What are the Generator Preferences for End-to-end Task-Oriented Dialog System?

Wanshi Xu, Xianwei Zhuang, Zhanpeng Chen, Zhihong Zhu, Xuxin Cheng, Yuexian Zou*
ADSPLAB, School of ECE, Peking University, China
{xwanshi, xwzhuang, troychen927, zhihongzhu, chengxx}@stu.pku.edu.cn
zouyx@pku.edu.cn

Abstract

Fully end-to-end task-oriented dialogue (EToD) systems have shown excellent performance, which requires the ability to retrieve entities accurately for generation. Existing methods improve the accuracy of entity retrieval and construct data flows between retrieval results and response generation, achieving promising results. However, most of them suffer from the following issues: (i) The entity is retrieved by directly interacting with the context at a coarse-grained level, so the similarity score may be disturbed by irrelevant attributes; (ii) The generator pays equal attention to retrieved entities and the context and does not learn the generation preferences for the current turn. In this paper, we propose a framework called **R**egulating **P**references of **G**enerator (RPG) based on retrieval results, which includes a generator preference extractor, an entity retriever, and a generator with the gate-controlled preference regulator. The generator preference extractor not only improves the entity retriever by filtering the interference of irrelevant attributes but also provides more focused guidance to the generator by performing inter-turn attribute prediction. Experiments and analyses on three standard benchmarks show that our framework outperforms existing methods and improves the quality of the dialogue.

1 Introduction

Task-oriented dialogue systems are designed to help users complete certain specific tasks, such as table booking, hotel booking, ticket booking, and online shopping (Zhang et al., 2020). Traditional task-oriented dialogue systems include several subtasks, such as dialogue state tracking (Kenton and Toutanova, 2019; Wu et al., 2019a), dialogue strategies (Takanobu et al., 2019), and natural language generation (Wen et al.).

*Corresponding author.

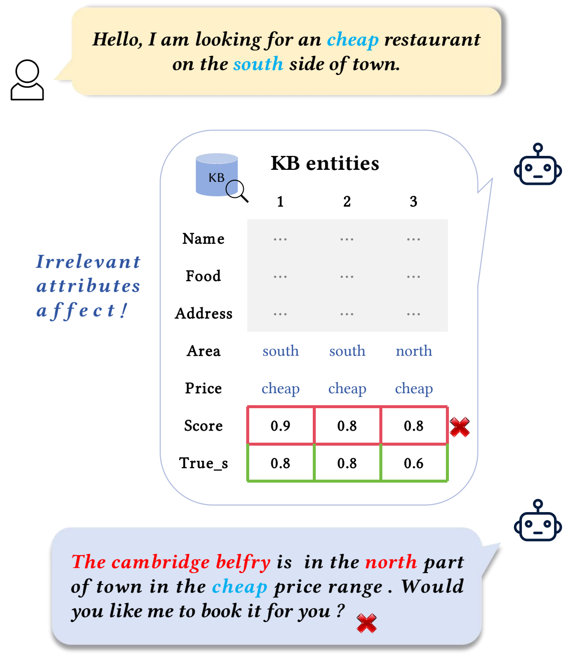


Figure 1: Visualisation of the impact of irrelevant attributes on entity retrieval. Grey attributes indicate attributes with low contextual similarity scores, which not only interfere with entity retrieval but also cause the generator's response to deviate from the user's intent, *e.g.*, the system is more likely to move on to the next query if more than one entity meets the requirements.

These modules require annotations to train and need to generate belief states as intermediate results for the generator to select entities in knowledge database (KB) and generate response. However, fully end-to-end task-oriented dialog system directly encodes KB and uses a neural network to query the KB in a differentiable manner, and it can directly generate system response given only dialogue history and the corresponding KB.

Due to advantages such as less dependency on additional annotations, current works increasingly focus on optimizing a fully end-to-end dialogue system. Based on the relationship between the knowledge retriever and the response generator, current EToD systems can be classified into

Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, pages 10992–11003
November 12-16, 2024 ©2024 Association for Computational Linguistics

two categories. The first one usually integrates the process of knowledge retrieval and response generation and is trained under the supervision of reference responses (Madotto et al., 2018; Qin et al., 2020; Raghu et al., 2021; Xie et al., 2022; Wu et al., 2022; Madotto et al., 2020; Huang et al., 2022). Still, this training method relies on the weak supervision signal returned by the generator output to supervise the retrieval process. The second category separates knowledge retrieval from response generation. It cleverly extracts supervisory signals from response generation to improve the retrieval process, alleviating the above problems to some extent (Tian et al., 2022; Wan et al., 2023; Shi et al., 2023; Shen et al., 2023).

Nevertheless, these methods only employ rough dot product interactions to compute the similarity between the dialogue context and the entities in KB, which is prone to interference from irrelevant attributes. It also leads to the generator not being able to obtain the correct signal. As shown in Fig. 1, multiple entities meet the requirements, the generator should make further inquiries, such as "*2 restaurants meet your needs. Do you have any other requirements?*" Such a query can greatly improve the success rate of conversations that end in a short turn, and improve the user experience. If the system replies directly to a particular entity, it may cause the user to lose interest. Therefore, even if the retrieval results are correct, whether the generator properly utilize the retrieval results to respond is also a problem we need to pay attention to. We need to judge whether the generator should give equal attention to the two components of the generator input in each dialogue turn and whether this weight signal can be extracted from the dialogue.

In this paper, we propose a framework called **R**egulating **P**references of **G**enerator (**RPG**), in which we first compute the attribute similarity scores of the context and entities for consecutive turns to obtain the *attribute mask matrix*, which is used to filter irrelevant attributes and perform relevant entity retrieval, and the *attribute shift matrix*, which is used to obtain the generator preference signal. This signal serves as a gating function for the generator preference regulator, which is responsible for synthesising specific representations. The specific representation is employed as input to the generator, which incorporates the weighted information of the top-K entities and the context in order to generate the target responses.

