Learning to Copy Coherent Knowledge for Response Generation

Jiaqi Bai[1], Ze Yang[2], Xinnian Liang[2], Wei Wang[3], Zhoujun Li[1,2]*

[1]School of Cyber Science and Technology, Beihang University, Beijing, China
[2]State Key Lab of Software Development Environment, Beihang University, Beijing, China
[3]China Resources Group, Shenzhen, China
{bjq, tobey, xnliang, lizj}@buaa.edu.cn, ww.cs.tj@gmail.com

Abstract

Knowledge-driven dialog has shown remarkable performance to alleviate the problem of generating uninformative responses in the dialog system. However, incorporating knowledge coherently and accurately into response generation is still far from being solved. Previous works dropped into the paradigm of non-goal-oriented knowledge-driven dialog, they are prone to ignore the effect of dialog goal, which has potential impacts on knowledge exploitation and response generation. To address this problem, this paper proposes a Goal-Oriented Knowledge Copy network, GOKC. Specifically, a goal-oriented knowledge discernment mechanism is designed to help the model discern the knowledge facts that are highly correlated to the dialog goal and the dialog context. Besides, a context manager is devised to copy facts not only from the discerned knowledge but also from the dialog goal and the dialog context, which allows the model to accurately restate the facts in the generated response. The empirical studies are conducted on two benchmarks of goal-oriented knowledge-driven dialog generation. The results show that our model can significantly outperform several state-of-the-art models in terms of both automatic evaluation and human judgments.

Introduction

Generating informative and attractive responses have been demonstrated as a long-standing challenge in open-domain dialog systems. Multifarious models are proposed based on the sequence-to-sequence structure (Li et al. 2016; Serban et al. 2017; Wu et al. 2018) and obtain promising results. However, these models are still subjected to generate generic and uninformative responses, such as "I am Okay"; "You are right". These bland responses will tend to degrade the experience of users.

Recently, large-scale knowledge-driven datasets have been proposed (Ghazvininejad et al. 2018; Dinan et al. 2018; Moon et al. 2019), to deal with the aforementioned problem, where the knowledge is linked with utterances to accelerate the research of knowledge-driven conversation models. Existing methods based on knowledge-driven datasets are either generative-based methods or retrieve-based methods.

*Corresponding Author

Knowledge Facts	Forrest Gump	Type	Feature Movie
		Honor	Academy Award for Best Picture
		Director	Robert Zemeckis
	Wish Dragon	Type	Cartoon
		Director	Chris Appelhans
Dialog	**User**	It's so boring. Let's talk about movies.	
	Bot	**#1:** Have you ever seen a feature movie directed by Chris Appelhans?	
		#2: Okay, have you ever seen a feature movie directed by Robert Zemeckis? This movie won the Academy Award for Best Picture!	

Figure 1: An illustrative example. #1 shows the response with conflict knowledge facts. #2 shows the response with coherent knowledge facts.

The retrieval methods select a best-matched response from the response candidates that are obtained from the database (Shang, Lu, and Li 2015; Sun et al. 2020), while the generative methods utilize the dialog context to select knowledge at first, then the selected knowledge will participate into the generation of responses (Long et al. 2017; Liu et al. 2018). In this paper, we focus on the generative-based methods since it shows much flexibility and efficiency on knowledge selection and response generation in knowledge-driven dialog systems.

It is crucial to discern appropriate knowledge facts for response generation. Although great progress has been made, common issues still exist in knowledge driven conversations. Specifically, the models are difficult to use coherent knowledge facts in response generation. Figure 1 shows an example to illustrate this problem. Here, the knowledge facts are given as the triplets form. The dialog is conducted between user and bot, the bot is required to select appropriate knowledge from knowledge facts, then generate responses to interact with users. However, if a model makes an inappropriate choice of conflict facts, the generated response will contain these contradict facts. In our example, the fact "feature movie" contradicts with the fact "Chris Appelhans" in generated response. To tackle this problem, previous works (Ghazvininejad et al. 2018; Dinan et al. 2018) mainly focused on using the encoded dialog context to discern knowledge. However, the dialog context may not always contain the appropriate information that is useful for selecting coherent knowledge, especially in the condition that the machine proactively guides the conversation (Wu et al. 2019). To tackle this problem, Recent studies (Wu et al. 2019; Liu et al. 2020) proposed the large-scale goal-oriented knowledge-driven conversation dataset, where the dialog goal is treated as the topic transition path and the topic is extracted from the knowledge facts of each dialog session. Existing works (Wu et al. 2019; Xu et al. 2020) encode dialog goal together with dialog context and knowledge facts along the same line, which neglects the effect of dialog goal for selecting coherent knowledge facts. However, in the goal-oriented knowledge-driven conversations, the dialog goal provides the vital information to help the dialog system filter the knowledge facts that are irrelevant to the information provided by the current goal, thus discern more coherent knowledge facts for response generation.

