Relevance Is a Guiding Light: Relevance-aware Adaptive Learning for End-to-end Task-oriented Dialogue System

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

Abstract

Retrieving accurate domain knowledge and providing helpful information are crucial in developing an effective end-to-end task-oriented dialogue system (E2ETOD). Existing approaches to this field follow a retrieve-then-generate paradigm and train their systems on one specific domain. However, existing approaches still suffer from the **Distractive Attributes Problem** (DAP): struggling to deal with false but similar knowledge (a.k.a hard negative entities), which is even more intractable when countless pieces of knowledge from different domains are blended in a real-world scenario. To alleviate DAP, we propose the **Re**levance-aware **A**daptive **L**earning (**ReAL**), a novel two-stage training framework that eliminates hard negatives step-by-step and aligns retrieval with generation. In the first stage, we introduce a top-k adaptive contrastive loss and utilize the divergence-driven feedback from the frozen generator to pre-train the retriever. In the second stage, we propose using the metric score distribution as an anchor to align retrieval with generation. Thorough experiments on three benchmark datasets demonstrate ReAL's superiority over existing methods, with extensive analysis validating its strong capabilities of overcoming in- and cross-domain distractions.

1 Introduction

The advent of an intelligent era has seen rapid advancement of task-oriented dialogue systems (TOD), demonstrating their potential and flexibility across various practical applications like health consulting, financial services, and home automation. Though conventional pipeline TOD systems model the task with interrelated modules and have achieved remarkable results, several limitations exist, such as error propagation (Qin et al., 2023), increased complexity (Goyal et al., 2022; Xin et al., 2021), and maintenance challenges (Qin et al., 2023), making the development and deployment of such systems resource-intensive and less scalable. Hence, end-to-end task-oriented dialogue systems (E2ETOD) have aroused increasing research attention for their ability to retrieve from external knowledge bases (KB) to generate the response end-to-end without any intermediate annotations.

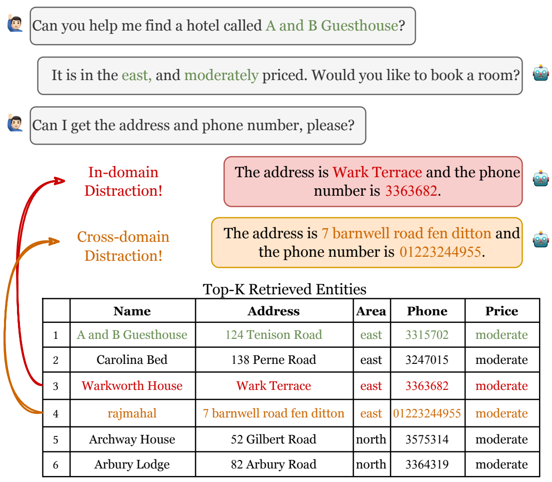


Figure 1: An example from MultiWOZ 2.1 (Eric et al., 2019). The top-k retrieved entities are severely distracted by in- and cross-domain similar but false entities, leading to inaccurate retrieval and retrieval-generation misalignment.

Retrieving domain knowledge precisely is a crucial part of E2ETOD (Wan et al., 2023; Shi et al., 2023a; Shen et al., 2023; Qin et al., 2019). Since knowledge bases are growing daily, old-fashioned methods combining knowledge retrieval and response generation into one single model (Madotto et al., 2018; Qin et al., 2020) are facing the issue of KB scalability. Drawing inspiration from retrieval-augmented generation (RAG) (Singh et al., 2021; Lewis et al., 2020; Guu et al., 2020), Q-TOD (Tian et al., 2022) employs a held-out retriever and decouples knowledge retrieval from response generation, which gets rid of the unscalable problem. Following the technical route, MAKER (Wan et al., 2023)

[*]Corresponding author.

Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, pages 5410–5420
November 12-16, 2024 ©2024 Association for Computational Linguistics

introduces a multi-grained knowledge retriever for large-scale cross-domain KBs. Shi et al. (2023a) propose a dual-feedback mechanism to provide the retriever with stronger supervision signals. Benefiting from the dual-encoder retriever architecture, these methods can train the retriever end-to-end and match semantics efficiently.

Despite these encouraging developments, the paradigm of decoupling knowledge retrieval from response generation can lead to **Distractive Attribute Problem** (DAP) in task-oriented dialogue systems: **(I) Inaccurate retrieval**: E2ETOD systems suffer from in- and cross-domain distractions while retrieving top-k entities from large-scale knowledge bases (Wan et al., 2023; Shi et al., 2023a). **(II) Retrieval-generation misalignment**: the top-k retrieved knowledge entities are similar in attributes and easily lead to retrieval-generation misalignment (Shen et al., 2023). **(III) Ambiguous retriever pre-training**: when determining the weakly labeled data, the common pre-training method (Qin et al., 2019) selects the entities with the most attribute value occurrences in the dialogue context and system response as pseudo-positive examples. However, these entities are possibly not the ground truth and lead to ambiguous retriever pre-training. As illustrated in Fig. 1, the ground truth entity is distracted by in- and cross-domain hard negatives during retrieval and generation.

In this paper, we propose a two-stage Relevance-aware Adaptive Learning (ReAL) framework to address three aspects of DAP jointly. In the first stage, we introduce a top-k adaptive contrastive loss and utilize the divergence-driven feedback from the frozen generator to mitigate ambiguity when pre-training the retriever. Instead of picking the entity with the most attribute value occurrences as the sole pseudo-positive, we construct top-k context-entity pairs and estimate the matching degree between the pairs. These relevance weights are assigned to different context-entity pairs to promote adaptive contrastive learning. In addition, to further avoid retrieving hard negatives, we measure the relevance between the entity and context with the generation probability of the frozen generator. Intuitively, the more relevant the entity and context are, the greater the generation probability. Thus, we utilize the KL divergence between the retrieval and relevance likelihood as the divergence-driven feedback to facilitate more precise retrieval. In the second stage, we propose using the metric score distribution as an anchor to align retrieval with generation. We significantly reduce the gap between retrieval and generation by minimizing the KL divergence between the metric score distribution and the likelihood of retrieving entities.