To the best of our knowledge, this is the first work to explore generator preference changes for multi-turn dialogues for the EToD system. Our main contributions can be summarised as: (i) We reduce the noise of irrelevant attributes on entity retrieval by obtaining fine-grained matching information between the context and the entity, and extracting the entity attribute shift matrix as a preference signal. (ii) We find that the generator should derive preferences from the search results instead of simply using them for generation, and design a generation preference regulator. (iii) The experimental results show that our system achieves state-of-the-art performance and abundant analyses demonstrate the validity of our method.

2 Method

Our system consists of three key components and implements a two-stage training framework, *cf.* Fig. 2. We first introduce the problem definition in Sec. 2.1, followed by the pre-trained encoder for generator preference extractor in Sec. 2.2, and the system training in Sec. 2.3, including the generator preference extractor, the entity retriever and the generator with the gate-controlled preference regulator.

2.1 Problem Definition

Given a dialog $\mathcal{D} = \{U_1, R_1, ..., U_T, R_T\}$ of T turns, where U_t and R_t are the t-th turn user utterance and system response, respectively. We use C_t to represent the dialog context of the t-th turn, where $C_t = \{U_1, R_1, ..., U_{t-1}, R_{t-1}, U_t\}$. An external knowledge database is provided as a set of entities, *i.e.*, $\mathcal{K} = \{E_1, E_2, ..., E_N\}$, where each entity E_i is composed of M attribute-value pairs, *i.e.*, $E_i = \{a^1, v_i^1, ..., a^N, v_i^M\}$. The goal of an end-to-end task-oriented dialogue system model is to learn a mapping that takes the dialogue context C_t and knowledge database $\mathcal{K}$ as input and generates an information response R_t.

$$R_t = f(C_t, \mathcal{K}). \tag{1}$$

2.2 Pre-trained Encoder

To increase the discrimination between different entities, we pre-train an encoder Enc_a, whose goal is to accurately predict the similarity scores between dialogue context and different attributes of the entity for a given turn. Unlike the attribute retriever in previous work (Wan et al., 2023),

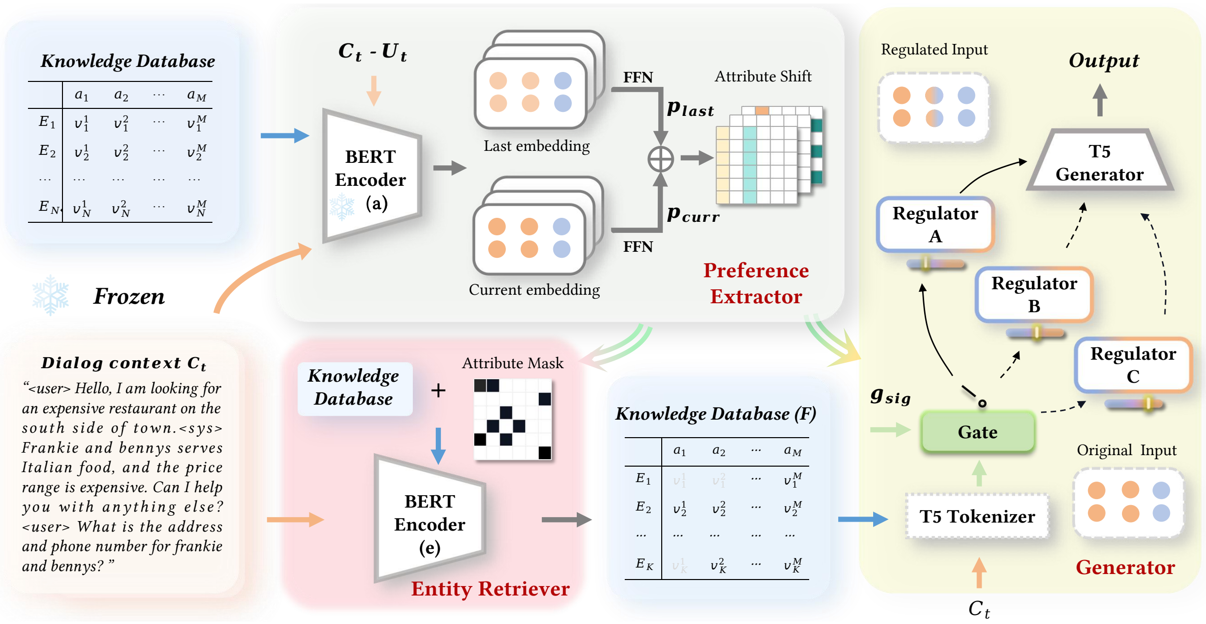


Figure 2: Overview of the proposed framework, which consists of a generator preference extractor, an entity retriever, and a response generator. The generator preference extractor is used to extract preference signals for the generator, and the generator is integrated with gate-controlled preference regulator.

our network does not predict the same attribute distributions for all relevant entities but focuses on exploiting the differences in attribute distributions between entities to increase the robustness of retrieval for false positive entities. It is mainly used to later infer preferences between turns.

Given E_i and C_t as input of the encoder, we can obtain the attribute score ${s_{t,i}}^a$ as:

$$ {s_{t,i}}^a = \mathrm{FFN}(\mathrm{Enc}_a([C_t; E_i])). \tag{2} $$

We use a multi-label classification task to train the model. The pseudo label is an N-dimensional vector $o_{t,i}$ constructed by checking whether any value of an attribute in E_i appears in the dialogue context C_t or system response R_t. A binary cross-entropy loss $\mathcal{L}_a$ for this task is:

$$ \mathcal{L}_a = \mathrm{BCELoss}({s_{t,i}}^a, o_{t,i}). \tag{3} $$

This model is used and frozen in the generator preference extractor in the system training.

