Inverse Prompt Response Reconstruction and Contextual Momentum Contrast for Few-Shot Dialogue Generation

1st Di Wu
Hebei University of Engineering School Information and Electronic Engineering
Handan, Hebei, P.R. China
wudiwudi@hebeu.edu.cn

2nd Fanming Meng*
Hebei University of Engineering School Information and Electronic Engineering
Handan, Hebei, P.R. China
18332552571@163.com

3rd Yao Peng
Hebei University of Engineering School Information and Electronic Engineering
Handan, Hebei, P.R. China
S24pengyao1035@hebeu.edu.cn

***Abstract*—Dialogue generation is an essential component of task-oriented dialogue systems, aiming to assist users in completing domain-specific tasks via natural language interaction. Existing few-shot dialogue generation models face challenges in perceiving domain dialogue state and generating responses that are inconsistent with the dialogue context. To address these challenges, we propose the IPCMC-TOD model, a Few-shot Dialogue Generation Model with Inverse Prompt Response Reconstruction and Contextual Momentum Contrast. Specifically, considering the critical role of the model's adaptability to domain dialogue state, we design an inverse prompt response reconstruction. An inverse-guided dual-prompt template is constructed. An inverse-prompt-trained slot extractor is employed to reconstruct the original system responses. To enhance the logical consistency of dialogue responses, we design a Contextual Momentum Contrast approach. Based on the posterior inference response, the complete set of samples is categorised into positive and negative sets. Momentum-based pseudo-samples are generated by the momentum encoder. We define a contrast loss function with margin constraints for positive and negative samples. This function significantly enhances the fine-grained discrimination ability of the response selector. Experimental results on the MultiWOZ 2.0 and MultiWOZ 2.1 datasets demonstrate the effectiveness of IPCMC-TOD over all baseline models.**

***Index Terms*—Task-Oriented Dialogue, Few-Shot Learning, Dialogue Response Generation, Contrastive Learning, Prompt Learning**

I. Introduction

Task-oriented dialogue (TOD), as a crucial branch of human-computer interaction, helps users accomplish domain-specific tasks via natural language interaction. As an essential component of TOD, dialogue generation maintains contextual semantic coherence throughout a conversation and produces responses aligned with user needs. Existing approaches rely on large-scale, well-labelled datasets for supervised training. However, annotating multi-turn dialogues in real-world settings is prohibitively costly [1]. This data scarcity has constrained development in task-oriented dialogue generation.

Few-shot dialogue generation aims to learn generalizable dialogue features, reducing reliance on large amounts of labelled data to mitigate data scarcity issues. Some works [2]–[5] adopted a multi-task pretraining paradigm. By sharing parameters and constraints across related tasks, this paradigm enhances the integration of general and complementary feature representations. Data are implicitly enhanced and the model's parameter size is reduced by this approach, improving its transferability and generalisation ability in scenarios with scarce annotated data. Other works [6], [7] characterised the mapping relationships of heterogeneous information, representing the alignment between dialogue states and user utterances. Database constraints are employed as discriminative signals by this approach, mitigating the reliance on explicit dialogue state annotations. By extracting supervisory information from scarce data, the approach improves data utilisation in few-shot scenarios. Although these approaches effectively mitigate the data scarcity problem, their performance remains constrained by two key challenges: (1) Insufficient awareness of dialogue state information in the target domain leads to hindered model adaptation to new domains. (2) Excessive reliance on local context in system responses leads to generated responses that are inconsistent with the dialogue context. An example of the challenges in few-shot dialogue generation is shown in Fig. 1.

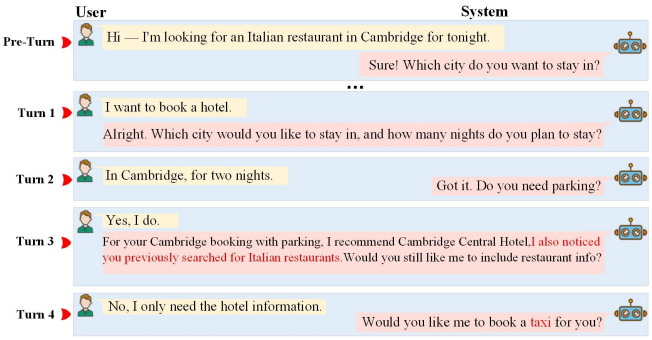


Fig. 1: An example of the challenges in few-shot dialogue generation.

In the third dialogue turn, the system mistakenly carries over the dialogue state from the restaurant domain to the hotel domain. This error leads to the introduction of redundant

information in the system's response, reducing the accuracy of response generation. In the fourth dialogue turn, the system generates a response for taxi booking, yet fails to fulfill the user's request for hotel recommendations. This results in a logical inconsistency between the system responses and the dialogue context, compromising the overall interaction quality.