Extensive experiments on multiple KB-attached task-oriented dialogue datasets demonstrate the great performance of ReAL in achieving accurate retrieval and fluent generation (Eric et al., 2019; Eric and Manning, 2017; Wen et al., 2016). To further illustrate the effectiveness of our adaptive retriever pre-training method, we replace the retrievers in previous methods with ours and show universal effectiveness.

In a nutshell, the main contributions of this paper are three-fold: (I) We propose a two-stage relevance-aware adaptive learning framework for E2ETOD to achieve accurate retrieval and generation. (II) We refine the pre-training of the retriever with an adaptive learning method, preventing mismatches caused by hard negative pairs. (III) Extensive experiments on three benchmark datasets demonstrate the effectiveness of the proposed ReAL.

2 Related Work

2.1 End-to-end Task-oriented Dialogue

From the perspective of knowledge embedding, recent years have witnessed remarkable development in E2ETOD systems. Drawing inspiration from pointer networks (Vinyals et al., 2015), Madotto et al. (2018), Qin et al. (2020), and Raghu et al. (2021) use memory networks to store knowledge explicitly and combine multi-hop attention over the memories to acquire relevant information. In contrast, Huang et al. (2022) and Ding et al. (2024) propose autoregressive entity generation to retrieve knowledge implicitly.

Ever since pre-trained language models (PLMs) take the lead in NLP tasks, the knowledge entities are linearized to be encoded by the PLM encoder (Xie et al., 2022; Tian et al., 2022), which are subsequent input for response generation. DialogKG (Rony et al., 2022) uses a graph neural network to select entities from the flattened records. MAKER (Wan et al., 2023) introduces multi-grained retrieval, involving both entity and attribute selection. As mentioned earlier, although the retrieve-then-generate framework has been successful, its paradigm of decoupling knowledge retrieval from response generation can lead to DAP in task-oriented dialogue systems. While previous researches (Shi et al., 2023a; Wan et al., 2023; Shen et al., 2023) only consider one part of DAP (either the retrieval distractions or retrieval-generation misalignment), we solve the problem jointly with our proposed framework.

2.2 Knowledge Retriever

Enhancing language models with pertinent information from diverse knowledge sources has proven effective in improving performance across various NLP tasks (Khandelwal et al., 2019; Borgeaud et al., 2022; Lewis et al., 2020; Xie et al., 2024). As one of the most successful retrieval structures, the dual-encoder architecture (Yih et al., 2011) encodes queries and passages separately. The relevance between a query-passage pair is computed through inner product or Euclidean distance. Based on this, DPR (Karpukhin et al., 2020) trains the retriever with in-batch documents and samples negative examples for contrastive learning, enabling the pre-trained retriever to perform well in open-domain question answering. REALM (Guu et al., 2020) and RAG (Lewis et al., 2020) consider the retrieved passages as latent variables and train the retriever-generator system jointly. Unlike prior retrieval-augmented language models, REPLUG (Shi et al., 2023b) simply treats the frozen LM as a black-box model and augments it with a tuneable retriever. Lei et al. (2023) reduce the false positive problem by pre-training the dense retriever with contrastive learning. Cheng et al. (2024) utilize an unbounded memory pool and employ a memory selector to choose a single output as the memory for the next generation round.

3 Methodology

3.1 Preliminaries

Given a dialog $\mathcal{D} = \{u_1, r_1, \cdots, u_T, r_T\}$ consisting of T turns, we denote the dialogue context of the t-th turn as $\mathcal{C}_t$, which encompasses all preceding user utterances and system responses up to that turn, i.e., $\mathcal{C}_t = \{u_1, r_1, \cdots, u_{t-1}, r_{t-1}, u_t\}$. u_t and r_t denote the user utterance and system response of t-th turn, respectively. To adapt to different domain-specific problems, an external knowledge base $\mathcal{K} = \{e_1, e_2, \cdots, e_B\}$ is provided as a set of entities. Here, each entity e_i comprises N attribute-value pairs, denoted as $e_i = \{a_1, v_{i,1}, \cdots, a_N, v_{i,N}\}$. End-to-end task-oriented dialogue systems leverage the dialogue context $\mathcal{C}$ and knowledge base $\mathcal{K}$ as input to generate responses $\mathcal{R}$.

3.2 Dual-encoder Knowledge Retriever

From the dialogue context $\mathcal{C}_t$ and knowledge base $\mathcal{K}$, the retriever aims to retrieve a small set of knowledge entities from $\mathcal{K}$, which are relevant to $\mathcal{C}_t$. To ensure simplicity and effectiveness, we follow the dual-encoder architecture widely used in open-domain QA and E2ETOD (Singh et al., 2021; Wan et al., 2023; Shen et al., 2023). By concatenating the user utterances and system responses, the dialogue context is encoded as the query, while the attribute-value pair of the i-th knowledge entity is concatenated and encoded as the external knowledge. The similarity between the encoded $\mathcal{C}_t$ and e_i is calculated as follows,

$$s_{t,i} = s(\mathcal{C}_t, e_i) = \boldsymbol{E}_c(\mathcal{C}_t)^\top \boldsymbol{E}_e(e_i). \tag{1}$$

The top-K entities with the highest similarity scores, *i.e.* $\hat{\mathcal{K}} = \{e_1, e_2, \cdots, e_K\}$, are retrieved as the candidate entities for response generation.