2.3 System Training

Generator Preference Extractor. Given a set of candidate entities $\mathcal{K} = \{E_1, E_2, ..., E_N\}$, and user context C_t, we feed them into the generator preference extractor, we first obtain $C_t - U_t$ as the last dialogue history ${C_t}^{last}$ relative to the current turn. If the current is the first turn, ${C_t}^{last}$ is set to 0. Therefore, we can obtain the scores of two groups of entity attributes $S^a_t = \{{s_{t,1}}^a, ..., s^a_{t,N}\}$ and $S^a_{last_t} = \{{s_{last_t,1}}^a, ..., s^a_{last_t,N}\}$ by Eq. (2), which both contain N M-dimensional vectors. By setting attribute values greater than a given threshold to 1 and less than that to 0, we can get two 0-1 distributions p^t_{curr} and p^t_{last}, which denote the entity attribute distribution of the current dialogue and the entity attribute distribution of the previous turn of the dialogue, respectively.

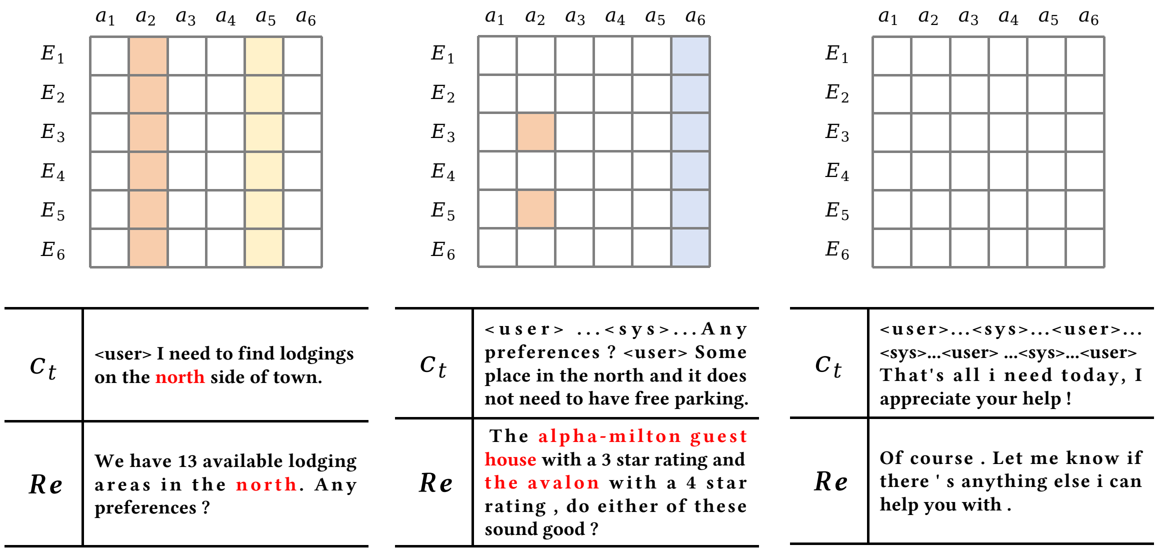
$$ {p_{curr,ij}}^t = \begin{cases} 1 & \text{if} {s_{t,ij}}^a > \theta, \\ 0 & \text{if} {s_{t,ij}}^a \leq \theta, \end{cases} \tag{4} $$

where θ denotes the threshold. Then, p_{curr} can be fed into the entity retriever along with the candidate entities as the attribute mask matrix. In addition, to represent the user's desired response preferences, we compute an attribute shift matrix p_{shift}:

$$ p^t_{shift} = p^t_{curr} - p^t_{last}. \tag{5} $$

Based on the characteristics and the statistical patterns, we obtain three types of attribute shift matrix, *cf.* Fig. 3:

- No targeted entities, but attribute: This is usually the case at the beginning of a dialogue, when the user makes a rough request, and the system usually needs to launch a further query.
- Targeted entities: When the user has asked for an obvious entity or an obvious attribute in the current turn, the system usually needs to answer based on the retrieved entity.



(a) No targeted entities, but attribute (b) Targeted entities (c) No entiy

Figure 3: Three typical attribute shift matrixes corresponding to the context preferred, entity preferred, and both preferred cases.

- No entity: This situation usually occurs at the end of a long dialogue, when we should prefer the model to focus on the context, and the user is not requesting new information about the entity and its attributes in the current turn.

After analyzing the typical attribute shift matrices, we find that the characteristics of the attribute shift matrix can be used as a reference for the generator preference signal. Progressively, the different characteristics of the three types can be reflected in their variance: When the variance is 0, it is obvious that the system does not point to any entity in particular and should output contextually relevant inheritance explanations or further queries, *etc.*, and vice versa.

Thus, we extract the variance of the attribute shift matrix and sum over the attribute dimensions to obtain the preference regulation signal of the generator gen^t_{sig}:

$$gen^t_{sig} = (std(p^t_{shift}), sum(p^t_{shift})). \quad (6)$$

Entity Retriever. The attribute mask matrix p^t_{curr} from the generator preference extractor, candidate entities, and contexts are fed into the entity encoder, and the top-K entities are selected by performing the cosine similarity score. The retrieval results of entities are able to dispense with the irrelevant attributes since the output entity embedding is obtained from the attribute mask matrix:

$$s^e_{t,i} = \text{Enc}(C_t)^T \text{Enc}(E_i, p^t_{curr}), \quad (7)$$

$$\mathcal{E}_t = \text{Top}K(s^e_{t,i}) = \{\hat{E_1}, ..., \hat{E_K}\}, \quad (8)$$

where $\hat{E_i}$ represents the entity with attribute attention mask.

