

# There Are a Thousand Hamlets in a Thousand People’s Eyes: Enhancing Knowledge-grounded Dialogue with Personal Memory

Tingchen Fu<sup>1†</sup>, Xueliang Zhao<sup>2†</sup>, Chongyang Tao<sup>3</sup>, Ji-Rong Wen<sup>1</sup>, Rui Yan<sup>1\*</sup>

<sup>1</sup>Gaoling School of Artificial Intelligence, Renmin University of China

<sup>2</sup>Wangxuan Institute of Computer Technology, Peking University

<sup>3</sup>Microsoft Corporation

{lucas.futingchen, zhaoxlpku, chongyangtao}@gmail.com

{jrwen, ruiyan}@ruc.edu.cn

## Abstract

Knowledge-grounded conversation (KGC) shows great potential in building an engaging and knowledgeable chatbot, and knowledge selection is a key ingredient in it. However, previous methods for knowledge selection only concentrate on the relevance between knowledge and dialogue context, ignoring the fact that age, hobby, education and life experience of an interlocutor have a major effect on his or her personal preference over external knowledge. Without taking the personalization issue into account, it is difficult to select the proper knowledge and generate persona-consistent responses. In this work, we introduce personal memory into knowledge selection in KGC to address the personalization issue. We propose a variational method to model the underlying relationship between one’s personal memory and his or her selection of knowledge, and devise a learning scheme in which the forward mapping from personal memory to knowledge and its inverse mapping is included in a closed loop so that they could teach each other. Experiment results show that our method outperforms existing KGC methods significantly on both automatic evaluation and human evaluation.

## 1 Introduction

Open-domain dialogue system often suffers from safe response (Li et al., 2015; Zhang et al., 2019) problem as they could only refer to the context when generating a response. To alleviate this, knowledge-grounded conversation (KGC) is proposed to introduce external fact and real-world commonsense as prior knowledge (Zhou et al., 2018a; Dinan et al., 2019; Zhao et al., 2020a), such that a dialogue system is able to ground the conversation with the provided knowledge and therefore

generate informative and engaging responses. As external knowledge supplements the background to the inputs and decides what to say, knowledge selection is a key ingredient in KGC.

Numerous methods have been developed to tackle the knowledge selection problem by sequential latent variables (Kim et al., 2020; Meng et al., 2020), reinforcement learning (Zhao et al., 2020b), or expectation maximization algorithm (Li et al., 2020). In spite of the progress in this task, knowledge selection remains an unsolved problem as the precision is still far from satisfactory in Wizard of Wikipedia (Dinan et al., 2019) and other benchmarks in KGC (Gopalakrishnan et al., 2019), which also hinders the optimization of subsequent response generation models. A crucial point is, they often make assumption that the golden knowledge is distinguishable as long as the dialogue context is known, yet this is not always held true because there exists a one-to-many relationship in conversation and the past utterance history in a dialogue session is insufficient to decide the knowledge selection or the future trend of a dialogue.

As is shown in Figure 1, personalization is a key to success in the task because knowledge selection is a personal or subjective process in nature. When people communicate with each other, their perception of dialogue context will evoke their past memory about relevant life experience, taste and values, which we refer to as *personal memory*. The aroused fragment of personal memory further guides their interest and preference for different knowledge. Motivated by this, we postulate a new task named personalized KGC, introducing personalization into knowledge-grounded dialogue to encourage more human-like knowledge selection.

Importing persona memory into knowledge selection is a non-trivial task. One of the challenge is concretization of personal memory. Personal memory is an abstract concept related to user-specific experience, which is difficult to depict or model.

<sup>†</sup>The first two authors contribute equally. Xueliang Zhao is responsible for the design of the methodology and algorithm. Tingchen Fu is responsible for the implementation and experiment. The order is decided by a coin flip.

\*Corresponding author: Rui Yan (ruiyan@ruc.edu.cn)

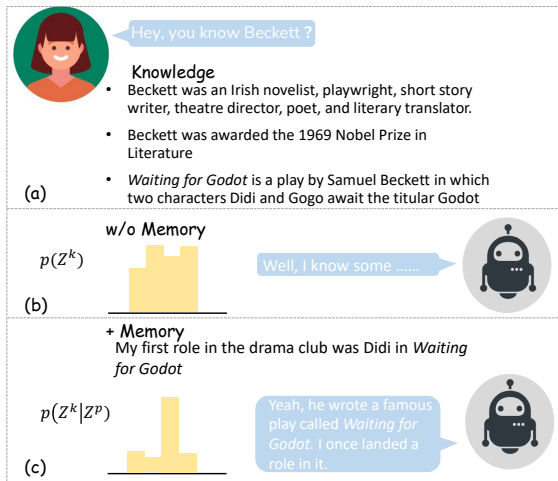


Figure 1: (a) The knowledge selection could not be certainly determined only based on dialogue context. (b) With out personal memory, the knowledge probability distribution is flat and is difficult to choose the proper knowledge. (c) Enhanced with personal memory, the knowledge probability distribution is sharper.

Though it has been discussed in open-domain dialogue (Li et al., 2016; Zhang et al., 2018), no previous research sheds light on the personalization issue in KGC and there exists no dialogue dataset featured with external facts and personal memory at the same time. Besides, there is no annotated label to indicate which knowledge candidate a person will choose based on his or her personal memory. Namely, the mapping between personal memory and knowledge selection is highly unconstrained without golden label. Intuitive resolution like treating personal memory as additional knowledge is sub-optimal because of dependency between knowledge and personal memory, as is shown in our experiments.

To address the above issue, we construct a KGC dataset featured with personalized memory repository, collecting user-specific utterance history under multiple types of context, which is a reflection of one’s personal memory. And to discover the underlying relationship between the dialogue context, personal memory and knowledge, we propose a variational method and introduce two latent variables  $Z^p$  and  $Z^k$  to indicate the fragment of personal memory to evoke and the knowledge candidate to select respectively. And to model the mapping from  $Z^p$  to  $Z^k$ , we introduce an inverse mapping as a dual task and employ dual learning to allow the two mappings to teach each other. The motivation behind this is intuitive: The reconstruction of personal memory from selected knowledge candidate is natural and easy if the mapping from

personal memory to knowledge is accurate. Extensive experiment shows that our methods outperform competitive baselines in both automatic evaluation and human evaluation, justifying the importance of introducing personal memory and the effect of the dual learning mechanism empirically.

The contributions of this work are three-fold:

(1) We explore the personalization issue of the knowledge selection task in KGC and construct a dataset featured with user-specific personal memory to benefit relevant research in the future. We are the first to explore the possibility of introducing personal memory into KGC.

(2) We propose a novel variational method and introduce two latent variables to model the interdependency between the persona and knowledge. Besides, we employ dual learning to optimize the relationship between the dialogue context, personal memory and knowledge in a unified framework.

(3) We conduct extensive experiments and verify the proposed methods empirically. Both the automatic and human evaluation evidence the efficacy of our proposed method.