To address these challenges, we propose the IPCMC-TOD model[1], a Few-Shot Dialogue Generation Model with Inverse Prompt Response Reconstruction and Contextual Momentum Contrast. We design an Inverse Prompt Response Reconstruction approach to enhance the model's awareness of domain-specific dialogue state. The approach employs an inverse-guided dual-prompt template and trains a slot extractor with inverse prompt to reconstruct the original system responses. We develop a Context-Momentum Contrast strategy to improve logical consistency between system responses and the dialogue context. The strategy employs a momentum encoder to generate pseudo-samples. A contrast loss with margin constraints is designed between the positive and negative sets derived from posterior inference. The fine-grained discriminative capability of the response selector is significantly enhanced. The main contributions of this paper are as follows:

- A novel few-shot dialogue generation model, IPCMC-TOD, is proposed to enhance domain-specific dialogue state adaptation and improve the logical consistency of responses.
- We design the approach of Inverse Prompt Response Reconstruction and Contextual Momentum Contrast. An inverse-guided dual-prompt template is constructed, and an inverse-prompt-trained slot extractor is employed to reconstruct original system responses. We design a contrast loss function with margin constraints for positive and negative samples. This function significantly enhances the fine-grained discrimination ability of the response selector.
- We validate the IPCMC-TOD model on the MultiWOZ 2.0 and MultiWOZ 2.1 datasets, achieving state-of-the-art performance. This performance demonstrates its effectiveness in few-shot dialogue generation.

II. Related Works

A. Few-shot dialogue generation

The pre-trained language models have been widely introduced into few-shot dialogue generation, owing to their strong capabilities in language understanding and contextual reasoning. Su et al. [8] proposed Galaxy, a plug-and-play unified dialogue model that uses semi-supervised self-training to explicitly learn dialogue policy. Li et al. [6] designed MDTOD, a multi-dimensional dual-learning framework that mines alignment information between uncertain utterances and deterministic dialogue state. Feng et al. [9] presented a reinforcement-learning framework called RewardNet, equipped with two learnable reward functions that supervise the training of dialogue decision-making and alleviate the delayed feedback of sparse signals. Liu et al. [2] introduced TAAT, a method that combines a turn-level auxiliary task with action-tree planning and sampling. Huang et al. [7] proposed OSTOD, a posterior-inference approach using database constraints as discriminative signals to reduce reliance on dialogue-state annotations. While these methods effectively mitigate data scarcity, they face challenges in effectively perceiving domain dialogue state and exhibit inconsistency in system response generation.

B. Contrastive Learning

Contrastive learning is widely applied in dialogue generation. Sun et al. [10] proposed Mars, a contrastive learning method designed to enhance the separability of dialogue-context representations in the semantic state space. Wang et al. [11] proposed Enco, an entity-based approach designed to exploit entity information in knowledge-grounded dialogue samples for constructing semantically relevant and irrelevant distractors. Gao et al. [12] introduced Winnie, a structure-aware contrastive learning method designed to mine intrinsic structural information from utterances across different categories by leveraging contrastive learning. Huangfu et al. [13] proposed NEC, a non-emotion-centric empathetic dialogue framework that constructs four types of negative samples and employs contrastive learning to reduce dependence on explicit emotion recognition.

While these methods improve response quality via contrastive learning, they still suffer from a scarcity of negative samples. Inspired by Momentum Contrast [14], we propose a Contextual Momentum Contrast approach. By maintaining a queue of negatives and generating momentum pseudo-samples, our method learns fine-grained context-response interactions. This significantly enhances contextual consistency.

C. Prompt Learning

Prompt learning can fully leverage the semantic knowledge and contextual information learned by pre-trained models, reducing the dependence on domain-specific data. Bai et al. [15] introduced I2-tuning, a hybrid prompt learning approach that employs attention-based semantic interface alignment and jointly optimizes the response generator. Huang et al. [16] proposed PERGM, a retrieval-augmented prompt learning method that decodes retrieval examples into prompt information to guide the model's focus on the interaction between personality and sentiment. Zhang et al. [17] introduced MPA, a modality-based prompt attention mechanism that designs specific prompt signals to guide the model's attention toward under-optimized modal features. He et al. [18] proposed BP4ER, a guided prompt method for explicit reasoning that employs a least-to-most prompting strategy to guide a large language model, improving the quality of responses in medical dialogue.

The above methods utilize prompt learning to enhance the data utilization in dialogue response processes. However, they do not adequately exploit the implicit information within the

[1]The source code and datasets are publicly available at: https://github.com/LLLLLearn/IPCMC-TOD

dialogue context. Inspired by inverse prompt learning [19], we propose an Inverse Prompt response Reconstruction approach. This approach learns deep slot-action semantics to reconstruct original system responses, supervising the fine-tuning process and significantly improving the model's perception of domain dialogue state.

III. METHODOLOGY

To address the limitations of perceiving domain dialogue state and the inconsistency between generated responses and the dialogue context in few-shot dialogue generation, we propose the IPCMC-TOD model, as shown in Fig. 2.

