Bridging the Gap between Prior and Posterior Knowledge Selection for Knowledge-Grounded Dialogue Generation

Xiuyi Chen[1,2,3], Fandong Meng[4], Peng Li[4], Feilong Chen[1,2,3], Shuang Xu[1], Bo Xu[1,2,3], Jie Zhou[4]
[1]Institute of Automation, Chinese Academy of Sciences (CASIA). Beijing, China
[2]Research Center for Brain-inspired Intelligence, CASIA
[3]University of Chinese Academy of Sciences, Beijing, China
{chenxiuyi2017,chenfeilong2018,shuang.xu,xubo}@ia.ac.cn
[4]Pattern Recognition Center, WeChat AI, Tencent Inc, Beijing, China
{fandongmeng,patrickpli,withtomzhou}@tencent.com

Abstract

Knowledge selection plays an important role in knowledge-grounded dialogue, which is a challenging task to generate more informative responses by leveraging external knowledge. Recently, latent variable models have been proposed to deal with the diversity of knowledge selection by using both prior and posterior distributions over knowledge and achieve promising performance. However, these models suffer from a huge gap between prior and posterior knowledge selection. Firstly, the prior selection module may not learn to select knowledge properly because of lacking the necessary posterior information. Secondly, latent variable models suffer from the exposure bias that dialogue generation is based on the knowledge selected from the posterior distribution at training but from the prior distribution at inference. Here, we deal with these issues on two aspects: (1) We enhance the prior selection module with the necessary posterior information obtained from the specially designed Posterior Information Prediction Module (PIPM); (2) We propose a Knowledge Distillation Based Training Strategy (KDBTS) to train the decoder with the knowledge selected from the prior distribution, removing the exposure bias of knowledge selection. Experimental results on two knowledge-grounded dialogue datasets show that both PIPM and KDBTS achieve performance improvement over the state-of-the-art latent variable model and their combination shows further improvement.

1 Introduction

Knowledge-grounded dialogue (Ghazvininejad et al., 2018) which leverages external knowledge to generate more informative responses, has become a popular research topic in recent years. Many researchers have studied how to effectively leverage the given knowledge to enhance dialogue understanding and/or improve dialogue generation (Zhao et al., 2019b; Sun et al., 2020; Madotto et al., 2018; Chen et al., 2019; Yavuz et al., 2019; Tang and Hu, 2019; Li et al., 2019; Zheng and Zhou, 2019; Niu et al., 2019; Meng et al., 2019; Ren et al., 2019; Ye et al., 2020). However, they usually use the pre-identified knowledge (Zhang et al., 2018; Moghe et al., 2018; Qin et al., 2019) which is not available in some real-world scenarios. And others leverage the retrieval system to get the knowledge which may contain noisy and irrelevant data (Chaudhuri et al., 2018; Parthasarathi and Pineau, 2018; Zhou et al., 2018; Gopalakrishnan et al., 2019). Recently, Dinan et al. (2019) propose to decompose this task into two subproblems: knowledge selection and response generation. This pipeline framework has been widely used for the open domain setting (Chen et al., 2017; Min et al., 2018; Nie et al., 2019) and shows promising performance with explicit use of knowledge in knowledge-grounded dialogue (Dinan et al., 2019).

Context	I just got a husky puppy.	
Knowledge Pool	#	**Knowledge Sentence**
	0	Husky is a general name for a sled type of dog used in northern regions, differentiated from other **sled-dog** types by **their fast pulling style**.
	1	Huskies are also today kept as pets, and groups work to find new pet homes for retired racing and adventure trekking dogs.
	2	Huskies are used in sled dog racing.
	3	The use of “husk” is recorded from 1852 for dogs kept by Inuit people
	...	...
	L	Child of the Wolves is a children’s **novel**, published in 1996, **about a Siberian husky puppy that joins a wolf pack**.
Response a	It sounds cute! Huskies are known amongst **sled-dogs for their fast pulling style**.	
Response b	It sounds cute! I have read a **novel about a husky puppy joining a wolf pack**. Is your husky puppy wolf-like in appearance?	

Table 1: An example shows the diversity of knowledge selection in knowledge-grounded dialogue. Here, we show two different responses with two possible selection decisions. For the same context, there may be diverse knowledge sentences to generate different responses which help their selection decisions in turn. The *prior knowledge selection* only depends on context while the *posterior knowledge selection* means selection with context and response (Lian et al., 2019).

Knowledge selection plays an important role in open-domain knowledge-grounded dialogue (Di-

Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pages 3426–3437,
November 16–20, 2020. ©2020 Association for Computational Linguistics

nan et al., 2019) since the inappropriate knowledge selection may prevent the model from leveraging the knowledge accurately (Lian et al., 2019), or even lead to an inappropriate response. The example in Table 1 shows two phenomena: (1) There may exist one-to-many relations between the dialogue context and the knowledge, resulting in the diversity of knowledge selection (Kim et al., 2020); (2) The posterior knowledge selection with context and response is much easier than the prior knowledge selection only depending on context. It is rather intuitive that we can dramatically reduce the scope of knowledge selection when we know the knowledge contained in the response, while such posterior information is not available at inference. Recently, latent variable models (Lian et al., 2019; Kim et al., 2020), using the posterior distribution to guide the prior distribution, have been proposed to deal with the diversity of knowledge selection. They can jointly model knowledge selection with response generation and achieve promising performance. Despite their success, latent variable models suffer from a huge gap between prior and posterior knowledge selection as discussed below.

We analyze the gap in latent variable models on two aspects: (1) Compared with the posterior selection module, the prior selection module has no access to the necessary posterior information. As a result, it is hard for the prior distribution to approximate the posterior distribution correctly at training, so that the prior selection module may not select knowledge properly at inference. (2) Response generation of latent variable models is based on the knowledge selected from the posterior distribution at training but from the prior distribution at inference. This discrepancy, also named exposure bias, leads to a gap between training and inference (Ranzato et al., 2015; Zhang et al., 2019; Zhao et al., 2019a), and therefore the decoder may have to generate a response with inappropriate knowledge selected from an unfamiliar prior distribution.

In this paper, we propose to bridge the gap between prior and posterior knowledge selection for knowledge-grounded dialogue generation. Firstly, we enhance the prior selection module with the necessary posterior information which is obtained by the specially designed Posterior Information Prediction Module (PIPM). Secondly, inspired by knowledge distillation (Hinton et al., 2015), we design a Knowledge Distillation Based Training Strategy (KDBTS) to train the decoder with the knowledge selected by the prior module, removing the exposure bias of knowledge selection. Experimental results show that both PIPM and KDBTS bring performance improvement on two knowledge-grounded dialogue datasets, i.e., Wizard of Wikipedia (Dinan et al., 2019) and Holl-E (Moghe et al., 2018). And the combination of PIPM and KDBTS obtains the new state-of-the-art performance with further improvement.