To this end, this paper proposed a *Goal-Oriented Knowledge Copy Network* (GOKC). Specifically, a goal-oriented knowledge discernment mechanism is designed to facilitate the dialog goal participating in the discernment of knowledge facts for generating prior knowledge distribution. Moreover, to further improve the coherence of discerned knowledge facts, we employ the ground-truth responses as the posterior information to supervise the training of prior knowledge distribution, which enhances the relevance between the discerned knowledge and the target response. Although the goal-oriented knowledge discernment mechanism provides an effective way to discern the coherent knowledge for generating appropriate responses, it still faces problems, i.e.: it is hard to explicitly restate the facts that appeared in input sources, and it tends to generate oov (out-of-vocabulary) words. Researchers have proposed various methods (See, Liu, and Manning 2017; Gu et al. 2016; Wang et al. 2019; Yavuz et al. 2019) to tackle these problems. However, very few of the existing works allow the model to copy tokens from multiple input sources (i.e.: dialog context, dialog goal, and knowledge facts) for response generation, while in the goal-oriented knowledge-driven conversations, the description words from knowledge and goal are usually an important component of dialog response. Therefore, we develop a context manager upon the model decoder, which can copy tokens from multi-sources. We show that generating responses using the context manager not only effectively alleviates the oov problem, but also accurately copy the facts that appeared in input sources. Our code is released at `https://github.com/jq2276/Learning2Copy`.

In summary, our main contributions are: (1) We propose a goal-oriented knowledge discernment mechanism, which can incorporate the dialog goal into the discernment of knowledge, through the supervising of posterior information, the model can generate responses with more knowledge-coherent facts. (2) We develop a context manager to copy tokens from multiple input sources, which not only maintain the accuracy of discerned knowledge that is used into the response generation but also alleviate the oov problem. (3) The proposed GOKC model combines the knowledge discernment mechanism and context manager to generate more coherent and fluent responses. (4) The experiment results on both human evaluation and automatic evaluation show that our model has superior performance than several competitive baseline models.

Related Work

Previous works on end-to-end conversation response generation (Wu et al. 2018; Zhang et al. 2018; Xing et al. 2017) benefit from the prosperity of sequence-to-sequence models on machine translation (Sutskever, Vinyals, and Le 2014), where the response generation is treated as the sequence generation problem to obtain the appropriate response by given the context from the previous dialog turn. However, these works drop into the paradigm of generating generic and uninformative responses since they lack the ability to effectively leverage external information.

To tackle the aforementioned problem, many knowledge-driven response generation models have been proposed (Zhao et al. 2019; Tian et al. 2020; Dinan et al. 2018). (Ghazvininejad et al. 2018) utilizes the memory network to store the knowledge and combine the widely used Seq2Seq model to generate responses. (Zhang, Ren, and de Rijke 2019) proposed to use context-aware knowledge preselection for guiding the knowledge selection in response generation. (Lian et al. 2019) firstly considered the target response as a part of the posterior information to participate in the knowledge selection for response generation, and obtain the impressive results on several knowledge-driven conversation datasets.

Recently, imposing goals on knowledge-driven conversation having attracted lots of research interests (Wu et al. 2019; Liu et al. 2020) since the conversation goal can provide potential guidance on knowledge selection and response generation. In this paper, we focus on goal-oriented knowledge-driven conversations. Our work is inspired by (Lian et al. 2019), where they leverage the posterior information to select the most relevant piece of knowledge for response generation, while we employ the dialog goal and dialog history as the prior information and combine with the posterior information (i.e. target response) to estimate the knowledge fact distribution, which provides a softer way to discern appropriate knowledge facts. Our work is also enlightened from the pointer generator network (PGN) (See, Liu, and Manning 2017). We extend PGN copy tokens from a single input source that can copy tokens from multiple input sources, which allows the model to accurately restate facts in generated responses.

