Modularized Pre-training for End-to-end Task-oriented Dialogue

Libo Qin, Xiao Xu, Lehan Wang, Yue Zhang, Wanxiang Che

***Abstract*—Pre-training for end-to-end task-oriented dialogue systems (EToDs) is a challenging task due to its unique knowledge base query (accuracy) need and lack of sufficient training data (fluency). In this paper, we try to mitigate the above challenges by introducing a modularized pre-training framework for EToDs, which achieves to effectively improve both accuracy and fluency of EToDs through a pre-training paradigm. The core insight is a modular design by decomposing EToDs into a *generation (fluency)* module and a *knowledge-retriever (accuracy)* module, which allows us to optimize each module by pre-training these two sub-modules with different well-designed pre-training tasks, respectively. In addition, such a modularized paradigm enables us to make full use of large amounts of KB-free dialogue corpus for the pre-training *generation* module, which can alleviate the insufficient training problem. Furthermore, we introduce a new *consistency-guided* data augmentation (CGDA) strategy to cope with the data scarcity problem to better pre-train the *knowledge-retriever* module. Finally, we fine-tune the pre-trained *generation* module and *knowledge-retriever* module jointly. Experimental results on three datasets show that our model achieve superior performance in terms of both fluency and accuracy. To our knowledge, this is the first work to explore modularized pre-training methods for EToDs.**

***Index Terms*—Task-oriented Dialogue System, Modularized Pre-training, consistency-guided data augmentation.**

I. Introduction

TASK-ORIENTED dialogue systems (ToDs) can complete user goals such as hotel bookings and restaurant reservations, which gains increasing attention. Traditional ToDs consists of modularly connected components for natural language understanding (NLU) [1], dialogue state tracking (DST) [2], dialogue policy (DP) [3] and natural language generation (NLG) [4] module. In recent years, end-to-end task-oriented dialogue systems (EToDs) has emerged in the literature, which use a unified sequence-to-sequence model to generate a response given a dialogue history and knowledge base (KB) [5]. For example, given the dialogue history *"Send me to the nearest gas station, I want to fuel my car."* in the first turn and the corresponding knowledge base in Figure 1, EToDs can directly produce the system response *"The nearest gas station is valero at 200 Alester Ave, 7 miles away setting directions now"*, which does not need any intermediate supervision.

Co-corresponding authors: Yue Zhang and Wanxiang Che.

Libo Qin is with Central South University, Changsha, Hunan 410017, China. (e-mail: lbqin@ir.hit.edu.cn)

Xiao Xu, Lehan Wang and Wanxiang Che, are with Research Center for Social Computing and Information Retrieval, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China (email: xxu, car@ir.hit.edu.cn).

Yue Zhang is with the School of Engineering, Westlake University, Hangzhou 310024, China, and also with the Institute of Advanced Technology, Westlake Institute of Advanced Study, Hangzhou 310024, China (e-mail: yue.zhang@wias.org.cn).

Knowledge Base (KB)

POI	POI type	Distance	Address	Traffic info
Mandarin Roots	chinese restaurant	2 miles	271 Springer Street	heavy traffic
Valero	*gas station*	*7 miles*	*200 Alester Ave*	*moderate traffic*
Safeway	grocery store	3 miles	452 Arcadia Pl	moderate traffic
Dominos	pizza restaurant	5 miles	776 Arastradero Rd	no traffic

Dialogue

Driver	Send me to the nearest gas station, I want to fuel my car.
Car	The nearest *gas station* is *Valero* at *200 Alester Ave*, *7 miles* away setting directions now.
Driver	Where is Valero and is there any traffic on the way?
Car	Valero is at *200 Alester Ave* and *moderate traffic* is being noted.

Fig. 1: An example of task-oriented dialogue from the SMD dataset [11]. Words with the blue color refer to the queried entity from the corresponding knowledge base.

Pre-trained language models, such as GPT-2 [6], have shown empirical success on *open-domain* dialogue direction [7], [8]. In addition, great progress has also been witnessed in and *task-oriented* dialogue direction, including dialogue state tracking [9], natural language generation [10]. However, it is relatively under-explored for EToDs due to the following challenges: **First**, pre-training for current EToDs requires large amounts of KB-grounded dialogues that is extremely hard to collect, which is insufficient for capturing task-oriented domain characteristic, resulting in low fluency; **Second**, more importantly, task-oriented dialogue systems require a KB retrieval module, which does not seamlessly integrate with the general pre-training stage, leading to low accuracy. As shown in Figure 1, the dialogue system is required to query a corresponding KB to retrieve the entity like "`Valero`" for the driver's query about `gas station`.

Motivated by this, we propose a new pre-training paradigm to solve the aforementioned challenges. To be more specific, we propose a modularized pre-training framework that decomposes the model into a *generation (fluency)* module and a *knowledge-retriever (accuracy)* module. With the help of modularized pre-training paradigm, *generation* module and *knowledge retriever* module are decoupled, which brings us at least two advantages: (1) it allows each sub-module to benefit from further pre-training to improve both fluency and accuracy, which makes it easier integrate them into a unified EToDs architecture; (2) it enables the model to make full use of a large amount of KB-free dialogues corpus, which is able to alleviate the insufficient data problem.

To improve fluency, as shown in Figure 2(a), we design two pre-training tasks for the *generation* module to mitigate data distribution shift issue between the pre-training stage and fine-tuning stage: (1) *Language Modeling Task (LMT):* inspired

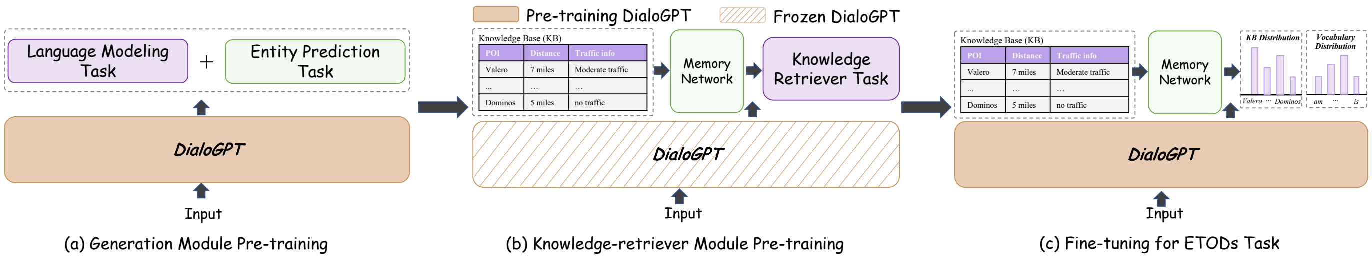


Fig. 2: The overall architecture of the proposed modularized pre-training framework. (a) We first pre-train the *generation* module and freeze the *knowledge retriever* module; (b) We pre-train the *knowledge-retriever* module and freeze the *generation* module; (c) Finally, we fine-tune the pre-trained modularized model on the end-to-end task-oriented dialogues.

by Wu *et al.* [12] , we apply LMT on collected task-oriented dialogue corpus to bridge the domain gap between the general domain and the EToDs domain; (2) *Entity Prediction Task (EPT)*: EPT is used for predicting whether each generated word is a knowledge entity during the decoding stage that aims to enhance the model with knowledge awareness, which is essential for task-oriented dialogue domain. It is worth noting that both LMT and EPT pre-training only need KB-free dialogues, which is much easier to collect.