Response Generator with Preference. Inspired by the Mixture-of-Experts (MoE) (Zhou et al., 2022), we set up multiple regulators to learn the preferred representations of the generator according to the signals obtained from the generator preference extractor. Unlike MoE, we do not need to train the gate network to obtain the expert weights. Instead, we directly map a finite range of discrete signals to activate the appropriate regulator and obtain the preferred representations through a learnable MLP. We first obtain the original input token H_t as formulated:

$$H_t = \text{Tokenizer}_g([C_t; \hat{E_1}; ...; \hat{E_K}]), \quad (9)$$

where Tokenizer_g represents the tokenizer of the response generator. We feed H_t into the activated regulator to get the preferred representation h_t:

$$\mathcal{R}_\theta = \Psi_\theta(\mathbb{1}(gen_{sig})), \; h_t = \mathcal{R}_\theta(H_t), \quad (10)$$

where Ψ_θ represents the gating function, which controls which regulator to activate, while $\mathcal{R}_\theta$ is the regulator chosen for obtaining h_t.

Taking h_t as input, the generator produces the response token by token. The probability distribution for each response token in R_t is defined as:

$$P(R_{t,i}) = \text{Gen}(R_{t,i}|R_{t,<i}, h_t). \quad (11)$$

Model	MWOZ		SMD		CamRest	
	BLEU	Entity F1	BLEU	Entity F1	BLEU	Entity F1
DF-Net (Qin et al., 2020)	9.40	35.10	14.40	62.70	-	-
GPT-2+KE (Madotto et al., 2020)	15.05	39.58	17.35	59.78	18.00	54.85
EER (He et al., 2020b)	13.60§	35.60§	17.20§	59.00§	19.20§	65.70§
FG2Seq (He et al., 2020a)	14.60§	36.50§	16.80§	61.10§	20.20§	66.40§
CDNET (Raghu et al., 2021)	11.90	38.70	17.80	62.90	21.80	68.60
GraphMemDialog (Wu et al., 2022)	14.90	40.20	18.80	64.50	22.30	64.40
ECO (Huang et al., 2022)	12.61	40.87	-	-	18.42	71.56
DialoKG (Rony et al., 2022)	12.60	43.50	20.00	65.90	23.40	**75.60**
T5-Base						
UnifiedSKG (Xie et al., 2022)	-	-	17.41	66.45	-	-
Q-TOD (Tian et al., 2022)	-	-	20.14	68.22	-	-
DF-TOD (Shi et al., 2023)	**18.26**	52.52	24.12	69.36	25.85	72.83
MK-TOD (Shen et al., 2023)	17.33	51.86	24.77	67.86	<u>26.76</u>	73.60
MAKER (Wan et al., 2023)	17.23	<u>53.68</u>	<u>24.79</u>	<u>69.79</u>	25.04	73.09
RPG(Ours)	<u>18.23</u>	**54.95**	**25.60**	**70.09**	**26.96**	<u>74.89</u>
T5-Large						
UnifiedSKG (Xie et al., 2022)	13.69*	46.04*	17.27	65.85	20.31*	71.03*
Q-TOD (Tian et al., 2022)	17.62	50.61	21.33	71.11	23.75	74.22
DF-TOD (Shi et al., 2023)	18.48	53.17	25.10	71.58	26.00	74.04
MK-TOD (Shen et al., 2023)	17.55	52.97	25.43	**73.31**	<u>26.20</u>	71.72
MAKER (Wan et al., 2023)	<u>18.77</u>	<u>54.72</u>	**25.91**	71.30	25.53	<u>74.36</u>
RPG(Ours)	**18.99**	**55.20**	<u>25.51</u>	<u>72.39</u>	**26.73**	**75.86**

Table 1: Overall results of EToD systems with dialog-level knowledge databases on MWOZ, SMD, and CamRest. The best scores are highlighted in bold, and the second-best scores are underlined. †, ‡, §, $*$ indicates that the results are cited from (Qin et al., 2019), (Qin et al., 2020), (Raghu et al., 2021), and (Tian et al., 2022), respectively.

We train the response generator by the standard cross-entropy loss as:

$$\mathcal{L}_{gen} = \sum_{i=1}^{|R_t|} -\log P(R_{t,i}), \tag{12}$$

where $|R_t|$ denotes the length of R_t.

3 Experiments

3.1 Dataset and Metric

We evaluate our system on three well-recognized datasets for multi-turn task-oriented dialogue system: MultiWOZ 2.1(MWOZ) (Eric et al., 2020), Stanford Multi-Domain (SMD) (Eric et al., 2017), and CamRest (Wen et al., 2017). Each dialogue in these datasets has a corresponding knowledge database that contains all entities that match the user goals for the dialogue. The statistics of the datasets and implementation details are presented in the Appendix.

Following previous work, BLEU (Papineni et al., 2002) and Entity F1 (Eric et al., 2017) are used as the evaluation metrics. BLEU measures the fluency of the generated response based on its n-gram overlaps with the gold response. Entity F1 measures whether the generated response contains correct knowledge by micro-averaging the precision and recall scores of attribute values in the generated response.

3.2 Main Results

We perform two experimental settings to obtain the main results. (i) **Dialog-level KB:** The database corresponding to the dialogues is the default setting of previous works, which usually includes a small number of correct entities. (ii) **Dataset-level KB:** Create the knowledge database by collecting all the entities contained in each domain of the dataset.

The results of our framework RPG and other competitive baselines for the MWOZ, SMD, and CamRest are shown in Tab. 1 and Tab. 2 respectively, from which we have the following observations: (i) It is obvious that our proposed RPG outperforms other baselines in almost all settings and datasets, mainly due to its higher entity retrieval accuracy and fine-grained guidance of the generator. (ii) In both dialog-level KB and dataset-level KB, our framework has a relatively stable improvement and further narrows the lift

Model	MWOZ		CamRest	
	BLEU	Entity F1	BLEU	Entity F1
FG2Seq	10.74	33.68	19.20	59.35
CDNET	10.90	31.40	16.50	63.60
Q-TOD	16.67	47.13	21.44	63.88
T5-Base				
MK-TOD	17.56	50.09	26.85	73.51
DK-TOD	**17.61**	51.61	27.39	70.74
MAKER	16.25	50.87	26.19	72.09
Ours	17.57	**51.97**	**28.09**	**74.07**
T5-Large				
DeepFM	17.40	53.26	**27.82**	71.98
GDCN-P	18.36	52.96	26.61	73.58
Ours	**18.57**	**54.02**	27.58	**74.93**

Table 2: Overall results of EToD systems with the total knowledge database on MWOZ and CamRest, respectively. The best scores are highlighted in bold, and the second-best scores are underlined.

