the training process of TOD. Also, we expound how to construct multijugate dual learning by paraphrasing user utterances and system responses with representational diversity based on deterministic dialogue states.

3.2.1 Dual Learning in TOD

We define the deterministic dialogue state $S_t = [B_t; D_t]$ consisting of two informational components: the belief state B_t and the database query results D_t .

As illustrated in Figure 2, dialogue states can be viewed as information with a unique manifestation of determinism (Zhang et al., 2020a) without regard to the order of dialogue actions. Utilizing dialogue state as a constraint, the natural language of context and response could be viewed as data with different representations of uncertainty. Therefore, we designed the dual task in TOD to learn

the mapping relationship between the utterance of linguistic forms and dialogue state representation.

Let $f_{cb} : C_t \mapsto B_t$ denote the forward learning objective of generating belief states according to the context referred to by Eq.2, and $f_{bc} : B_t \mapsto C_t$ denote the reverse learning objective of reconstructing the context according to the belief states, then the dual learning task between user utterance and dialogue state is defined as maximizing the following logarithmic probability:

$$\log \sum_{i \in N} \sum_{t \in T_i} P_\theta(S_t^i|C_t^i; f_{cb})(C_t^i|S_t^i; f_{bc}). \quad (4)$$

Similarly, let $f_{cr} : C_t \mapsto R_t$, $f_{rc} : R_t \mapsto C_t$ denote the dual learning task between the dialogue context C_t and the system response R_t :

$$\log \sum_{i \in N} \sum_{t \in T_i} P_\theta(R_t^i|C_t^i; f_{cr})(C_t^i|R_t^i; f_{rc}). \quad (5)$$

Accordingly, the loss function of the total dual learning objective is the sum of the above two components:

$$\begin{aligned} \mathcal{L}_{\text{Dual}} = & \mathbb{E}_{\substack{i \sim N \\ t \sim T_i}} -(\log P_\theta(S_t^i, R_t^i|C_t^i; f_{cr}, f_{cb}) \\ & + \log P_\theta(C_t^i|S_t^i; f_{bc}) + \log P_\theta(C_t^i|R_t^i; f_{rc})). \end{aligned} \quad (6)$$

Furthermore, the two dual learning objectives share a set of model parameters in a multi-task paradigm, thus ensuring knowledge transfer between the dual tasks.

3.2.2 Construction of Multijugate Relations

Dual learning enhances data usage efficiency by acquiring additional entity alignment information between heterogeneous data, but it does not lessen the effect of spurious correlations on model generalization. Leveraging the deterministic properties of dialogue states and the uncertainty of linguistic representations, we expand the original one-to-one dual learning to multijugate dual learning by paraphrases. Theoretically, several semantically identical but inconsistently expressed contexts or system responses exist for a deterministic dialogue state. Consequently, given (S_t, C_t) or (S_t, R_t) , we rephrase the context C_t and the response R_t restricted by the entities in dialogue state S_t with the following constraint generation method:

$$\tilde{C}_t \sim \mathcal{P}(C_t, S_t), \tilde{R}_t \sim \mathcal{P}(S_t, R_t). \quad (7)$$

Specifically, we utilize an off-the-shelf paraphrasing model with the dialogue context C_t as the model input. Also the value in the dialogue state S_t will be treated as a constraint to limit the decoding. Then, beam search is employed in generation to obtain K different contexts \tilde{C}_t or responses \tilde{R}_t as the result of paraphrase generation.

Moreover, since the context C_t of the current turn depends on the dialogue history $(\dots, C_{t-1}, S_{t-1}, R_{t-1})$ of the previous turn, rewriting the context or responses of each turn results in a combinatorial explosion. Therefore, a heuristic was adopted whereby the dialogue context C_t and system response R_t would only be rewritten once every dialogue turns. The method for producing the final paraphrase is:

$$\tilde{C}_t^{ij} \sim \sum_{i=1}^N \sum_{t=1}^{T_i} \sum_{j=1}^M \mathcal{P}(C_t^{ij}, S_t^{ij}), \quad (8)$$

$$\tilde{R}_t^{ij} \sim \sum_{i=1}^N \sum_{t=1}^{T_i} \sum_{j=1}^M \mathcal{P}(S_t^{ij}, R_t^{ij}), \quad (9)$$

where M represents the number of single samples to be rewritten. In practice, as the proportion of training data increases, the number of M decreases. In addition, paraphrasing was preferred over word substitution or addition/deletion-based techniques (Wei and Zou, 2019) because word substitution is based on a particular probability of word-level alterations, preventing the modification of phrases with false correlation. Moreover, section 4.4.3 approved paraphrasing produces more diverse and high-quality augmented content, alleviating the risk of spurious relevance more effectively.