Approach

Problem Formalization

Suppose we have a goal-oriented dialogue corpus $\mathcal{D} = \{(U_i, K_i, G_i, Y_i)\}_{i=1}^{N}$, where $\forall(U_i, K_i, G_i, Y_i) \in \mathcal{D}$, $U_i = (u_{i,1}, \cdots, u_{i,n})$ represents the dialogue history, $K_i = \{k_{i,j}\}_{j=1}^{N_K}$ is a set of knowledge facts that correspond to this conversation and each element $k_{i,j}$ could be in an arbitrary format such as a passage or a triple. The response

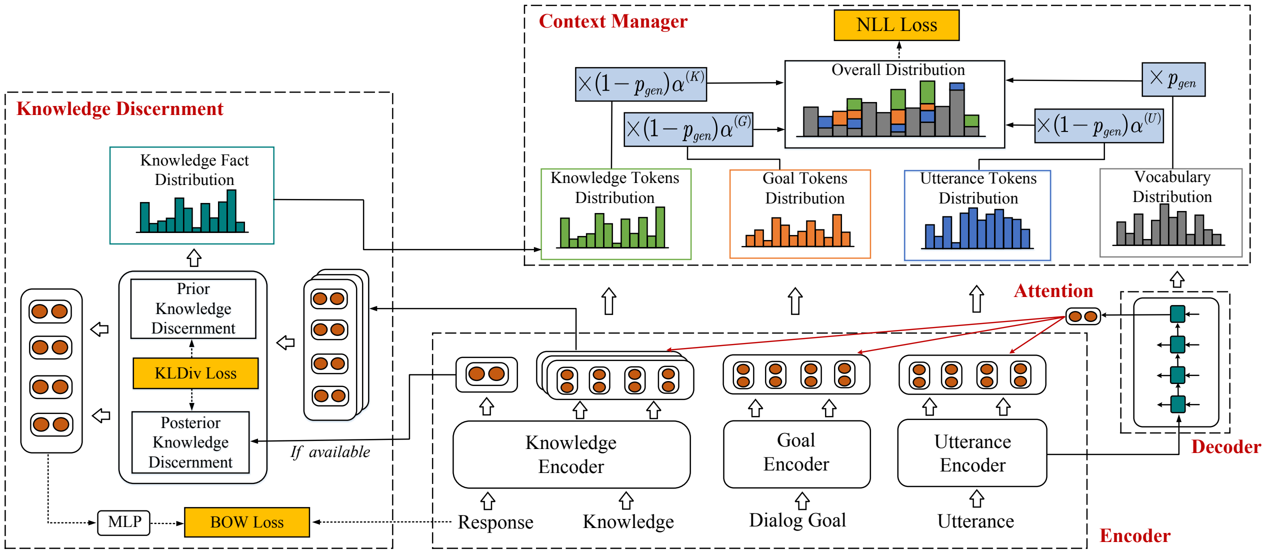

Figure 2: The architecture of the proposed GOKC model.

$Y_i = (y_{i,1}, \cdots, y_{i,m})$ is produced on the basis of the provided dialogue goal $G_i = (g_{i,1}, \cdots, g_{i,l})$ which is constructed upon the knowledge set K_i. Here, n, m, l denote the sequence lengths of U_i, Y_i, G_i respectively. Our goal is to learn a response generation model $P(Y|U, K, G)$ with $\mathcal{D}$ when given a new dialogue history U paired with the related knowledge set K, one can generate an appropriate response Y that achieved the given dialogue goal G. Thus the dialogue goal could be used to lead the dialogue session from one topic to another smoothly.

Architecture Overview

A high-level architecture overview of the proposed GOKC is shown in Figure 2. The model consists of four parts: encoder, decoder, knowledge discernment module, and context manager. (1) The encoder part encodes the knowledge K, the dialog goal G, and the dialog history U by knowledge encoder, goal encoder, and utterance encoder respectively. Please note that the response Y is also encoded by the knowledge encoder during the training stage. (2) The knowledge discernment module receives the information encoded by the encoder and obtains the knowledge facts distribution by leveraging the knowledge discernment mechanism. (3) The decoder generates hidden states at each time t, which is used to attend over each representation of input tokens and generate distributions over these tokens for copying them from input sources. Meanwhile, the decoder outputs a distribution over a fixed vocabulary. (4) The context manager combines these distributions to generate an overall distribution for decoding tokens at each time t.

Encoder

The utterance encoder, goal encoder and knowledge encoder are all built upon the bi-directional recurrent neural network (Bi-RNN) with gated recurrent unit (GRU) (Cho et al. 2014). For simplicity, denote the input sequence as $X = (x_1, x_2, ..., x_N)$, at step t, the forward RNN receives the current input x_t and the previous forward hidden state h_{t-1}^{fw} to generate current forward state h_t^{fw}. Meanwhile the backward RNN generates the backward hidden state h_t^{bw} by encoding the input x_t and h_{t-1}^{bw}. The overall output state at time t is formulated as:

$$h_t = [h_t^{fw}; h_t^{bw}] = [\overrightarrow{\mathrm{GRU}}(x_t, h_{t-1}^{fw}); \overleftarrow{\mathrm{GRU}}(x_t, h_{t-1}^{bw})] \quad (1)$$

Where $[a; b]$ means the concatenation operation between a and b. We define $o = (h_1, h_2, ..., h_N) \in \mathbb{R}^{d_e \times N}$ as the output hidden states at all time steps, and $s = [h_N^{fw}; h_0^{bw}] \in \mathbb{R}^{d_e \times 1}$ as final hidden state. Intuitively, s is a concatenation of forward hidden state and backward hidden state at their final steps.