Our contributions are summarized as follows:[1]

- We clearly point out the gap between prior and posterior knowledge selection of latent variable models and propose to enhance the prior selection module with the necessary posterior information. Moreover, we explore several variants of posterior information.

- We focus on the exposure bias of knowledge selection for knowledge-grounded dialogue, and design a knowledge distillation based training strategy to deal with it.

- Experimental results show that both PIPM and KDBTS bring performance improvement, and their combination achieves the state-of-the-art performance with further improvement.

2 Background

2.1 Task Formulation

For a dialogue with T turns, each turn is a pair of message x_t and response y_t. Besides, each turn is associated with a knowledge pool $\mathcal{K}_t = \{k_t^l\} = \{k_t^1, \cdots, k_t^L\}$ consisting of L sentences. The context ctx_t consists of the dialogue history $hist_t = [x_1, y_1, \cdots, x_{t-1}, y_{t-1}]$ and the message x_t.

Given the context ctx_t, we firstly select a knowledge sentence $k_t^{\text{sel}} \in \mathcal{K}_t$ from the knowledge pool, then leverage the selected knowledge to generate an informative response y_t. The selection history at the t-th turn is $kh_t = [k_1^{\text{sel}}, \cdots, k_{t-1}^{\text{sel}}]$.

2.2 Latent Knowledge Selection For Response Generation

To obtain the likelihood of response y_t, latent variable models treat knowledge k as the latent variable and marginalize over all possible knowledges $\mathcal{K}_t$:

$$p\left(y_t|ctx_t\right) = \mathbb{E}_{k\sim\pi_\theta(\mathcal{K}_t)} p_\theta\left(y_t|ctx_t, k\right), \qquad (1)$$

[1] The code is available at https://github.com/youngornever/bridge_latent_knowledge_selection_gap_for_conversation.

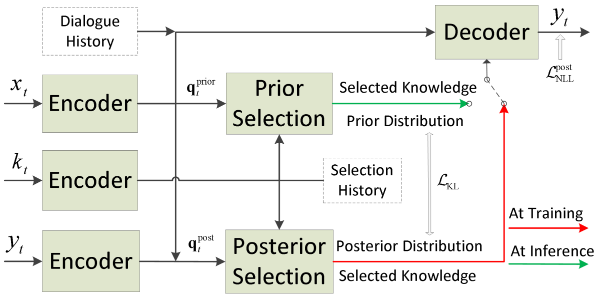

Figure 1: Latent knowledge selection for response generation. We train the decoder with the selected knowledge from posterior distribution with the red line, while we have to infer with the knowledge selected by the prior selection module without access to the response.

where $p_\theta(y_t|ctx_t,k)$ is the decoder network, and $\pi_\theta(\mathcal{K}_t)$, short for $\pi_\theta(\mathcal{K}_t|ctx_t,kh_t)$, is the prior distribution over the knowledge pool $\mathcal{K}_t$ based on the context ctx_t and selection history kh_t. The evidence lower bound (ELBO) is written as:

$$\begin{aligned}\mathcal{L}_{\text{ELBO}} &= -\mathcal{L}_{\text{NLL}}^{\text{post}} - \mathcal{L}_{\text{KL}} \leq \log p\left(y_t|ctx_t\right)\\ \mathcal{L}_{\text{NLL}}^{\text{post}} &= -\mathbb{E}_{k\sim q_\phi(\mathcal{K}_t)}\left[\log p_\theta\left(y_t|ctx_t,k\right)\right],\\ \mathcal{L}_{\text{KL}} &= D_{\text{KL}}\left(q_\phi\left(\mathcal{K}_t\right)\|\pi_\theta\left(\mathcal{K}_t\right)\right)\end{aligned}\tag{2}$$

where $q_\phi(\mathcal{K}_t)$, short for $q_\phi(\mathcal{K}_t|ctx_t,y_t,kh_t)$, is an inference network to approximate the true posterior distribution $p(\mathcal{K}_t|ctx_t,y_t,kh_t)$.

The Gap between Prior and Posterior Knowledge Selection. Firstly, compared with the posterior selection module, the prior selection module has no access to the posterior information as shown in Figure 1. As a result, it is hard for the prior distribution to approximate the posterior distribution correctly by minimizing the KL divergence in Equation 2 at training. Hence, the prior selection module may not select knowledge properly at inference. Secondly, comparing the two expectation terms in Equation 1 and 2, we see that the selected knowledge for response generation at training and inference is drawn from different distributions, i.e., the posterior distribution $k\sim q_\phi(\mathcal{K}_t)$ at training and the prior distribution $k\sim\pi_\theta(\mathcal{K}_t)$ at inference. Figure 1 clearly shows this discrepancy which will cause the decoder to have to generate with knowledge selected from the unfamiliar prior distribution. These issues lead to the gap between prior and posterior knowledge selection, which we try to deal with in this paper.

2.3 Sequential Knowledge Transformer

Recently, Kim et al. (2020) propose Sequential Knowledge Transformer (SKT), the state-of-the-art latent variable model for knowledge selection. Here, we briefly describe SKT, based on which we validate the effectiveness of our approach.

Sentence Encoding. For any sentence $sent_t$ with N_w words at t-th turn, SKT uses a shared BERT (Devlin et al., 2019) to obtain the context aware word representations $\mathbf{H}_t^{sent}$ with d dims and then converts them into the sentence representation $\mathbf{h}_t^{sent}$ by mean pooling (Cer et al., 2018):

$$\begin{aligned}\mathbf{H}_t^{sent} &= \text{BERT}\left(sent_t\right)\in\mathbb{R}^{N_w\times d}\\ \mathbf{h}_t^{sent} &= \text{Mean}\left(\mathbf{H}_t^{sent}\right)\in\mathbb{R}^{d}\end{aligned}.\tag{3}$$

As a result, we obtain $\mathbf{H}_t^x$ and $\mathbf{h}_t^x$ for the message x_t, $\mathbf{H}_t^y$ and $\mathbf{h}_t^y$ for the response[2] y_t, and $\mathbf{H}_t^{k_t^l}$ and $\mathbf{h}_t^{k_t^l}$ for any knowledge sentence $k_t^l\in\mathcal{K}_t$.