3.2.3 Multijugate Dual Learning for Training

By acquiring paraphrase-enhanced samples, the original one-to-one dual learning can be augmented with multijugate dual learning, allowing the model to completely leverage the entity alignment information between heterogeneous data while maintaining appropriate generalization. The overall framework of our method is illustrated in Figure 2. Consequently, the final loss function for multijugate dual learning of TOD is as follows:

$$\begin{aligned} \tilde{\mathcal{L}}_{\text{Dual}} = & \mathbb{E}_{\substack{i \sim N \\ t \sim T_i \\ j \sim M}} -(\log P_{\theta}(S_t^{ij}, R_t^{ij} | C_t^{ij}; f_{cr}, f_{cb}) \\ & + \log P_{\theta}(C_t^{ij} | S_t^{ij}; f_{bc})(C_t^{ij} | R_t^{ij}; f_{rc})). \end{aligned} \quad (10)$$

4 Experiments

In the context of an end-to-end dialogue scenario, we examine the comprehensive performance of multijugate dual learning on several dialogue datasets, including performance on dialogue state tracking and end-to-end task completion. In addition, evaluation studies were conducted in a scenario with limited resources to assess how effectively dual learning utilizes the knowledge contained within the data. In addition, the impact of several dual learning components and rewriting procedures on the method’s overall performance is investigated.

4.1 Datasets and Evaluation Metrics

MultiWOZ2.0 (Budzianowski et al., 2018), MultiWOZ2.1 (Eric et al., 2020), and KVRET (Eric et al., 2017), three of the most extensively investigated datasets in the task-oriented dialogue domain, were analyzed. MultiWOZ2.0 is the first proposed dialogues dataset across seven domains, and MultiWOZ2.1 is the version with several MultiWOZ2.0 annotation problems fixed. Following earlier research, we simultaneously evaluate both datasets to assess the robustness of the model against mislabeling. KVRET is a multi-turn TOD dataset containing three domains: calendar scheduling, weather query, and navigation. Detailed statistics of the three datasets are illustrated in Table 7.

For the selection of metrics under the end-to-end dialogue task, we use the standard and widely used Inform, Success, BLEU, and Combined score, where Inform measures whether the system’s responses refer to the entity requested by the user, Success measures whether the system has answered all of the user’s requests, BLEU measures the quality of the model generation. The Combined score indicates the overall performance of the task-oriented system. It is calculated using the formula: Combined Score = (Inform + Success) * 0.5 + BLEU. For the dialogue state tracking task, the Joint Goal Accuracy (JGA) is applied to quantify the fraction of total turns where the model predicts that all slots in one turn are correct.

MultiWOZ 2.0												
Model	5% Training set				10% Training set				20% Training set			
	Inform	Success	BLEU	Comb.	Inform	Success	BLEU	Comb.	Inform	Success	BLEU	Comb.
MD-Sequicity	49.40	19.70	10.30	44.85	58.10	34.70	11.40	57.80	64.40	42.10	13.00	66.25
DAMD	52.50	31.80	11.60	53.75	55.30	30.30	13.00	55.80	62.60	44.10	14.90	68.25
SOLOIST	69.30	52.30	11.80	72.60	69.90	51.90	14.60	75.50	74.00	60.10	15.25	82.29
MinTL	75.48	60.96	13.98	82.20	78.08	66.87	15.46	87.94	82.48	68.57	13.00	88.53
UBAR	73.04	60.28	16.03	82.89	79.20	68.70	16.09	90.04	82.50	66.60	17.72	92.26
T5-Base	77.80	63.30	14.56	84.94	81.00	67.00	15.17	89.17	84.20	72.70	17.71	96.16
BORT	69.80	45.90	11.00	68.90	74.50	60.60	15.50	83.10	82.10	65.60	14.30	88.10
PPTOD	79.86	63.48	14.89	86.55	84.42	68.36	15.57	91.96	84.94	71.70	17.01	95.32
MTTOD	82.00	64.00	14.48	87.49	82.10	71.10	16.21	92.81	89.50	78.50	15.53	99.53
MDTOD	85.65 (±2.35)	62.20 (±2.70)	15.24 (±1.04)	89.16 (±1.48)	86.30 (±0.90)	71.50 (±0.60)	14.47 (±1.19)	93.37 (±1.04)	90.25 (±0.55)	80.90 (±0.42)	16.40 (±1.15)	101.97 (±0.73)

Table 1: The performance of MDTOD is evaluated at 5%, 10%, and 20% of the data size. Comb. denotes Combined Score.

4.2 Baselines

We did comparison experiments with the following potent baselines. (1) **DAMD** (Zhang et al., 2020a): addresses the one-to-many issue in dialogue by extending dialogue states to many system actions. (2) **SimpleTOD** (Hosseini-Asl et al., 2020): A language model serves as the foundation for end-to-end TOD tasks by generating sequential dialogue states, dialogue actions, and dialogue responses. (3) **DoTS** (Jeon and Lee, 2021): tackles the problem of higher memory consumption owing to lengthy conversation histories by reducing the context and adding domain states as contexts. (4) **SOLOIST** (Peng et al., 2021): further pre-training on heterogeneous dialogue data and transfer learning for dialogue tasks downstream. (5) **MinTL**: employs a copy method to carry over past dialogue states and introduces Levenshtein belief spans to generate a minimal amount of dialogue states efficiently. (6) **UBAR** (Yang et al., 2020): considers belief states, system actions, and system responses as dialogue contexts, hence optimizing the utilization of the dataset’s content. (7) **PPTOD** (Su et al., 2022): A T5-based backbone network with additional pre-training on numerous dialogue datasets and simultaneous multitasking of several dialogue tasks with prompt learning. (8) **MTTOD** (Lee, 2021): Using T5 as the backbone model, two decoders were employed to create dialogue states and system responses, and an additional span prediction task was introduced on the encoder side. (9) **BORT** (Sun et al., 2022): utilizing denoised reconstruction to recover noisy dialogue states and system responses.