3.3 Adaptive Retriever Pre-training

To learn the representations of the dialogue context and the knowledge entities, we employ a pre-trained language model (PLM) to extract the final [CLS] token to represent them. However, the PLMs initialized with their pre-trained weights show a bad retrieval performance, which may result in the "collapsed representations" mentioned in previous studies (Shi et al., 2023a; Wan et al., 2023). Consequently, we propose a top-K adaptive contrastive loss and utilize the divergence-driven feedback from the frozen PLM generator (Θ) to pre-train the retriever.

Top-K Adaptive Contrastive Learning Since the knowledge base does not annotate the gold retrieved entities, Qin et al. (2019) utilize distant supervision and design a set of heuristics to extract training data for the retriever. When determining the weakly labeled data, the vanilla method selects the entity with the most attribute value occurrences in the dialogue context and system response as pseudo-positive examples. For a positive context-entity pair $(\mathcal{C}_t, e^+)$ of the t-th turn, the vanilla contrastive loss is computed by:

$$\mathcal{L}_{info} = -\log \frac{\exp\left(s\left(\mathcal{C}_t, e^+\right)/\tau\right)}{\sum_{j=1}^{N} \exp\left(s\left(\mathcal{C}_t, e_j\right)/\tau\right)}, \tag{2}$$

where τ and N denote the temperature parameter and the mini-batch size, respectively. Ideally, a

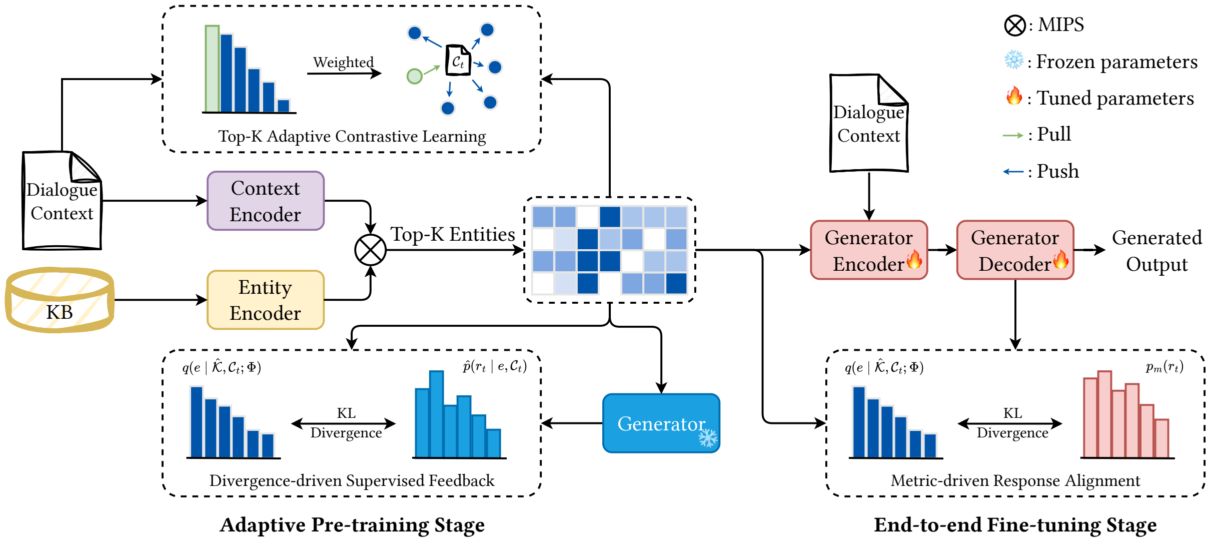


Figure 2: The framework of our proposed ReAL. During the adaptive pre-training stage, the retriever is guided by the top-k contrastive loss and divergence-driven supervised feedback to eliminate distractions step-by-step. The subsequent end-to-end fine-tuning stage aligns retrieval with generation through the metric-driven KL divergence. The generator parameters are frozen during the former stage and tuned during the latter one.

good retriever can identify the close relationship between the corresponding dialogue context and the labeled entity in the representation space. However, the entity with the most occurrences of its attributes is not necessarily the ground truth entity. The vanilla contrastive learning method may mislead the model to pull similar but unrelated entities toward the context in the embedding space and further harm the validity of representations.

Therefore, we propose a simple yet effective Top-K adaptive contrastive loss. We construct top-K context-entity pairs and estimate the matching degree between the pairs, conducted using the intermediate trained retriever Φ. According to the estimated relevance, we assign the weights to different pairs adaptively and improve the vanilla contrastive loss as follows,

$$\mathcal{L}_{adapt} = \sum_{i=1}^{K} \frac{s_\Phi\left(\mathcal{C}_t, e_i^+\right)}{\sum_{j=1}^{K} s_\Phi\left(\mathcal{C}_t, e_j^+\right)} \mathcal{L}_{info}\left(\mathcal{C}_t, e_i^+\right). \tag{3}$$

With the adaptive loss enabled, the retriever can concentrate adaptively on true positive pairs, preventing mismatches caused by false positive pairs.

Divergence-driven Supervised Feedback Inspired by Sachan et al. (2023), we acquire the divergence-driven supervised feedback from the frozen PLM generator to further avoid the retrieval of the hard negative entities. We first compute the likelihood of retrieving the entities conditioned on the query:

$$q(e_i \mid \hat{\mathcal{K}}, \mathcal{C}_t; \Phi) = \frac{\exp\left(s\left(\mathcal{C}_t, e_i\right)/\alpha\right)}{\sum_{j=1}^{K} \exp\left(s\left(\mathcal{C}_t, e_j\right)/\alpha\right)}, \tag{4}$$

where α is a temperature hyperparameter. Since it is impractical to marginalize over all the entities in the knowledge base, we only consider the retrieved top-K entities for estimation. In addition, the hard negative entities always show up with the true positives in the top-K candidates. Therefore, the estimation over the retrieved $\hat{\mathcal{K}}$ would lay more emphasis on this specific problem. Subsequently, we utilize the generation probability of the frozen generator to measure the relevance between the entity and the dialogue context, which has a clear intuition that the more relevant the entity and context are, the greater the generation probability. The relevance score is computed as follows,

$$\hat{p}(r_t \mid e_i, \mathcal{C}_t) = \frac{\exp\left(\log p(r_t \mid e_i, \mathcal{C}_t; \Theta)\right)}{\sum_{j=1}^{K} \exp\left(\log p(r_t \mid e_j, \mathcal{C}_t; \Theta)\right)}, \tag{5}$$

where the generation probability is calculated as:

$$p(r_t \mid e_i, \mathcal{C}_t; \Theta) = \prod_{j=1}^{|r_t|} p(r_{t,j} \mid r_{t,<j}, e_i, \mathcal{C}_t; \Theta). \tag{6}$$