Knowledge Selection. To utilize the dialogue history and selection history, two GRUs (Cho et al., 2014) are used to summarize them as corresponding states $\mathbf{s}_t^{hist}$ and $\mathbf{s}_t^{kh}$ with zero initialization:

$$\begin{aligned}\mathbf{s}_t^{hist} &= \text{GRU}_{\text{dial}}\left(\left[\mathbf{h}_t^x;\mathbf{h}_t^y\right],\mathbf{s}_{t-1}^{hist}\right)\in\mathbb{R}^d\\ \mathbf{s}_t^{kh} &= \text{GRU}_{\text{sel}}\left(\mathbf{h}_t^{k_t^{\text{sel}}},\mathbf{s}_{t-1}^{kh}\right)\in\mathbb{R}^d\end{aligned},\tag{4}$$

where $\mathbf{h}_t^x$, $\mathbf{h}_t^y$ and $\mathbf{h}_t^{k_t^{\text{sel}}}$ are sentence vectors of message x_t, response y_t and the selected knowledge k_t^{sel} and $[\cdot;\cdot]$ denotes concatenation. Then, the prior query $\mathbf{q}_t^{\text{prior}}$ and posterior query $\mathbf{q}_t^{\text{post}}$ are obtained:

$$\begin{aligned}\mathbf{q}_t^{\text{prior}} &= \left[\mathbf{s}_{t-1}^{kh};\mathbf{s}_{t-1}^{hist};\mathbf{h}_t^x\right]\\ \mathbf{q}_t^{\text{post}} &= \left[\mathbf{s}_{t-1}^{kh};\mathbf{s}_t^{hist}\right]\end{aligned}.\tag{5}$$

Note that $\mathbf{s}_{t-1}^{kh}$ summarizes the selection history $kh_t=[k_1^{\text{sel}},\cdots,k_{t-1}^{\text{sel}}]$, and $\mathbf{s}_t^{hist}$ contains current response information, i.e., $\mathbf{h}_t^y$, not available at inference. The distribution $\mathbf{a}_t^{\text{P}}(\mathcal{K}_t)$ over knowledge pool $\mathcal{K}_t$ is obtained by the dot-product attention:

$$\begin{aligned}\mathbf{a}_t^{\text{P}}\left(\mathcal{K}_t\right) &= \text{softmax}\left(\mathbf{v}_t^{\text{P}}\left[\mathbf{h}_t^{k_t^1},\cdots,\mathbf{h}_t^{k_t^L}\right]\right)\\ \mathbf{v}_t^{\text{P}} &= \mathbf{W}^{\text{P}}\mathbf{q}_t^{\text{P}}\in\mathbb{R}^d\end{aligned},\tag{6}$$

where "P" denotes either prior or posterior, $\mathbf{W}_{\text{P}}$ is the projection matrix and $\mathbf{a}_t^{\text{P}}(\mathcal{K}_t)\in\mathbb{R}^L$.

[2] Response is only used for the posterior selection module at the training stage.

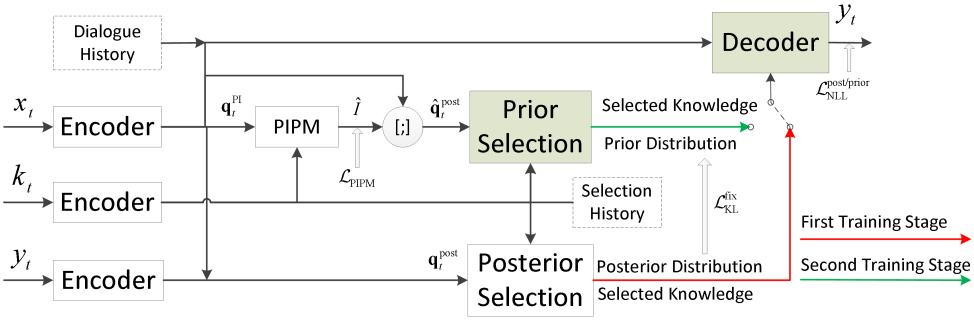

Figure 2: Our Framework. Firstly, we enhance the prior selection module with the necessary posterior information $\hat{I}$ obtained from the Posterior Information Prediction Module (PIPM). Secondly, we design a two-stage Knowledge Distillation Based Training Strategy (KDBTS). At the first stage, KDBTS trains the posterior selection module (the teacher) with the red line. At the second stage, response generation is based on the knowledge selected by the prior module, which is guided by the well-trained teacher, and we only update the green blocks at this stage.

Finally, the knowledge k_t^{sel} is selected by sampling from the posterior distribution $q_\phi\left(\mathcal{K}_t\right) = \mathbf{a}_t^{\text{post}}\left(\mathcal{K}_t\right)$ at training while selected with the highest probability over the prior distribution $\pi_\theta\left(\mathcal{K}_t\right) = \mathbf{a}_t^{\text{prior}}\left(\mathcal{K}_t\right)$ at inference.

Generation with Knowledge. SKT takes the concatenation of message x_t and selected knowledge sentence k_t^{sel} as input and generates responses by the Transformer decoder (Vaswani et al., 2017) with copy mechanism (Gu et al., 2016). Though there are various models studying how to improve the generation quality based on the given knowledge, here, we simply follow the decoder of SKT and mainly focus on the knowledge selection issue.

3 Approach

In this section, we show how to bridge the gap between prior and posterior knowledge selection in knowledge-grounded dialogue. Firstly, we design the Posterior Information Prediction Module (PIPM) to enhance the prior selection module with the necessary posterior information. Secondly, we design a Knowledge Distillation Based Training Strategy (KDBTS) to train the decoder with the knowledge selected from the prior distribution, removing the exposure bias of knowledge selection.

3.1 Posterior Information Prediction Module

As shown in Figure 2, we design a Posterior Information Prediction Module (PIPM) to predict the necessary posterior information. The main motivation is that we want to enhance the prior selection module with the necessary posterior information, so that it could approximate the posterior distribution better at training and leverage the posterior information for knowledge selection at inference. Following the typical setting in latent variable models (Lian et al., 2019; Kim et al., 2020), we use the response in bag-of-words (BOW) format (Zhao et al., 2017) as the posterior information. Here we take the dialogue context and the knowledge pool as input to generate the posterior information.