4.3 Overall Results

4.3.1 Performance in Low-resource Setting

MultiWOZ To investigate the generalizability of multijugate dual learning with limited resources, we assessed the model on the MultiWOZ2.0 dataset for dialogue sizes of 5%, 10%, and 20%. As shown in Table 1, MDTOD received the highest combined score compared to baselines for all data sizes. MDTOD obtains a 1.67-point improvement in the combined score at 5% of the training data compared to the previous best result. Our strategy produces the highest results for Inform and Success, which are task completion metrics, when applied to 10% and 20% of the data, respectively. In addition, our method obtains highly competitive results compared to PPTOD with additional dialogue data for pre-training and MTTOD with 50% more parameters. Thus, the results above imply that paraphrasing augmented multijugate dual learning that leverages implicit information embedded within the data is more effective in settings with limited resources.

KVRET We also evaluate the impact of multijugate dual learning on the performance improvement of TOD on the KVRET dataset. We use T5-base as the backbone network, where T5+DL indicates the addition of dual learning on T5 and MDTOD indicates the combination of multijugate dual learning on T5. From the experimental results in Table 2, it can be concluded that after applying the dual learning objective under the low resource setting, the model achieves a significant improvement in Success when given different proportions of training samples, indicating that the dual learning can

KVRET												
	10% Training set				20% Training set				50% Training set			
	Inform	Success	BLEU	Comb.	Inform	Success	BLEU	Comb.	Inform	Success	BLEU	Comb.
T5	75.82 (3.42)	18.30 (6.74)	10.51 (0.77)	57.57 (5.14)	80.25 (3.08)	50.81 (8.71)	15.72 (1.75)	81.25 (6.26)	83.42 (2.57)	70.45 (3.13)	17.26 (1.27)	94.20 (2.15)
T5+DL	73.82 (1.29)	33.11 (9.10)	11.55 (1.53)	65.02 (6.36)	82.25 (0.68)	59.58 (3.76)	16.18 (0.90)	87.09 (2.62)	81.07 (5.16)	74.05 (1.18)	18.59 (0.90)	96.15 (2.94)
MDTOD	78.89 (±0.94)	56.49 (±4.62)	14.60 (±0.99)	82.30 (±2.97)	78.71 (±3.36)	64.03 (±6.36)	16.57 (±0.64)	87.94 (±4.98)	84.15 (±1.97)	71.80 (±2.44)	19.06 (±0.79)	97.03 (±1.45)

Table 2: The performance is evaluated at 10%, 20%, and 50% of the data size. The numbers in parentheses indicate the variance of the four runs.

Model	Training Set			
	1%	5%	10%	20%
SimpleTOD	7.91 \pm 1.07	16.14 \pm 1.48	22.37 \pm 1.17	31.22 \pm 2.32
MinTL	9.25 \pm 2.33	21.28 \pm 1.94	30.32 \pm 2.14	35.96 \pm 1.25
SOLOIST	13.21 \pm 1.97	26.53 \pm 1.62	32.42 \pm 1.13	38.68 \pm 0.98
PPTOD _{base}	29.72 \pm 0.61	40.20 \pm 0.39	43.35 \pm 0.64	46.96 \pm 0.40
MDTOD	21.22 \pm 2.86	40.90 \pm 0.20	45.10 \pm 1.40	47.89 \pm 0.55

Table 3: DST evaluated at different proportions of low resources. The results are the means and standard deviations of the four runs.

further learn the alignment information between entities and thus improve the success rate of the task. Meanwhile, T5+DL achieves higher values on BLEU with different proportions of training data, indicating that the dual learning objective between user utterance and system response is also beneficial for improving the quality of text generation. In addition, MDTOD with multijugate dual learning achieves better results, indicating that controlled rephrasing can further enhance the effect of dual learning.

4.3.2 Dual Learning in Dialogue State Tracking

To further investigate the effectiveness of the dual learning task between user utterance and dialogue state on the gain of TOD in multijugate dual learning, we conducted experiments on the MultiWOZ2.0 dataset for dialogue state tracking in low-resource scenarios. We set four different quantitative training sizes of 1%, 5%, 10% and 20% to represent different degrees of low-resource scenarios.

We can infer from the experimental results in Table 3 that MDTOD had the greatest accuracy at three different magnitudes, 5%, 10%, and 20%. MDTOD is lower than PPTOD at 1% magnitude

Model	MultiWOZ 2.0			
	Inform	Success	BLEU	Comb.
Full	85.27	71.07	15.26	93.43
-w/o Para	85.12	70.93	15.09	93.12 (↓0.31)
-w/o DU-DL	85.23	71.23	13.48	91.71 (↓1.72)
-w/o RU-DL	84.70	70.70	13.86	91.56 (↓1.87)
-w/o Both-DL	83.20	70.80	14.42	91.41 (↓2.02)

Table 4: Different setting of multijugate dual learning.

due to that PPTOD performs further pre-training on a large amount of additional dialogue data and thus can achieve relatively better results in extremely low-resource scenarios. Conversely, MDTOD does not perform any additional pre-training, but still achieves the highest accuracy in the case of the other three magnitudes of data, indicating that multijugate dual learning between user utterances and dialogue states is an important component that makes the overall approach effective.