## 2 Related Work

There is a substantial literature in the field of knowledge-grounded conversation. With the grounding of external knowledge in format of knowledge graph (Zhou et al., 2018a; Wu et al., 2019), document (Ghazvininejad et al., 2018; Zhou et al., 2018b; Zhao et al., 2019) or visual background (Das et al., 2017), it is regarded as a critical method towards intelligent dialogue system. Nowadays, existing methods in KGC often share a paradigm that decomposes the task into two related sub-problems, namely knowledge selection and utterance generation (Kim et al., 2020). In this work, we mainly focus on the knowledge selection task. To this end, a great deal of methods have been proposed to retrieve the most relevant knowledge by memory network (Ghazvininejad et al., 2018), sequential latent variables (Kim et al., 2020; Meng et al., 2020), reinforcement learning (Zhao et al., 2020b) and so on. A recent work gives attention to the expression style of knowledge (Zhao et al., 2021). However, they only focus on the decoding phase and no methods shed light on the personalization issue of knowledge selection, to our best knowledge.

Our work is related to dual learning as well. First proposed in neural machine translation by He et al.

(2016), dual learning is a semi-supervision learning scheme aiming at utilizing large-scale unlabeled data. Together with its newly appeared variants in recent years (Xia et al., 2017, 2018; Wang et al., 2019), dual learning has been successfully applied in neural machine translation (Xia et al., 2017; He et al., 2017), image-image-translation (Yi et al., 2017; Lin et al., 2018), sentiment analysis (Xia et al., 2017), automatic speech recognition (Ren et al., 2019), question answering (Tang et al., 2017), and knowledge-grounded dialogue (Meng et al., 2020). Our work is related to dual learning as well. First proposed in neural machine translation by He et al. (2016), dual learning is a semi-supervision learning scheme aiming at utilizing the large scale unlabeled data. In this work, we apply dual learning to model the inter-dependency relationship between one’s personal memory and his or her choice of knowledge.

### 3 Methodology

#### 3.1 Problem Formulation

Suppose we have a KGC dataset  $\mathcal{D}$  with  $N$  cases, and every case is in format of  $(C, \mathcal{K}, R)$ , where  $C = [u_1, u_2, \dots, u_{l_C}]$  is the context of the dialogue with  $l_C$  tokens in total,  $\mathcal{K} = \{K_1, K_2, \dots, K_{|\mathcal{K}|}\}$  is a set of  $|\mathcal{K}|$  knowledge candidates. And  $R = [r_1, r_2, \dots, r_{l_R}]$  is a response in this conversation corresponding to a specific user with unique user id. Different from the original KGC task, we have a memory repository  $\mathcal{M}$ . For every interlocutor corresponding to the response, a set of his or her personal memory  $\mathcal{P} = \{P_1, P_2, \dots, P_{|\mathcal{P}|}\}$  composed of  $|\mathcal{P}|$  customized utterance history could be retrieved from the memory repository. Our goal is to learn a probabilistic model  $p(R|C, \mathcal{K}, \mathcal{P})$  that could generate a personalized and informative response based on personal memory and knowledge.

#### 3.2 Model Overview

Figure 2 gives a graphical model of our methods. As is shown, the core of our proposed method is five probabilistic models to calculate the prior and posterior distribution of  $Z^p$ ,  $Z^k$  and an auxiliary distribution of  $Z^p$ . During training, we devise an unsupervised learning scheme, in which we optimize the distribution of two latent variables  $Z^p$  and  $Z^k$  by dual learning. To be more specific, we first sample a  $\tilde{Z}^p$  from the posterior distribution  $q_\phi(Z^p|C, R)$ , and then calculate the forward map-

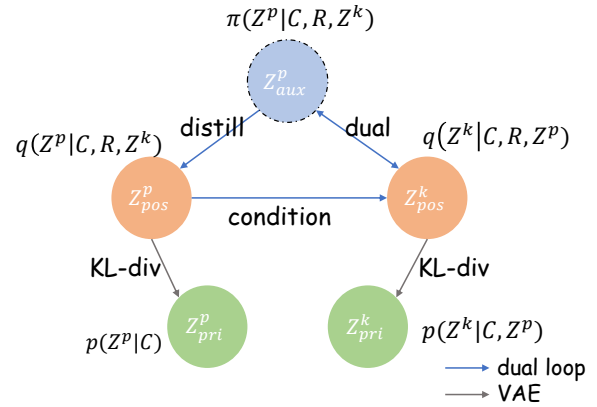


Figure 2: A graphical representation of our proposed method. It depicts the dependency and interaction between  $Z^p$  and  $Z^k$ .

ping from memory to knowledge  $q_\phi(Z^k|C, R, \tilde{Z}^p)$ , from which we sample a  $\tilde{Z}^k$ . The reward is designed as the probability of reconstructing the selected memory fragment by the auxiliary distribution  $\pi_\psi(Z^p = \tilde{Z}^p|C, R, \tilde{Z}^k)$ . By maximizing the reward, the primal task and the auxiliary task could benefit each other. The gains of the auxiliary distribution is distilled to  $q_\phi(Z^p|C, R)$ , such that the two posterior distribution and the auxiliary distribution form a closed loop. Besides, the prior distribution is forced to get close to the posterior distribution via KL-divergence.

In the inference phase, the prior distribution of  $Z^p$  is calculated at first, from which we sample and activate a personal memory fragment. After that, the woken memory fragment is used to decide the prior knowledge distribution  $p_\theta(Z^k|C)$ . Finally, the knowledge sampled from  $Z^k$  together with the memory fragment is sent into a generator to synthesize a response. Note that the golden response is only involved in the training phase.  $\pi$ ,  $\phi$  and  $\psi$  are all learnable parameters.

#### 3.3 Neural parameterization

To make the latent variables interpretable, we set the latent space of  $Z^p$  and  $Z^k$  as the number of memory fragments or knowledge candidates to choose from, and each sampling corresponds to a single piece of memory fragment or a knowledge candidate. Furthermore, motivated by human cognitive process, the aroused personal memory fragment implies one’s preference for different external knowledge, which influences the likelihood of choosing different knowledge. In light of this,

the prior distribution of  $(Z^p, Z^k)$  is factorized as:

$$p(Z^p, Z^k) = p(Z^k|Z^p)p(Z^p) \quad (1)$$

And to calculate their probability distribution, we adopt BERT (Devlin et al., 2018) as the backbone of our method to obtain a dense representation of dialogue context, response, candidate knowledge sentence or personal memory fragment. Take the calculation of the prior distribution  $p_\theta(Z^k|C, Z^p)$  as an example. We first concatenate the context  $C$ , the memory fragment  $P$  indicated by the sampled  $Z^p$ , and the  $i$ -th candidate knowledge  $K_i$  together as a long sequence. A special [CLS] token is prepended at the beginning of the sequence and [SEP] is inserted to separate different utterances:

$$\mathcal{I} = u_1, u_2, \dots, u_{l_C}, p_1, p_2, \dots, p_{l_P}, k_1, k_2, \dots, k_{l_{K_i}}, \quad (2)$$

where  $l_C$ ,  $l_P$  and  $l_{K_i}$  are the number of tokens in the context, memory facet and knowledge candidate respectively. Then the embedding layer will convert  $\mathcal{I}$  into input representations, which is the sum of the corresponding token embedding and position embedding. Thereafter, the BERT encoder performs multi-head attention on the input representation to obtain a dense representation. There are  $n$  identical layers in the BERT encoder, and for each layer, the multi-head attention could be formulated as

$$\mathbf{H}^l = \text{FFN}(\text{MultiHead}(\mathbf{Q}^{l-1}, \mathbf{K}^{l-1}, \mathbf{V}^{l-1})), \quad (3)$$

where  $\text{FFN}(\cdot)$  is a feed-forward network and we use  $\mathbf{Q}^{l-1}$ ,  $\mathbf{K}^{l-1}$ , and  $\mathbf{V}^{l-1}$  to denote the query matrix, key matrix and value matrix after the  $l-1$ -th layer respectively. For self-attention, we have

$$\mathbf{Q}^{l-1} = \mathbf{K}^{l-1} = \mathbf{V}^{l-1} = \mathbf{H}^{l-1}, \quad (4)$$

where  $\mathbf{H}^l$  means the hidden state at the  $l$ -th layer. Specially,  $\mathbf{H}^0$  is the input embedding and  $\mathbf{H}^n$  is the final output of the BERT.