We firstly summarize the context as the query of this module $\mathbf{q}_t^{\text{PI}} = \left[\mathbf{s}_{t-1}^{hist}; \mathbf{h}_t^{x}\right]$ and use it to get the attention distribution $\mathbf{a}_t^{\text{PI}}\left(\mathcal{K}_t\right)$ over the knowledge pool $\mathcal{K}_t$ by Equation 6. Then, we summarize the knowledge representation in the knowledge pool with the weights in $\mathbf{a}_t^{\text{PI}}\left(\mathcal{K}_t\right)$ considered:

$$\mathbf{h}_t^{\text{PI}} = \left[\mathbf{h}_t^{k_t^1}, \cdots, \mathbf{h}_t^{k_t^L}\right] \cdot \mathbf{a}_t^{\text{PI}}\left(\mathcal{K}_t\right) \in \mathbb{R}^d. \quad (7)$$

Secondly, we take the summarization of the dialogue context and the knowledge pool as input and use a position wise feed-forward network (FFN) (Vaswani et al., 2017) to generate the posterior information $\hat{I}_t$ in BOW format as:

$$\hat{I}_t = \text{softmax}\left(\text{FFN}\left(\left[\mathbf{q}_t^{\text{PI}}; \mathbf{h}_t^{\text{PI}}\right]\right)\right) \in \mathbb{R}^{|V|}. \quad (8)$$

Finally, we use the generated posterior information $\hat{I}_t$ to obtain the updated prior query $\hat{\mathbf{q}}_t^{\text{post}}$ as:

$$\hat{\mathbf{q}}_t^{\text{post}} = \left[\mathbf{q}_t^{\text{prior}}; \hat{I}_t \cdot \mathbf{E}\right], \quad (9)$$

where $\mathbf{E} \in \mathbb{R}^{|V| \times d}$ is the embedding matrix and $|V|$ is the vocabulary size. Compared with $\mathbf{q}_t^{\text{prior}}$ in Equation 5, $\hat{\mathbf{q}}_t^{\text{post}}$ additionally contains the predicted posterior information, and we replace $\mathbf{q}_t^{\text{prior}}$ with $\hat{\mathbf{q}}_t^{\text{post}}$ for knowledge selection in Equation 6. We supervise this module by an addition loss $\mathcal{L}_{\text{PIPM}}$ with the grounded posterior information I, i.e., the bag of tokens in the golden response:

$$\mathcal{L}_{\text{PIPM}} = -1/|I| \sum\nolimits_{w \in I} \log\left(\hat{I}_t^w\right). \tag{10}$$

Note that we remove the context information from I because this information is already contained in the prior query $\mathbf{q}_t^{\text{prior}}$.

3.2 Knowledge Distillation Based Training Strategy

Current latent variable models (Lian et al., 2019; Kim et al., 2020) suffer from the exposure bias of knowledge selection as shown in Figure 1. Therefore, the decoder may have to generate a response with inappropriate knowledge selected from an unfamiliar prior distribution. Inspired by knowledge distillation (Hinton et al., 2015), we design the Knowledge Distillation Based Training Strategy (KDBTS) to deal with this exposure bias. KDBTS is a two-stage training strategy that we firstly train the posterior selection module as the teacher and then leverage the well-trained posterior module to train the prior selection module as the student.

First Training Stage. We train a teacher at this stage, which is used to guide the student at the next stage. We can obtain a teacher as the by-product, i.e., the posterior selection module, from the training process of latent variable models in Figure 1. However, we find the posterior selection module is affected by the prior distribution when minimizing the KL term in Equation 2, and experiments in Section 5.2 show that it is usually not good enough. As a result, we introduce a “fix” operation to make sure that the posterior selection module can not be affected by the prior distribution, and replace the KL term $\mathcal{L}_{\text{KL}}$ with the fixed KL term $\mathcal{L}_{\text{KL}}^{\text{fix}}$:

$$\mathcal{L}_{\text{KL}}^{\text{fix}} = D_{\text{KL}}\left(\text{fix}\left(q_\phi\left(\mathcal{K}_t\right)\right) \parallel \pi_\theta\left(\mathcal{K}_t\right)\right). \tag{11}$$

The total loss for training the teacher is as follows:

$$\mathcal{L}_1 = \mathcal{L}_{\text{NLL}}^{\text{post}} + \mathcal{L}_{\text{KL}}^{\text{fix}} - \log q_\phi\left(k_t^a\right) + \lambda \mathcal{L}_{\text{PIPM}}, \tag{12}$$

where $\mathcal{L}_{\text{NLL}}^{\text{post}}$ is defined in Equation 2, $k_t^a \in \mathcal{K}_t$ is the golden selected knowledge for knowledge loss $-\log q_\phi\left(k_t^a\right)$ which is proposed by (Kim et al., 2020) and λ is a hyperparameter.

Second Training Stage. Once we obtain the teacher, we could leverage the posterior distribution from the well-trained teacher to deal with the diversity of knowledge selection. At this training stage, we feed the knowledge selected by the prior module into the decoder as shown in Figure 2, which is the same as the inference process. In this way, KDBTS removes the exposure bias of knowledge selection. Here, we only update the prior selection module and the decoder (the green blocks in Figure 2) because the encoder is shared by the student and the teacher. The total loss for training the student at this stage is:

$$\mathcal{L}_2 = \mathcal{L}_{\text{NLL}}^{\text{prior}} + \mathcal{L}_{\text{KL}}^{\text{fix}} - \log \pi_\theta\left(k_t^a\right), \tag{13}$$

where $\mathcal{L}_{\text{NLL}}^{\text{prior}} = -\mathbb{E}_{k \sim \pi_\theta(\mathcal{K}_t)}\left[\log p_\theta\left(y_t | ctx_t, k\right)\right]$ and $-\log \pi_\theta\left(k_t^a\right)$ is the knowledge loss.

Compared with $\mathcal{L}_{\text{NLL}}^{\text{post}}$ defined in Equation 2, $\mathcal{L}_{\text{NLL}}^{\text{prior}}$ optimizes the decoder with knowledge selected from the prior distribution. Figure 2 clearly shows that KDBTS removes the exposure bias of knowledge selection by feeding the knowledge selected from prior distribution into the decoder.

4 Experiments

4.1 Dataset

We adopt two multi-turn knowledge-grounded dialogue datasets for experiments.

Wizard of Wikipedia (Dinan et al., 2019) contains $18,430$ training dialogues, $1,948$ validation dialogues and $1,933$ test dialogues on 1365 topics. And test set is split into two subsets according to topics, which are Test Seen with 965 dialogues and Test Unseen with 968 dialogues whose topics are never seen in training and validation set. There are about 61 sentences on average in the knowledge pool per turn, which are retrieved from Wikipedia based on the context.