4.4 Analysis

4.4.1 Dismantling multijugate dual learning

To investigate the effect of different dual learning components and paraphrase augmentation on the proposed technique, we conducted ablation experiments by omitting various components using a 10% data size setting. In Table 4, Para represents the approach of paraphrase augmentation, DU-DL represents dual learning between context and dialogue state, and RU-DL indicates dual learning between context and system response.

As shown in Table 4, the model’s performance decreases slightly when only dual learning is retained and the paraphrase enhancement is removed, indicating that multijugate dual learning can partially mitigate the overfitting problem caused by pairwise learning and thereby improve the model’s generalization capability. Among the various dual

KVRET				
Domains	$X_{/schedule} \rightarrow schedule$		$X_{/weather} \rightarrow weather$	
Para. Num	Goal Score	BLEU	Goal Score	BLEU
0	25.84 _{1.63}	10.59 _{0.05}	10.88 _{2.01}	5.80 _{0.68}
1	26.26 _{1.17}	10.01 _{0.50}	13.40 _{3.59}	5.02 _{0.05}
2	26.70 _{0.72}	11.30 _{1.05}	15.09 _{2.29}	5.88 _{0.37}

Table 5: The outcomes of the cross-domain evaluation. $X_{/*} \rightarrow *$ denotes that the $*$ domain is excluded from the training set and only the $*$ domain is tested.

learning components, removing dual learning between context and system responses resulted in a 1.87-point performance decrease, indicating that fully exploiting the implicit alignment information between context and system responses was more effective at enhancing the model’s overall performance. Additionally, deleting both dual learning components resulted in a 2.02 points decrease in the combined score, demonstrating that both dual learning objectives are effective for this strategy.

4.4.2 Mitigating Spurious Correlation for Generalization

This section explores the generalizability of dual learning across domains when different numbers of paraphrases are tested, i.e., on a domain that does not appear in the training process, to examine the effect of rephrasing enhanced multijugate dual learning for mitigating spurious correlations of entities and improving generalization. In the In-Car dataset, we explore the ability of MDTOD to generalize to both the scheduling and weather domains separately.

The Goal Score is calculated as (inform + success) * 0.5 to signify task accomplishment. As indicated in Table 5, the model exhibits some improvement in task completion rate and text generation performance in both new domains when using rephrased augmented multijugate dual learning. Further, when the number of paraphrases is 2, a boost of 4.21 points is obtained on the Goal Score compared to no additional rephrasing mechanism. This improvement indicates that the multiple conjunctions further alleviate the shallow spurious correlations among entities captured by the model, thus improving the task completion rate.

4.4.3 Effect of Different Paraphrases

To investigate the impact of various rephrasing techniques on the construction of multijugate dual learning, we examined the impact of easy data aug-

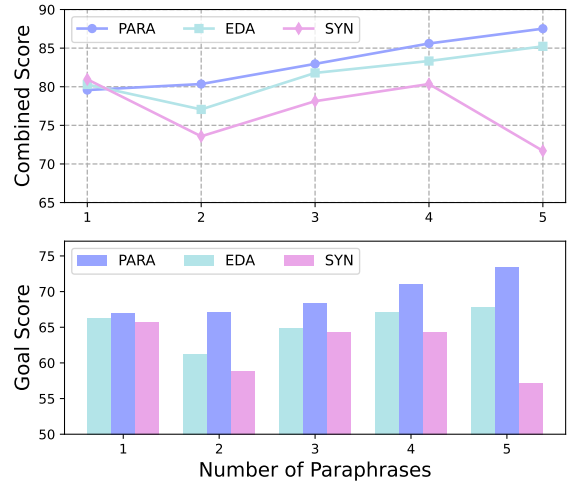


Figure 3: To investigate the impact of various rephrasing strategies on multijugate dual learning.

mentation (EDA) (Wei and Zou, 2019), synonym replacement (SYN), and paraphrasing (PARA) to generate augmented data with limited resources. As demonstrated in the upper part of Figure 3, both PARA and EDA demonstrate minor improvements as the number of augmented data increases, with PARA exceeding EDA. The results indicate that PARA generates higher-quality augmented data, whereas SYN increases noise.

The results in Figure 3 indicate that increasing the number of PARA leads to an increase in the completion rate of dialogue goals. In contrast, EDA and SYN provide a minor boost or decrease in the model’s performance. This analysis reveals that a rephrasing strategy enables better discourse rewriting under dialogue state constraints, alleviating the spurious correlation issue and enhancing the model’s generalizability.

5 Conclusion

We propose a novel multijugate dual learning for task-oriented dialogues in low-resource scenarios. Exploiting the duality between deterministic dialogue states and uncertain utterances enables the entity alignment information in heterogeneous data to be fully exploited. Meanwhile, paraphrase-enhanced multijugate dual learning alleviates the spurious correlation of shallow pattern statistics. Experiments on several TOD datasets show that the proposed method achieves state-of-the-art results in both end-to-end response generation and dialogue state tracking in low-resource scenarios.