Dataset	Entity accuracy	Attribute accuracy
MWOZ	92.1	98.7
CamRest	91.9	97.9

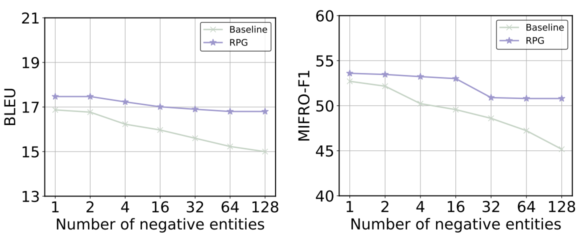
Table 3: Results of the pre-trained model for the generator preference extractor on MWOZ and CamRest. Entity accuracy refers to the proportion of all attributes in an entity that are predicted correctly; Attribute accuracy refers to the proportion of all attributes predicted correctly.

gap of the previous approaches in both database configurations. With the generator preference, the system enhances the utilize of accurate results under the condition of excellent retrieval performance and improves the performance of the generator.

4 Analysis

4.1 Generator Preference Extractor

Results of Pre-training Tasks. In order to reduce the retrieval noise caused by irrelevant attributes and to output the generated preference signals



(a) BLEU (b) Entity F1

Figure 4: Performance of different numbers of negative entities in BLEU (a) and Entity F1 (b) scores.

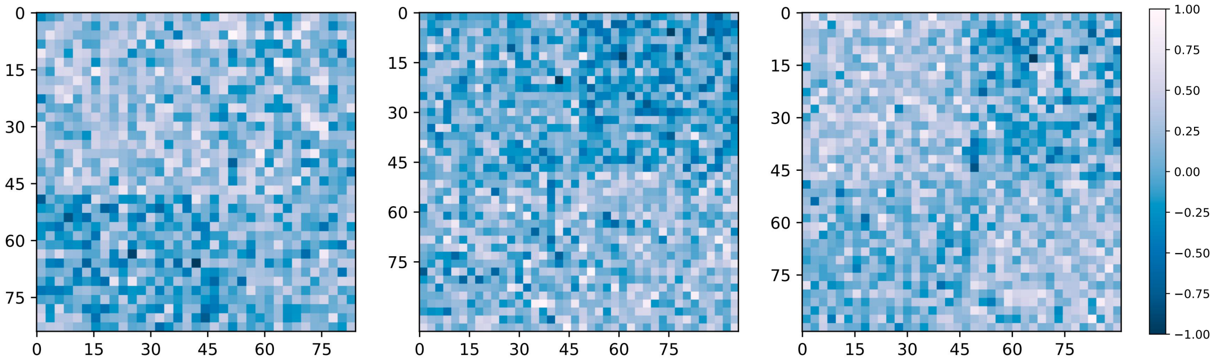
with accurate guidance, it is necessary to train the encoder in the generator preference extractor. We train it on two datasets MWOZ and CamRest, *cf.* Tab. 3.

Robustness to Noisy Entities. In order to compare the robustness of our proposed framework for noisy entity retrieval, we exponentially increase the number of negative entities in the database and test the robustness of the baseline model (Wan et al., 2023) versus our model when there are different numbers of erroneous entities in the knowledge database on two datasets, *cf.* Fig. 4. We find that the performance of our method is more stable in the presence of more noisy entities, without causing large fluctuations or degradation in comparison to the baseline.

Distribution of Attribute Shift Matrix. To statistically characterize the attribute shift matrices, we use the signals output from the generator preference extractor as low-dimensional features and visualize their distribution on both two datasets, referring to Appendix C. We find that the distribution of the attribute shift matrix in the space is in roughly three categories, which can be assigned to the three categories in Fig. 3. Therefore, it makes sense to set three regulators for the generator to re-represent the inputs.

4.2 Ablation Study

We conduct the ablation study using dialog-level and dataset-level knowledge databases on two datasets, *i.e.*, MWOZ (Eric et al., 2020) and CamRest (Wen et al., 2017), *cf.* Tab. 4. To verify that the attribute mask matrix obtained by the preference extractor facilitates entity retrieval by filtering irrelevant attributes, and thus improves the generation quality, we design the following ablations: (i) We cancel the attribute mask matrix fed into the entity retrieval to investigate whether irrelevant attributes affect the selection of entities (*w/o* attribute mask matrix). The experimental results demonstrate a significant performance degradation on two types of knowledge databases, with MWOZ exhibiting a BLEU score of $0.37/1.14$ and an entity F1 score of $1.16/1.79$, and CamRest exhibiting a BLEU score of $0.87/0.15$ and an entity F1 score of $1.07/1.70$. (ii) We further investigate the impact of removing the regulator from the system (*w/o* regulator). The results demonstrate that the gate-controlled regulators play an important role in the overall system. The introduction of regulators and the gating

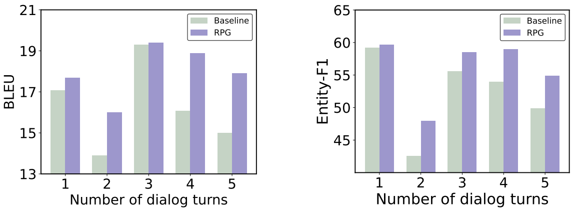


(a) Context Preferred (b) Entity Preferred (c) Both Preferred

Figure 5: Visualisation of the similarity of the inputs and outputs of the regulator. We randomly selected several sets of input and output samples from the three regulators and processed them to obtain the average cosine similarity. Lighter colors represent higher similarity.