We define the output states at all time steps of U, G, K and Y as o_U, o_G, $\{o_{K,j}\}_{j=1}^{N_K}$ and o_Y, respectively. The final states of U, G, K, Y can be denoted as s_U, s_G, $\{s_{K,j}\}_{j=1}^{N_K}$, s_Y, respectively[1].

Knowledge Discernment

The knowledge discernment module is used to discern the knowledge facts that are highly correlated to the dialog goal G and dialog context U. The knowledge discernment module first receives the encoded information of G and U and then generates a knowledge facts distribution, which is used to represent the discernment weights over the knowledge facts K for generating an appropriate response. The knowledge discernment module consists of two sub-modules: (1) The prior knowledge discernment module. (2) The posterior knowledge discernment module.

The prior knowledge discernment module obtains the prior knowledge distribution by calculating the semantic

[1]The knowledge encoder receives one piece of knowledge fact at a time. A knowledge fact is concatenated as a sequence if it is in the form of a triple.

similarity between the prior information and each knowledge fact k_i:

$$P(k_j|U,G)=\frac{\exp(s_{K,j}\cdot d_{prior})}{\sum_{i=1}^{N_K}\exp(s_{K,i}\cdot d_{prior})} \tag{2}$$

Here, d_{prior} is the prior information, which can be obtained by $d_{prior}=\tanh(\beta\odot s_U+(1-\beta)\odot s_G)$. $\odot$ means dot product. β is a gated fusion unit adapted from (Yao et al. 2017) which is used to balance the contribution between s_U and s_G. It can be obtained by $\beta=\sigma(W_p[\tanh(W_U s_U);\tanh(W_G s_G)])$, where $W_U, W_G\in\mathbb{R}^{d_e\times d_e}$ and $W_p\in\mathbb{R}^{1\times 2d_e}$ are three parameters matrices. Intuitively, the dialog goal will serve as the prior information to estimate the prior knowledge distribution when $\beta=0$, which is especially suitable for the scenario that the machine proactively leading the conversation, in which case the dialog history will be hardly considered for the discernment of knowledge facts, while the dialog goal is dominant.

The posterior knowledge discernment module obtains the posterior knowledge distribution by:

$$P(k_i|Y)=\frac{\exp(s_{K,i}\cdot d_{post})}{\sum_{j=1}^{N_K}\exp(s_{K,j}\cdot d_{post})} \tag{3}$$

Where $d_{post}=\tanh(W_{post}s_Y)$ is the posterior information and $W_{post}\in\mathbb{R}^{d_e\times d_e}$ is a parameter matrix. The introduction of posterior information can provide auxiliary guidance on knowledge discernment since the actual knowledge used in the target response is considered.

KLDiv Loss. The posterior knowledge distribution $P(k_i|Y)$ is hard to estimate at the inference time since the ground truth Y is unknown. Therefore, we employ the Kullback-Leibler Divergence loss (Kullback and Leibler 2006) to minimize the KL-distance between prior knowledge distribution $P(k_i|Y)$ and posterior knowledge distribution $P(k_i|U,G)$, which is optimized at the training stage so that $P(k_i|U,G)$ could approximate $P(k_i|Y)$ with its guidance. Hence, our model could maintain the coherence of the discerned knowledge facts at inference time. The KLDiv loss is defined as:

$$L_{KL}(\theta)=\sum_{i=1}^{N}P(k_i|\,Y)\left(\frac{P(k_i|\,Y)}{P(k_j|\,U,G)}\right) \tag{4}$$

BOW Loss. To further improve the knowledge estimation accuracy, we adapt the BOW Loss proposed by (Zhao, Zhao, and Eskenazi 2017) to ensure the relevancy between the estimated knowledge distribution and the response. Specifically, we define s_k is the weighted sum over all the knowledge representations $\{s_{k,j}\}_{j=1}^{N_K}$, which can be obtained by $s_k{=}\sum_{i=1}^{N_K}\delta_{k,i}\cdot s_{k,i}$, where $\delta_{k,i}$ is the probability distribution of knowledge fact k_i, $\delta_{k,i}{=}P(k_i|\,Y)$ if response is available, otherwise $\delta_{k,i}{=}P(k_j|\,U,G)$. The BOW loss can be obtained by:

$$L_{BOW}(\theta)=-\frac{\sum_{y_t\in\mathrm{B}}\log\varphi(y_t|\,s_k)}{|\mathrm{B}|} \tag{5}$$

Where B is the bag of words in Y, $\varphi(\cdot)$ is a two layer MLP activated by *softmax* function, which outputs the probability distribution over the fixed vocabulary.