Limitations

Multijugate dual learning improves the model’s performance in TOD tasks in low-resource scenarios, but the introduction of the dual training objects increases the required graphics memory and training steps. In addition, the rephrasing mechanism necessitates an additional paraphraser to rewrite the training samples; hence, the number of training samples increases according to the number of paraphrases. Despite this, we find that the higher training cost associated with multijugate dual learning is preferable to employing a large quantity of dialogue data for further pre-training or manually labeling data.

Considered from a different angle, the scenario described above presents possibilities for future research, such as the development of higher-quality rephrasing algorithms to filter the augmented text. In the meantime, multijugate dual learning is a learning objective between structured and unstructured texts. Therefore it may be extended to any task involving heterogeneous data, such as generative information extraction, and data-to-set generation.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (No.2020AAA0108700) and National Natural Science Foundation of China (No.62022027).

References

- Pawel Budzianowski, Tsung-Hsien Wen, Bo-Hsiang Tseng, Iñigo Casanueva, Stefan Ultes, Osman Ramadan, and Milica Gasic. 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, Brussels, Belgium, October 31 - November 4, 2018*, pages 5016–5026. Association for Computational Linguistics.
- Zhi Chen, Lu Chen, Yanbin Zhao, Su Zhu, and Kai Yu. 2020. [Dual learning for dialogue state tracking](#). *CoRR*, abs/2009.10430.
- Shaobo Cui, Rongzhong Lian, Di Jiang, Yuanfeng Song, Siqi Bao, and Yong Jiang. 2019. Dal: Dual adversarial learning for dialogue generation. *arXiv preprint arXiv:1906.09556*.
- Mihail Eric, Rahul Goel, Shachi Paul, Abhishek Sethi, Sanchit Agarwal, Shuyang Gao, Adarsh Kumar, Anuj Kumar Goyal, Peter Ku, and Dilek Hakkani-Tür. 2020. [Multiwoz 2.1: A consolidated multi-domain dialogue dataset with state corrections and state tracking baselines](#). In *Proceedings of The 12th Language Resources and Evaluation Conference, LREC 2020, Marseille, France, May 11-16, 2020*, pages 422–428. European Language Resources Association.
- Mihail Eric, Lakshmi Krishnan, François Charette, and Christopher D. Manning. 2017. [Key-value retrieval networks for task-oriented dialogue](#). In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue, Saarbrücken, Germany, August 15-17, 2017*, pages 37–49. Association for Computational Linguistics.
- Jianfeng Gao, Michel Galley, and Lihong Li. 2018. [Neural approaches to conversational AI](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts*, pages 2–7, Melbourne, Australia. Association for Computational Linguistics.
- Qipeng Guo, Zhijing Jin, Xipeng Qiu, Weinan Zhang, David Wipf, and Zheng Zhang. 2020. Cyclegt: Unsupervised graph-to-text and text-to-graph generation via cycle training. *arXiv preprint arXiv:2006.04702*.
- Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu, Tie-Yan Liu, and Wei-Ying Ma. 2016. [Dual learning for machine translation](#). In *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, pages 820–828.
- Wanwei He, Yinpei Dai, Yinhe Zheng, Yuchuan Wu, Zheng Cao, Dermot Liu, Peng Jiang, Min Yang, Fei Huang, Luo Si, Jian Sun, and Yongbin Li. 2022. [GALAXY: A generative pre-trained model for task-oriented dialog with semi-supervised learning and explicit policy injection](#). In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pages 10749–10757. AAAI Press.
- Chris Hokamp and Qun Liu. 2017. [Lexically constrained decoding for sequence generation using grid beam search](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1535–1546, Vancouver, Canada. Association for Computational Linguistics.
- 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: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*.

- Hyunmin Jeon and Gary Geunbae Lee. 2021. Domain state tracking for a simplified dialogue system. *arXiv preprint arXiv:2103.06648*.
- Sungdong Kim, Sohee Yang, Gyuwan Kim, and Sang-Woo Lee. 2020. [Efficient dialogue state tracking by selectively overwriting memory](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 567–582, Online. 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, Virtual Event / Punta Cana, Dominican Republic, 16-20 November, 2021*, pages 1296–1303. Association for Computational Linguistics.
- Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. [BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 7871–7880. Association for Computational Linguistics.
- Jinpeng Li, Yingce Xia, Rui Yan, Hongda Sun, Dongyan Zhao, and Tie-Yan Liu. 2021. Stylized dialogue generation with multi-pass dual learning. *Advances in Neural Information Processing Systems*, 34:28470–28481.
- 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, EMNLP 2020, Online, November 16-20, 2020*, pages 3391–3405. Association for Computational Linguistics.
- Fei Mi, Yasheng Wang, and Yitong Li. 2022. [CINS: comprehensive instruction for few-shot learning in task-oriented dialog systems](#). In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelfth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pages 11076–11084. AAAI Press.
- Baolin Peng, Chunyuan Li, Jinchao Li, Shahin Shayan-deh, Lars Liden, and Jianfeng Gao. 2021. [SOLOIST: building task bots at scale with transfer learning and machine teaching](#). *Trans. Assoc. Comput. Linguistics*, 9:907–824.
- Xipeng Qiu, Tianxiang Sun, Yige Xu, Yunfan Shao, Ning Dai, and Xuanjing Huang. 2020. [Pre-trained models for natural language processing: A survey](#). *CoRR*, abs/2003.08271.
- Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.
- 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](#). *J. Mach. Learn. Res.*, 21:140:1–140:67.
- 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.
- Yixuan Su, Lei Shu, Elman Mansimov, Arshit Gupta, Deng Cai, Yi-An Lai, and Yi Zhang. 2022. [Multi-task pre-training for plug-and-play task-oriented dialogue system](#). In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 4661–4676. Association for Computational Linguistics.
- 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: NAACL 2022, Seattle, WA, United States, July 10-15, 2022*, pages 2156–2170. Association for Computational Linguistics.
- Xin Tian, Yingzhan Lin, Mengfei Song, Siqi Bao, Fan Wang, Huang He, Shuqi Sun, and Hua Wu. 2022. [Q-TOD: A query-driven task-oriented dialogue system](#). *CoRR*, abs/2210.07564.
- Jason Wei and Kai Zou. 2019. [EDA: Easy data augmentation techniques for boosting performance on text classification tasks](#). 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 6382–6388, Hong Kong, China. 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.
- Yingce Xia, Tao Qin, Wei Chen, Jiang Bian, Nenghai Yu, and Tie-Yan Liu. 2017. Dual supervised learning. In *International conference on machine learning*, pages 3789–3798. PMLR.
- Min Yang, Wenting Tu, Qiang Qu, Zhou Zhao, Xiaojun Chen, and Jia Zhu. 2018. [Personalized response generation by dual-learning based domain adaptation](#). *Neural Networks*, 103:72–82.