To improve accuracy, we first freeze the *generation* module and further pre-train the *knowledge-retriever* module independently. Specifically, we propose a *Knowledge Retrieve Task (KRT)* that is used for selecting which knowledge entities in KB will be retrieved by the *knowledge-retriever* module, which improves the ability to retrieve correct knowledge entities from KB for dialogue generation, as shown in Figure 2(b). The pre-trained *knowledge-retriever* module can be easily integrated with the *generation* module due to our modularization framework. Nevertheless, unlike the *generation* module pre-training, the *knowledge-retriever* module has to rely on a considerable amount of KB-grounded dialogue for pre-training, which is much difficult to acquire than KB-free dialogues collection. To tackle the data scarcity problem, we further propose a *consistency-guided* data augmentation strategy, which enlarges pseudo KB-grounded dialogues corpus from existing data by replacing specific knowledge entities without breaking the dialogue context consistency and KB consistency (see §III-B).

As shown in Figure 2(c), after separately pre-training the *generation* module and *knowledge-retriever* module, we fine-tune the pre-trained system on three public EToDs datasets including SMD [11], CamRest [13] and an extension of Multi-WOZ 2.1 [14]. Experimental results demonstrate the effectiveness of our proposed modularized pre-training framework by obtaining promising performance on both accuracy and fluency in EToDs.

The contribution of this work can be summarized as: (1) We propose a modularized pre-training framework for EToDs by decoupling the *generation* module and the *knowledge-retriever* module, allowing the model to optimize the two sub-modules to improve accuracy and fluency, respectively. In addition, such modularized paradigm can enable model to utilize KB-free dialogue; (2) We devise a new *consistency-guided* data augmentation strategy (CGDA) to alleviate the data insufficiency problem for pre-training *knowledge-retriever* module while keeping the dialogue context consistency and KB consistency; (3) Results on three public datasets show that our framework achieves superior performance. Extensive analysis verifies that the well-designed pre-training tasks improve both fluency and accuracy in EToDs.

To facilitate the further research, our code are publicly available at https://github.com/LooperXX/MPEToDs.

II. Background

In this section, we provide a brief introduction to end-to-end task-oriented dialogue systems (EToDs).

a) Dialogue History: Given a user u and system s, we follow Qin *et al.* [15] to represent the n-turned dialogue utterances as $(u_1, s_1), (u_2, s_2), ..., (u_n, s_n)$. In addition, at the i-th turn of the dialogue, we flatten dialogue context $(u_1, u_2, \ldots, u_i)$ and denote $X = (x_1, x_2, ..., x_m)$ as the dialogue history where m stands for the number of words in the dialogue history.

b) Knowledge Base: As described in Qin *et al.* [15], the corresponding knowledge base (KB) is a relational-database-like KB B, which contains $|\mathcal{R}|$ rows and $|\mathcal{C}|$ columns. The value of entity in the i-th row and the j-th column is represented as $v_{i,j}$.

c) End-to-end Task-oriented Dialogue Systems: Following Eric *et al.* [11] and Qin *et al.* [14] , the end-to-end task-oriented dialogue generation task can be defined as predicting the most likely response Y given the dialogue history $X = (x_1, x_2, \ldots, x_m)$ and KB B. Formally, this process is defined as:

$$P(Y \mid X, B) = \prod_{t=1}^{n} p(y_t \mid y_1, \ldots, y_{t-1}, X, B), \tag{1}$$

where y_t denotes an output word; m and n are the length of dialog history and response, respectively.

III. Modularized Pre-training Framework

This section illustrates the overflow of Modularized Pre-training Framework. Specifically,

It consists of three main stages:

(1) first pre-train generation module (§ III-A);
(2) then freeze the pre-trained *generation* module and pre-train *knowledge-retriever* module (§ III-B);

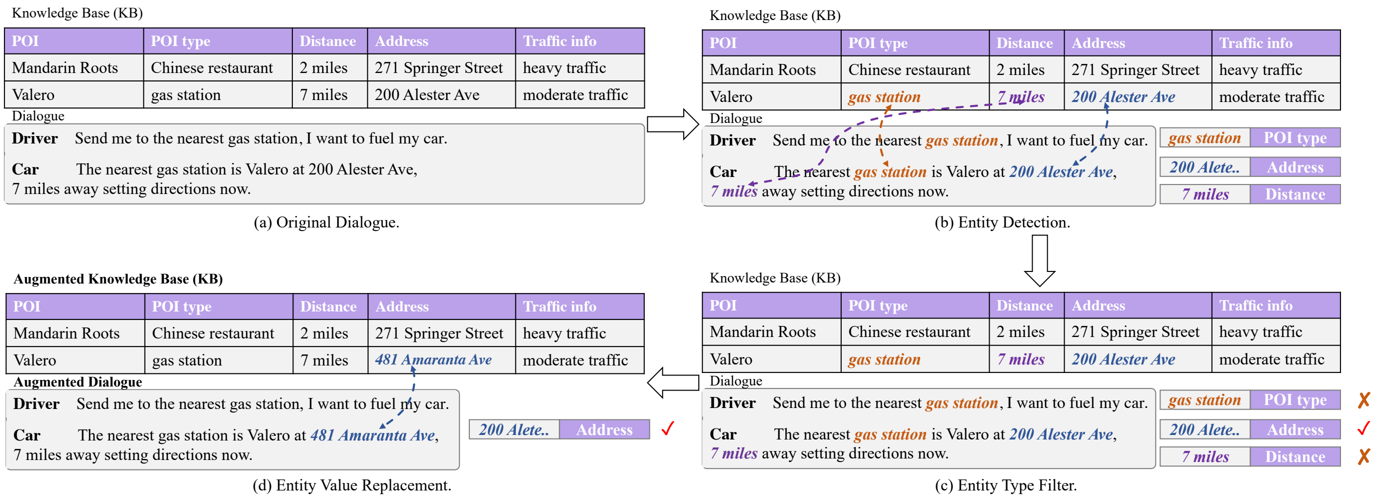


Fig. 3: Illustration of consistence-guided data augmentation (CGDA). (a) shows the original dialogue; (b) illustrates the entity detection process and (c) shows the entity type filter process; (d) illustrates the entity value replacement process and the final augmented dialogues and knowledge base (KB).

(3) finally fine-tune the pre-trained *generation* and *knowledge-retriever* for EToDs tasks (§ III-C).

In our framework, we employ `DialoGPT` as the backbone since it has incorporated open-domain dialogue features.

A. Generation Module Pre-Training

a) Data collection: Following Wu *et al.* [12] , we collect and combine nine human-human task-oriented dialogue corpora to further pre-train `DialoGPT`. In total, there are around 102K dialogues with 1.7M utterances. The statistics of data used in the generation module pre-training are shown in Table I.

Name	# Dialogue	# Utterance
Schema [16]	22,825	463,284
Taskmaster-1 [17]	13,215	303,066
Taskmaster-2 [17]	17,289	341,801
Taskmaster-3 [17]	23,757	480,784
MSR-E2E [18]	10,087	74,686
WOZ [19]	1,200	5,012
Training set of SMD [11]	2,425	12,580
Training set of CamRest676 [13]	405	3,332
Training set of MWOZ [20]	8,438	115,424

TABLE I: Data Statistics for Task-Oriented Dialogue Datasets in Generation Module Pre-Training.

b) Language Modeling Task (LMT): Though achieving promising performance on open-domain dialogue, `DialoGPT` leads to a large gap in domain distribution between open domain dialogue (OOD) and EToDs. Therefore, we explore Language Modeling Task (LMT) to further pre-train `DialoGPT`, achieving to capture domain feature of EToDs.