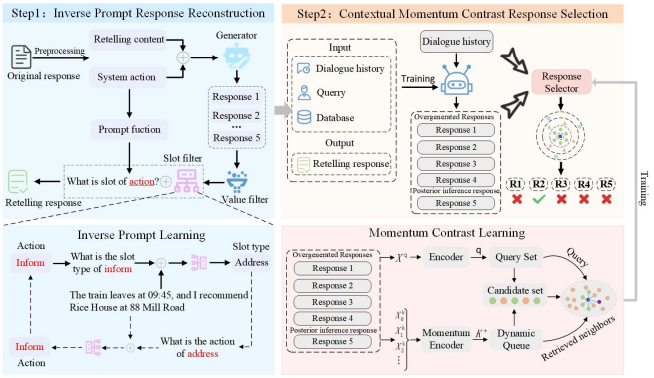


Fig. 2: The overall architecture of the proposed IPCMC-TOD model.

A. Preliminaries

In task-oriented dialogue generation, given a T-turn dialogue history $H = \{u_1, y_1, ..., u_T, y_T\}$, where u_t and y_t denotes the user utterance and the system response at the t-th turn ($1 \leq t \leq T$), respectively. We define the dialogue context for t-th turn as $c_t = \{u_1, y_1, ..., u_{t-1}, y_{t-1}\}$, which is a subset of the dialogue history H, and u_t is the current user query. In addition, $D = \{ds\_nores, ds\_0, ds\_1, ds\_2, ds\_3\}$ denote the number of retrieval entities from the database. $[ds\_0]$ denotes no results. $[ds\_1]$ denotes 0 to 5 results. $[ds\_2]$ denotes 6 to 10 results. $[ds\_3]$ denotes more than 10 results. The overall task is thus designed as learning a mapping that takes (c_t, u_t, D) as input and generates the natural language response y_t.

B. Inverse Prompt Response Reconstruction

In task-oriented dialogue generation, dialogue state is represented as domain slot value triplets, constraining database queries. This representation limits the model's cross-domain generalisation capability [7]. To address the limitations of domain-aware dialogue state, we propose an inverse prompt response reconstruction approach. Two types of generative model T5 [20] are employed for generating and filtering the retelling response, respectively.

To fully exploit deep semantic information in the dialogue, we design an inverse prompt domain slot filtering approach. As shown in Fig. 2, the primary task predicts slot information conditioned on the dialogue action. We introduce an inverse prompt mechanism [19] as an auxiliary task that inversely supervises the filter T5_{fil}. We employ a prompt-based question template. The equation of prompt input sequence X_i for the primary task is shown below:

$$X_i = \textit{prefix} \oplus c_t \oplus f(a) \tag{1}$$

where *prefix* denotes the predefined prompt, set to "answer the question:"; c_t denotes the dialogue context; $f(a)$ is the prompting function instantiated as $f(a)$ = "what is the slot of $[a]$?", and a denotes the domain action.

The equation of the loss function L_1 for the primary task is shown below:

$$L_1 = -\sum \log P\left(S \mid \textit{prefix}, c_t, f(a)\right) \tag{2}$$

where $P\left(S \mid \textit{prefix}, c_t, f(a)\right)$ denotes the probability of generating the gold slot S.

To assist in training the primary task by distinguishing slot types across different domain actions, we employ an inverse prompt auxiliary task that reverses the original question-answer pairs. The question is specified by the slot type and the answer by the corresponding action. This auxiliary task encourages the model to capture deeper semantic information about the system response, improving learning for the slot-prediction objective. The equation of prompt input sequence Y_i for the auxiliary task input is shown below:

$$Y_i = \textit{prefix} \oplus c_t \oplus I(s) \tag{3}$$

where $I(s)$ is the inverse prompting function instantiated as $I(s)$ = "What is the act of $[s]$?", and s denotes the domain slot.

The equation of the loss function L_2 for the auxiliary task is shown below:

$$L_2 = -\sum \log P\left(A \mid \textit{prefix}, c_t, I(s)\right) \tag{4}$$

where $P\left(A \mid \textit{prefix}, c_t, I(s)\right)$ denotes the probability of generating the gold action A.

The total loss function incorporates two loss functions from the primary and auxiliary tasks. The equation of the total loss function L is shown below:

$$L = L_1 + W * L_2 \tag{5}$$

where W is the auxiliary task weight that controls its contribution to the total loss.

By fusing the current turn domain dialogue state information with the retrieved entity names from the original response, we construct the retelling content. By combining the retelling content with the system action, we construct a retelling system action. Conditioned on this retelling system action, a pretrained generator T5_{gen} produces multiple fluent candidate retelling system responses.

Finally, using the dialogue state information contained in the retelling content, we perform slot-value filtering. The primary task prompting function $f(a)$ is fed into the inverse prompt-trained filter T5_{fil}, which extracts slot information from the candidate retelling system responses. In this way, we obtain retelling responses with matched slot-value pairs.

C. Contextual Momentum Contrast Response Selection

To address the inconsistency between generated responses and the dialogue context, we propose a contextual momentum contrast response selection approach. It distinguishes fine-grained differences among candidate responses. Inspired by prior work [21], given candidate set $R = \{r_1, r_2, ..., r_m\}$, where m denotes the number of responses, for each $r_k \in R$ $(1 \leq k \leq m)$ we compute its similarity to the ground-truth response r_{gt}, yielding a score $S_k = S(r_k, r_{\text{gt}})$. Specifically, the posterior-inferred response is generated by producing delexicalized candidate responses for each database state and filtering them based on state consistency [7]. According to the score S_{post} of this response, the candidate set R is partitioned into a high-score subset R_{high} and a low-score subset R_{low}.