- Shiquan Yang, Xinting Huang, Jey Han Lau, and Sarah M. Erfani. 2022. [Robust task-oriented dialogue generation with contrastive pre-training and adversarial filtering](#). *CoRR*, abs/2205.10363.
- Yunyi Yang, Yunhao Li, and Xiaojun Quan. 2020. [UBAR: towards fully end-to-end task-oriented dialog systems with GPT-2](#). *CoRR*, abs/2012.03539.
- Hainan Zhang, Yanyan Lan, Jiafeng Guo, Jun Xu, and Xueqi Cheng. 2018. [Reinforcing coherence for sequence to sequence model in dialogue generation](#). In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden*, pages 4567–4573. [ijcai.org](#).
- Yichi Zhang, Zhijian Ou, and Zhou Yu. 2020a. [Task-oriented dialog systems that consider multiple appropriate responses under the same context](#). In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 9604–9611. AAAI Press.
- Yizhe Zhang, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, and Bill Dolan. 2020b. [DIALOGPT : Large-scale generative pre-training for conversational response generation](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, ACL 2020, Online, July 5-10, 2020*, pages 270–278. Association for Computational Linguistics.
- Li Zhou and Kevin Small. 2019. [Multi-domain dialogue state tracking as dynamic knowledge graph enhanced question answering](#). *CoRR*, abs/1911.06192.

Parameters	MultiWOZ2.0	MultiWOZ2.1	KVRET
Optimizer	AdamW	AdamW	AdamW
LR Scheduler	Linear	Linear	Linear
LR	{2e-4, 4e-4, 6e-4, 8e-4}		{6e-4, 8e-4}
Warmup ratio	0.2	0.2	{0.2, 0.3, 0.4}
Epoch	10	10	6
Top-p	0.7	0.7	0.7
Input Length	512	512	512
Output Length	200	200	200

Table 6: Hyper-parameters used for MultiWOZ2.0, MultiWOZ2.1 and In-Car.

Metric	MWOZ2.0	MWOZ2.1	KVRET
Train	8438	8438	2425
Dev	1000	1000	302
Test	1000	1000	304
Avg. #turns per dialog	13.46	13.46	5.25
Avg. #tokens per turn	13.13	13.13	8.02

Table 7: Statistics of evaluated datasets.

A Implementation Details

A.1 Setup for Experiments

All of our experiments utilize Huggingface’s checkpoints. The backbone network of the end-to-end dialogue model is T5-base. For the generation of paraphrases, we adopt `tuner007/pegasus_paraphrase`¹ directly and construct multiple paraphrases with beam search in decoding. The AdamW optimizer was applied to train the dialogue model and adjusted using linear scheduling with a warmup technique. For the entire dataset in MultiWOZ, we trained 10 epochs with a batch size of 3. Training epochs were relatively increased in the scenario with limited resources. All trials were executed on NVIDIA GeForce RTX 3090 GPU (24G) or NVIDIA A800 (80G). Without additional specifications, the average of three runs with different random seeds was taken as the final result for all experiments.

B Experiments with Full Training Data

B.1 End-to-End Evaluation

Table 9 demonstrates that, given the entire dataset, our proposed technique beats the comparable baselines on both datasets. The combined score on MultiWOZ2.1 has increased by 1.48 points compared to the previous highest result. Notably, our approach does not use more dialogue data for further pre-training, nor does it introduce additional param-

¹https://huggingface.co/tuner007/pegasus_paraphrase.