The task of LM is to predict a distribution of the next word given the previous words. Formally, given the hidden outputs $\{\mathbf{h}_1, \ldots, \mathbf{h}_{m+n}\}$ of `DialoGPT`, the prediction probability for next word (t timestep) is computed as:

$$y_t^{\mathrm{LM}} = \mathrm{Softmax}\left(\mathbf{U}\mathbf{h}_t\right), \tag{2}$$

where $\mathbf{U} \in \mathbb{R}^{|\mathcal{V}|\times d}$ is the output embedding matrix; d denotes the dimension size; $\mathcal{V}$ denotes the vocabulary size.

c) Entity Prediction Task (EPT): It's hard for `DialoGPT` to distinguish whether the generated word is a common word or an entity, since it's just pre-trained on open-domain dialogue corpus. To facilitate `DialoGPT` awareness of the knowledge that is essential for EToDs, we propose the entity prediction task, which is used for judging whether the generated word is an entity or not. Formally, the task is defined as:

$$y_i^{\mathrm{EP}} = \mathrm{Sigmoid}(\mathbf{u}^\top \mathbf{h}_i + b)), \tag{3}$$

where y_i^{EP} denotes the probability that the i-th generated word is an entity. For each word in the system response, we use 1 to indicate the word is a knowledge entity, and 0 otherwise.

We argue that if model can successfully predict the entity in the system response, the model is capable of capturing knowledge awareness.

B. Knowledge-retriever Module Pre-training

A *Knowledge Retrieve Task (KRT)* is introduced to pre-train the *knowledge-retriever* module to improve the KB-retriever ability. Specifically, KRT is used for selecting which knowledge entities in KB will be retrieved by the *generation* module, which enhances the ability to retrieve correct knowledge entities from KB for dialogue generation.

a) Knowledge Retriever Task (KRT): Following Wu *et al.* [21] and Qin *et al.* [14] , we adopt memory-network to store knowledge. Since knowledge entities can either come from the KB or the dialog history, we use an external knowledge memory to encode both the KB $\mathbf{B}$ and the dialogue history $\mathbf{X}$. The entities in external knowledge memory can be denoted as a triple format *(subject, relation, object)*, which are represented as $\mathbf{M} = [\mathbf{B}; \mathbf{X}] = (m_1, \ldots, m_b, m_{b+1}, \ldots, m_{b+T})$, where m_i stands for triplet in $\mathbf{M}$; b and T denotes the number of triplets in KB and dialog history, respectively. A bag-of-word representation is used to represent each entity (triple) in $\mathbf{M}$. For a K-hop memory network, the external knowledge includes a set of trainable embedding matrices $\mathbf{C} = (\mathbf{C}^1, \ldots, \mathbf{C}^{K+1})$.

Task Formulation: KRT aims to predict knowledge retriever filter $\mathbf{G} = (g_1, \ldots, g_{b+T})$, which is used for filtering irrelevant knowledge information in the knowledge entity generation process. $\mathbf{G}$ can be calculated with a multi-hop memory network. We use the last hidden state of dialog history, $\mathbf{h}_m = \texttt{DialoGPT}(\mathbf{e}_m, \mathbf{h}_{m-1})$ where $\mathbf{e}_m$ denotes the input representation of m word, as the initial query vector $\mathbf{q}^1$. The k hop process is defined:

$$p_i^k = \text{Softmax}((\mathbf{q}^k)^\top \mathbf{c}_i^k), \tag{4}$$

$$\mathbf{o}^k = \sum\nolimits_i p_i^k \mathbf{c}_i^{k+1}, \tag{5}$$

$$\mathbf{q}^{k+1} = \mathbf{q}^k + \mathbf{o}^k, \tag{6}$$

where p_i^k is logits at the i-th position, $\mathbf{c}_i^k$ is embeddings in i-th memory position using $\mathbf{C}^k$, $\mathbf{o}^k$ denotes the weighted sum over $\mathbf{c}^{k+1}$ and $\mathbf{q}^{k+1}$ represents the updated query vector.

At the last K hop, each pointer value can be obtained by:

$$g_i = \text{Sigmoid}((\mathbf{q}^K)^\top \mathbf{c}_i^K), \tag{7}$$

where g_i represents the possibility of the i-th object word existing in the system response.

The intuition behind KRT pre-training is that a better $\mathbf{G}$ and external knowledge trainable parameters obtained from the pre-training stage can be better used to correctly retrieve knowledge in the KB for EToDs.

b) Consistency-Guided Data Augmentation: Unlike the pre-training process of the *generation* module that can use KB-free dialogues, only KB-grounded task-oriented dialogues can be used for pre-training the *knowledge-retriever* module, which is extremely hard to collect. To solve the issue of insufficient training data problem, we propose a novel *consistency-guided* data augmentation (CGDA) method to generate a considerable amount of pseudo knowledge-grounded dialogues.

The main idea of CGDA is to replace some specific knowledge entities existing in the dialogues and KB with other different knowledge entities to construct new pseudo dialogues. Unlike data augmentation in other areas, we require the augmented dialogues to follow two important general logical rules: (1) *dialog context consistency* and (2) *KB consistency*. Next, we describe the process of data augmentation and how to keep the *consistency*. As shown in Figure 3, the augmentation process includes three steps: (1) *Entity Detection*, (2) *Entity Type Filter* and (3) *Entity Value Replacement*:

- *Entity Detection*. We first detect all the entities and their entity types in the dialogue history. For example, the `gas station` and its entity type `POI type` are detected, which is shown in Figure 3(b).
- *Entity Type Filter*. We replace the selected entities. Here, we should not randomly select what entities to replace, since this may cause *dialog context inconsistency*. For example, if we replace the selected `gas station` with `a Chinese restaurant`, there is a contradiction between dialog context. This is because the replaced entity `Chinese restaurant` has a conflict with the driver's utterance *I want to fuel my car*. To alleviate this issue, we maintain an entity type set $\mathcal{E}_{\text{type}}$ [1] to ensure *dialog context consistency*.
- *Entity Value Replacement*. We perform entity value replacement. To keep *KB consistency*, we ensure that the entities used for replacement do not exist in the current KB. As shown in Figure 3(d), we replace `200 Alester Ave` with `481 Amaranta Ave`. However, *KB consistency* is violated when we replace `200 Alester Ave` with `271 Springer Street`, since it causes two different `POIs` to have the same `Address 271 Springer Street` in the augmented KB. This violates the common sense that the address attribute of each `POI` is unique in the world.

Algorithm 1: Consistence-guided data augmentation (CGDA) in knowledge-retriever module pre-training

Input: Seed KB-grounded dialogues: $\mathcal{D}_S$;
  Entity set $\mathbf{E}$;
  Candidate entity type set: $\mathcal{E}_{\text{type}}$;
  Entity dictionary: $\mathcal{E}_{\text{ent}}$: {entity type: entity set};
Output: Augmented KB-grounded Dialogues: $\mathcal{D}_K$.