Method	MWOZ		CamRest	
	BLEU	Entity F1	BLEU	Entity F1
Ours$_{\text{dialog-level}}$	18.23	54.95	26.73	75.86
w/o mask	17.86 (↓ **0.37**)	53.79 (↓ **1.16**)	25.86 (↓ **0.87**)	74.79 (↓ **1.07**)
w/o regulator	17.59 (↓ **0.64**)	53.41 (↓ **1.54**)	24.99 (↓ **1.74**)	73.99 (↓ **1.87**)
Ours$_{\text{dataset-level}}$	18.57	54.02	27.58	74.93
w/o mask	17.43 (↓ **1.14**)	52.23 (↓ **1.79**)	27.43 (↓ **0.15**)	73.23 (↓ **1.70**)
w/o regulator	17.59 (↓ **0.98**)	53.01 (↓ **1.01**)	26.59 (↓ **0.99**)	73.01 (↓ **1.92**)

Table 4: Results of ablation study on MWOZ and CamRest with T5-large, where "w/o" means without, "mask" denotes attribute mask matrix for entity retrieval.



(a) BLEU (b) Entity F1

Figure 6: Performance of the different number of dialog turns in BLEU (a) and Entity F1 (b) scores.

mechanism not only enables the generator to make more efficient use of the retrieval results, but also allows for a better differentiation of the user's response preferences. Furthermore, in case of long dialogue turns, the truncation of the generator is directly detrimental to the system. The proposed method offers a promising solution to this issue. Quantitatively, the removal of the regulator also results in a significant decline in performance. The MWOZ exhibits a BLEU score of $0.64/0.98$ and an entity F1 score of $1.54/1.01$, while CamRest exhibits a BLEU score of $1.74/0.99$ and an entity F1 score of $1.87/1.92$.

4.3 Analysis on Long-turn Dialog

We statistically compare our method to the baseline (Wan et al., 2023) in terms of dialogue length for different turns, *cf.* Fig. 6. We find that the improvement of our method becomes more significant as the number of dialogue turns increases, which may be due to the following three reasons: Firstly, our preference signal extractor accurately identifies the entity attribute information that is useful for the current dialogue and blocks out the interference of the noisy signals; Secondly, our generator preference regulator selectively focuses on the more important parts of the input according to the preference signal; Last but not least, the generator's inputs usually face the automatic truncation problem when the dialogue turns become longer, in which case the preferred representations we obtain by the regulator carries the preference information of the user's desired answers and avoid length truncation.

4.4 Visualisation of Regulator

To verify the ability of the generator preference regulators to learn the features of different types of preferences and to obtain the corresponding generator preferred representations, we visualize the cosine similarity of the inputs and outputs of the three regulators. The results obtained are shown in Fig. 5. The horizontal axis represents the index of the output token, while the vertical axis represents the index of the input token. The input token is a concatenation of context and entities, with the context preceding the entities. We observe that in the first sub-figure, the output of the regulator is more similar to the first half of the input, which indicates that the regulator will output a more contextualized representation. Similarly, in the second sub-graph, the output of the regulator is more similar to the second half of the input, which indicates that the regulator will output a more entity-biased representation.

5 Related Work

5.1 End-to-End Task-Oriented Dialog

End-to-end trainable methods combining external knowledge database as network input to generate replies in task-oriented dialogue systems have received increasing attention. Some studies have used memory networks to encode knowledge database and use attention weights between dialogue context and memory units to select KB records (Qin et al., 2019; Wu et al., 2019b; Raghu et al., 2021). Some studies explore the use of knowledge database and dialogue contexts in series and take them as input to pre-trained language models (Xie et al., 2022; Rony et al., 2022). Besides, some works stord the KB in the model parameters for implicit retrieval (Madotto et al., 2020; Huang et al., 2022). However, these approaches perform implicit entity retrieval during the training of generation, which leads to poor model interpretability and sub-optimal retrieval performance when provided with large-scale knowledge database.

Later approaches separate knowledge retrieval from the corresponding generation, and explicitly construct generator inputs from the retrieval results, which improve the retrieval process by extracting supervisory signals from response generation. Tian et al. (2022) extracts basic information from the dialogue context as a query and further uses it to retrieve relevant knowledge records for generating responses. Wan et al. (2023) introduces multi-level retrieval through entity and attribute selection. Shi et al. (2023) proposes a dual-feedback network to obtain supervised signals from the corresponding generators. Shen et al. (2023) proposes maximal marginal likelihood to train a perceptive retriever by utilizing signals from response generation for supervision and introduces meta knowledge to enhance the entity retrieval.

5.2 Dense Retrieval

Recently, dense retrieval (Karpukhin et al., 2020) has been widely used in first-stage retrieval that efficiently recalls candidate documents from the large corpus. One of the mainstream approaches to information retrieval is to build retrievers using dual-encoder architecture (Yih et al., 2011) to build. Gillick et al. (2019) employs the dual encoder architecture for separately encoding mentions and entities into high-dimensional vectors for entity retrieval. However, fine-grained information about entities in the modeling external database is often ignored through this coarse interaction metric calculation (Xu et al., 2023; Liu et al., 2023). Retrieval-Augmented Generation (RAG) is a technique that combines a retrieval system with a generator model to improve the accuracy and richness of generated content (Lewis et al., 2020). Although existing end-to-end task-oriented dialogue systems have followed this paradigm, they have explored less on how to balance the retrieval results with the user query.

6 Conclusion

In this work, we propose an RPG framework that effectively improves the performance of end-to-end task-oriented dialogue systems. RPG consists of two key enhancements. Firstly, it reduces the interference of irrelevant attributes on entity retrieval by obtaining fine-grained matching information between contexts and entities and extracting the entity attribute shift matrix as preference signals. Secondly, it employs multiple generator preference regulators to process the different types of dialogues. Our RPG focuses on the different attention that should be allocated to retrieval results and user input in different situations, providing a more fine-grained guidance for the generator. We hope that this exploration will encourage the community to think about the problems of generator preference.