In the response selection task, each data sample is a triplet (c, r, l) with $l \in \{0, 1\}$, indicating whether the candidate r is a high-quality or a low-quality response. For each training sample $(c^{(i)}, r_{\text{gt}}^{(i)})$, the model produces a candidate set $R^{(i)} = \{r_1^{(i)}, r_2^{(i)}, ..., r_k^{(i)}\}$, which also includes the posterior inferred response $r_{\text{post}}^{(i)}$. We compute the cosine similarity score $S_k^{(i)}$ between each candidate $r_k^{(i)}$ and the groundtruth response $r_{\text{gt}}^{(i)}$. Using the score $S_{\text{post}}^{(i)}$ of the posterior inferred response, as a threshold, the set $R^{(i)}$ is split into a positive subset $R_{\text{high}}^{(i)}$ and a negative subset $R_{\text{low}}^{(i)}$. Specifically, if $S_k^{(i)} \geq S_{\text{post}}^{(i)}$, then $r_k^{(i)} \in R_{\text{high}}^{(i)}$; otherwise $r_k^{(i)} \in R_{\text{low}}^{(i)}$. To balance the two subsets of mismatched sizes, we downsample the larger one to match the scale of the smaller, retaining $\min\left(|R_{\text{high}}^{(i)}|, |R_{\text{low}}^{(i)}|\right)$ candidates.

For each $r_k^{(i)} \in R_{\text{high}}^{(i)}$, we construct a positive sample $(c^{(i)}, r_k^{(i)}, 1)$. For each $r_k^{(i)} \in R_{\text{low}}^{(i)}$, we construct a negative sample $(c^{(i)}, r_k^{(i)}, 0)$. To perform the response selection task, we pretrain a response selector in an unsupervised manner using contrastive learning. This training drives the positives $R_{\text{high}}^{(i)}$ to form a compact cluster in the embedding space, well separated from the negatives $R_{\text{low}}^{(i)}$. Finally, we cast the response selection task as a sentence-similarity problem.

We introduce a momentum queue mechanism to continuously store negative features produced by a momentum encoder, alleviating the issue of limited negative samples in each training mini-batch. The equation of the update rule for the momentum encoder parameters θ_v is shown below:

$$\theta_v = l * \theta_v + (1 - l) * \theta_{\text{p}} \tag{6}$$

where θ_v and θ_p denote the parameters of the momentum encoder and the primary encoder, respectively, and the momentum coefficient l is set to 0.999.

As illustrated in Fig. 2, we design a margin constraints contrast loss that pulls the anchor sample toward its positives in feature space and pushes it away from negatives. This objective enhances the model's ability to discriminate between positive and negative samples. The equation of the definition for the positive index set $P(i)$ and the negative index set $N(i)$ for the *i*-th sample in a mini-batch is shown below:

$$P(i) = \{\, j \in \{1, \ldots, Q\} \mid g_j = g_i \,\} \tag{7}$$

$$N(i) = \{\, j \in \{1, \ldots, Q\} \mid g_j \neq g_i \,\} \tag{8}$$

where Q denotes the momentum-queue size and g_i is the label of the *i*-th sample.

The equation of the definition for the minimum distances from sample i to all positive samples $d_{\text{pos}}(i)$ and negative samples $d_{\text{neg}}(i)$ is shown below:

$$d_{\text{pos}}(i) = \begin{cases} \min\limits_{j \in P(i)} d(z_i, m_j), & |P(i)| > 0, \\ 0, & |P(i)| = 0 \end{cases} \tag{9}$$

$$d_{\text{neg}}(i) = \begin{cases} \min\limits_{j \in N(i)} d(z_i, m_j), & |N(i)| > 0, \\ \delta, & |N(i)| = 0 \end{cases} \tag{10}$$

where z_i denotes the embedding vector of the *i*-th sample, m_j denotes the feature vector of the j-th sample in the momentum queue, and $d(\cdot, \cdot)$ is the Euclidean distance.

The equation of the average loss over positive samples L_{pos} is shown below:

$$L_{\text{pos}} = \frac{\sum_{i=1}^{B} \mathbf{1}_A(|P(i)| > 0) \cdot d_{\text{pos}}(i)}{\sum_{i=1}^{B} \mathbf{1}_A(|P(i)| > 0)} \tag{11}$$

where B denotes the current batch size and $\mathbf{1}_A(\cdot)$ is the indicator function of event A, which returns 1 if A is true and 0 otherwise.

To enforce a margin between sample i and all negative samples, we introduce a margin parameter $\delta > 0$ and define the average negative-sample loss L_{neg}. The equation of the average negative-sample loss L_{neg} is shown below:

$$L_{\text{neg}} = \frac{\sum_{i=1}^{B} \mathbf{1}_A(|N(i)| > 0) \cdot \max(0, \delta - d_{\text{neg}}(i))}{\sum_{i=1}^{B} \mathbf{1}_A(|N(i)| > 0)} \tag{12}$$

The equation of L_{total} , defined as the sum of positive and negative sample losses, is shown below:

$$L_{\text{total}} = L_{\text{pos}} + L_{\text{neg}} \tag{13}$$

We utilize a response selector to extract multiple candidate responses, using a portion of the data as the anchor set. This captures the fine-grained differences between candidate responses. The equation of the final response $\hat{r}$ is shown below:

$$\hat{r} = \arg\max(e) \tag{14}$$

where e represents the embedding obtained from the response selector, with $e = \text{enc}(c, r)$, c representing the dialogue context, and r denoting the candidate response.