1 $\mathcal{D}_K = \emptyset$;
2 **for** $\{\mathbf{X}, \mathbf{B}\}$ *in* $\mathcal{D}_S$ **do**
3   **if** $\exists$ *entity* e *in* D **then**
4     $\mathbf{E} \leftarrow$ Entity-Detection$(\mathbf{X}, \mathcal{E}_{\text{ent}})$;
5     $\tilde{\mathbf{E}} \leftarrow$ Entity-Type-Filter$(\mathbf{E}, \mathcal{E}_{\text{type}})$;
6   **end**
7   **if** $\tilde{\mathbf{E}} \neq \emptyset$ **then**
8     $\{\hat{\mathbf{X}}, \hat{\mathbf{B}}\} \leftarrow \{\mathbf{X}, \mathbf{B}\}$ ▷ *Init Augmented Data*
9   **for** e *in* $\tilde{\mathbf{E}}$ **do**
10      $t \leftarrow$ Entity-Type$(e, \mathcal{E}_{\text{ent}})$;
11      $\hat{e} \leftarrow$ Entity-Select$(\mathcal{E}_{\text{ent}}[t], \hat{\mathbf{B}})$;
12      $\{\hat{\mathbf{X}}, \hat{\mathbf{B}}\} \leftarrow$ Entity-Replace$(\hat{\mathbf{X}}, \hat{\mathbf{B}}, e, \hat{e})$;
13   **end**
14    $\mathcal{D}_K \leftarrow \mathcal{D}_K \cup \{\hat{\mathbf{X}}, \hat{\mathbf{B}}\}$;
15 **end**
16 **end**

Algorithm 1 shows the pseudocode for the consistence-guided data augmentation process, where lines 4-7 denote the entity detection step and entity type filter step, lines 8-16 denote the entity value replacement step.

The details of the mentioned function are as follows:

- The Entity-Detection function returns the entities $\mathbf{E}$ when they exist in the dialog history $\mathbf{X}$ based on the $\mathcal{E}_{\text{ent}}$, where $\mathcal{E}_{\text{ent}}$ represents the whole Knowledge entities. (E is a set). Specifically, we utilize the predefined entity dictionary to automatically extract the entities.
- The Entity-Type-Filter function filters entities in $\mathbf{E}$ if they violate *dialog context consistency* based on the $\mathcal{E}_{\text{type}}$.
- The Entity-Type function returns the entity type t of the input entity e based on the entity dictionary $\mathcal{E}_{\text{ent}}$.
- The Entity-Select function selects a random entity $\hat{e}$ from entity set $\mathcal{E}_{\text{ent}}[t]$, where the selected entity should not exist in the augmented knowledge base $\hat{\mathbf{B}}$ to maintain the *KB consistency*.

[1] In our work, entity type set $\mathcal{E}_{\text{type}}$ are {*party*, *room*, *agenda*, *location*, *address*}, {*address*, *area*, *id*, *location*, *phone*, *postcode*}, {*address*, *area*, *phone*, *postcode*, *ref*, *stars*} for SMD, Camrest676, and Multi-WOZ 2.1, respectively.

Dataset	Domains	Dialogues	Utterances
SMD	Navigate, Weather, Schedule	18,710	51,720
CamRest	Restaurant	4,040	16,620
MultiWOZ2.1	Restaurant, Attraction, Hotel	15,830	72,720

TABLE II: Data Statistics for Task-Oriented Dialogue Augmentation Data in Knowledge-retriever Module Pre-Training.

- The $\mathrm{Entity\text{-}Replace}$ function replaces the original entity e in the augmented dialog history $\hat{\mathbf{X}}$ and augmented knowledge base $\hat{\mathbf{B}}$ with the selected entity $\hat{e}$.

Finally, the augmented pseudo dialogues are used for pre-training the *knowledge-retriever* module.

The data statistics of the augmentation data obtained by our consistency-guided data augmentation method are shown in Table II. Each dialogue obtained through data augmentation has replaced all replaceable entities based on our consistency-guided data augmentation method.

C. EToDs Task Fine-tuning

Finally, after separately pre-training the *generation* module and *knowledge-retriever* module, we show how to fine-tune the pre-trained system on the EToDs tasks. Following Wu *et al.* [21] and Qin *et al.* [14], we use a sketch tag to control whether the model generates the common word or knowledge entity. Sketch tag denotes the possible slot types that start with a special token (e.g., *@Distance* stands for all the `2 miles`). When a sketch tag is yielded by the *generation* module, we query the knowledge entities from the KB.

1) Generation Module: We fine-tune the further pre-trained `DialoGPT` to generate response word by word, calculating:

$$\mathbf{h}_t = \mathtt{DialoGPT}(\mathbf{e}_t, \mathbf{h}_{t-1}), \tag{8}$$

where $\mathbf{h}_t$ is the hidden representations at t timestep.

$\mathbf{h}_t$ is used to generate the next token y_t:

$$\mathbf{o}_t = \mathbf{U}\,\mathbf{h}_t, \tag{9}$$

$$p(y_t) = \mathrm{Softmax}(\mathbf{o}_t), \tag{10}$$

where $\mathbf{U}$ represents the trainable projection matrix and $p(y_t)$ denotes the probability of token y_t.

2) Knowledge-retriever Module: When a sketch tag is generated, $\mathbf{h}_t$ is also used as query to retrieve knowledge from KB, which can be calculated by:

$$\mathbf{q}^1 = \mathbf{h}_t, \tag{11}$$

$$p_i^k = \mathrm{Softmax}((\mathbf{q}^k)^\top \mathbf{c}_i^k g_i). \tag{12}$$

Following Wu *et al.* [21] and Qin *et al.* [14] , at the last hop, $\mathbf{P}_t = (p_1^K, \ldots, p_{b+T}^K)$ is treated as the probabilities of knowledge entities at t timestep.

Hyperparameter Name	SMD	Camrest676	Multi-WOZ 2.1
Epoch	10	10	10
Batch Size	16	16	16
Hidden Size	256	128	128
Dropout Ratio	0.1	0.1	0.1
Warmup Steps	500	60	500
Learning Rate	7e-4	1e-3	1e-3
GM Learning Rate	4e-5	4e-5	7e-5

TABLE III: Hyperparameters used for all datasets. “GM” is short for Generation Module

D. Training Objective

1) Generation Module Pre-Training: LM and EP tasks are adopted to pre-train our *generation* module jointly and the loss is:

$$\mathcal{L}_G = \mathcal{L}_{LM} + \mathcal{L}_{EP}, \tag{13}$$

$$\mathcal{L}_{LM} = -\sum_{t=1}^{m+n} \hat{y}_t^{\mathrm{LM}} \log\left(y_t^{\mathrm{LM}}\right), \tag{14}$$

$$\begin{aligned}\mathcal{L}_{EP} = -\sum_{i=1}^{n} (\hat{y}_i^{\mathrm{EP}} \cdot \log\left(y_i^{\mathrm{EP}}\right) + \\ \left(1 - \hat{y}_i^{\mathrm{EP}}\right) \cdot \log\left(1 - y_i^{\mathrm{EP}}\right)),\end{aligned} \tag{15}$$

where $\hat{y}_t^{\mathrm{LM}}$ and $\hat{y}_t^{\mathrm{EP}}$ denote gold labels.

2) Knowledge-retriever Module Pre-training: The loss of the knowledge retriever task is:

$$\mathcal{L}_{GP} = -\sum_{i=1}^{b+T} (\hat{g}_i \cdot \log g_i + (1 - \hat{g}_i) \cdot \log(1 - g_i)), \tag{16}$$

where $\hat{g}_i$ denotes gold label and $\hat{g}_i$ = 1 if the i-th object word in the memory existing in the system response, 0 otherwise.