Limitations

With the rapid growth of generation capacity, how to provide accurate and effective guidance for generator has become an important topic nowadays, and our work aims to guide the generator to generate results that are more in line with the user's expectations by exploring the balance between the retrieval results and user inputs. However, our work still has limitations. Firstly, the tuning method of the generator can be more diverse; Secondly, how to obtain preference representation and map it to a signal to the generator can be further optimized; Thirdly, in the current EToD systems, the evaluation metrics of preference responses are not well-developed, and we hope that our work will stimulate further exploration of this area by the NLP community.

Acknowledgement

We would like to thank all reviewers for their insightful comments. This paper was partially supported by NSFC (No: 62176008). Special acknowledgements are given to the AOTO-PKUSZ Joint Research Center for its support.

Ethics Statement

All the experimental procedures in this study have been performed exclusively on publicly accessible datasets that do not contain any sensitive or private information. The datasets used in our research are publicly available and do not infringe upon individual privacy rights. Our work strictly adheres to ethical guidelines and does not involve the analysis or consideration of personal identity characteristics, including but not limited to gender and race. We emphasize that our research is focused solely on the scientific aspects of the problem at hand and does not engage in any form of discriminatory practices or bias based on gender, race, or any other protected attributes. The primary objective of our study is to contribute to the advancement of knowledge and technology in a fair, unbiased, and inclusive manner.

References

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

Mihail Eric, Rahul Goel, Shachi Paul, Abhishek Sethi, Sanchit Agarwal, Shuyang Gao, Adarsh Kumar, Anuj Goyal, Peter Ku, and Dilek Hakkani-Tur. 2020. MultiWOZ 2.1: A consolidated multi-domain dialogue dataset with state corrections and state tracking baselines. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 422–428, Marseille, France. European Language Resources Association.

Mihail Eric, Lakshmi Krishnan, Francois 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*, pages 37–49, Saarbrücken, Germany. Association for Computational Linguistics.

Dan Gillick, Sayali Kulkarni, Larry Lansing, Alessandro Presta, Jason Baldridge, Eugene Ie, and Diego Garcia-Olano. 2019. Learning dense representations for entity retrieval. In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, pages 528–537.

Zhenhao He, Yuhong He, Qingyao Wu, and Jian Chen. 2020a. Fg2seq: Effectively encoding knowledge for end-to-end task-oriented dialog. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8029–8033. IEEE.

Zhenhao He, Jiachun Wang, and Jian Chen. 2020b. Task-oriented dialog generation with enhanced entity representation. In *INTERSPEECH*, pages 3905–3909.

Guanhuan Huang, Xiaojun Quan, and Qifan Wang. 2022. Autoregressive entity generation for end-to-end task-oriented dialog. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 323–332.

Vladimir Karpukhin, Barlas Oguz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih. 2020. Dense passage retrieval for open-domain question answering. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6769–6781.

Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of naacL-HLT*, volume 1, page 2.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33:9459–9474.

Yi Liu, Yuan Tian, Jianxun Lian, Xinlong Wang, Yanan Cao, Fang Fang, Wen Zhang, Haizhen Huang, Denvy Deng, and Qi Zhang. 2023. Towards better entity linking with multi-view enhanced distillation. *arXiv preprint arXiv:2305.17371*.

Ilya Loshchilov and Frank Hutter. 2018. Decoupled weight decay regularization. In *International Conference on Learning Representations*.

Andrea Madotto, Samuel Cahyawijaya, Genta Indra Winata, Yan Xu, Zihan Liu, Zhaojiang Lin, and Pascale Fung. 2020. Learning knowledge bases with parameters for task-oriented dialogue systems. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 2372–2394.

Andrea Madotto, Chien-Sheng Wu, and Pascale Fung. 2018. Mem2Seq: Effectively incorporating knowledge bases into end-to-end task-oriented dialog systems. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1468–1478, Melbourne, Australia. Association for Computational Linguistics.

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

Libo Qin, Yijia Liu, Wanxiang Che, Haoyang Wen, Yangming Li, and Ting Liu. 2019. Entity-consistent end-to-end task-oriented dialogue system with KB retriever. 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 133–142, Hong Kong, China. Association for Computational Linguistics.

Libo Qin, Xiao Xu, Wanxiang Che, Yue Zhang, and Ting Liu. 2020. Dynamic fusion network for multi-domain end-to-end task-oriented dialog. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6344–6354, Online. Association for Computational Linguistics.

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

Dinesh Raghu, Atishya Jain, Sachindra Joshi, et al. 2021. Constraint based knowledge base distillation in end-to-end task oriented dialogs. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 5051–5061.

Md Rashad Al Hasan Rony, Ricardo Usbeck, and Jens Lehmann. 2022. Dialokg: Knowledge-structure aware task-oriented dialogue generation. *arXiv preprint arXiv:2204.09149*.

Vishal Vivek Saley, Rocktim Jyoti Das, Dinesh Raghu, et al. 2024. Synergizing in-context learning with hints for end-to-end task-oriented dialog systems. *arXiv preprint arXiv:2405.15585*.

Weizhou Shen, Yingqi Gao, Canbin Huang, Fanqi Wan, Xiaojun Quan, and Wei Bi. 2023. Retrieval-generation alignment for end-to-end task-oriented dialogue system. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 8261–8275.

Tianyuan Shi, Liangzhi Li, Zijian Lin, Tao Yang, Xiaojun Quan, and Qifan Wang. 2023. Dual-feedback knowledge retrieval for task-oriented dialogue systems. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 6566–6580.

Ryuichi Takanobu, Hanlin Zhu, and Minlie Huang. 2019. Guided dialog policy learning: Reward estimation for multi-domain task-oriented dialog. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 100–110.

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. *arXiv preprint arXiv:2210.07564*.

Fanqi Wan, Weizhou Shen, Ke Yang, Xiaojun Quan, and Wei Bi. 2023. Multi-grained knowledge retrieval for end-to-end task-oriented dialog. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 11196–11210, Toronto, Canada. Association for Computational Linguistics.