Holl-E (Moghe et al., 2018) contains $7,228$ training dialogues, 930 validation dialogues and 913 test dialogues. There are two versions of the test set: one with a single golden reference, the other with multiple golden references. Each dialogue is assigned with a document of about 60 sentences on average as the knowledge pool. Here, we use the modified version (Kim et al., 2020) which fits for knowledge selection.

4.2 Baseline Models

We compare our approach with a set of competitive baselines:

TMN is short for E2E Transformer MemNet (Dinan et al., 2019). TMN uses a Transformer memory network for knowledge selection and a Transformer decoder with copy mechanism for utterance prediction. Knowledge selection is trained based on the knowledge label without posterior distribution.

TMN$_{\text{BERT}}$, is short for TMN+BERT, implemented by (Kim et al., 2020). TMN$_{\text{BERT}}$ replaces the Transformer memory network with a pre-trained BERT.

PostKS (Lian et al., 2019) only uses the posterior knowledge distribution as a pseudo-label for knowledge selection. PostKS uses GRU-based encoder and decoder without copy mechanism and does not use the knowledge label at the training stage.

TMN$_{\text{BERT+PostKS+CP}}$ is short for TMN+BERT+PostKS+Copy, implemented by (Kim et al., 2020). Compared with TMN$_{\text{BERT}}$, it additionally uses the the posterior distribution in PostKS.

SKT (Kim et al., 2020) is the current state-of-the-art latent variable model. Compared with TMN$_{\text{BERT}}$+PostKS$_{\text{cp}}$, SKT leverages the posterior distribution by sequential latent modeling.

4.3 Implementation Details

We validate the effectiveness of our approach based on current state-of-the-art model SKT (Kim et al., 2020), using the same datasets and data-processing codes[3]. A shared encoder initialized with BERT$_{\text{BASE}}$ (Devlin et al., 2019) is used to encode dialogue and knowledge sentences. Then, a 5-layer Transformer decoder with copy mechanism is used to generate the response. We use a FFN with 512 hidden dims to generate the posterior information in BOW formats. The hidden size d is 768 and the vocabulary size $|V|$ is $30,522$. Each batch consists of dialogues rather than individual turns, and the batch size is 1. The hyperparameter λ in Equation 13 is set to 0.5 for all experiments without searching. The "fix" operation in Equation 11 is implemented by gradient stoping.

Models are trained end-to-end using the Adam optimizer (Kingma and Ba, 2014) with gradient clipping at 0.4 and the learning rate is 0.00002. And we apply label smoothing (Pereyra et al., 2017) and set 0.1 for knowledge selection and 0.05 for response generation. We approximate the expectation in Equation 1 and 2 by drawing one sample with Gumbel-Softmax function (Jang et al., 2016) with temperature $\tau = 0.1$. We train our teacher with 5 epochs, then select the teacher to teach the student according to the prior knowledge selection accuracy rather than posterior selection accuracy, because the shared encoder and decoder may be optimized overly for the posterior selection module, which is not a good initialization for the student. We train other models 20 epoches and select them according to the R-1 score as the final goal is to generate high-quality responses.

4.4 Evaluation

Automatic Evaluation. Following (Dinan et al., 2019; Kim et al., 2020), we use accuracy (Acc) to evaluate the knowledge selection and use perplexity (PPL), unigram F1 (R-1) and bigram F1 (R-2) to evaluate the quality of response generation automatically. Following (Kim et al., 2020), we remove all the punctuation and (a, an, the) before computing the R-1 and R-2 scores. Note that lower perplexity and higher R-1 and R-2 indicate better generation quality.

Human Evaluation. We firstly select 100 samples from each test set on the Wizard of Wikipedia dataset for human evaluation. Then, we ask three annotators to judge whether the response makes sense (Sensibleness) or is specific (Specificity) with the dialogue context. Finally, we obtain Sensibleness and Specificity Average (SSA) scores, which could penalize boring and vague responses and is suitable for the goal of knowledge-grounded dialogue (Adiwardana et al., 2020). Moreover, compared with Engagingness and Knowledgeability used in (Dinan et al., 2019; Kim et al., 2020), SSA, evaluated in 0/1 format, is more objective and easier to conduct (Adiwardana et al., 2020).

5 Results and Analysis

5.1 Main Results

Quantitative Results. We report automatic results on the Wizard of Wikipedia dataset in Table 2 and we have the following observations: (1) From row 4 and 5 we can see that the PIPM indeed provides some necessary posterior information which is helpful for knowledge selection. (2) Comparing row 4 and 6, we see that the KDBTS brings about significant improvement on generation quality by removing the exposure bias of knowledge selection. (3) The combination of PIPM and KDBTS achieves further improvement on most metrics, except the knowledge selection accuracy. We think the reason is that there may exist several similar knowledge sentences leading to the same response. As a result, SKT+PIPM+KDBTS may select a reasonable

[3] We thank the authors for releasing their code and datasets at https://github.com/bckim92/sequential-knowledge-transformer.

#	Method	Test Seen				Test Unseen			
		Acc ↑	PPL ↓	R-1 ↑	R-2 ↑	Acc ↑	PPL ↓	R-1 ↑	R-2 ↑
0	TMN	22.5	63.5	16.9	NA	12.2	97.3	14.4	NA
1	PostKS	4.8	79.1	13.0	1.0	4.2	193.8	13.1	1.0
2	TMN$_{\text{BERT}}$	23.7	53.5	16.8	4.5	13.6	105.7	13.5	2.2
3	TMN$_{\text{BERT+PostKS+CP}}$	25.5	52.2	19.0	6.5	14.4	83.4	15.6	3.9
4	SKT	26.8	52.0	19.3	6.8	18.3	81.4	16.1	4.2
5	SKT+PIPM (Ours)	**27.9**	53.1	19.6	7.0	19.6	78.5	17.0	4.9
6	SKT+KDBTS (Ours)	27.3	46.0	**19.9**	**7.3**	**20.6**	68.8	17.2	5.1
7	SKT+PIPM+KDBTS (Ours)	27.7	**42.7**	**19.9**	**7.3**	19.4	**65.7**	**17.6**	**5.4**

Table 2: Quantitative results on the Wizard of Wikipedia dataset. Both PIPM and KDBTS bring about improvement over the state-of-the-art model (SKT) and their combination obtains further improvement on most metrics.