3) Fine-tuning Training Objective: Specifically, given the system response Y, we can get local memory pointer label sequence $\hat{\mathbf{L}} = (\hat{l}_1, \ldots, \hat{l}_n)$ as follows:

$$\hat{l}_t = \begin{cases} \max(z) & \text{if } \exists z \text{ s.t. } y_t = \mathrm{Object}(m_z) \\ b + T + 1 & \text{otherwise} \end{cases}, \tag{17}$$

where $\mathrm{Object}(\cdot)$ function is used for acquiring the object word from a triplet.

Based on the $\hat{\mathbf{L}}$ and $\mathbf{P}_t = (p_1^K, \ldots, p_{b+T}^K)$, we can calculate the standard cross-entropy loss $\mathcal{L}_{LP}$ as follows:

$$\mathcal{L}_{LP} = \sum_{t=1}^{n} -\log(\mathbf{P}_t(\hat{l}_t)). \tag{18}$$

The final training objective $\mathcal{L}$ used in the fine-tuning process is the weighted-sum of four losses:

$$\mathcal{L} = \gamma_{lm}\mathcal{L}_{LM} + \gamma_{ep}\mathcal{L}_{EP} + \gamma_{gp}\mathcal{L}_{GP} + \gamma_{lp}\mathcal{L}_{LP}, \tag{19}$$

where γ_{lm}, γ_{ep}, γ_{gp} and γ_{lp} are hyperparameters.

IV. Experiments

We evaluate our framework by fine-tuning the pre-trained *generation* and *knowledge-retriever* module for the end-to-end task-oriented dialogue system.

Model	SMD					Camrest676		Multi-WOZ 2.1				
	BLEU	F1	Navigate F1	Weather F1	Calendar F1	BLEU	F1	BLEU	F1	Restaurant F1	Attraction F1	Hotel F1
Non pre-trained Models												
Mem2Seq† [22]	12.6	33.4	20.0	32.8	49.3	16.6	42.4	5.8	14.4	19	18.5	8.2
GLMP† [21]	13.9	60.7	54.6	56.5	72.5	17.4	54.7	6.9	32.4	38.4	24.4	28.1
TTOS† [23]	17.4	55.4	45.9	64.1	63.5	20.5	61.5	-	-	-	-	-
DDMN† [24]	15.8	60.7	53.2	**64.7**	69.3	18.7	59.1	11.5	34.2	38.5	34.1	31.1
DF-Net† [14]	14.4	62.7	57.9	57.6	73.1	18.8	59.8	9.4	35.1	40.9	28.1	30.6
Pre-trained Models												
DialoGPT‡ [7]	14.6	47.0	32.2	55.6	52.8	12.3	44.4	7.0	15.9	20.9	13.9	11.6
DialoGPT+KB‡ [25]	16.5	59.2	51.2	53.1	73.3	12.7	51.6	8.4	29.0	32.9	35.6	24.7
Our framework	**18.8***	**63.8***	**59.1***	58.4	**75.0***	**22.0***	**63.9***	**13.6***	**36.3***	**41.5***	**36.2**	**31.2**

TABLE IV: Main results. The bolded number indicates the best performance and "-" indicates the original paper does not report results in the same dataset. Results with * indicate that the improvement of our framework is statistically significant with $p < 0.05$ under t-test. † results are cited from Qin *et al.* [14], He *et al.* [23] and Madotto *et al.* [26]. ‡ results: we adopt their open-sourced code to get the results.

A. Experimental Settings

We conduct experiments on three datasets, including SMD [11], CamRest [13] and an extension of Multi-WOZ 2.1 [14]. We follow same partition as Qin *et al.* [14],

The batch size we use in our framework is 16 and the dropout ratio is 0.1. We use AdamW to optimize the parameters in our model and adopt the suggested hyper-parameters for optimization. We adopt `DialoGPT-Medium` [7] architecture. All hyperparameters are selected according to the performance of the validation set. All experiments are conducted at Tesla A100 and V100. The detailed hyperparameters used for all datasets are shown in Table III.

B. Baselines

We compare our model with the following strong baselines: (1) `Mem2Seq` [22]: the model adopts a memory network to encode knowledge entities; (2) `GLMP` [21]: the framework adopts the global-to-local pointer mechanism to query the KB; (3) `TTOS` [23]: the model proposes a teacher-student framework to improve the generation and KB query ability; (4) `DDMN` [24]; the model uses a dual memory network to select better knowledge; (5) `DF-Net` [14]: the model considers domain features to promote the multi-domain EToDs, which obtains the best results; (6) `DialoGPT` [7]: the model directly yields response given the dialog history with pre-trained model `DialoGPT`; (7) `DialoGPT+KB`: the model first linearizes a table into a sequence [25]. Then it concatenates the linearized KB and the dialog history as input and directly models the task-oriented dialog task as the language modeling task with `DialoGPT` or `GPT2`;

V. Main Results

A. Automatic Evaluation

Following Wu *et al.* [21] and Qin *et al.* [14] , we adopt the *Micro-Entity F1* and *BLEU* [27] to evaluate the knowledge querying and fluent response generation ability, respectively.

The results are listed in Table IV. We have the following observations:

(1) On *BLEU*, the DialogGPT-based pre-traied models outperform some strong models without pretraining (i.e., `Mem2Seq`, `GLMP`, `DDMN` and `DF-Net`) revealing that pre-training from large amounts of dialogues can improve *fluency* performance, which is consistent with the observation in Peng *et al.* [28];

(2) On *Micro-Entity F1*, some non pre-trained models (e.g., `DF-Net`) outperform `DialoGPT-based` models (e.g., 62.7% vs. 59.2%). This indicates that simply fine-tuning the `DialoGPT-based` pre-trained models does not effectively integrate the KB query ability;

(3) Our framework not only outperforms the non pre-trained models but also the `DialoGPT-based` pre-trained models in both metrics. Compared with the best non pre-trained model `DF-Net`, our framework obtains 4.4 point and 1.1% improvements on BLEU and F1 scores, respectively. Our model also outperforms the best pre-trained model `DialoGPT + KB` by 2.3 point and 4.6% on BLEU and F1, respectively. The same trend is witnessed on other two datasets, which verifies the effectiveness of the modularized pre-training framework that can integrate fluent generation capability and accurate KB query capability.

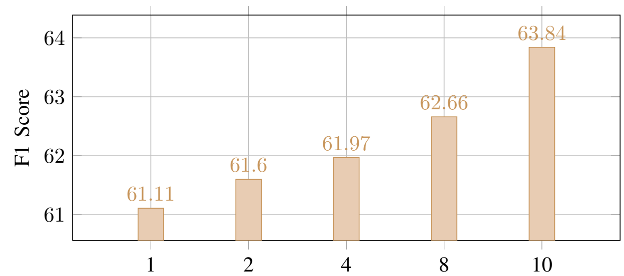


Fig. 4: Performance of different pseudo data sizes on SMD.