Tsung-Hsien Wen, Milica Gašic, Nikola Mrkšic, Pei-Hao Su, David Vandyke, and Steve Young. Semantically conditioned lstm-based natural language generation for spoken dialogue systems.

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

Chien-Sheng Wu, Andrea Madotto, Ehsan Hosseini-Asl, Caiming Xiong, Richard Socher, and Pascale Fung. 2019a. 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.

Chien-Sheng Wu, Richard Socher, and Caiming Xiong. 2019b. Global-to-local memory pointer networks for task-oriented dialogue. In *International Conference on Learning Representations*.

Jie Wu, Ian G Harris, and Hongzhi Zhao. 2022. Graphmemdialog: Optimizing end-to-end task-oriented dialog systems using graph memory networks.

Tianbao Xie, Chen Henry Wu, Peng Shi, Ruiqi Zhong, Torsten Scholak, Michihiro Yasunaga, Chien-Sheng Wu, Ming Zhong, Pengcheng Yin, Sida I Wang, et al. 2022. Unifiedskg: Unifying and multi-tasking structured knowledge grounding with text-to-text language models. *arXiv preprint arXiv:2201.05966*.

Shicheng Xu, Liang Pang, Huawei Shen, and Xueqi Cheng. 2023. Berm: Training the balanced and extractable representation for matching to improve generalization ability of dense retrieval. *arXiv preprint arXiv:2305.11052*.

Wen-tau Yih, Kristina Toutanova, John C Platt, and Christopher Meek. 2011. Learning discriminative projections for text similarity measures. In *Proceedings of the fifteenth conference on computational natural language learning*, pages 247–256.

Zheng Zhang, Ryuichi Takanobu, Qi Zhu, MinLie Huang, and XiaoYan Zhu. 2020. Recent advances and challenges in task-oriented dialog systems. *Science China Technological Sciences*, 63(10):2011–2027.

Yanqi Zhou, Tao Lei, Hanxiao Liu, Nan Du, Yanping Huang, Vincent Zhao, Andrew M Dai, Quoc V Le, James Laudon, et al. 2022. Mixture-of-experts with expert choice routing. *Advances in Neural Information Processing Systems*, 35:7103–7114.

A Statistics of Datasets

The statistics of the datasets are shown in Tab. 6.

B Implementation Details

We employ BERT (Devlin et al., 2019) as the encoder of our entity selector and attribute selector and employ T5 (Raffel et al., 2020) to implement the response generator. All these models are fine-tuned using AdamW optimizer (Loshchilov and Hutter, 2018) with a batch size of 64. We train these models for 15k gradient steps with a linear decay learning rate of 10^{-4}. We conduct all experiments on a single 24G NVIDIA RTX 3090 GPU and select the best checkpoint based on model performance on the validation set.

C Analysis

To statistically characterize the attribute shift matrices, we use the signals output from the generator preference extractor as low-dimensional features and visualize their distribution on both two datasets, referring to Fig. 7 and Fig. 8. The three cases are more uniformly distributed. This is intuitive because the characteristics of the data from multi-turn dialogues changes with the number of dialogue turns.

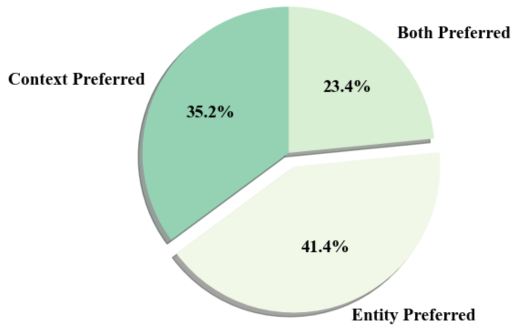


Figure 7: Percentage of distribution of different classes of attribute shift matrices in space on the MWOZ dataset.

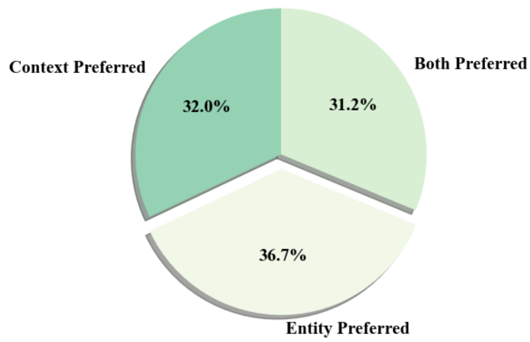


Figure 8: Percentage of distribution of different classes of attribute shift matrices in space on the CamRest dataset.

D Comparison with LLMs

We also compare our results with the baseline used LLMs(Saley et al., 2024), as shown in Tab. 5,

Model	MWOZ		SMD	
	BLEU	Entity F1	BLEU	Entity F1
Zero-shot (*ChatGPT*)	3.39	28.16	6.91	60.11
Few-shot (*ChatGPT*)	8.83	40.25	17.21	70.58
Few-shot (*GPT4*)	6.25	36.47	10.08	63.57
RAG (*ChatGPT*)	8.89	40.2	16.71	70.25
RAG (*GPT4*)	7.64	41.14	13.44	71.02
SyncTOD (*ChatGPT*)	14.33	52.99	22.08	71.60
SyncTOD (*GPT4*)	13.01	**54.99**	19.08	**72.99**
Ours (T5-Base)	<u>18.23</u>	<u>54.95</u>	**25.60**	**70.09**
Ours (T5-Large)	**18.99**	**55.20**	**25.51**	**72.39**

Table 5: Comparison with LLM-based Methods.

Dataset	Domains	#Dialogues		
		Train	Val	Test
MWOZ (Eric et al., 2020)	Restaurant, Attraction, Hotel	1839	117	141
SMD (Eric et al., 2017)	Navigate, Weather, Schedule	2425	302	304
CamRest (Wen et al., 2017)	Restaurant	406	135	135

Table 6: Statistics of the three datasets.

we see that although LLMs have strong ability of comprehension and generation, they still performs poorly without training.