In practice, the entity and the context are concatenated before feeding into the generator. Consequently, the retriever is trained by minimizing the

KL divergence between the retrieval likelihood and the relevance likelihood:

$$\mathcal{L}_{div} = \sum_{i=1}^{K} KL\left(q(e_i \mid \hat{\mathcal{K}}, \mathcal{C}_t; \Phi) \mid\mid \hat{p}(r_t \mid e_i, \mathcal{C}_t)\right). \tag{7}$$

The overall adaptive pre-training of the retriever is conducted by optimizing the two losses as follows,

$$\mathcal{L}_{pre} = \mathcal{L}_{adapt} + \mathcal{L}_{div}. \tag{8}$$

3.4 Metric-driven Response Alignment

Following Shi et al. (2023a) and Wan et al. (2023), we build the generator based on the Fusion-in-Decoder (FiD) (Izacard and Grave, 2020) method. Before encoding, each input $h_{t,i}$ is constructed by concatenating the dialogue context $\mathcal{C}_t$ and a candidate entity e_i from $\hat{\mathcal{K}}$. We then acquire the joint representation from the encoder:

$$\mathcal{H}_t = \boldsymbol{E}_g(h_{t,1}) \oplus \cdots \oplus \boldsymbol{E}_g(h_{t,K}), \tag{9}$$

where $\oplus$ denotes the concatenation. The generator decoder takes $\mathcal{H}_t$ as input and generates the response autoregressively based on the probability $p(r_t \mid \mathcal{H}_t; \Theta)$ defined as Eq. 6. The negative log-likelihood (NLL) loss is incorporated to train the generator during the end-to-end training stage:

$$\mathcal{L}_{nll} = -\log p(r_t \mid \mathcal{H}_t; \Theta). \tag{10}$$

However, as stated in Sec. 1, training the generator only with the NLL loss still results in a misalignment between generation and retrieval. Drawing inspiration from Cheng et al. (2024), we propose utilizing the BLEU score distribution as an anchor to align the retrieval and generation. To be specific, we define the BLEU score distribution as follows,

$$p_m(r_t^{(i)}) = \frac{\exp\left(\delta(r_t^{(i)}, \hat{r}_t)/\beta\right)}{\sum_{j=1}^{K} \exp\left(\delta(r_t^{(j)}, \hat{r}_t)/\beta\right)}, \tag{11}$$

where $r_t^{(i)}$ and $\delta(,)$ denote the response of the t-th turn with the entity e_i and the BLEU score, respectively. β is a temperature parameter. $\hat{r}_t$ is the gold response of the t-th turn. To bridge the gap between generation and retrieval, we quantify it using the KL divergence:

$$\mathcal{L}_{align} = \sum_{i=1}^{K} KL\left(q(e_i \mid \hat{\mathcal{K}}, \mathcal{C}_t; \Phi) \mid\mid p_m(r_t^{(i)})\right). \tag{12}$$

Dataset	Domains	# Turns Train/Val/Test
MWOZ	Restaurant, Attraction, Hotel	9943/576/711
SMD	Navigate, Weather, Schedule	6291/777/808
CamRest	Restaurant	2095/675/643

Table 1: Statistics of the three datasets.

Finally, we set the generator to trainable mode and fine-tune the dialogue system end-to-end with a combination of the two losses:

$$\mathcal{L}_{fine} = \mathcal{L}_{nll} + \mathcal{L}_{align}. \tag{13}$$

4 Experimental Setup

4.1 Datasets and Evaluation Metrics

We conduct experiments on three KB-attached task-oriented dialogue datasets: MultiWOZ 2.1 (MultiWOZ) (Eric et al., 2019), Stanford Multi-Domain (SMD) (Eric and Manning, 2017), and CamRest (Wen et al., 2016). The knowledge bases are condensed with all the entities that meet the user goal of the current dialogue. The statistics of the benchmark datasets are listed in Table 1. Each dialogue in these datasets is linked to a condensed knowledge base containing all entities that meet the user's goal for that dialogue. For MultiWOZ, each condensed knowledge base includes 7 entities. For SMD and CamRest, the size of these knowledge bases varies: from 0 to 8 entities with an average of 5.95 for SMD, and from 0 to 57 entities with an average of 1.93 for CamRest.

Following previous work (Tian et al., 2022; Xie et al., 2022; Wu et al., 2022), we adopt BLEU (Papineni et al., 2002) and Entity F1 (Eric and Manning, 2017) as our primary evaluation metrics. BLEU assesses the fluency of a generated response by measuring its n-gram overlap with a reference response, while Entity F1 evaluates the accuracy of embedded knowledge by micro-averaging precision and recall scores of attribute values in the generated response.