VI. Analysis

We conduct in-depth analysis to better understand our framework. Specifically, we first explore the effect of generation module pre-training mechanism. Next, we study the effect of the knowledge-retriever module pre-training mechanism. Then, we further explore the impact of modularized pre-training strategy and training data. We also investigate the

Model	*Generation* Module Pre-training	*Knowledge-retriever* Module Pre-training	SMD BLEU	SMD F1	SMD Navigate F1	SMD Weather F1	SMD Calendar F1	Camrest676 BLEU	Camrest676 F1	Multi-WOZ 2.1 BLEU	Multi-WOZ 2.1 F1	Multi-WOZ 2.1 Restaurant F1	Multi-WOZ 2.1 Attraction F1	Multi-WOZ 2.1 Hotel F1
Full model	✓	✓	**18.8**	**63.8**	**59.1**	**58.4**	**75.0**	**22.0**	**63.9**	**13.6**	**36.3**	**41.5**	36.2	31.2
w/o GP	✗	✓	18.1	61.6	53.1	58.2	74.4	20.3	62.0	13.0	34.5	37.2	38.2	31.4
w/o KRP	✓	✗	18.2	61.2	54.7	58.1	71.6	21.3	59.7	13.4	33.5	37.8	38.8	28.9
w/o MP	✗	✗	18.0	60.2	53.5	55.9	72.8	18.6	56.3	13.0	32.8	36.0	36.1	29.3

TABLE V: Ablation study. *GP*, *KRP* and *MP* are the shortcuts for *w/o Generation Module Pre-training*, *w/o Knowledge-retriever Module Pre-training* and *w/o Modularized Pre-training*.

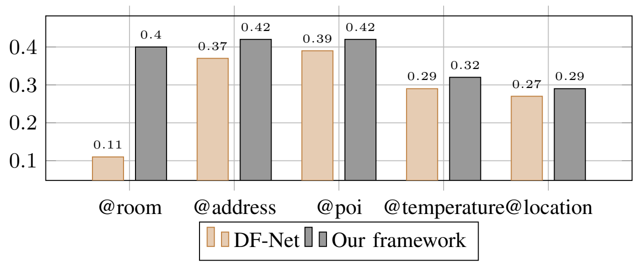


Fig. 5: Qualitative Analysis.

Model	BLEU	F1
DialoGPT + KB (w/ Pre-training Data)	14.6	59.2
Our framework (w/ Pre-training Data)	18.8	63.8

TABLE VI: Influence of Pre-training Data.

impact of CGDA method and give qualitative analysis to understand the proposed framework. Finally, we explore the effectiveness of our framework in low-resource setting.

A. Generation Module Pre-training Improves Fluency

To verify the effectiveness of the *generation* module pre-training, instead of further pre-training the `DialoGPT` model, we directly load the initial `DialoGPT` weights and then conduct the *knowledge-retriever* module pre-training. Finally, the pre-trained model is fine-tuned on the downstream task dataset.

The results are illustrated in Table V (*w/o GP* row). It can be seen that without *generation* module pre-training, the performance decreases especially on *BLEU* (-0.7 point on SMD, -1.7 point on Camrest676 and -0.6 point on Multi-WOZ2.1), indicating that further pre-training on the task-oriented dialogue can encourage our model to learn the characteristic of EToDs domain to improve fluency.

B. Knowledge-retriever Module Pre-training Boosts Accuracy

To analyze the *knowledge-retriever* module pre-training stage, we remove the *knowledge-retriever* module pre-training stage and utilize the model pre-trained with the *generation* module to fine-tune EToDs.

Table V (*w/o KRP* row) presents the results. We observe performance drops, especially on F1 scores, which demonstrates the *knowledge-retriever* module pre-training can learn better representation to enhance the ability to query KB to improve accuracy.

Model	Correct	Fluent	Human-like
DF-Net	4.07	4.29	4.26
DialoGPT + KB	4.22	4.33	4.27
Our framework	4.30	4.40	4.38
Agreement	65%	72%	73%

TABLE VII: Human evaluation on SMD. The agreement is calculated by Fleiss' Kappa [29].

C. Investigation of Modularized Pre-training

To analyze the effectiveness of the proposed modularized pre-training mechanism, we directly fine-tune our model on EToDs without the modularization pre-training strategy.

Table V shows the results. We observe without each pre-training module, the results drop a lot in terms of BLUE or F1 scores, which demonstrates the proposed modularized pre-training boosts both fluency and accuracy.

D. Impact of Training data

In addition, a natural question that arises is whether the collected pre-training data rather than the proposed modularized pre-training framework contributes to the final performance. To answer the question, we first pre-train the `DialoGPT+KB` model on the same KB-free task-oriented datasets. Then we fine-tune the model on the pseudo KB-grounded dialogues and SMD dataset.

As shown in Table VI, when these two models use the same data, our framework significantly outperforms the `DialoGPT+KB` model by 4.2 point and 4.6% on BLEU and F1. This further demonstrates that performance improvement comes from the proposed modularized pre-training rather than the pre-training data.

E. Impact of CGDA

To investigate the impact of CGDA, we conduct experiments with different amounts of pseudo dialogues (vary from 1 to 10 times) generated by CGDA.

From Figure 4, we can observe that more pseudo dialogues, the higher the F1 score, which indicates CGDA is essential for generating sufficient augmented training data to pre-train the *knowledge-retriever* module.

F. Qualitative Analysis

We calculate Entity F1 for each entity type to provide a qualitative analysis to help intuitively understand the effectiveness of the *knowledge-retriever* module pre-training.

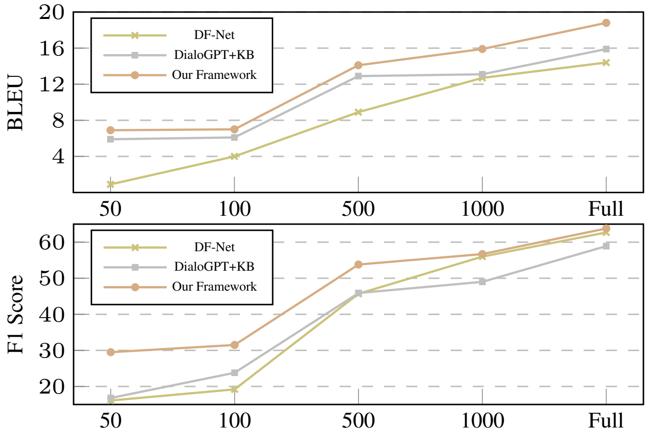


Fig. 6: Results of different data sizes on SMD dataset.

As shown in Figure 5, we select five entity types with the most significant performance improvement among all the entity types. In particular, the entity types "`room`", '`address`" and "`location`" are in the candidate entity type set $\mathcal{E}_{\text{type}}$, which further verifies that our *consistency-guided* data augmentation can effectively improve the accuracy performance of our model.

G. Exploration on Low-resource Settings

We explore whether our framework is effective in the low-resource settings, by randomly selecting dialogues in the training set varying from [50, 100, 500, 1000] to simulate the low-resource setting.

The results are shown in Figure 6. We find that our framework consistently outperforms `DF-Net` and `DialoGPT+KB` on all sizes, which verifies the robustness of our framework. Besides, our framework gives the largest improvement on extreme low-resource settings (50 dialogues), which makes it more scalable in real scenarios.

H. Human Evaluation

Following Qin *et al.* [14] , we perform human evaluation. by hiring human experts to judge the quality of the responses according to the correctness, fluency, and human-likeness on a scale from 1 to 5[2]. We randomly selected 100 dialogue histories on the SMD test data and evaluated the generated responses on our model, `DF-Net` and `DialoGPT + KB`. The experts judge the response generated by the anonymous system based on the given dialogue history, knowledge base and golden response.