We use the vector corresponding to the position of the special [CLS] token in  $\mathbf{H}^n$  as the representation of the  $i$ -th knowledge candidate, which is referred to as  $\mathbf{h}_i$ . Then the distribution of  $Z^k$  is calculated as

$$p_\theta(Z^k = i|C, Z^p) = \frac{\exp(f(\mathbf{h}_i))}{\sum_j^{|K|} \exp(f(\mathbf{h}_j))}, \quad (5)$$

where  $f(\cdot)$  is a multi-layer perceptron. The prior and posterior distribution of  $Z^k$  and  $Z^p$  are calculated in a similar way. The only difference lies

in the constitution of input sequence  $\mathcal{I}$ : For the prior distribution of  $Z^p$ ,  $\mathcal{I}$  is the concatenation of dialogue context and a candidate personal memory facet:

$$\mathcal{I} = u_1, u_2, \dots, u_{l_C}, p_1, p_2, \dots, p_{l_P} \quad (6)$$

And to calculate the posterior distribution, we insert the response tokens behind the dialogue context tokens as the response usually contains clue indicating the selected knowledge and memory. Namely, to compute  $q_\phi(Z^p|C, R)$ , the posterior of  $Z^p$ , the input is:

$$\mathcal{I} = u_1, u_2, \dots, u_{l_C}, r_1, r_2, \dots, r_{l_R}, p_1, p_2, \dots, p_{l_P} \quad (7)$$

And for  $q_\phi(Z^k|C, R, Z^p)$ :

$$\mathcal{I} = u_1, u_2, \dots, u_{l_C}, r_1, r_2, \dots, r_{l_R}, p_1, p_2, \dots, p_{l_P}, k_1, k_2, \dots, k_{l_K} \quad (8)$$

Normally, the generator  $g$  of our method could be specified as any large-scale pre-trained language model. Here we define the generator as GPT-2 (Radford et al., 2019). Previous methods often synthesize a response merely based on the dialogue context and the selected knowledge, taking no consideration of the persona of the interlocutor, which may lead to an inconsistency in persona. Different from that, we input the sampled personal memory fragment and the sampled knowledge candidate into GPT-2 all together with the dialogue context. Intuitively, personal memory fragment implies why the knowledge is paid attention to and underlying relevance between the persona of the interlocutor and the knowledge, which endows the generator to generate persona-consistent and knowledgeable responses:

$$\begin{aligned} g(R) &= g(R|C, Z^p, Z^k) \\ &= \prod_{i=1}^{l_R} g(r_i|C, Z^p, Z^k, r_{<i}) \end{aligned} \quad (9)$$

### 3.4 Learning Details

Directly maximizing the marginal log-likelihood of generating the correct response  $g(R|C, Z^p, Z^k)$  requires integrating over all possibilities of  $Z^k$  and  $Z^p$ , which is more than time-consuming. Inspired by variational inference, we introduce a variational posterior as the true posterior is intractable. Thereby, instead of directly optimizing

---

**Algorithm 1** The proposed learning algorithm.

---

- 1: **Input:** Training KGC dataset  $\mathcal{D}$ , memory repository  $\mathcal{M}$
  - 2: Warm up  $p_\theta(Z^p)$ ,  $p_\theta(Z^k|Z^p)$ ,  $q_\phi(Z^p|R)$  and  $q_\phi(Z^k|R, Z^p)$  on  $\mathcal{D}$ .
  - 3: **while** not converge **do**
  - 4:   Sample a mini-batch  $\{(C, \mathcal{K}, R)\}$  from  $\mathcal{D}$ .
  - 5:   Retrieve the user-specific personal memory  $\mathcal{P}$  from the memory repository.
  - 6:   Calculate the prior personal memory distribution  $p_\theta(Z^p)$  with  $C$ .
  - 7:   Sample a  $Z^p$  and then calculate the prior distribution of knowledge  $p_\theta(Z^k|Z^p)$ .
  - 8:   Calculate the posterior memory distribution  $q_\phi(Z^p|R)$  based on  $C$  and  $R$ , and then sample a  $\tilde{Z}^p$  from that.
  - 9:   Calculate the posterior knowledge distribution  $q_\phi(Z^k|R, \tilde{Z}^p)$ , and then sample a  $\tilde{Z}^k$  from that. {The primal task}
  - 10:   Compute the reward  $Re_1$  as the Reconstruct probability  $\pi_\psi(Z^p = \tilde{Z}^p|Z^k)$ .
  - 11:   Update  $\phi$  according to Eq. 16.
  - 12:   Calculate the auxiliary memory distribution  $\pi_\psi(Z^p|R, \tilde{Z}^k)$  based on the pseudo knowledge label  $\tilde{Z}^k$ , and sample a  $\tilde{Z}^p$  from that. {The dual task}
  - 13:   Compute the reward  $Re_2$  as  $q_\phi(Z^k = \tilde{Z}^k|\tilde{Z}^p)$ .
  - 14:   Update  $\psi$  according to Eq. 15.
  - 15:   Update  $\theta$  according to Eq. 10.
  - 16:   Update  $\phi$  according to Eq. 17.
  - 17: **end while**
  - 18: **return** The prior distribution  $p_\theta(Z^p)$  and  $p_\theta(Z^k|Z^p)$
- 

the marginal log-likelihood, we derive an evidence lower bound objective to maximize:

$$\begin{aligned} \mathcal{L}_{ELBO} = & \mathbb{E}_{q_\phi(Z^k|Z^p)q_\phi(Z^p)}g(R|C, Z^p, Z^k) \\ & - \mathbb{E}_{q_\phi(Z^p)}KL(q_\phi(Z^k|Z^p)||p_\theta(Z^k|Z^p)) \quad (10) \\ & - KL(q_\phi(Z^p)||p_\theta(Z^p)) \end{aligned}$$

where  $q_\phi(Z^k|Z^p)$ ,  $q_\phi(Z^p)$ ,  $p_\theta(Z^p)$ ,  $p_\theta(Z^k|Z^p)$  are shorthand for  $q_\phi(Z^k|C, R, Z^p)$ ,  $q_\phi(Z^p|C, R)$ ,  $p_\theta(Z^p)$  and  $p_\theta(Z^k|C, Z^p)$  respectively. A step-wise derivation could be found in the supplementary materials.

The forward mapping from personal memory to knowledge candidates is relatively implicit and obscure, partially because the customized utterance history contains unwanted noise. As a result, there is a tendency that  $Z^p$  is ignored and  $p_\theta(Z^k|Z^p, C)$  is degenerated into  $p_\theta(Z^k|C)$ , which we refer to as the *vanishing memory*.