IV. Experiments

A. Datasets and Metrics

We evaluate the performance of the IPCMC-TOD model on the MultiWOZ 2.0 [1] and MultiWOZ 2.1 [22] datasets. To simulate the scarcity of labelled data, we construct the few-shot environment by randomly sampling 5%, 10%, 20%, and 50% of training data at the dialogue level to form labeled subsets, and treat the remaining dialogues as unlabeled data.

We use **Inform**, **Success**, and **BLEU** to evaluate the IPCMC-TOD model. Specifically, Inform evaluates whether the system, under the user's constraints, returns an entity of the correct type. Success measures whether the system fulfills the user's request. BLEU reflects the fluency of the generated response. We report a **Combined** $((\text{Inform} + \text{Success}) * 0.5 + \text{BLEU})$ metric for comprehensive evaluation.

B. Implementation

All experiments are conducted in a cloud server environment. We build and train the IPCMC-TOD model on a single NVIDIA V100 GPU (32 GB). The experimental environment requires Python 3.8.10 programming language, PyTorch 1.8.1 deep learning framework, and CUDA 11.1. We use T5-small as both the filter and the generator, and also as the backbone of the IPCMC-TOD model. For the response selection task, we compute similarity scores of candidate responses with all-mpnet-base-v2 [23], and employ BERT [24] as the backbone of the response selector. The Top-K value was set to 5 for MultiWOZ 2.0 and 4 for MultiWOZ 2.1. For all other hyperparameters across both datasets, we set the learning rate to 5e-5 and the batch size to 16. The model was trained for 20 epochs, and the random seed was fixed at 62.

C. Comparison Experiments

To evaluate the performance of the proposed IPCMC-TOD model, we compare it with eleven baseline models for dialogue response generation and analyse the results.

From TABLE I, on the MultiWOZ 2.0 and MultiWOZ 2.1 datasets, the IPCMC-TOD model using only 60M parameters and without any further pre-training achieves strong results on Inform, Success, BLEU, and the Combined metric. Experiment analysis shows a particularly prominent gain on *Success*: relative to the best baseline model, OSTOD, the scores increase by 1.2% on MultiWOZ 2.0 and 1.0% on MultiWOZ 2.1. The improvement primarily stems from the proposed contextual momentum contrast, which enhances fine-grained semantic discrimination. This improves the match between candidate responses and the context. Meanwhile, the IPCMC-TOD model delivers strong *Inform* performance, surpassing the best baseline UBAR by 0.1% on MultiWOZ 2.0 and by 1.4% on MultiWOZ 2.1, respectively. This benefit is attributed to inverse prompt response reconstruction, which enhances bidirectional association learning between domain slots and system acts. As a result, it improves the model's perception and retelling of domain dialogue state.

D. Few-shot Experiment

To validate the performance of the proposed IPCMC-TOD model for task-oriented dialogue in few-shot settings, we conduct experiments on the MultiWOZ 2.0 dataset using training sets of 5%, 10%, 20%, and 50%. As shown in in TABLE II, the IPCMC-TOD model delivers consistently strong results across all data scales. With 10% and 20% data, the combined metric surpasses all baseline models, whereas the Inform and Success metrics are relatively weaker. The primary cause is the imbalance of the training sample distribution. This imbalance induces a reliance on high-frequency dialogue patterns; it also limits the model's ability to represent structurally complex or infrequent dialogue situations. In addition, under limited samples, the IPCMC-TOD model cannot fully learn the mapping from complex dialogue state to system actions, reducing the accuracy of the generated response and the effectiveness of task completion. In summary, the proposed inverse prompt response reconstruction together with contextual momentum contrast for response selection demonstrates a more pronounced advantage in low-resource conditions.

E. Ablation Experiments

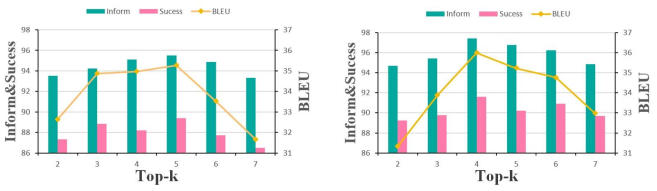
To verify the effectiveness of each module, we conduct ablation experiments on the MultiWOZ 2.0 dataset by selectively removing the Inverse Prompt response reconstruction (IP) and the Contextual Momentum Contrast (CMC). The ablation results for the IPCMC-TOD model are reported in TABLE III.

After removing the Inverse Prompt response reconstruction module, the Inform, Success, and BLEU metrics decrease by 0.2%, 0.6%, and 0.16%, respectively. This result indicates that the module primarily addresses redundancy in domain-slot value triplets within dialogue state. This is achieved by extracting retelling content from the original system response and constructing a retelling system action, thereby effectively enhancing the model's dialogue state awareness.