Results are illustrated in Table VII. We observe our framework beats other baselines on all metrics, which is consistent with automatic evaluation.

VII. Related Work

Existing EToDs work can be classified into two main categories: (1) EToDs without any intermediate supervision and (2) EToDs without intermediate supervision. This section will describe the related work in detail.

[2]Annotators evaluate the correctness of the predicted response based on the ground-truth response and entity labels.

A. End-to-end Task-oriented Dialog without any Intermediate Supervision

The first strand of end-to-end task-oriented dialogue systems (EToDs) directly integrate a unified sequence-to-sequence model for generating system response given the dialogue history and the corresponding knowledge base. This line of work can reduce the efforts of annotating manually designed pipeline modules and easily be adapted to a new domain, which has attracted more and more attention. Our work follows this line of literature. Recently, some work use attention-based models [11], [30] to query the knowledge entity. Eric *et al.* [11] performs attention over KB to generate entities. Lei *et al.* [30] considers track dialogue beliefs in the end-to-end task-oriented dialog. Another series of work [21]–[24], [31]–[34] consider the memory network [35] to encode knowledge base. Madotto *et al.* [22] first combines end-to-end memory network [35] to consider KB query. Gangi *et al.* [32] proposes a multi-level memory architecture for end-to-end task oriented dialogue system. Wu *et al.* [21] proposes a global-to-locally pointer mechanism to query KB. Qin *et al.* [14] further incorporate domain-aware features for EToDs, achieving promising performance. Different from the above work, we explore the pre-training paradigm for EToDs and focus on how to improve both fluency and accuracy with a modularized pre-training framework while their work mainly considers improving performance in fine-tuning stage. To our knowledge, we are the first to explore modularized pre-training methods for EToDs.

B. End-to-end Task-oriented Dialog with Intermediate Supervision

This second strand of end-to-end task-oriented dialogue systems investigates to model of all pipeline sub-tasks dialogue state tracking, action, and response generation as a single sequence in a unified way [36]–[38]. While appealing, this line of literature requires intermediate supervision for training the whole system. Specifically, Peng *et al.* [28] introduce a transformer-based auto-regressive model to parameterize classical modular task-oriented dialog systems, which requires belief state annotation for supervision. Yang *et al.* [37] propose a GPT-based end-to-end task-oriented dialogue to generate belief states, the system acts and responses simultaneously. Su *et al.* [39] propose a unified model that is capable of supporting both task-oriented dialogue understanding and response generation. To this end, they introduce a novel multi-task pre-training strategy to enable the model to learn from different heterogeneous dialog corpora (e.g., NLU and DST). He *et al.* [40] take the first step to explicitly learn dialog policy by introducing a dialogue act prediction task, which achieves promising performance on the end-to-end dialogue tasks. The above work requires intermediate supervision (e.g., belief states or dialogue act annotation), which is not within the scope of our discussion. Compared with their work, we focus on pre-training for EToDs that directly operate on dialog history and interact with KB without any intermediate supervision.

VIII. Conclusion

In this paper, we proposed a modularized pre-training framework for EToDs, with the *generation* module and *knowledge-retriever* module. With modularization, our framework can devise various pre-training tasks to enhance the the two modules, achieving high fluency and accuracy. In addition, we proposed a simple yet effective consistency-guided data augmentation strategy to generate sufficient KB-grounded dialogues without breaking the dialogue context consistency and KB consistency. Experimental results on three datasets show that our framework achieves superior performance on both fluency and accuracy metrics. To the best of our knowledge, we are the first to explore the modularized pre-training approach for EToDs.

Acknowledgements

This work was supported by the National Key R&D Program of China via grant 2020AAA0106501 and the National Natural Science Foundation of China (NSFC) via grant 61976072 and 62176078.

References

[1] L. Qin, T. Xie, W. Che, and T. Liu, "A survey on spoken language understanding: Recent advances and new frontiers," in *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, Z.-H. Zhou, Ed. International Joint Conferences on Artificial Intelligence Organization, 8 2021, pp. 4577–4584, survey Track. [Online]. Available: https://doi.org/10.24963/ijcai.2021/622

[2] V. Balaraman, S. Sheikhalishahi, and B. Magnini, "Recent neural methods on dialogue state tracking for task-oriented dialogue systems: A survey," in *Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2021, pp. 239–251.

[3] L. Chen, Z. Chen, B. Tan, S. Long, M. Gasic, and K. Yu, "Agentgraph: Toward universal dialogue management with structured deep reinforcement learning," *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, vol. 27, no. 9, p. 1378–1391, sep 2019. [Online]. Available: https://doi.org/10.1109/TASLP.2019.2919872

[4] Y. Li, K. Yao, L. Qin, W. Che, X. Li, and T. Liu, "Slot-consistent NLG for task-oriented dialogue systems with iterative rectification network," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics, Jul. 2020, pp. 97–106. [Online]. Available: https://aclanthology.org/2020.acl-main.10

[5] M. Eric and C. Manning, "A copy-augmented sequence-to-sequence architecture gives good performance on task-oriented dialogue," in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Valencia, Spain: Association for Computational Linguistics, Apr. 2017, pp. 468–473. [Online]. Available: https://aclanthology.org/E17-2075

[6] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving language understanding by generative pre-training," 2018.

[7] Y. Zhang, S. Sun, M. Galley, Y.-C. Chen, C. Brockett, X. Gao, J. Gao, J. Liu, and B. Dolan, "DIALOGPT : Large-scale generative pre-training for conversational response generation," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. Online: Association for Computational Linguistics, Jul. 2020, pp. 270–278. [Online]. Available: https://aclanthology.org/2020.acl-demos.30

[8] S. Roller, E. Dinan, N. Goyal, D. Ju, M. Williamson, Y. Liu, J. Xu, M. Ott, E. M. Smith, Y.-L. Boureau, and J. Weston, "Recipes for building an open-domain chatbot," in *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, Apr. 2021.

[9] Z. Lin, B. Liu, A. Madotto, S. Moon, Z. Zhou, P. Crook, Z. Wang, Z. Yu, E. Cho, R. Subba, and P. Fung, "Zero-shot dialogue state tracking via cross-task transfer," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 7890–7900. [Online]. Available: https://aclanthology.org/2021.emnlp-main.622

[10] B. Peng, C. Zhu, C. Li, X. Li, J. Li, M. Zeng, and J. Gao, "Few-shot natural language generation for task-oriented dialog," in *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online: Association for Computational Linguistics, Nov. 2020, pp. 172–182. [Online]. Available: https://aclanthology.org/2020.findings-emnlp.17

[11] M. Eric, L. Krishnan, F. Charette, and C. D. Manning, "Key-value retrieval networks for task-oriented dialogue," in *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*. Saarbrücken, Germany: Association for Computational Linguistics, Aug. 2017, pp. 37–49. [Online]. Available: https://aclanthology.org/W17-5506

[12] C.-S. Wu, S. C. Hoi, R. Socher, and C. Xiong, "TOD-BERT: Pre-trained natural language understanding for task-oriented dialogue," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics, Nov. 2020, pp. 917–929. [Online]. Available: https://aclanthology.org/2020.emnlp-main.66

[13] T.-H. Wen, D. Vandyke, N. Mrkšić, M. Gašić, L. M. Rojas-Barahona, P.-H. Su, S. Ultes, and S. Young, "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*, Apr. 2017.