To address this issue, inspired by the idea of dual learning (He et al., 2016), we introduce an inverse mapping from knowledge candidate to personal memory as a dual task, which is depicted by the auxiliary distribution  $\pi_\psi(Z^p|C, R, Z^k)$ . Intuitively, there is a natural duality between the mapping from personal memory to knowledge and the inverse mapping. Therefore, if the forward mapping makes a good inference about the knowledge

to choose, the inverse mapping is able to map it back to personal memory, which means that the memory is not vanishing.

And before the dual learning procedure, the primal task and the dual task are warmed up to speed up convergence and alleviate error accumulation in the dual learning process, following the idea of He et al. (2016) and Meng et al. (2020). Namely, we construct pseudo knowledge label  $\bar{P}$  and persona label  $\bar{K}$  based on their similarity to the response.

$$\begin{aligned} \bar{K} &= \max_{K_i \in \mathcal{K}} Sim(K_i, R) \\ \bar{P} &= \max_{P_i \in \mathcal{P}} Sim(P_i, R) \quad (11) \end{aligned}$$

Then, both the primal task and the dual task are warmed up with a traditional maximum likelihood estimation objective.

After the warm-up procedure, for each iteration, we first sample a  $\tilde{Z}^p$  according to its posterior distribution  $q_\phi(Z^p|C, R)$ . Then the forward mapping calculates the probability distribution  $q_\phi(Z^k|C, R, \tilde{Z}^p)$ , from which we sample a  $\tilde{Z}^k$ . The reward for the forward mapping is defined as the probability that the auxiliary distribution recovers the  $\tilde{Z}^p$ . Mathematically, we have

$$Re_1 = \pi_\psi(Z^p = \tilde{Z}^p|C, R, \tilde{Z}^k) \quad (12)$$

Symmetrically, the reward for the auxiliary distribution is the prediction probability of the golden knowledge by the forward mapping:

$$Re_2 = q_\phi(Z^k = \tilde{Z}^k|C, R, Z^p), \quad (13)$$

where  $\tilde{Z}^k$  is corresponding to the pseudo knowledge label.

And the objective of the dual learning is to maximize the reward:

$$\mathcal{L}_{dual} = \mathbb{E}_{\mathcal{D}}[Re_1 + Re_2] \quad (14)$$

For reward maximization, we optimize the parameter through policy gradient method (Sutton et al., 2000):

$$\nabla_\psi \mathcal{L}_{dual} = \nabla_\psi \log \pi_\psi(Z^p = \tilde{Z}^p|C, R, \tilde{Z}^k) Re_2. \quad (15)$$

$$\nabla_\phi \mathcal{L}_{dual} = \nabla_\phi \log q_\phi(Z^k = \tilde{Z}^k|C, R, \tilde{Z}^p) Re_1. \quad (16)$$

Finally, the gains of the dual task is distilled into the posterior distribution of  $Z^p$  via a cross-entropy loss:

$$\begin{aligned} \mathcal{L}_{dis} = & -KL(\pi_\psi^T(Z^p|C, R, Z^k)||q_\phi^T(Z^k|C, R, Z^p)) \\ & + \alpha \log q_\phi(Z^p = \tilde{Z}^p|C, R, Z^k), \quad (17) \end{aligned}$$

where  $\alpha$  is a hyper-parameters to balance the weights of two parts and the superscript  $T$  means that the distribution is normalized at temperature  $T$ . Thus, the three probabilistic models form a closed loop in which each component is trained alternatively. The full procedure of our proposed learning algorithm is concluded in Algorithm 1.

## 4 Experiment

### 4.1 Dataset

Since existing dataset like CMU\_DoG (Zhou et al., 2018b) or Holl-E (Moghe et al., 2018) do not contain information about personal memory, we establish a new KGC dataset equipped with a memory repository. The dataset is constructed based on Reddit (Baumgartner et al., 2020).

In detail, we download the conversational data on the PushShift dump of Reddit ranging from 2011 to the first half of 2015 and divide them into a training set, a validation set and a test set according to the date. To construct a memory repository, we maintain a dictionary where the key is a long string hashed from the user account name and the value is a set of utterances of the user. Since it is a repository for user-specific utterances, it may inevitably contain false beliefs or subjective opinions. We shall leave this issue for future work. Elaborated data filtering is conducted to ensure: (1) We only keep utterances from users that have at least 5 utterances in the memory repository; (2) Utterances that are too long or too short are filtered; (3) Paraphrase tool (Damodaran, 2021) is applied on every utterances to avoid tracing the utterances back to real reddit users.

The statistics of our dataset is shown in Table 1. And the code is available at <https://github.com/Lucasftc/PersonaKGC>. A few examples is shown in Appendix A.3. To benefit future research and meanwhile avoid possible malicious abuse, the dataset is available upon request from the authors<sup>1</sup>.

### 4.2 Compared Methods

To verify the effectiveness of the proposed methods, we compare our methods with baselines in KGC. Meanwhile, since our proposed method makes use of personal memory to generate persona-consistency response, we also compare our methods with baselines in personalized dialogue.

<sup>1</sup>Please contact [lucas.futingchen@gmail.com](mailto:lucas.futingchen@gmail.com) for the dataset.

	Train	Valid	Test
# Dialogues	217,095	11,186	6,236
# Utterances	1,442,975	74,480	41,519
# Knowledges	5,459,744	290,349	148,057
# User	48,858	5,603	3,281
# Memory facets	490,460	70,494	38,354
AvG.Len (# words):			
Utterance	34.15	33.95	33.60
Knowledge	54.54	52.39	53.17
Memory facet	42.10	40.21	40.60

Table 1: The statistics of the dataset.

- *Generative Profile Memory Network (GPMN)* (Zhang et al., 2018) is a method in personalized dialogue which employs Memory Network along with persona information.
- *Transformer Memory Network (TMN)* (Dinan et al., 2019) adopts the traditional Memory Network with transformer architecture and introduces the knowledge selection loss.
- *Transfertransfo* (Wolf et al., 2019) is a combination of a transfer learning based training scheme and a high-capacity transformer model and achieves the best results in the Conversational Intelligence Challenge 2.
- *Sequential Knowledge Transformer (SKT)* (Kim et al., 2020) utilizes sequential latent variables for knowledge selection. We use the pseudo knowledge labels for the golden knowledge label in implementation.
- *KnowledGPT* (Zhao et al., 2020b) puts the knowledge selector and the response generator in a framework and employ reinforcement learning and curriculum learning to accomplish the state-of-the-art performance in KGC.
- *KnowledGPT+M*, a variant of KnowledGPT where we treat personal memory as knowledge candidates as well and input them to the knowledge selector.
- *P<sup>2</sup>BOT* (Liu et al., 2020) is a transmitter-receiver based framework explicitly modeling the perception between the interlocutors and achieves the state-of-the-art in personalized dialogue.
- *BoB* (Song et al., 2021) is a newly published method that disentangles personalized dialogue into persona understanding and personalized generation.

For more implementation details about the baselines and our method, please refer to appendix A.2.