#	Method	Single Reference				Multi Reference			
		Acc ↑	PPL ↓	R-1 ↑	R-2 ↑	Acc ↑	PPL ↓	R-1 ↑	R-2 ↑
0	TMN	22.7	140.6	20.1	10.3	32.3	83.6	24.3	12.8
1	PostKS	1.5	196.6	15.2	6.0	3.2	114.1	19.2	7.9
2	TMN$_{\text{BERT}}$	28.2	112.6	25.9	18.3	37.5	66.9	31.1	22.7
3	TMN$_{\text{BERT+PostKS+CP}}$	27.8	47.4	29.2	22.3	37.8	27.9	35.9	29.0
4	SKT	29.2	48.9	29.8	23.1	39.2	28.5	36.5	29.7
5	SKT+PIPM (Ours)	**31.0**	47.5	30.2	23.8	**41.1**	27.0	37.4	**31.0**
6	SKT+KDBTS (Ours)	30.3	39.7	**30.8**	23.8	40.1	23.2	37.4	30.5
7	SKT+PIPM+KDBTS (Ours)	30.6	**39.2**	**30.8**	**23.9**	40.6	**23.1**	**37.7**	30.7

Table 3: Quantitative results on the Holl-E dataset.

Method	Test Seen	Test Unseen
SKT	48.2 (52.7&43.7)	45.5 (50.7&40.3)
Ours	60.8 (65.0&56.7)	55.2 (61.3&50.0)
Human	81.3 (81.7&81.0)	80.7 (81.0&80.3)

Table 4: Qualitative results on Wizard of Wikipedia. Ours denotes SKT+PIPM+KDBTS. The scores in each cell are "SSA (Sensibleness&Specificity)".

Name	**Details**
$I = y_{\text{BOW}}$	bag-of-word information in the response, i.e., the words.
$I = y^x_{\text{BOW}}$	y_{BOW} - x_{BOW}
$I = yk^x_{\text{BOW}}$	y_{BOW} + k^a_{BOW} - x_{BOW}
$I = y_{\text{Seq}}$	the response

Table 5: Several variants of posterior information.

knowledge but not the golden one to generate informative and fluent responses, because models are selected according to the generation quality rather than the selection accuracy (Kim et al., 2020).

Results in Table 3 lead to consistent observations on the Holl-E dataset that both PIPM and KDBTS bring performance improvement over the state-of-the-art latent variable model SKT, and their combination achieves further improvement, resulting in the new state-of-the-art performance.

Qualitative Results. We report human evaluation results of the generated responses in Table 4. We see that our approach brings about consistent improvement over the state-of-the-art model SKT. Our approach can leverage the selected knowledge appropriately to generate more sensible and specific responses, which are fluent and informative.

5.2 Analysis

PIPM. We explore several variants of posterior information in Tabel 5 to better study this module. Besides the default y^x_{BOW}, we also explore two variants in BOW format: (1) y_{BOW} does not remove the context information; (2) yk^x_{BOW} additionally considers another source of posterior information, i.e., the golden selected knowledge k^a_t. And we consider the sequence information in y_{Seq}, as the BOW format discards the word order. Note that we use the FFN in Equation 8 to obtain the posterior information in BOW formats, and we take $\mathbf{H}^x_t + \mathbf{h}^{\text{PI}}_t$ as input and use a 3-layer Transformer decoder to generate y_{Seq}. Moreover, we also perform an ablation study to investigate the effectiveness

#	Method	Test Seen				Test Unseen				# Parameters
		Acc ↑	PPL ↓	R-1 ↑	R-2 ↑	Acc ↑	PPL ↓	R-1 ↑	R-2 ↑	
0	SKT	26.8	52.0	19.3	6.8	18.3	81.4	16.1	4.2	174056704
1	SKT+PIPM (y_{Seq})	27.9	51.9	19.1	6.7	18.8	77.5	16.0	4.0	192361216
2	SKT+PIPM (y_{BOW})	27.3	51.5	19.3	7.0	18.4	78.0	16.1	4.3	192664890
3	SKT+PIPM (yk^{x}_{BOW})	27.6	52.2	19.0	6.8	19.8	78.1	16.3	4.6	192664890
4	SKT+PIPM (y^{x}_{BOW})	27.9	53.1	19.6	7.0	19.6	78.5	17.0	4.9	192664890
5	SKT+PIPM (w/o update)	26.9	51.8	19.3	7.2	19.0	77.0	16.6	4.7	192075066
6	SKT† — teacher	32.8	47.4	20.4	7.5	20.2	75.0	17.0	4.8	174056704
7	SKT (KLfix)† — teacher	52.0	43.1	24.1	10.4	38.3	61.6	20.3	7.1	174056704
8	SKT (KLfix)	28.0	73.3	18.5	6.5	20.6	99.3	15.8	4.4	174056704
9	SKT+KDBTS	27.4	45.9	20.0	7.3	20.5	68.6	17.2	5.1	174056704

Table 6: Quantitative results for model analysis on the Wizard of Wikipedia dataset. The default setting of PIPM is y^{x}_{BOW}. Models with † are teachers which are not comparable as they use the posterior knowledge selection at inference, and their results can be regarded as the upper bound. KLfix means that we replace $\mathcal{L}_{\text{KL}}$ with $\mathcal{L}^{\text{fix}}_{\text{KL}}$.

of generated posterior information for prior knowledge selection. We remove the predicted posterior information $\hat{I}$ in Equation 9, but still use $\mathcal{L}_{\text{PIPM}}$ in Equation 10 for comparison.

These results are reported in the upper part (row $1\sim5$) of Table 6 and the observations are stated as follows: (1) Compared with SKT in row 0, the variants of posterior information in row $1\sim4$ bring improvement on selection accuracy, though some generation metrics are slightly lower because of the exposure bias of knowledge selection. (2) From row 1 and 2, we see that the sequence information in y_{Seq} contributes to the knowledge selection. However, it is inefficient to generate y_{Seq} word by word, and y_{Seq} is not better than y^{x}_{BOW} in row 4. And there is no significant difference between y^{x}_{BOW} and yk^{x}_{BOW} which combines another source of posterior information. (3) We report the ablation result in row 5 to investigate the effectiveness of generated posterior information for knowledge selection. We see that this information improves the selection accuracy (compared with row 4).