Generation-based Methods		
Model	Joint Goal Accuracy	
	2.0	2.1
TRADE (Wu et al., 2019)	48.62	46.00
COMER (Ren et al., 2019)	48.79	-
DSTQA (Zhou and Small, 2019)	51.44	51.17
SOM-DST (Kim et al., 2020)	51.38	52.57
dual-DST (Chen et al., 2020)	-	49.88
T5-Base (Raffel et al., 2020)	52.16	52.08
SimpleTOD† (Hosseini-Asl et al., 2020)	51.37	50.14
SOLOIST† (Peng et al., 2021)	53.20	53.36
PPTOD† (Su et al., 2022)	53.57	51.68
MTTOD (Lee, 2021)	53.56	53.44
BORT (Sun et al., 2022)	54.00	-
MDTOD	54.41	53.85

Table 8: Results of the performance comparison between MDTOD and other generative models, using MultiWOZ 2.0 and 2.1 datasets, for the dialogue state tracking. †: The results provided in the publications of these approaches could not be reproduced in MultiWOZ2.1 or with an unfair evaluation script, so we corrected these results based on their open source code.

eters or use a more powerful pre-training model for dialogue. Despite this, Dual-Dialog earns the highest results, proving that dual learning can more thoroughly exploit the information included in the original data and enhance the performance of task-oriented dialogue systems despite the vast amount of data. Our proposed strategy likewise achieves the greatest BLEU on MultiWOZ2.0, showing that the quality of the model’s generated responses has been substantially enhanced.

B.2 Dialogue State Tracking

To further investigate the influence of bipartite modeling between uncertain user utterances and deterministic belief states in dual learning on TOD systems, we compared MDTOD with different generating paradigm baselines while performing the belief state tracking task. According to Table 8, MDTOD obtained up-to-date results for both datasets in the belief state tracking challenge. On MultiWOZ 2.0 and 2.1, our suggested technique achieves a 0.41 JGA improvement above the previous highest BORT and MTTOD. Dual learning between dialogue states and user utterances can learn entity alignment information in the data, resulting in improved performance in belief state tracking.

C Case Analysis

We present partial selections of paraphrases in Table 10 to demonstrate the effect of the rephraser.

Model	MultiWOZ 2.0				MultiWOZ 2.1			
	Inform	Success	BLEU	Comb.	Inform	Success	BLEU	Comb.
DAMD (Zhang et al., 2020a)	76.33	60.40	16.60	84.97	-	-	-	-
SimpleTOD (Hosseini-Asl et al., 2020)	84.40	70.10	15.01	92.26	85.00	70.50	15.23	92.98
DoTS (Jeon and Lee, 2021)	86.59	74.14	15.06	95.43	86.65	74.18	15.90	96.32
SOLOIST (Peng et al., 2021)	85.50	72.90	16.54	95.74	-	-	-	-
MinTL (Lin et al., 2020)	84.88	74.91	17.89	97.79	-	-	-	-
UBAR† (Yang et al., 2020)	85.10	71.02	16.21	94.27	86.20	70.32	16.48	94.74
PPTOD (Su et al., 2022)	89.20	79.40	18.62	102.92	87.09	79.08	19.17	102.26
GALAXY (w/o pretrain) (He et al., 2022)	93.10	81.00	18.44	105.49	93.50	81.70	18.32	105.92
MTTOD‡ (Lee, 2021)	91.80	83.80	19.56	107.36	90.40	81.70	20.15	106.20
MDTOD	92.70	85.00	19.72	108.57	92.70	84.60	19.03	107.68

Table 9: Full dataset comparison results between MDTOD and baselines under end-to-end settings. †: the results in (Su et al., 2022) are utilized. ‡: reproduced results operating the author’s open-source code.

As shown in the first example, when the constraints are set to the entities "hail" and "los angeles", the rephraser still produces paraphrases that are fluent and satisfy the constraints.

In addition, we illustrate a sample of the dialog generated by MDTOD in Table 11 . The dialogue begins with the user seeking an Indian restaurant in the center of town, and the model correctly extracts the values of the slots "food" and "area". When the conversation proceeds to turn 2, MDTOD generates more belief states than oracle’s belief states, but the model generates the correct results. The reason is that there are some labeling errors in MultiWOZ2.0, while MDTOD can still generate correct belief states, which shows the robustness of MDTOD. When the conversation progressed to turn 5, MDTOD still predicted the correct belief state despite the user changing the reservation time from 13:30 to 12:30, indicating that the model understood the semantic meaning of the current input sentences rather than simply repeating the belief state from the previous turn.

Examples	
Constraints	[weather] [value_weather_attribute] hail [value_location] los angeles
Original Utterance	is there going to be hail in los angeles this weekend ?
Original Response	on Sunday hail is predicted to fall in san mateo
Paraphrased Utterance	will hail hit los angeles this weekend?
Paraphrased Response	on sunday hail is foreshadow to fall in san mateo
Constraints	[schedule] [value_event] dentist appointment
Original Utterance	give me the date and time of my dentist appointment
Original Response	your dentist appointment is at [value_time] on [value_date] .
Paraphrased Utterance	tell me the date and time of the dentist appointment
Paraphrased Response	your tooth doctor appointment is at [value_time] on [value_date] .
Constraints	[schedule] [value_party] martha [value_event] dinner [value_time] 11am
Original Utterance	schedule dinner with Martha later this week at 11 am.
Original Response	what day of the week do you want the reminder set for?
Paraphrased Utterance	plan a dinner with Martha later this week at 11 am.
Paraphrased Response	what day of the workweek do you want the reminder set for?