Models	BLEU				ROUGE			Distinct		METEOR
	B-1	B-2	B-3	B-4	R-1	R-2	R-3	D-1	D-2	
GPMN	3.87	1.41	0.43	0.13	4.25	0.23	3.94	0.06	0.15	2.30
TMN	1.05	0.31	0.12	0.02	8.91	1.38	7.88	0.10	0.28	2.60
Transfertransfo	6.09	1.57	0.62	0.34	9.31	0.73	7.34	8.33	43.43	3.79
SKT	3.48	0.85	0.28	0.10	7.95	0.94	6.95	3.41	14.35	2.75
KnowledGPT	5.22	1.76	0.77	0.39	10.68	1.71	9.12	6.65	28.64	4.09
KnowledGPT+M	7.81	3.55	2.46	2.02	10.79	2.82	9.32	7.37	35.13	4.58
P2bot	5.95	1.61	0.57	0.24	7.54	0.72	6.54	4.98	17.74	3.20
BoB	4.69	1.57	0.65	0.31	10.68	1.57	9.30	4.94	17.06	3.97
Ours	<b>13.09</b>	<b>6.22</b>	<b>4.23</b>	<b>3.33</b>	<b>13.60</b>	<b>3.73</b>	<b>10.64</b>	<b>8.97</b>	39.29	<b>6.65</b>

Table 2: Automatic evaluation results. Numbers in bold mean that the improvement to the best performing baseline is statistically significant (t-test with  $p$ -value  $< 0.05$ ).

	Fluency	Coherence	Faithfulness	Kappa
Transfertransfo	1.65	1.73	1.68	0.72
KnowledGPT+M	1.71	1.72	1.77	0.69
BoB	1.67	1.62	1.70	0.70
Ours	<b>1.77</b>	<b>1.79</b>	<b>1.82</b>	0.69

Table 3: Human evaluation result. Numbers in bold mean that the improvement to the best performing baseline is statistically significant (t-test with  $p$ -value  $< 0.05$ ).

### 4.3 Evaluation Metrics

We choose distinctness, BLEU(Papineni et al., 2002), ROUGE(Lin, 2004)<sup>2</sup> and METEOR(Denkowski and Lavie, 2014)<sup>3</sup> to be our automatic metrics. Focusing on the exact n-gram co-occurrence in hypothesis and reference, BLEU and ROUGE evaluate the appropriateness of the proposed model. Distinctness is calculated as the ratio of unique unigrams and bigrams, paying more attention to the diversity of generated text. METEOR measures the alignment, or the exact, stem, synonym, and paraphrase matches between the hypothesis and reference.

Apart from automatic evaluation, we conduct human evaluation. Specifically, 200 examples are randomly sampled from the test set and well-educated native speakers are recruited to assess the quality of the generation from different models with their source hidden. Each annotators are required to give a score in  $\{0 : \text{bad}, 1 : \text{fair}, 2 : \text{good}\}$  for three independent aspects: (1) *fluency*: whether the reply is fluent; (2) *coherence*: whether the reply is coherent with the context; and (3) *faithfulness*: whether the reply is well-grounded and faithful to the selected knowledge sentence and memory fragment. The agreement of annotators is measured via Fleiss’ kappa (Fleiss, 1971).

<sup>2</sup><https://github.com/bckim92/language-evaluation>

<sup>3</sup><https://github.com/Maluuba/nlg-eval>

### 4.4 Experiment Results

We first report the experimental result in automatic evaluation. As is shown in Table 2, our method outperforms the state-of-the-art baselines in KGC and personalized dialogue in most metrics, verifying the effectiveness of our model empirically. Among non-pretrained methods, TMN and GPMN are low in diversity, since their generator is not pre-trained on large corpus before. SKT improves distinctness but shows low appropriateness, possibly because that it highly relies on the golden knowledge label, which is costly and not always available. In pre-trained based methods, Transfertransfo attains impressive results on distinctness. It also achieves competitive appropriateness results, but not as good as ours. We gauge the performance of the model to the large document-level training corpus, a critical choice for pre-trained language model, which may boost the diversity of generated text. Besides, the performance of the BoB, a recently published baseline, is less satisfactory compared with others. The premise of BoB is the disentanglement between contextual coherence and persona consistency, which is not always achievable especially when we use user-specific dialogue history for personal memory information. And it is notable from the table that there is a significant gap between the baseline methods in KGC or personalized dialogue and ours, validating that neither simply projecting personal information into dialogue nor purely grounding on knowledge is an acceptable solution to the KGC task. It is necessary to combine personal memory and external knowledge together. The comprehensive improvement of KnowledGPT+M in contrast with the original KnowledGPT also reveals this viewpoint. Additionally, the considerable advantage of our proposed method over KnowledGPT+M illustrates the fact

Models	BLEU				ROUGE			Distinct		METEOR
	B-1	B-2	B-3	B-4	R-1	R-2	R-3	D-1	D-2	
BoB	4.69	1.57	0.65	0.31	10.68	1.57	9.30	4.94	17.06	3.97
w/o. <i>know</i>	6.37	2.13	1.07	0.69	9.68	1.41	8.06	7.19	25.87	3.87
w/o. <i>mem</i>	6.79	1.90	0.65	0.23	9.79	1.16	8.11	5.17	16.95	3.91
w/o. <i>dual</i>	8.58	3.74	2.42	1.84	12.05	2.80	9.87	8.74	34.23	4.97
w/o. <i>dep</i>	8.64	3.29	1.90	1.35	10.78	1.96	8.65	9.03	36.01	4.57

Table 4: Ablation results (RQ1).

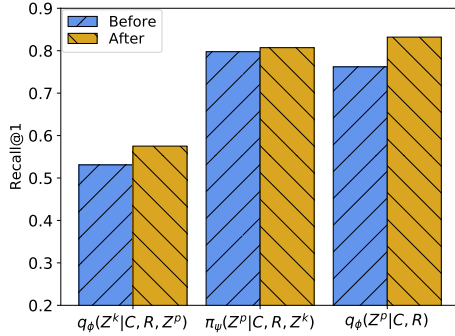


Figure 3: The Recall@1 of knowledge (or personal memory) before and after the closed dual loop.

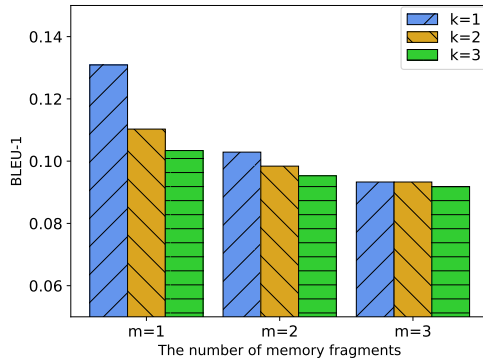


Figure 4: The performance of our model in terms of BLEU-1 under different number of personal memory fragment and knowledge.

that treating personal memory as knowledge is not enough. The dependency between personal memory and the knowledge should not be ignored.

We also present the result of human evaluation since no automatic metric is perfect in this task (Dinan et al., 2019). Since human evaluation is time-consuming and expensive, only competitive baselines are involved. As shown in Table 3, our proposed model outperforms the baseline methods and there is an evident improvement.

#### 4.5 Analysis

Apart from the main results, we are especially interested in some research questions:

- (RQ1) How does each component contributes to the performance of our model?
- (RQ2) How many knowledge sentences and

memory fragments to select?