After removing the Contextual Momentum Contrast module, the Inform, Success, and BLEU metrics decrease by 0.15%, 0.30%, and 0.11%, respectively. This result indicates that the module mitigates the inconsistency between the system response and the dialogue context. Lacking the ability to discriminate fine-grained differences among candidate responses, the model struggles to select a system response that is consistent with the context from multiple semantically similar responses.

F. Hyperparameter Experiment

We investigate how the number of generated responses (Top-K) affects model performance. Specifically, we vary Top-K from 2 to 7 and evaluate the performance of the IPCMC-TOD model under different numbers of generated responses on the MultiWOZ 2.0 and MultiWOZ 2.1 datasets. The experimental results are shown in Fig. 3.



(a) The MultiWOZ 2.0 dataset (b) The MultiWOZ 2.1 dataset

Fig. 3: Results of the Hyperparameter Experiment on the IPCMC-TOD model

TABLE I: The Performance of Benchmark Models

Model	backbone	Parameter	MultiWOZ 2.0				MultiWOZ 2.1			
			Inform	Success	BLEU	Combined	Inform	Success	BLEU	Combined
SimpleTOD [25]	DistilGPT2	82M	84.40	70.10	15.01	92.26	85.00	70.50	15.23	92.98
SOLOIST [26]	GPT2	124M	85.50	72.90	16.54	95.74	–	–	–	–
BORT [27]	T5-small	60M	93.80	85.80	18.50	108.30	–	–	–	–
MTTOD [28]	T5-base	220M	91.00	82.60	21.60	108.30	91.00	82.10	21.00	107.50
MinTL [29]	BART-large	400M	84.88	74.91	17.89	97.79	–	–	–	–
UBAR [23]	DistilGPT2	82M	95.40	80.70	17.00	105.05	95.70	81.80	16.50	105.25
GALAXY [8]	UniLM-base	110M	93.10	81.00	18.44	105.49	93.50	81.70	18.32	105.92
RewardNet [9]	BART-large	400M	92.63	84.32	18.35	106.83	–	–	–	–
MDTOD [6]	T5-base	220M	92.70	85.00	19.72	108.57	92.70	84.60	19.03	107.68
Mars [10]	T5-small	60M	94.50	85.30	19.07	108.97	–	–	–	–
OSTOD [7]	T5-small	60M	95.00	88.20	34.97	126.57	96.00	90.60	35.36	128.66
InstructTODS [30]	GPT-4	–	–	–	–	–	90.70	76.20	3.94	87.39
Ours	T5-small	60M	**95.50**	**89.40**	**35.27**	**127.72**	**97.40**	**91.60**	**35.99**	**130.49**

Bold indicates the model's best result on the dataset, while a hyphen (-) denotes that the model did not report results for the corresponding dataset/metric.

TABLE II: Results of the Few-Shot Experiment

Model	5%				10%				20%				50%			
	Inform	Success	BLEU	Combined	Inform	Success	BLEU	Combined	Inform	Success	BLEU	Combined	Inform	Success	BLEU	Combined
SOLIST [26]	69.30	52.30	11.80	72.60	69.90	51.90	14.60	75.50	74.00	60.10	15.24	82.29	–	–	–	–
MinTL [29]	75.48	60.96	13.98	82.2	78.08	66.87	15.46	87.93	82.48	68.57	13.00	88.53	90.10	78.60	17.90	102.25
PPTOD [31]	79.86	63.48	14.89	86.56	84.42	68.36	15.57	91.96	84.94	71.70	17.01	95.33	–	–	–	–
UBAR [23]	73.04	60.28	16.03	82.69	79.20	68.70	16.09	90.04	82.50	66.60	17.72	92.27	91.35	78.20	17.05	101.90
GALAXY [8]	80.59	67.43	17.39	91.40	87.00	75.00	17.65	98.65	89.55	75.85	17.54	100.24	93.35	82.35	18.37	106.22
RewardNet [9]	82.90	69.61	14.26	90.52	89.67	77.48	14.80	98.38	90.15	78.70	15.81	100.24	–	–	–	–
MDTOD [6]	85.65	62.20	15.24	89.17	86.30	71.50	14.47	93.37	90.25	80.90	16.40	101.98	–	–	–	–
OSTOD [7]	80.00	52.90	11.31	77.76	**93.40**	83.30	13.49	101.84	**94.90**	**86.10**	14.36	104.86	95.10	88.50	14.22	106.02
Ours	**89.10**	**80.40**	**28.05**	**112.80**	93.10	**83.90**	**31.40**	**119.9**	93.40	86.00	**32.85**	**122.55**	**95.30**	**89.00**	**34.17**	**126.52**

Bold indicates the model's best result on the dataset, while a hyphen (-) denotes that the model did not report results for the corresponding dataset/metric.