4.2 Implementation Details

We employ BERT (Devlin et al., 2018) as the encoder of our entity selector and attribute selector and employ T5 (Raffel et al., 2020) to implement the response generator. Our model is trained using AdamW optimizer (Loshchilov and Hutter, 2017) with a batch size of 64. We conduct all experiments on 4 24G NVIDIA RTX 3090 GPUs and select the best checkpoint based on model performance

Methods	MultiWOZ		SMD		CamRest	
	BLEU	Entity F1	BLEU	Entity F1	BLEU	Entity F1
FG2Seq	14.60[§]	36.50[§]	16.80[§]	61.10[§]	20.20[§]	66.40[§]
CDNET	11.90	38.70	17.80	62.90	21.80	68.60
GraphMemDialog	14.90	40.20	18.80	64.50	22.30	64.40
ECO	12.61	40.87	-	-	18.42	71.56
DialoKG	12.60	43.50	20.00	65.90	23.40	75.60
Uni-TOD	12.30	44.30	-	-	24.7	**77.80**
T5-Base						
UnifiedSKG	-	-	17.41	66.45	-	-
Q-TOD	-	-	20.14	68.22	-	-
DF-TOD	<u>18.26</u>	52.52	24.12	69.36	25.85	72.83
MK-TOD$_{ctr}$	17.33	51.86	<u>24.77</u>	67.86	**26.76**	73.60
MAKER	17.23	<u>53.68</u>	**24.79**	<u>69.79</u>	25.04	73.09
Ours	**18.51**	**56.13**[†]	24.64	**72.23**[†]	<u>26.61</u>	<u>76.11</u>
T5-Large						
UnifiedSKG	13.69[‡]	46.04[‡]	17.27[‡]	65.85[‡]	20.31[‡]	71.03[‡]
Q-TOD	17.62	50.61	21.33	71.11	23.75	74.22
DF-TOD	18.48	53.17	25.10	71.58	26.00	74.04
MK-TOD$_{ctr}$	17.55	52.97	25.43	<u>73.31</u>	<u>26.20</u>	71.72
MAKER	<u>18.77</u>	<u>54.72</u>	<u>25.91</u>	71.30	25.53	74.36
Ours	**19.03**	**57.26**[†]	**26.33**	**74.90**[†]	**27.02**	<u>76.34</u>

Table 2: Performance on three condensed benchmark datasets. The best scores are in **bold** and the second-best ones are <u>underlined</u>. † denotes our model significantly outperforms baselines with $p < 0.05$ under t-test. ‡ and § indicates the results cited from Tian et al. (2022) and Raghu et al. (2021), respectively.

Methods	MultiWOZ		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
T5-Base				
MK-TOD$_{ctr}$	15.96	51.35	26.85	<u>73.51</u>
DF-TOD	<u>17.61</u>	<u>51.61</u>	**27.39**	70.74
MAKER	16.25	50.87	26.19	72.09
Ours	**17.83**	**54.43**[†]	<u>26.89</u>	**75.74**[†]
T5-Large				
Q-TOD	16.67	47.13	21.44	63.88
MK-TOD$_{ctr}$	17.40	<u>53.26</u>	**27.82**	71.98
DF-TOD	<u>18.36</u>	52.96	26.61	<u>73.58</u>
MAKER	18.23	52.12	25.34	72.43
Ours	**18.64**	**56.84**[†]	<u>27.65</u>	**75.53**[†]

Table 3: Performance on two large-scale benchmark datasets. The best scores are in **bold** and the second-best ones are <u>underlined</u>. † denotes our model significantly outperforms baselines with $p < 0.05$ under t-test.

on the validation set. All experiments' results are obtained by averaging the scores over five runs with different random seeds. Following Wan et al. (2023), we pre-train the retriever on the MultiWOZ and CamRest datasets. Due to the SMD dataset's knowledge base being specific to each dialogue, it is not possible to compile a global knowledge base from the dialogues. Consequently, pre-training is not conducted on the SMD dataset.

4.3 Comparative Baselines

We perform a comprehensive comparative study against ReAL by considering the baselines with different retrieval strategies. FG2Seq (He et al., 2020), CDNET (Raghu et al., 2021), GraphMemDialog (Wu et al., 2022), ECO (Huang et al., 2022), and Uni-TOD (Ding et al., 2024) integrate knowledge retrieval and response generation in one single model. In contrast, DigloKG (Rony et al., 2022), UnifiedSKG (Xie et al., 2022), Q-TOD (Tian et al., 2022), DF-TOD (Shi et al., 2023a), MK-TOD (Shen et al., 2023), and MAKER (Wan et al., 2023) decouple the retrieval and generation process.

5 Results and Analysis

5.1 Quantitative Analysis

Condensed Knowledge Base As shown in Table 2, we can observe that our proposed method achieves excellent performance among all the benchmark datasets with either backbone. Specifically, with T5-Large as the backbone model, ReAL surpasses previous state-of-the-art methods on MultiWOZ and SMD. On MultiWOZ, ReAL demonstrates improvements of 0.26 points in BLEU and

Methods	Condensed		Large-scale	
	BLEU	Entity F1	BLEU	Entity F1
ReAL	18.11	55.13	17.83	54.43
w/o $\mathcal{L}_{adapt}$	17.97 (↓ **0.14**)	54.04 (↓ **1.09**)	17.72 (↓ **0.11**)	52.89 (↓ **1.54**)
w/o $\mathcal{L}_{div}$	17.24 (↓ **0.87**)	54.34 (↓ **0.79**)	17.16 (↓ **0.67**)	53.23 (↓ **1.20**)
w/o $\mathcal{L}_{pre}$	16.22 (↓ **1.89**)	53.17 (↓ **1.96**)	16.59 (↓ **1.24**)	51.38 (↓ **3.05**)
w/o $\mathcal{L}_{align}$	17.04 (↓ **1.07**)	54.10 (↓ **1.03**)	16.87 (↓ **0.96**)	53.29 (↓ **1.14**)

Table 4: Results of ablation study on MultiWOZ under the condensed and large-scale setting with T5-base, where "*w/o*" means without. When ablating $\mathcal{L}_{adapt}$ and $\mathcal{L}_{pre}$, we substitute $\mathcal{L}_{adapt}$ with the vanilla contrastive loss $\mathcal{L}_{info}$.