To answer the first question, we conduct ablation study and compare the full model with several variants: (1) w/o. *know*. the external knowledge base to grounding the dialogue is removed; (2) w/o. *mem*. personal memory is removed and this variant is a standard KGC model essentially; (3) w/o. *dual*. the dual task is removed, so there is no dual learning and distillation in this variant; (4) w/o. *dep*. the dependency of the two latent variables is removed so  $Z^p$  and  $Z^k$  are calculated independently. The ablation result is shown in Table 4, from which we could have the following observations: (1) w/o. *know* and w/o. *mem* exhibit a degeneration at a great extent, further justifying the necessity of introducing knowledge and personal memory into a dialogue system, respectively. (2) w/o. *dep* also shows an obvious deterioration. This is in line with our expectation since w/o. *dep* model  $Z^k$  and  $Z^p$  as two independent latent variables, ignoring the underlying dependence between them. Comparatively speaking, w/o. *dual* achieves a better result, but not as good as the full model due to the destroy of the closed dual loop.

And to have a intuitive perception about the effect of the closed dual loop, we examine the promotion brought to the  $q_\phi(Z^k|C, R, Z^p)$ ,  $\pi_\psi(Z^p|C, R, Z^k)$  and  $q_\phi(Z^p|C, R)$  in terms of Recall@1 of knowledge or personal memory. The result is shown in Figure 3. From the figure we could see that there is an obvious improvement after trained with our proposed learning algorithm.

For the (RQ2), we first explore it by varying the amount of selected personal memory fragments and observe how the knowledge selection procedure is influenced. In detail, we vary the number of personal memory fragments  $m$  sampled by  $p_\theta(Z^p|C)$  from 1 to 4 and evaluate the performance of  $p_\theta(Z^k|C, Z^p)$  in terms of Recall@n ( $n \in \{1, 2, 5, 10\}$ ).

As is shown in Table 5, we could find that the best performance is reached when  $m = 2$ . There is



	Recall@1	Recall@2	Recall@5	Recall@10
m=1	0.173	0.286	0.505	0.720
m=2	0.176	<b>0.289</b>	<b>0.513</b>	0.730
m=3	<b>0.177</b>	0.289	0.509	0.730
m=4	0.176	0.288	0.508	0.730

Table 5: The performance of  $p_\theta(Z^k|C, Z^p)$  under different  $m$ . The numbers in bold are the best results.

a fluctuation or slight drop when  $m$  continues to increase possibly owing to the distraction mixed with the redundant personal memory. Besides, we are also curious about the final generation performance under different numbers of knowledge and personal memory fragment. It could be seen from Figure 4 that there appears a decline when we increase the number of knowledge and personal memory fragment, which we attribute to the unwanted noise mixed with personal memory and knowledge.

## 5 Conclusion

In this work, we explore personalized KGC by introducing personal memory into knowledge selection task. Two latent variables are introduced to select knowledge and personal memory respectively. Besides, dual learning scheme is employed to allow the two selection task to teach each other. For future work, we would like to extend the personalized knowledge-grounded dialogue to personalized conversational recommendation system for application in online shopping.

## Ethical Considerations

**Intended Use** The chief purpose of our dataset is to examine a dialogue model’s capacity in selecting proper knowledge with the help of personal memory. The dataset is mainly for research propose and it is not supposed to be directly used to train a production system. And researchers should be aware of the possible ethic issues before exploiting our dataset.

**Data Collection** All the examples in our dataset are in English and no human annotators are involved in the data collection process. As mentioned in Sec.4.1, our dataset is built on the basis of the Reddit dumps from Pushshift (Baumgartner et al., 2020), which is a publicly available resource widely used in more than a hundred peer-reviewed publications. Our data collection is in consistent with the term of use and the research is granted ethical approval by an external institutional review board. To avoid potential abuse, the dataset is avail-

able upon request to the authors. Contact the authors (by email) and clearly state your intended use if you believe the dataset might be helpful in your research.

**User Privacy** Although our dataset includes user-specific utterance history as personal memory, no user account names will be revealed or inferred from the dataset. Besides, the utterance histories are paraphrased during our procession of the dataset such that they can not be traced back to the real users in Reddit. In conclusion, There is no *personally identifiable information* in our dataset or underlying leakage of personal information.

## Acknowledgement

Thanks for the reviewers for their valuable suggestions. This work was supported by National Natural Science Foundation of China (NSFC Grant No. 62122089 & No. 61876196 & No. 61832017), Beijing Outstanding Young Scientist Program NO. BJJWZYJH012019100020098, and Intelligent Social Governance Platform, Major Innovation & Planning Interdisciplinary Platform for the "Double-First Class" Initiative, Renmin University of China. We also wish to acknowledge the supports provided and contributions made by Public Policy and Decision-making Research Lab of RUC, and the Public Computing Cloud, Renmin University of China. Rui Yan is also supported by Beijing Academy of Artificial Intelligence (BAAI).

## References

- Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The pushshift reddit dataset. In *Proceedings of the international AAAI conference on web and social media*, volume 14, pages 830–839.
- Prithviraj Damodaran. 2021. Parrot: Paraphrase generation for nlu.
- Abhishek Das, Satwik Kottur, Khushi Gupta, Avi Singh, Deshraj Yadav, José MF Moura, Devi Parikh, and Dhruv Batra. 2017. Visual dialog. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 326–335.
- Michael Denkowski and Alon Lavie. 2014. Meteor universal: Language specific translation evaluation for any target language. In *Proceedings of the EACL 2014 Workshop on Statistical Machine Translation*.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep

- bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. 2019. Wizard of wikipedia: Knowledge-powered conversational agents. In *ICLR*.
- Joseph L Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76(5):378.
- 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 *AAAI*.
- Karthik Gopalakrishnan, Behnam Hedayatnia, Qinlang Chen, Anna Gottardi, Sanjeev Kwatra, Anu Venkatesh, Raefer Gabriel, Dilek Hakkani-Tür, and Amazon Alexa AI. 2019. Topical-chat: Towards knowledge-grounded open-domain conversations. *Proc. Interspeech 2019*, pages 1891–1895.
- Di He, Hanqing Lu, Yingce Xia, Tao Qin, Liwei Wang, and Tie-Yan Liu. 2017. Decoding with value networks for neural machine translation. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 177–186.
- Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu, Tie-Yan Liu, and Wei-Ying Ma. 2016. Dual learning for machine translation. *Advances in neural information processing systems*, 29:820–828.
- Byeongchang Kim, Jaewoo Ahn, and Gunhee Kim. 2020. Sequential latent knowledge selection for knowledge-grounded dialogue. *arXiv preprint arXiv:2002.07510*.
- Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2015. A diversity-promoting objective function for neural conversation models. *NAACL*, pages 110–119.
- Jiwei Li, Will Monroe, Alan Ritter, Dan Jurafsky, Michel Galley, and Jianfeng Gao. 2016. Deep reinforcement learning for dialogue generation. In *EMNLP*, pages 1192–1202.
- Linxiao Li, Can Xu, Wei Wu, Yufan Zhao, Xueliang Zhao, and Chongyang Tao. 2020. Zero-resource knowledge-grounded dialogue generation. *arXiv preprint arXiv:2008.12918*.
- Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81.
- Jianxin Lin, Yingce Xia, Tao Qin, Zhibo Chen, and Tie-Yan Liu. 2018. Conditional image-to-image translation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5524–5532.
- Qian Liu, Yihong Chen, Bei Chen, Jian-Guang Lou, Zixuan Chen, Bin Zhou, and Dongmei Zhang. 2020. You impress me: Dialogue generation via mutual persona perception. *arXiv preprint arXiv:2004.05388*.
- Chuan Meng, Pengjie Ren, Zhumin Chen, Weiwei Sun, Zhaochun Ren, Zhaopeng Tu, and Maarten de Rijke. 2020. Dukenet: A dual knowledge interaction network for knowledge-grounded conversation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1151–1160.
- Nikita Moghe, Siddhartha Arora, Suman Banerjee, and Mitesh M Khapra. 2018. Towards exploiting background knowledge for building conversation systems. *arXiv preprint arXiv:1809.08205*.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 311–318. Association for Computational Linguistics.
- Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners.
- Yi Ren, Xu Tan, Tao Qin, Sheng Zhao, Zhou Zhao, and Tie-Yan Liu. 2019. Almost unsupervised text to speech and automatic speech recognition. In *International Conference on Machine Learning*, pages 5410–5419. PMLR.
- Haoyu Song, Yan Wang, Kaiyan Zhang, Wei-Nan Zhang, and Ting Liu. 2021. BoB: BERT over BERT for training persona-based dialogue models from limited personalized data. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 167–177, Online. Association for Computational Linguistics.
- Richard S Sutton, David A McAllester, Satinder P Singh, and Yishay Mansour. 2000. Policy gradient methods for reinforcement learning with function approximation. In *Advances in neural information processing systems*, pages 1057–1063.
- Duyu Tang, Nan Duan, Tao Qin, Zhao Yan, and Ming Zhou. 2017. Question answering and question generation as dual tasks. *arXiv preprint arXiv:1706.02027*.
- Yiren Wang, Yingce Xia, Tianyu He, Fei Tian, Tao Qin, ChengXiang Zhai, and Tie-Yan Liu. 2019. Multi-agent dual learning. In *Proceedings of the International Conference on Learning Representations (ICLR) 2019*.
- Thomas Wolf, Victor Sanh, Julien Chaumond, and Clement Delangue. 2019. Transfertransfo: A transfer learning approach for neural network based conversational agents. *arXiv preprint arXiv:1901.08149*.