KDBTS. We investigate the training strategy in the lower part (row $6\sim9$) of Table 6 and the observations are stated as follows: (1) Comparing row 6 and 7, we see that SKT (KLfix)† is a good teacher with a much higher selection accuracy because $\mathcal{L}^{\text{fix}}_{\text{KL}}$ in Equation 11 guarantees the posterior selection module not affected by the prior distribution. When using $\mathcal{L}_{\text{KL}}$ in Equation 2, SKT† achieves the lower accuracy (32.8 vs 52.0) and the lower KL divergence (0.31 vs 1.41), which indicates that the posterior module is affected by the prior distribution. (2) Despite doing well in knowledge selection, SKT (KLfix) in row 8 performs worse in generation than SKT. Because SKT (KLfix) has a larger KL divergence (1.41 vs 0.31) than SKT[4], it has to generate the response with the knowledge selected from a much more unfamiliar prior distribution. (3) As our KDBTS in row 9 does not suffer from the exposure bias of knowledge selection, the generation quality is improved significantly. We are amazed at finding that the selection accuracy is also improved and we argue that the KDBTS drives the prior distribution to approximate the posterior distribution at most by minimizing the KL term with the posterior distribution fixed.

5.3 Case Study

Table 7 shows two example response on the Wizard of Wikipedia dataset. We see that the posterior information $\hat{I}$, e.g., "rock band" and "formed", provided by the PIPM, is helpful for knowledge selection, although there is some noise information, e.g., "american". And quantitative results in Section 5.1 show that this module is effective although not perfect. We see that the responses generated by SKT have repeated words in case 1 and do not reply directly in case 2. Our method can leverage the selected knowledge sufficiently to generate a fluent and informative response because our KDBTS removes the exposure bias of knowledge selection by training the decoder with knowledge selected in the same way as the inference process.

[4] Note that SKT and SKT† are the same network tested with prior and posterior knowledge selection, respectively. Hence, the KL divergence is same.

Case 1	**Context**	I used to listen to the rolling stones a lot when i was a child.	
	Golden Knowledge	The rolling stones are an English **rock band** **formed** in london, england in 1962.	
	Human	Me too. I can't believe they formed in London as far back as 1962! what's your favorite song?	
	SKT	The rolling stones are an English rock band formed in London in London.	
	Ours	Posterior Information in BOW format (Top 20)	[PAD] in the **rock band** , they **formed** is are ? of a you i ' ! from and american
		Response	I love the rolling stones! They are an English rock band formed in 1962.
Case 2	**Context**	How do teams score points?	
	Golden Knowledge	**Points** are primarily **scored** by advancing the **ball** into the opposing **team's** end zone for a touchdown or kicking the **ball** through the opponent.	
	Human	Points are scored by advancing the ball into the opposing teams end zone for a touchdown or kicking it through the opponents goalposts for a field goal. good question!	
	SKT	The object is to score by getting the ball into the opposing goal.	
	Ours	Posterior Information in BOW format (Top 20)	the players [PAD] of , is goal a **ball** and to **team** on it each in **score points** most field
		Response	Points are scored by advancing the ball into the opposing team's end zone.

Table 7: Examples of generated responses on the Wizard of Wikipedia dataset.

6 Related Work

We mainly focus on knowledge selection in open domain knowledge-grounded dialogue, and there are several work studying this issue (Dinan et al., 2019; Lian et al., 2019; Kim et al., 2020). Here, we point out the gap between prior and posterior knowledge selection and try to deal with this gap.

The PIPM has some relations with several work. Deliberation Decoder (Xia et al., 2017; Wang et al., 2019; Li et al., 2019) leverages two decoders for two-pass generation. We also use two decoders in PIPM (y_{Seq}), but the first decoder is used to generate posterior information for the knowledge selection rather than the second-pass generation. BOW loss, proposed by (Zhao et al., 2017), is adopted to supervise the posterior module (Lian et al., 2019). Here, we have three different aspects: (1) We use the BOW loss for the prior module rather than the posterior module; (2) We use posterior information in BOW format to enhance the prior selection module; (3) We explore several BOW variants.

Our KDBTS is inspired by knowledge distillation (Hinton et al., 2015). Instead of using more complex structure as the teacher for model compression, we treat the posterior selection module with additional input information (e,g., the response) as the teacher, and deal with the exposure bias of knowledge selection. Lite ELBO (Zhao et al., 2019a) is proposed to remove the exposure bias at latent space for a different task. However, Lite ELBO naturally does not leverage the posterior distribution as it sets the posterior module the same as the prior module. Our KDBTS is a two-stage training strategy that uses the posterior distribution as the teacher to guide the prior selection module and uses the knowledge selected from the prior distribution to train the decoder.

7 Conclusion

In this paper, we firstly analyze the gap between prior and posterior knowledge selection for open-domain knowledge-grounded dialogue. Then, we deal with it on two aspects: (1) We enhance the prior selection module with the posterior information obtained by the PIPM and we explore several variants of posterior information. (2) We design the KDBTS to train the decoder with knowledge selected from the prior distribution, removing the exposure bias of knowledge selection. Experiments show that both PIPM and KDBTS improve the state-of-the-art latent variable model and their combination achieves further improvement. In the future, we would explore three aspects: (1) more efficient posterior information representation and corresponding prediction module, (2) the interpretability of knowledge selection and (3) knowledge selection without knowledge label.

Acknowledgments

We would like to thank the anonymous reviewers for their insightful comments. This work was supported by the Key Programs of the Chinese Academy of Sciences (ZDBS-SSW-JSC006) and the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB32070000).

References

Daniel Adiwardana, Minh-Thang Luong, David R So, Jamie Hall, Noah Fiedel, Romal Thoppilan, Zi Yang, Apoorv Kulshreshtha, Gaurav Nemade, Yifeng Lu, et al. 2020. Towards a human-like open-domain chatbot. *arXiv preprint arXiv:2001.09977*.

Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, et al. 2018. Universal sentence encoder. *arXiv preprint arXiv:1803.11175*.

Debanjan Chaudhuri, Agustinus Kristiadi, Jens Lehmann, and Asja Fischer. 2018. Improving Response Selection in Multi-Turn Dialogue Systems by Incorporating Domain Knowledge. In *Proceedings of the 22nd Conference on Computational Natural Language Learning*, pages 497–507.

Danqi Chen, Adam Fisch, Jason Weston, and Antoine Bordes. 2017. Reading Wikipedia to Answer Open-Domain Questions. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1870–1879.

Xiuyi Chen, Jiaming Xu, and Bo Xu. 2019. A working memory model for task-oriented dialog response generation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2687–2693, Florence, Italy. Association for Computational Linguistics.

Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734.

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

Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. 2019. Wizard of Wikipedia: Knowledge-Powered Conversational Agents. In *International Conference on Learning Representations*.