Retriever	BLEU	Entity F1	Recall@7
Oracle*	16.17	51.45	100.00
Ours	**17.51**	**53.26**	**92.44**
MAKER*	17.18	49.05	86.47
Vanilla $\mathcal{L}_{info}$*	16.67	48.77	82.71
Frequency*	16.60	48.00	75.94
BM25*	16.21	45.56	26.32

Table 5: Comparison of different retrievers under the large-scale setting of MultiWOZ. *Oracle* means directly using the condensed knowledge base. *Vanilla* $\mathcal{L}_{info}$ denotes using the vanilla contrastive pre-trained retriever. * indicates the results cited from Wan et al. (2023).

2.54 points in Entity F1. On SMD, the improvements are 0.42 in BLEU and 1.09 in Entity F1. However, the best Entity F1 result on CamRest is kept by the strong baseline Uni-TOD, though its results on MultiWOZ are comparatively much weaker. It is worth noting that each condensed knowledge base for CamRest only contains 1.93 entities on average (Wan et al., 2023), which is beneficial to autoregressive models like DialoKG and Uni-TOD. When the number comes to 7 for MultiWOZ, our model takes the lead naturally.

Through our proposed effective relevance-aware adaptive learning method, our model with the T5-Base backbone even exceeds the ones with the T5-Large generator. On MultiWOZ and CamRest, ReAL shows superiority in Entity F1 and leads by a big margin, which is attributed to the refined top-k adaptive contrastive loss and divergence-driven feedback.

Large-scale Knowledge Base Though the results in Table 2 are outstanding, the real-world task-oriented dialogue systems face a more intractable situation in which countless pieces of knowledge from different domains are blended without orders. Thus, we follow previous researches (Wan et al., 2023; Shi et al., 2023a) to conduct experiments under a large-scale setting. As shown in Table 3, we gather the entities of all dialogs in MultiWOZ and CamRest respectively for evaluation. On one hand, the existing approaches experience significant performance degradation without exception when utilizing large-scale knowledge bases. The Entity F1 score of MAKER and DF-TOD drop 2.81/2.6 and 0.91/0.21 points on MultiWOZ, respectively. In contrast, our method shows much greater stability with only a decline of 0.7/0.42 points while keeping a competitive performance. On the other hand, our ReAL system demonstrates superior performance compared to all baselines, despite a minor deficit in BLEU on CamRest. Notably, our system, even with large-scale knowledge bases, consistently outperforms other systems that rely on condensed knowledge bases, which are easier to retrieve. These findings highlight the exceptional capability of our system in handling large-scale knowledge bases and its practicality for real-world applications.

5.2 Ablation Study

To better understand how each part of ReAL affects its performance, we conduct an ablation study on MultiWOZ under the condensed and large-scale setting with T5-base backbone, and the results are shown in Table 4. When substituting the top-k adaptive contrastive loss $\mathcal{L}_{adapt}$ with the vanilla $\mathcal{L}_{info}$, the performance of our model drops noticeably. This drop is more severe under the large-scale setting, with a reduction of 0.11 in BLEU and 1.54 in Entity F1 scores, compared to the condensed setting, where the reductions are 0.14 in BLEU and 1.09 in Entity F1. The larger drop in the large-scale setting suggests that the poorly-trained retriever suffers significantly from DAP without adaptive pre-training. Removing the divergence-driven supervised feedback also causes a performance reduction, though it is less severe than removing $\mathcal{L}_{adapt}$. However, the ablation of $\mathcal{L}_{pre}$ demonstrates the significant coordination between the two losses.

Methods	MultiWOZ		CamRest	
	BLEU	Entity F1	BLEU	Entity F1
		T5-Base		
DF-TOD*	17.53	51.46	27.13	70.48
w/ Ours	17.48 (↓ **0.05**)	52.97 (↑ **1.51**)	27.24 (↑ **0.11**)	72.08 (↑ **1.6**)
MAKER*	16.24	50.76	25.83	71.66
w/ Ours	16.37 (↑ **0.13**)	52.03 (↑ **1.27**)	25.71 (↓ **0.12**)	73.50 (↑ **1.84**)
		T5-Large		
DF-TOD*	18.23	52.96	26.57	73.40
w/ Ours	18.24 (↑ **0.01**)	54.21 (↑ **1.25**)	26.33 (↓ **0.24**)	75.17 (↑ **1.77**)
MAKER*	18.12	51.98	25.33	72.47
w/ Ours	17.88 (↓ **0.24**)	53.03 (↑ **1.05**)	25.46 (↑ **0.13**)	74.22 (↑ **1.75**)

Table 6: Performance of applying our adaptive pre-trained retriever to advanced models under the large-scale benchmark setting. * denotes our re-implementation.

Abandoning $\mathcal{L}_{pre}$ results in the largest performance reduction, with decreases of 1.89 in BLEU and 1.96 in Entity F1 in the condensed setting, and 1.24 in BLEU and 3.05 in Entity F1 in the large-scale setting. This substantial drop certifies the effectiveness of eliminating hard negatives step-by-step and highlights the critical role of the proposed adaptive pre-training in maintaining high performance. In addition, the ablation results of the metric-driven loss $\mathcal{L}_{align}$ show its validity in aligning retrieval and generation. Therefore, the proposed method ReAL is effective as each of its components contributes to its overall performance, and removing any of these components leads to a noticeable decline in both metrics.