- Wenquan Wu, Zhen Guo, Xiangyang Zhou, Hua Wu, Xiyuan Zhang, Rongzhong Lian, and Haifeng Wang. 2019. Proactive human-machine conversation with explicit conversation goals. *arXiv preprint arXiv:1906.05572*.
- Yingce Xia, Tao Qin, Wei Chen, Jiang Bian, Nenghai Yu, and Tie-Yan Liu. 2017. Dual supervised learning. In *International Conference on Machine Learning*, pages 3789–3798. PMLR.
- Yingce Xia, Xu Tan, Fei Tian, Tao Qin, Nenghai Yu, and Tie-Yan Liu. 2018. Model-level dual learning. In *International Conference on Machine Learning*, pages 5383–5392. PMLR.
- Zili Yi, Hao (Richard) Zhang, Ping Tan, and Minglun Gong. 2017. **Dualgan: Unsupervised dual learning for image-to-image translation**. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 2868–2876. IEEE Computer Society.
- 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? *arXiv preprint arXiv:1801.07243*.
- Yizhe Zhang, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, and Bill Dolan. 2019. Dialogpt: Large-scale generative pre-training for conversational response generation. *arXiv preprint arXiv:1911.00536*.
- Xueliang Zhao, Tingchen Fu, Chongyang Tao, Wei Wu, Dongyan Zhao, and Rui Yan. 2021. Learning to express in knowledge-grounded conversation.
- Xueliang Zhao, Chongyang Tao, Wei Wu, Can Xu, Dongyan Zhao, and Rui Yan. 2019. A document-grounded matching network for response selection in retrieval-based chatbots. In *IJCAI*, pages 5443–5449.
- Xueliang Zhao, Wei Wu, Chongyang Tao, Can Xu, Dongyan Zhao, and Rui Yan. 2020a. Low-resource knowledge-grounded dialogue generation. *arXiv preprint arXiv:2002.10348*.
- Xueliang Zhao, Wei Wu, Can Xu, Chongyang Tao, Dongyan Zhao, and Rui Yan. 2020b. Knowledge-grounded dialogue generation with pre-trained language models. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3377–3390.
- Hao Zhou, Tom Young, Minlie Huang, Haizhou Zhao, Jingfang Xu, and Xiaoyan Zhu. 2018a. Commonsense knowledge aware conversation generation with graph attention. In *IJCAI*, pages 4623–4629.
- Kangyan Zhou, Shrimai Prabhumoye, and Alan W Black. 2018b. A dataset for document grounded conversations. *arXiv preprint arXiv:1809.07358*.

## A Appendix

### A.1 The derivation of ELBO

$$\begin{aligned}
\mathcal{L}_{elbo} &= \log p(R) \\
&= \log \int_{Z^k} \int_{Z^p} p(R, Z^p, Z^k) dZ^p dZ^k \\
&= \log \int_{Z^k} \int_{Z^p} g(R|Z^p, Z^k) p(Z^p, Z^k) dZ^p dZ^k \\
&= \log \int_{Z^k} \int_{Z^p} g(R|Z^p, Z^k) p_\theta(Z^k|Z^p) p_\theta(Z^p) dZ^p dZ^k \\
&= \log \int_{Z^k} \int_{Z^p} g(R|Z^p, Z^k) p_\theta(Z^k|Z^p) p_\theta(Z^p) \\
&\quad \cdot \frac{q(Z^p, Z^k)}{q(Z^p, Z^k)} dZ^p dZ^k \\
&= \log \int_{Z^k} \int_{Z^p} g(R|Z^p, Z^k) p_\theta(Z^k|Z^p) p_\theta(Z^p) \\
&\quad \cdot \frac{q(Z^k|Z^p)q(Z^p)}{q(Z^k|Z^p)q(Z^p)} dZ^p dZ^k \\
&= \log \mathbb{E}_{q(Z^k|Z^p)q(Z^p)} g(R|Z^p, Z^k) \frac{p_\theta(Z^k|Z^p)p_\theta(Z^p)}{q(Z^k|Z^p)q(Z^p)} \\
&\geq \mathbb{E}_{q(Z^k|Z^p)q(Z^p)} \log g(R|Z^p, Z^k) \frac{p_\theta(Z^k|Z^p)p_\theta(Z^p)}{q(Z^k|Z^p)q(Z^p)} \\
&= \mathbb{E}_{q(Z^k|Z^p)q(Z^p)} \log g(R|Z^p, Z^k) \\
&+ \mathbb{E}_{q(Z^k|Z^p)q(Z^p)} [\log p_\theta(Z^k|Z^p) - \log q(Z^k|Z^p)] \\
&+ \mathbb{E}_{q(Z^k|Z^p)q(Z^p)} [\log p_\theta(Z^p) - \log q(Z^p)]
\end{aligned} \tag{18}$$

For the first term, it could be decomposed as:

$$\begin{aligned}
&\mathbb{E}_{q(Z^k|Z^p)q(Z^p)} \log g(R|Z^p, Z^k) \\
&= \mathbb{E}_{q(Z^k|Z^p)q(Z^p)} \sum_{i=1}^{l_R} \log g(R|Z^p, Z^k, r_{<i})
\end{aligned} \tag{19}$$