[14] L. Qin, X. Xu, W. Che, Y. Zhang, and T. Liu, "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*, Jul. 2020.

[15] L. Qin, Y. Liu, W. Che, H. Wen, Y. Li, and T. Liu, "Entity-consistent end-to-end task-oriented dialogue system with KB retriever," in *Proc. of EMNLP*, 2019.

[16] A. Rastogi, X. Zang, S. Sunkara, R. Gupta, and P. Khaitan, "Towards scalable multi-domain conversational agents: The schema-guided dialogue dataset," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, 2020, pp. 8689–8696.

[17] B. Byrne, K. Krishnamoorthi, C. Sankar, A. Neelakantan, B. Goodrich, D. Duckworth, S. Yavuz, A. Dubey, K.-Y. Kim, and A. Cedilnik, "Taskmaster-1: Toward a realistic and diverse dialog dataset," 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)*. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 4516–4525. [Online]. Available: https://aclanthology.org/D19-1459

[18] X. Li, S. Panda, J. J. Liu, and J. Gao, "Microsoft dialogue challenge: Building end-to-end task-completion dialogue systems," in *SLT 2018*, July 2018.

[19] N. Mrkšić, D. Ó Séaghdha, T.-H. Wen, B. Thomson, and S. Young, "Neural belief tracker: Data-driven dialogue state tracking," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Vancouver, Canada: Association for Computational Linguistics, Jul. 2017, pp. 1777–1788. [Online]. Available: https://aclanthology.org/P17-1163

[20] 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," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics, Oct.-Nov. 2018, pp. 5016–5026. [Online]. Available: https://aclanthology.org/D18-1547

[21] C. Wu, R. Socher, and C. Xiong, "Global-to-local memory pointer networks for task-oriented dialogue," in *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. [Online]. Available: https://openreview.net/forum?id=ryxnHhRqFm

[22] A. Madotto, C.-S. Wu, and P. Fung, "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)*. Melbourne, Australia: Association for Computational Linguistics, Jul. 2018, pp. 1468–1478. [Online]. Available: https://aclanthology.org/P18-1136

[23] W. He, M. Yang, R. Yan, C. Li, Y. Shen, and R. Xu, "Amalgamating knowledge from two teachers for task-oriented dialogue system with adversarial training," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics, Nov. 2020, pp. 3498–3507. [Online]. Available: https://aclanthology.org/2020.emnlp-main.281

[24] J. Wang, J. Liu, W. Bi, X. Liu, K. He, R. Xu, and M. Yang, "Dual dynamic memory network for end-to-end multi-turn task-oriented dialog

systems," in *Proceedings of the 28th International Conference on Computational Linguistics*, Dec. 2020.

[25] W. Chen, H. Wang, J. Chen, Y. Zhang, H. Wang, S. Li, X. Zhou, and W. Y. Wang, "Tabfact: A large-scale dataset for table-based fact verification," in *International Conference on Learning Representations (ICLR)*, Addis Ababa, Ethiopia, April 2020.

[26] A. Madotto, S. Cahyawijaya, G. I. Winata, Y. Xu, Z. Liu, Z. Lin, and P. Fung, "Learning knowledge bases with parameters for task-oriented dialogue systems," in *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online: Association for Computational Linguistics, Nov. 2020, pp. 2372–2394. [Online]. Available: https://aclanthology.org/2020.findings-emnlp.215

[27] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: a method for automatic evaluation of machine translation," in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. Philadelphia, Pennsylvania, USA: Association for Computational Linguistics, Jul. 2002, pp. 311–318. [Online]. Available: https://aclanthology.org/P02-1040

[28] 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," *Transactions of the Association for Computational Linguistics*, vol. 9, pp. 807–824, 2021. [Online]. Available: https://aclanthology.org/2021.tacl-1.49

[29] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.

[30] W. Lei, X. Jin, M.-Y. Kan, Z. Ren, X. He, and D. Yin, "Sequicity: Simplifying task-oriented dialogue systems with single sequence-to-sequence architectures," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Melbourne, Australia: Association for Computational Linguistics, Jul. 2018, pp. 1437–1447. [Online]. Available: https://aclanthology.org/P18-1133

[31] H. Wen, Y. Liu, W. Che, L. Qin, and T. Liu, "Sequence-to-sequence learning for task-oriented dialogue with dialogue state representation," in *Proc. of COLING*, 2018.

[32] R. Gangi Reddy, D. Contractor, D. Raghu, and S. Joshi, "Multi-level memory for task oriented dialogs," 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)*. Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 3744–3754. [Online]. Available: https://aclanthology.org/N19-1375

[33] Z. He, Y. He, Q. Wu, and J. Chen, "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)*, 2020.

[34] A. Madotto, Z. Lin, C.-S. Wu, J. Shin, and P. Fung, "Attention over parameters for dialogue systems," *arXiv preprint arXiv:2001.01871*, 2020.

[35] S. Sukhbaatar, A. Szlam, J. Weston, and R. Fergus, "End-to-end memory networks," in *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*, ser. NIPS'15. Cambridge, MA, USA: MIT Press, 2015, p. 2440–2448.

[36] D. Ham, J.-G. Lee, Y. Jang, and K.-E. Kim, "End-to-end neural pipeline for goal-oriented dialogue systems using GPT-2," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics, Jul. 2020, pp. 583–592. [Online]. Available: https://aclanthology.org/2020.acl-main.54

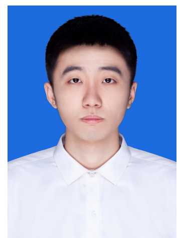
[37] Y. Yang, Y. Li, and X. Quan, "Ubar: Towards fully end-to-end task-oriented dialog system with gpt-2," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 16, 2021, pp. 14 230–14 238.

[38] E. Hosseini-Asl, B. McCann, C. Wu, S. Yavuz, and R. Socher, "A simple language model for task-oriented dialogue," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., 2020, pp. 20 179–20 191.

[39] 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 *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 4661–4676. [Online]. Available: https://aclanthology.org/2022.acl-long.319

[40] W. He, Y. Dai, Y. Zheng, Y. Wu, Z. Cao, D. Liu, P. Jiang, M. Yang, F. Huang, L. Si *et al.*, "Galaxy: A generative pre-trained model for task-oriented dialog with semi-supervised learning and explicit policy injection," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 10, 2022, pp. 10 749–10 757.

Libo Qin received his Ph.D. degree in computer science from the Harbin Institute of Technology (HIT), China. He is a Professor at Central South University. His current research interests include natural language processing and dialogue systems.



Xiao Xu is a Ph.D. candidate at the Harbin Institute of Technology (HIT), China. He received his B.S. degree in computer science from Northeastern University, in 2020. His research fields include natural language processing and dialog systems.

Lehan Wang is an undergraduate student in Harbin Institute of Technology (HIT), China. She worked as a research intern at Research Center for Social Computing and Information Retrieval (HIT-SCIR) from HIT. Her research fields include natural language processing and dialog systems.

Yue Zhang is currently a Professor at Westlake University. He has been working on statistical parsing, text synthesis, natural language synthesis, machine translation, information extraction, sentiment analysis and stock market analysis. His research interests include natural language processing and computational finance.

Wanxiang Che received his Ph.D. degree in computer science from the Harbin Institute of Technology (HIT), China, in 2008. He is a Full Professor in the School of Computer Science and Technology, HIT. His current research interests include natural language processing and dialogue systems.