5.3 Comparison of Retrievers

To further validate the effectiveness of our ReAL retriever, we compare the performance of different retrievers under the large-scale setting of MultiWOZ. To ensure a fair comparison, we follow the procedure of Wan et al. (2023) and employ the generator with the same backbone. Since the details are not mentioned, we utilize T5-Base as the backbone for response generation due to its similar performance. The results in Table 5 demonstrate that our method achieves the highest BLEU and Entity F1 scores among all methods, indicating superior precision in text generation and entity retrieval. Though directly acquiring all entities, the oracle method fails to perform the best due to similar entities, which leads to the retrieval-generation misalignment. In contrast, our proposed approach effectively mitigates the problem and maintains the best performance with a smaller Recall@7.

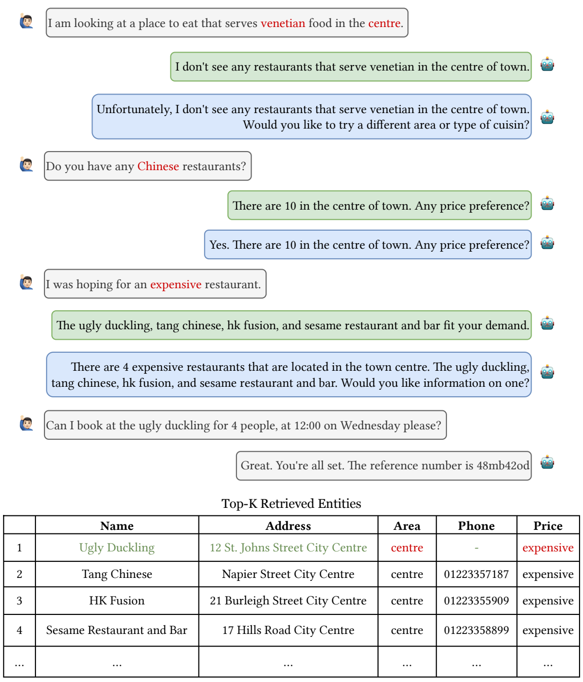
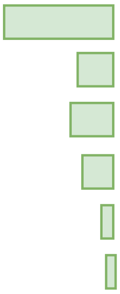


Figure 3: An example from MultiWOZ 2.1 (Eric et al., 2019). The sentences in green are generated by ReAL and the ones in blue are gold responses.

5.4 Generalization Ability of ReAL Retriever

To prove the strong universality and generalization ability of our proposed retriever pre-training method, we apply our pre-trained retriever to two advanced models. The results are shown in Table 6. The adaptive pre-trained retriever enhances the Entity F1 score across both datasets and for both models, indicating superior entity recognition capability. The BLEU score shows minor changes, with slight improvements or decreases, suggesting that the retriever's impact on overall text generation quality is minimal but generally positive. This indicates that the proposed method effectively im-



	Name	Address	Area	Phone	Price
1	A and B Guesthouse	124 Tenison Road	east	3315702	moderate
2	Carolina Bed	138 Perne Road	east	3247015	moderate
3	Warkworth House	Wark Terrace	east	3363682	moderate
4	rajmahal	7 barnwell road fen ditton	east	01223244955	moderate
5	Archway House	52 Gilbert Road	north	3575314	moderate
6	Arbury Lodge	82 Arbury Road	north	3364319	moderate

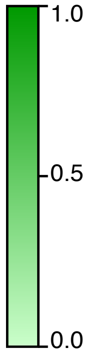


Figure 4: KB score distribution. The distribution is the timestep when generating entity attributes `124 Tenison Road` and `3315702` for the response "The address is 124 Tenison Road and the phone number is 3315702".

proves entity recognition while maintaining overall text generation quality.

5.5 Qualitative Analysis

As shown in Figure 3, we present a dialogue example from the MultiWOZ 2.1 dataset. For a given user utterance, our system successfully retrieves entities that meet the user's goal while excluding irrelevant distractions. It then generates appropriate system responses. Notably, when the user's goal changes, such as in the second turn when the user requests any Chinese restaurant, our retriever adapts and retrieves the relevant entity accordingly.

5.6 Visualization

To better understand the impact of our retriever on the overall KB score distribution, we visualized the KB entity probabilities at the decoding positions where we generate the entity attributes `124 Tenison Road` and `3315702`. As shown in Figure 4, the first row, and the second and fourth columns have the highest probabilities for generation, confirming accurate retrieval and demonstrating the effectiveness of our adaptive retriever pre-training.

6 Conclusion

In this paper, we propose a two-stage Relevance-aware Adaptive Learning (ReAL) framework to eliminate hard negatives step-by-step and align retrieval with the generation. With the first adaptive pre-training stage and the subsequent end-to-end fine-tuning stage, our model effectively mitigates DAP. Extensive experiments and analysis demonstrate ReAL's superiority.

Limitations

Despite the notable superiority of our proposed method over existing SOTA approaches, it is imperative to acknowledge and address several challenges in future research endeavors. Firstly, when the number of entities contained in the knowledge base increases, the performance of the model is still not ideal and stable, which is reflected in the large gap between MultiWOZ's Entity F1 and the other two datasets. Secondly, existing TOD systems still perform poorly in the case of long context and extreme multi-turn conversations. These limitations present a critical area for further investigation in our subsequent research efforts.

References

Sebastian Borgeaud, Arthur Mensch, Jordan Hoffmann, Trevor Cai, Eliza Rutherford, Katie Millican, George Bm Van Den Driessche, Jean-Baptiste Lespiau, Bogdan Damoc, Aidan Clark, et al. 2022. Improving language models by retrieving from trillions of tokens. In *International conference on machine learning*, pages 2206–2240. PMLR.

Xin Cheng, Di Luo, Xiuying Chen, Lemao Liu, Dongyan Zhao, and Rui Yan. 2024. Lift yourself up: Retrieval-augmented text generation with self-memory. *Advances in Neural Information Processing Systems*, 36.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Zeyuan Ding, Zhihao Yang, Ling Luo, Yuanyuan Sun, and Hongfei Lin. 2024. From retrieval to generation: A simple and unified generative model for end-to-end task-oriented dialogue. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 17907–17914.