TABLE III: Results of the Ablation Experiments on the IPCMC-TOD Model

Ablation Experiments	Inform	Success	BLEU	Combine
IPCMC-TOD	**95.50**	**89.40**	**35.27**	**127.72**
IPCMC-TOD$_{\text{w/o CMC}}$	95.35	89.10	35.16	127.39
IPCMC-TOD$_{\text{w/o IP}}$	95.30	88.80	35.11	127.16

From Fig. 3, as the number of generated responses (Top-K) increases, the IPCMC-TOD model exhibits a rise-then-fall trend on the Inform, Success, and BLEU metrics. Specifically, as shown in Fig. 3a, performance peaks at Top-K=5 on MultiWOZ 2.0, while for MultiWOZ 2.1, Fig. 3b indicates that the peak occurs at Top-K=4. These results indicate that moderately expanding the candidate set can enhance response quality and improve information coverage. However, overly large Top-K tends to introduce irrelevant or redundant content, creating challenges for semantic alignment and contextual consistency.

G. Case Study

To directly evaluate the performance of the IPCMC-TOD model in few-shot dialogue generation. we conduct a case analysis based on the dialogue example pmul0550 from the MultiWOZ 2.0 dataset. The detailed results are presented in TABLE IV.

TABLE IV: Case Study of IPCMC-TOD on the MultiWOZ 2.0 Dataset

User	Hey, I am looking for a train from Cambridge to bishop's stortford, mind helping out?
System response	There are many trains from Cambridge to bishop's stortford, what time would you like to leave?
User	I want to leave on monday and arrive by 18:45.
System response (original)	I've already found a suitable train for you: TR4283.Would you like to make reservations?
System response (retell)	Train TR4283 leaves Cambridge on monday at 5:29 and arrives in bishop's stortford at 18:07, would you like me to book it for you?
IPCMC-TOD	Train TR4283 departs Cambridge at 5:29 on monday and arrives in Bishop's Stortford at 18:07. Shall I book it for you?
System response1	There are rooms available at 18:45 on monday, what time would you like to check in?
System response2	Would you like me to reserve a table for 18:45?
System response3	Down Hall Hotel has availability at 18:45 on monday, would you like me to book it for you?
System response4	Train TR4283 departs Cambridge at 5:29 on monday and arrives in Bishop's Stortford at 18:07. Shall I book it for you?
System response5	Wishing you a smooth journey.

In this case, the user first requests a train inquiry from

Cambridge to Bishop's Stortford. The IPCMC-TOD model performs inverse prompt response reconstruction. This mechanism effectively integrates the dialogue state information of the current train domain, including departure day such as Monday and time constraints like no later than 18:45. It thereby constructs a reconstructed system response that contains specific train details, including train number TR4283, departure time 5:29, and arrival time 18:07. Furthermore, the IPCMC-TOD model leverages the contextual momentum contrast response selection module to distinguish fine-grained semantic information among five candidate system responses. It ultimately selects System 4, which demonstrates the highest contextual semantic consistency, as the final output. This response correctly associates the user's query intent with the specific train information. In contrast, the other candidates mistakenly introduce irrelevant domain content, such as hotel or restaurant reservations, revealing their bias in contextual understanding.

V. Conclusion

We propose the IPCMC-TOD, a few-shot dialogue generation model with inverse prompt response reconstruction and contextual momentum contrast, to address the limitations of domain dialogue state perception and the inconsistency of system response generation in few-shot task-oriented dialogue. Specifically, the inverse prompt-based response reconstruction mechanism enhances the model's capability of understanding domain dialogue state. Furthermore, the contextual momentum contrast strategy is introduced to strengthen contextual semantic consistency. On both MultiWOZ 2.0 and MultiWOZ 2.1 datasets, the IPCMC-TOD model establishes state-of-the-art results on Inform, Success, BLEU, and Combined metric, with a pronounced advantage in low-resource conditions. In summary, this work provides a novel methodological perspective for few-shot dialogue generation.

References

[1] P. Budzianowski, T.-H. Wen, B.-H. Tseng, I. Casanueva, S. Ultes, O. Ramadan, and M. Gašić, "Multiwoz-a large-scale multi-domain wizard-of-oz dataset for task-oriented dialogue modelling," pp. 5016–5026, 2018.

[2] L. Liu, X. Li, and Y. Feng, "Ta&at: enhancing task-oriented dialog with turn-level auxiliary tasks and action-tree based scheduled sampling," vol. 38, no. 17, pp. 18 671–18 679, 2024.

[3] H. Sun, J. Bao, Y. Wu, and X. He, "Comet: Dialog context fusion mechanism for end-to-end task-oriented dialog with multi-task learning," pp. 10 541–10 553, 2025.

[4] Y. Lee, "Improving end-to-end task-oriented dialog system with a simple auxiliary task," pp. 1296–1303, 2021.

[5] M. Yang, S.-K. Ng, and J. Fu, "Omnidialog: An omnipotent pre-training model for task-oriented dialogue system," *IEEE Transactions on Audio, Speech and Language Processing*, 2025.

[6] S. Li, X. Zhang, Y. Zheng, L. Li, and X. Qiu, "Multijugate dual learning for low-resource task-oriented dialogue system," pp. 11 037–11 053, 2023.