Marjan Ghazvininejad, Chris Brockett, Ming-Wei Chang, Bill Dolan, Jianfeng Gao, Wen-tau Yih, and Michel Galley. 2018. A knowledge-grounded neural conversation model. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Karthik Gopalakrishnan, Behnam Hedayatnia, Qinlang Chen, Anna Gottardi, Sanjeev Kwatra, Anu Venkatesh, Raefer Gabriel, and Dilek Hakkani-Tür. 2019. Topical-Chat: Towards Knowledge-Grounded Open-Domain Conversations. *Proc. Interspeech 2019*, pages 1891–1895.

Jiatao Gu, Zhengdong Lu, Hang Li, and Victor O.K. Li. 2016. Incorporating copying mechanism in sequence-to-sequence learning. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1631–1640, Berlin, Germany. Association for Computational Linguistics.

Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. 2015. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*.

Eric Jang, Shixiang Gu, and Ben Poole. 2016. Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144*.

Byeongchang Kim, Jaewoo Ahn, and Gunhee Kim. 2020. Sequential Latent Knowledge Selection for Knowledge-Grounded Dialogue. *arXiv: Computation and Language*.

Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Zekang Li, Cheng Niu, Fandong Meng, Yang Feng, Qian Li, and Jie Zhou. 2019. Incremental Transformer with Deliberation Decoder for Document Grounded Conversations. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 12–21.

Rongzhong Lian, Min Xie, Fan Wang, Jinhua Peng, and Hua Wu. 2019. Learning to select knowledge for response generation in dialog systems. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 5081–5087. AAAI Press.

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

Chuan Meng, Pengjie Ren, Zhumin Chen, Christof Monz, Jun Ma, and Maarten de Rijke. 2019. RefNet: A reference-aware network for background based conversation. *arXiv preprint arXiv:1908.06449*.

Sewon Min, Victor Zhong, Richard Socher, and Caiming Xiong. 2018. Efficient and Robust Question Answering from Minimal Context over Documents. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1725–1735.

Nikita Moghe, Siddhartha Arora, Suman Banerjee, and Mitesh M Khapra. 2018. Towards Exploiting Background Knowledge for Building Conversation Systems. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2322–2332.

Yixin Nie, Songhe Wang, and Mohit Bansal. 2019. Revealing the Importance of Semantic Retrieval for Machine Reading at Scale. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 2553–2566.

Zheng-Yu Niu, Hua Wu, Haifeng Wang, et al. 2019. Knowledge Aware Conversation Generation with Explainable Reasoning over Augmented Graphs. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 1782–1792.

Prasanna Parthasarathi and Joelle Pineau. 2018. Extending Neural Generative Conversational Model using External Knowledge Sources. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 690–695.

Gabriel Pereyra, George Tucker, Jan Chorowski, Łukasz Kaiser, and Geoffrey Hinton. 2017. Regularizing neural networks by penalizing confident output distributions. *arXiv preprint arXiv:1701.06548*.

Lianhui Qin, Michel Galley, Chris Brockett, Xiaodong Liu, Xiang Gao, Bill Dolan, Yejin Choi, and Jianfeng Gao. 2019. Conversing by Reading: Contentful Neural Conversation with On-demand Machine Reading. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5427–5436.

Marc'Aurelio Ranzato, Sumit Chopra, Michael Auli, and Wojciech Zaremba. 2015. Sequence level training with recurrent neural networks. *arXiv preprint arXiv:1511.06732*.

Pengjie Ren, Zhumin Chen, Christof Monz, Jun Ma, and Maarten de Rijke. 2019. Thinking Globally, Acting Locally: Distantly Supervised Global-to-Local Knowledge Selection for Background Based Conversation. *arXiv preprint arXiv:1908.09528*.

Yajing Sun, Yue Hu, Luxi Xing, Jing Yu, and Yuqiang Xie. 2020. History-adaption Knowledge Incorporation Mechanism for Multi-turn Dialogue System. In *AAAI*.

Xiangru Tang and Po Hu. 2019. Knowledge-Aware Self-Attention Networks for Document Grounded Dialogue Generation. In *International Conference on Knowledge Science, Engineering and Management*, pages 400–411. Springer.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.

Yiren Wang, Yingce Xia, Fei Tian, Fei Gao, Tao Qin, Cheng Xiang Zhai, and Tie-Yan Liu. 2019. Neural Machine Translation with Soft Prototype. In *Advances in Neural Information Processing Systems*, pages 6313–6322.

Yingce Xia, Fei Tian, Lijun Wu, Jianxin Lin, Tao Qin, Nenghai Yu, and Tie-Yan Liu. 2017. Deliberation networks: Sequence generation beyond one-pass decoding. In *Advances in Neural Information Processing Systems*, pages 1784–1794.

Semih Yavuz, Abhinav Rastogi, Guan-Lin Chao, Dilek Hakkani-Tür, and Amazon Alexa AI. 2019. DEEPCOPY: Grounded Response Generation with Hierarchical Pointer Networks. In *20th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, page 122.

Hao-Tong Ye, Kai-Lin Lo, Shang-Yu Su, and Yun-Nung Chen. 2020. Knowledge-grounded response generation with deep attentional latent-variable model. *Computer Speech & Language*, page 101069.

Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. Personalizing Dialogue Agents: I have a dog, do you have pets too? In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2204–2213.

Wen Zhang, Yang Feng, Fandong Meng, Di You, and Qun Liu. 2019. Bridging the gap between training and inference for neural machine translation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4334–4343, Florence, Italy. Association for Computational Linguistics.

Tiancheng Zhao, Kaige Xie, and Maxine Eskenazi. 2019a. Rethinking Action Spaces for Reinforcement Learning in End-to-end Dialog Agents with Latent Variable Models. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 1208–1218.

Tiancheng Zhao, Ran Zhao, and Maxine Eskenazi. 2017. Learning Discourse-level Diversity for Neural Dialog Models using Conditional Variational Autoencoders. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 654–664.

Xueliang Zhao, Chongyang Tao, Wei Wu, Can Xu, Dongyan Zhao, and Rui Yan. 2019b. A document-grounded matching network for response selection in retrieval-based chatbots. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 5443–5449. AAAI Press.

Wen Zheng and Ke Zhou. 2019. Enhancing Conversational Dialogue Models with Grounded Knowledge. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 709–718.

Hao Zhou, Tom Young, Minlie Huang, Haizhou Zhao, Jingfang Xu, and Xiaoyan Zhu. 2018. Commonsense Knowledge Aware Conversation Generation with Graph Attention. In *IJCAI*, pages 4623–4629.