Mihail Eric, Rahul Goel, Shachi Paul, Adarsh Kumar, Abhishek Sethi, Peter Ku, Anuj Kumar Goyal, Sanchit Agarwal, Shuyang Gao, and Dilek Hakkani-Tur. 2019. Multiwoz 2.1: A consolidated multi-domain dialogue dataset with state corrections and state tracking baselines. *arXiv preprint arXiv:1907.01669*.

Mihail Eric and Christopher D Manning. 2017. Key-value retrieval networks for task-oriented dialogue. *arXiv preprint arXiv:1705.05414*.

Abhinav Goyal, Anupam Singh, and Nikesh Garera. 2022. End-to-end speech to intent prediction to improve e-commerce customer support voicebot in hindi and english. *arXiv preprint arXiv:2211.07710*.

Kelvin Guu, Kenton Lee, Zora Tung, Panupong Pasupat, and Mingwei Chang. 2020. Retrieval augmented language model pre-training. In *International conference on machine learning*, pages 3929–3938. PMLR.

Zhenhao He, Yuhong He, Qingyao Wu, and Jian Chen. 2020. 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.

Guanhuan Huang, Xiaojun Quan, and Qifan Wang. 2022. Autoregressive entity generation for end-to-end task-oriented dialog. *arXiv preprint arXiv:2209.08708*.

Gautier Izacard and Edouard Grave. 2020. Leveraging passage retrieval with generative models for open domain question answering. *arXiv preprint arXiv:2007.01282*.

Vladimir Karpukhin, Barlas Oğuz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih. 2020. Dense passage retrieval for open-domain question answering. *arXiv preprint arXiv:2004.04906*.

Urvashi Khandelwal, Omer Levy, Dan Jurafsky, Luke Zettlemoyer, and Mike Lewis. 2019. Generalization through memorization: Nearest neighbor language models. *arXiv preprint arXiv:1911.00172*.

Yibin Lei, Liang Ding, Yu Cao, Changtong Zan, Andrew Yates, and Dacheng Tao. 2023. Unsupervised dense retrieval with relevance-aware contrastive pre-training. *arXiv preprint arXiv:2306.03166*.

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.

Ilya Loshchilov and Frank Hutter. 2017. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.

Andrea Madotto, Chien-Sheng Wu, and Pascale Fung. 2018. Mem2seq: Effectively incorporating knowledge bases into end-to-end task-oriented dialog systems. *arXiv preprint arXiv:1804.08217*.

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. *arXiv preprint arXiv:1909.06762*.

Libo Qin, Wenbo Pan, Qiguang Chen, Lizi Liao, Zhou Yu, Yue Zhang, Wanxiang Che, and Min Li. 2023. End-to-end task-oriented dialogue: A survey of tasks, methods, and future directions. *arXiv preprint arXiv:2311.09008*.

Libo Qin, Xiao Xu, Wanxiang Che, Yue Zhang, and Ting Liu. 2020. Dynamic fusion network for multi-domain end-to-end task-oriented dialog. *arXiv preprint arXiv:2004.11019*.

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

Dinesh Raghu, Atishya Jain, Sachindra Joshi, et al. 2021. Constraint based knowledge base distillation in end-to-end task oriented dialogs. *arXiv preprint arXiv:2109.07396*.

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

Devendra Singh Sachan, Mike Lewis, Dani Yogatama, Luke Zettlemoyer, Joelle Pineau, and Manzil Zaheer. 2023. Questions are all you need to train a dense passage retriever. *Transactions of the Association for Computational Linguistics*, 11:600–616.

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. *arXiv preprint arXiv:2310.08877*.

Tianyuan Shi, Liangzhi Li, Zijian Lin, Tao Yang, Xiaojun Quan, and Qifan Wang. 2023a. Dual-feedback knowledge retrieval for task-oriented dialogue systems. *arXiv preprint arXiv:2310.14528*.

Weijia Shi, Sewon Min, Michihiro Yasunaga, Minjoon Seo, Rich James, Mike Lewis, Luke Zettlemoyer, and Wen-tau Yih. 2023b. Replug: Retrieval-augmented black-box language models. *arXiv preprint arXiv:2301.12652*.

Devendra Singh, Siva Reddy, Will Hamilton, Chris Dyer, and Dani Yogatama. 2021. End-to-end training of multi-document reader and retriever for open-domain question answering. *Advances in Neural Information Processing Systems*, 34:25968–25981.

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*.

Oriol Vinyals, Meire Fortunato, and Navdeep Jaitly. 2015. Pointer networks. *Advances in neural information processing systems*, 28.

Fanqi Wan, Weizhou Shen, Ke Yang, Xiaojun Quan, and Wei Bi. 2023. Multi-grained knowledge retrieval for end-to-end task-oriented dialog. *arXiv preprint arXiv:2305.10149*.

Tsung-Hsien Wen, David Vandyke, Nikola Mrksic, Milica Gasic, Lina M Rojas-Barahona, Pei-Hao Su, Stefan Ultes, and Steve Young. 2016. A network-based end-to-end trainable task-oriented dialogue system. *arXiv preprint arXiv:1604.04562*.

Jie Wu, Ian G Harris, and Hongzhi Zhao. 2022. Graphmemdialog: Optimizing end-to-end task-oriented dialog systems using graph memory networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 11504–11512.

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*.

Yuxin Xie, Zhihong Zhu, Xianwei Zhuang, Liming Liang, Zhichang Wang, and Yuexian Zou. 2024. Gpa: Global and prototype alignment for audio-text retrieval. In *Interspeech 2024*, pages 5078–5082.

Doris Xin, Hui Miao, Aditya Parameswaran, and Neoklis Polyzotis. 2021. Production machine learning pipelines: Empirical analysis and optimization opportunities. In *Proceedings of the 2021 international conference on management of data*, pages 2639–2652.

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.