Table 10: A partial sample of the rephrased user utterances and system responses generated by paraphraser, where the constraints are marked in green.

SNG0586		Dialogue
Turn 0	User Utterance	hi, i am looking for a restaurant in the centre of town serving north indian food .
	Oracle Belief State	[restaurant] {food : north indian, area : centre}
	Generated Belief State	[restaurant] {food : north indian, area : centre}
	Oracle Reference	we don't have any specific restaurants in that category. let's try something else ?
	Delexicalized Response	i am sorry , there are no restaurants serving [value_food] food in the centre . would you like to try a different cuisine ?
	Lexicalized Response	i am sorry , there are no restaurants serving north indian food in the centre . would you like to try a different cuisine ?
Turn 1	User Utterance	that 's too bad . how about chinese ?
	Oracle Belief State	[restaurant] {food : chinese, area : centre}
	Generated Belief State	[restaurant] {food : chinese, area : centre}
	Oracle Reference	okay sure what do you have ?
	Delexicalized Response	there are [value_choice] [value_food] restaurants in the [value_area] . do you have a price range preference ?
	Lexicalized Response	there are 3 chinese restaurants in the centre . do you have a price range preference ?
Turn 2	User Utterance	okay when you find something i would like to make a reservation for 2 people on a sunday at 13:30 .
	Oracle Belief State	[restaurant] {food : chinese, area : centre}
	Generated Belief State	[restaurant] {food : chinese, area : centre, time : 13:30, day : sunday, people : 2}
	Oracle Reference	do you have a price range in mind ?
	Delexicalized Response	i was able to book you at [value_name] . your reference number is [value_reference] . is there anything else i can help you with ?
	Lexicalized Response	i was able to book you at charlie chan . your reference number is 4wngilmf . is there anything else i can help you with ?
Turn 3	User Utterance	i don't mind about the price range .
	Oracle Belief State	[restaurant] {food : chinese, area : centre}
	Generated Belief State	[restaurant] {food : chinese, area : centre, time : 13:30, day : sunday, people : 2}
	Oracle Reference	okay , how about lan hong house ? it is a chinese resturant in the centre area and is in the moderate price range . will that work for you ?
	Delexicalized Response	i would recommend [value_name] .
	Lexicalized Response	i would recommend lan hong house .
Turn 4	User Utterance	that 's fine . can you try booking it ?
	Oracle Belief State	[restaurant] {food : chinese, area : centre, time : 13:30, day : sunday, people : 2}
	Generated Belief State	[restaurant] {food : chinese, area : centre, time : 13:30, day : sunday, people : 2}
	Oracle Reference	i am sorry , but i was n't able to book the restaurant for that date and time . would you like to try a different time ?
	Delexicalized Response	booking was successful . the table will be reserved for 15 minutes . reference number is : [value_reference] . anything else i can help with ?
	Lexicalized Response	booking was successful . the table will be reserved for 15 minutes . reference number is : f1ntkwzo . anything else i can help with ?
Turn 5	User Utterance	how about 12:30 ?
	Oracle Belief State	[restaurant] {food : chinese, area : centre, name : lan hong house, time : 12:30, day : sunday, people : 2}
	Generated Belief State	[restaurant] {food : chinese, area : centre, time : 12:30, day : sunday, people : 2}
	Oracle Reference	you are booked for 12:30 on sunday for 2 people . your reference number is f1ntkwzo . is there something else i can help you with ?
	Delexicalized Response	booking was successful . the table will be reserved for 15 minutes . reference number is : [value_reference] . anything else i can help with ?
	Lexicalized Response	booking was successful . the table will be reserved for 15 minutes . reference number is : f1ntkwzo . anything else i can help with ?
Turn 6	User Utterance	thanks so much . have a nice day . bye .
	Oracle Belief State	[restaurant] {food : chinese, area : centre, name : lan hong house, time : 12:30, day : sunday, people : 2}
	Generated Belief State	[restaurant] {food : chinese, area : centre, time : 12:30, day : sunday, people : 2}
	Oracle Reference	thank you , if you need anything else let me know . good bye .
	Delexicalized Response	thank you for using our system !
	Lexicalized Response	thank you for using our system !

Table 11: A dialogue sample generated by MDTOD on MultiWOZ 2.0.

ACL 2023 Responsible NLP Checklist

A For every submission:

- A1. Did you describe the limitations of your work?
Section 6
- A2. Did you discuss any potential risks of your work?
Section 6
- A3. Do the abstract and introduction summarize the paper’s main claims?
Section 1
- A4. Have you used AI writing assistants when working on this paper?
Left blank.

B Did you use or create scientific artifacts?

Section 3,4

- B1. Did you cite the creators of artifacts you used?
Section 3,4
- B2. Did you discuss the license or terms for use and / or distribution of any artifacts?
Section 3,4
- 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)?
Section 4.1
- 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?
Section 4.1
- B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?
Section 4.1
- 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.
Section 4.1

C Did you run computational experiments?

Section 4.3, 4.4

- C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?
Section A.1

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?

Section A.1

- 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?

Section 4.3, 4.4

- 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.)?

Section 4.1

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.