For the second term and the third term, they could be further simplified:

$$\begin{aligned}
&\mathbb{E}_{q(Z^k|Z^p)q(Z^p)} [\log p_\theta(Z^p) - \log q(Z^p)] \\
&= -KL(q_\phi(Z^p) || p_\theta(Z^p))
\end{aligned} \tag{20}$$

$$\begin{aligned}
&\mathbb{E}_{q(Z^k|Z^p)q(Z^p)} [\log p_\theta(Z^k|Z^p) - \log q(Z^k|Z^p)] \\
&= -\mathbb{E}_{q_\phi(Z^p)} KL(q_\phi(Z^k|Z^p) || p_\theta(Z^k|Z^p))
\end{aligned} \tag{21}$$

### A.2 Implementation Details

We choose BERT<sub>base</sub> (Devlin et al., 2018)<sup>4</sup> and GPT-2 (Radford et al., 2019)<sup>5</sup> as the pre-trained language model, and implement our methods with the code in Hugging Face. To tag the pseudo knowledge label and personal memory label, the similarity score function used in Eq. 11 is implemented as unigram F1 (Dinan et al., 2019) with the code

shared at ParlAI<sup>6</sup>. In the warm up phase, we pre-train the primal task and dual task for 5000 steps and set the batch size and learning rate to be 16 and  $1e-5$  respectively. The posterior distribution of  $Z^p$  is optimized for 1000 steps with a learning rate of  $1e-5$  and a batch size of 16. In the dual learning phase, the algorithm 1 runs for 1000 steps with a batch size of 16 and a learning rate of  $1e-6$ . All modules are learned with Adam on a GTX 1080, and we set the hyperparameter of Adam to be  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$  respectively. Cosine learning schedule is applied to adjust the learning rate during training. We set the minimum learning rate to be 0 in cosine learning schedule. Gradient clip is set to 2.0 to avoid the explosion of gradient. When decoding, beam search is applied with a beam width of 5 and the minimum generated length is 10. The repetition penalty and the length penalty is set to be 1.0 and 0.0 respectively.

### A.3 Data Examples

In Table 6, We present several examples of our constructed dataset.

### A.4 Case Study

To further analyse the model’s features, a case in test set is provided in Table 7. As is shown, baseline methods in personalized dialogue has no access to external knowledge and facts, thus their generation result tend to be a little generic. And it seems that the ordinary KGC methods usually give a plain response like KnowledGPT. Our proposed method generates a more human-like response, which is in line with our expectation.

<sup>4</sup><https://huggingface.co/bert-base-uncased>

<sup>5</sup><https://huggingface.co/gpt2>

<sup>6</sup><https://github.com/facebookresearch/ParlAI/blob/master/parlai/core/metrics.py>

<b>Context</b>	<ul style="list-style-type: none"> <li>Til in 1985, Ryan white was refused re-entry to his school due to him having aids. 117 parents and 50 teachers petitioned for his ban.</li> <li>Keep in mind a lot of people were in full blown panic mode at the time . It was a fairly new disease that people didn't know much about or didn't trust what they heard.</li> <li>.....</li> <li>I don't remember , it's in a sealed box in storage . had a blonde on the cover.</li> </ul>
<b>Persona</b>	<ul style="list-style-type: none"> <li>I am sorry for that. I do believe that is dangerous thinking.</li> <li>While the choice is up to the mother, if she choses to wait that long.</li> <li>The fetus should have some rights . after a certain period of time its just not right.</li> </ul>
<b>Knowledge</b>	<ul style="list-style-type: none"> <li>In 1988 , Ryan white spoke before the president's commission on the HIV epidemic.</li> <li>Ryan white was born at St . joseph memorial hospital in kokomo, indiana, to hubert wayne and jeanne elaine white.</li> </ul>
<b>Response</b>	Oh yea I know that one . it had that girl in it too.
<b>Context</b>	<ul style="list-style-type: none"> <li>Til that during the first few minutes of the hunt for red October the film switches from Russian to English . The switch occurs on the word armageddon , which is the same in both languages.</li> <li>What , you're telling me that you can't learn a foreign language by just listening to it with no context?</li> <li>.....</li> <li>That scene is obviously when it clicks. They've been traveling for months...</li> </ul>
<b>Persona</b>	<ul style="list-style-type: none"> <li>Did you not read my comment or are you just spouting non sequiturs ?</li> <li>You don't have an elephant trunk, but how is that relevant to what I said?</li> <li>That does not make people a shill just because they disagree with others. Go back to r conspiracy with that logic .</li> </ul>
<b>Knowledge</b>	<ul style="list-style-type: none"> <li>As of January 2014 , the film held an 86 % rating at the review aggregator website rotten tomatoes, based on 66 critics.</li> <li>The hunt for red October is a 1990 American espionage thriller film produced by mace Neufeld , directed by John Mctiernan , that stars Sean Connery, Alec Baldwin, Scott Glenn, James earl jones, and Sam Neill.</li> </ul>
<b>Response</b>	That's not something that happens.

Table 6: Examples of our constructed dataset.

<b>Knowledge</b>	
	<ul style="list-style-type: none"> <li>• ...</li> <li>• Advertisement the recent experiment, however, addressed this concern head-on, while also demonstrating the engine's potential to work in space.</li> <li>• He em drive's thrust was due to the quantum vacuum behaving like propellant ions behave in a magnet ohydrodynamics drive for spacecraft propulsion.</li> <li>• Advertisement serious inquiry, indeed. it's crucial now that these tests be analyzed, replicated, and confirmed elsewhere. A peer-review and formal paper would also seem to be in order lest we get too carried away with these results. But wow. Just wow.</li> <li>• It's still early days, but the implications are mind-boggling to say the least. A full-fledged em drive could be used on everything from satellites working in low earth orbit, to missions to the moon, mars, and the outer solar system.</li> <li>• ...</li> </ul>
<b>Personal Memory</b>	
	<ul style="list-style-type: none"> <li>• ...</li> <li>• Frankly, i'd expect the constitution to win in a fight against a north korean submarine.</li> <li>• Modern diesel subs can be pretty spectacular. The israeli dolphin class are top notch, and they can be quieter than nuclear submarines if they want to be.</li> <li>• Russia'll sell new kilo class subs to pretty much anyone who'll pay.</li> <li>• ...</li> </ul>
<b>Context</b>	
<b>U1:</b>	New test suggests nasa's impossible em drive will work in space the em appears to violate conventional physics and the law of conservation of momentum the engine converts electric power to thrust without the need for any propellant by bouncing microwaves within a closed container.
<b>U2:</b>	In fairness it usually turns out to be false .
<b>U3:</b>	It's an extraordinary claim, and we don't have extraordinary proof yet. It'll be a while before we have a yes no on the emdrive.
<b>Response</b>	
GPMN	I't have a lot of the same point view.
TMN	I'm not sure it's not a lot of people convinced by that.
Transfertransfo	The police would a be a better case scenario, as the officers the people claiming to be the best cops would have police.
SKT	What the fuck is wrong with the fact that the guy is trying to prove to have extraordinary proof?
KnownedGPT	It generates thrust when it's in a vacuum.
KnownedGPT+M	it's not rocket science . it's physics .
BoB	i'm not saying it can be done, just that it can be done in a way
P <sup>2</sup> BOT	it's not like we're going to be able to get rid of this.
Ours	Yes, but it's not a scientific breakthrough. It's an extraordinary claim, and we don't have extraordinary proof yet.
Human	No one reputable is willing yet to rule out experimental error of some sort. The vacuum test rules out one of the outstanding possibilities, but it's by no means the final word. as the announcement says, the question of where the thrust is coming from deserves serious inquiry.

Table 7: A case from the test set.