[7] H. Huang, P. Yang, W. Wei, S. Shi, and X.-L. Mao, "Ostod: One-step task-oriented dialogue with activated state and retelling response," *Knowledge-Based Systems*, vol. 293, p. 111677, 2024.

[8] W. He, Y. Dai, Y. Zheng, and Wu, "Galaxy: A generative pre-trained model for task-oriented dialog with semi-supervised learning and explicit policy injection," vol. 36, no. 10, pp. 10 749–10 757, 2022.

[9] Y. Feng, S. Yang, S. Zhang, J. Zhang, C. Xiong, M. Zhou, and H. Wang, "Fantastic rewards and how to tame them: A case study on reward learning for task-oriented dialogue systems," *arXiv preprint arXiv:2302.10342*, 2023.

[10] H. Sun, J. Bao, Y. Wu, and X. He, "Mars: Modeling context & state representations with contrastive learning for end-to-end task-oriented dialog," pp. 11 139–11 160, 2023.

[11] J. Wang, J. Qu, K. Wang, Z. Li, W. Hua, X. Li, and A. Liu, "Improving the robustness of knowledge-grounded dialogue via contrastive learning," in *AAAI*, vol. 38, no. 17, 2024, pp. 19 135–19 143.

[12] K. Gao, T. Wang, and Ma, "Winnie: task-oriented dialog system with structure-aware contrastive learning and enhanced policy planning," in *AAAI*, vol. 38, no. 16, 2024, pp. 18 021–18 029.

[13] Y. Huangfu, P. Li, Y. Fan, and Q. Zhu, "Non-emotion-centric empathetic dialogue generation," pp. 989–999, 2025.

[14] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," pp. 9729–9738, 2020.

[15] J. Bai, Z. Yan, S. Zhang, J. Yang, H. Guo, and Z. Li, "Infusing internalized knowledge of language models into hybrid prompts for knowledgeable dialogue generation," *Knowledge-Based Systems*, vol. 296, p. 111874, 2024.

[16] Z. Huang, P. Liu, G. de Melo, L. He, and L. Wang, "Generating persona-aware empathetic responses with retrieval-augmented prompt learning," pp. 12 441–12 445, 2024.

[17] X. Zhang, W. Wei, and S. Zou, "Modal feature optimization network with prompt for multimodal sentiment analysis," in *Computational Linguistics*, 2025, pp. 4611–4621.

[18] Y. He, Y. Zhang, S. He, and J. Wan, "Bp4er: Bootstrap prompting for explicit reasoning in medical dialogue generation," in *LREC-COLING*, 2024, pp. 2480–2492.

[19] M. Gu, Y. Yang, C. Chen, and Z. Yu, "State value generation with prompt learning and self-training for low-resource dialogue state tracking," pp. 390–405, 2024.

[20] C. Raffel, N. Shazeer, A. Roberts, and Lee, "Exploring the limits of transfer learning with a unified text-to-text transformer," *Journal of machine learning research*, vol. 21, no. 140, pp. 1–67, 2020.

[21] S. Hu, I. Vulić, F. Liu, and A. Korhonen, "Reranking overgenerated responses for end-to-end task-oriented dialogue systems," pp. 13 970–13 991, 2024.

[22] M. Eric, R. Goel, S. Paul, A. Sethi, S. Agarwal, S. Gao, A. Kumar, A. Goyal, P. Ku, and D. Hakkani-Tur, "Multiwoz 2.1: A consolidated multi-domain dialogue dataset with state corrections and state tracking baselines," pp. 422–428, 2020.

[23] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks," pp. 3982–3992, 2019.

[24] J. D. M.-W. C. Kenton and L. K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," in *NAACL-HLT*, vol. 1. Minneapolis, Minnesota, 2019, pp. 4171–4286.

[25] E. Hosseini-Asl, B. McCann, C.-S. Wu, S. Yavuz, and R. Socher, "A simple language model for task-oriented dialogue," vol. 33, 2020, pp. 20 179–20 191.

[26] B. Peng, C. Li, J. Li, S. Shayandeh, L. Liden, and J. Gao, "Soloist: Building task bots at scale with transfer learning and machine teaching," vol. 9. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ..., 2021, pp. 807–824.

[27] Y. W. X. H. H. Sun, J. Bao, "Bort: Back and denoising reconstruction for end-to-end task-oriented dialog," in *NAACL*, 2022, pp. 2156–2170.

[28] Y. Lee, "Improving end-to-end task-oriented dialog system with a simple auxiliary task," in *EMNLP*, 2021, pp. 1296–1303.

[29] Y. Yang, Y. Li, and X. Quan, "Ubar: Towards fully end-to-end task-oriented dialog system with gpt-2," vol. 35, no. 16, pp. 14 230–14 238, 2021.

[30] W. Chung, S. Cahyawijaya, B. Wilie, and Lovenia, "Instructtods: Large language models for end-to-end task-oriented dialogue systems," in *Natural Language Interfaces*, 2023, pp. 1–21.

[31] Y. Su, L. Shu, E. Mansimov, A. Gupta, D. Cai, Y.-A. Lai, and Y. Zhang, "Multi-task pre-training for plug-and-play task-oriented dialogue system," in *ACL*, 2022, pp. 4661–4676.