Decoder

The decoder is composed of a forward RNN encoder with gated recurrent units. At each time t, it receives the embedding vector y_{t-1} of the word predicted at time-step $t-1$ as well as the previous decoder state h_{t-1}, and emits current hidden state $h_t\in\mathbb{R}^{d_h\times 1}$, which is formally defined by:

$$h_t=\mathrm{GRU}(y_{t-1},h_{t-1}) \tag{6}$$

Where h_t can be used to obtain the generation probability $P_{vocab}(w_t)$ over the fixed vocabulary obtained from the training set. The $P_{vocab}(w_t)$ can be obtained by:

$$P_{vocab}(w_t)=MLP(h_t) \tag{7}$$

Where $MLP(\cdot)$ is a two-layer MLP activated by *softmax* function.

Context Manager

Copying from Multi-Sources. Our model allows copying tokens from multiple input sources. Specifically, we use Φ to represent one of the input sources, where $\Phi\in\{U,G,k_1,k_2,...,k_{N_K}\}$. Given the decoder state h_t and the encoder output state o_Φ. We apply attention[2] to the o_Φ at decoder step t, which can be defined by:

$$d_t^\Phi, c_t^\Phi=Attention(o_\Phi,h_t) \tag{8}$$

Where $d_t^\Phi\in\mathbb{R}^{N_\Phi\times 1}$ is the attention distribution over each tokens appeared in Φ, and $c_t^\Phi\in\mathbb{R}^{d_e}$ is the context vector of Φ. The d_t^Φ is then aggregated to obtain the probability distribution $P_\Phi(w_t)$ over context tokens w_t, which can be computed by:

$$P_\Phi(w_t)=\sum_{\{l:\varphi_l=w_t\}}d_{t,l}^\Phi \tag{9}$$

Where φ_l is the token appeared in Φ, $d_{t,l}^\Phi$ is the attention weights corresponding to the l^{th} token in Φ. The probability of copying token w_t from the dialog history U and dialog goal G can be defined as $P_U(w_t)$ and $P_G(w_t)$ respectively. While the probability of copying tokens from knowledge K is a weighted sum of copying tokens from all the knowledge facts over the knowledge fact distribution, which is formulated as:

$$\begin{aligned}P_K(w_t)&=\sum_{i=1}^{N_K}P(k_i)\cdot P(w_t|\,k_i)\\&=\sum_{i=1}^{N_K}\delta_{k,i}\cdot\sum_{\{l:k_i^l=w_t\}}d_{t,l}^{k,i}\end{aligned} \tag{10}$$

Recall that $\delta_{k,i}$ is the probability distribution of knowledge fact k_i.

Sources Fusion. We now present the mechanism to fuse the sources by incorporating their distributions $P_U(w_t)$, $P_K(w_t)$, $P_G(w_t)$, as well as $P_{vocab}(w_t)$. We first obtain the overall knowledge representations c_t^K by:

$$c_t^K=\sum_{i=1}^{N_K}\delta_{k,i}\cdot c_t^{k,i} \tag{11}$$

[2]We have omitted the description of attention. Please refer to (Bahdanau, Cho, and Bengio 2015) for the detail.

Where $c_t^{k,i}$ is the context vector of knowledge fact k_i. Then the decoder state h_t attends over the dialog history representation c_t^U, the knowledge representation c_t^K and the goal representation c_t^G by:

$$\alpha_t, c_t = Attention\left(\left[c_t^U, c_t^K, c_t^G\right]^T, h_t\right) \tag{12}$$

Where $c_t \in \mathbb{R}^{d_e \times 1}$ is the overall representation of input sources. $\alpha_t = \left[\alpha_t^{(U)}, \alpha_t^{(K)}, \alpha_t^{(G)}\right]^T$, $\alpha^t \in \mathbb{R}^{3\times 1}$ is used to combine the distributions of input sources as shown in equation 13. We also use a *generation probability* $p_t^{gen} \in [0,1]$ described in (See, Liu, and Manning 2017) to balance the contribution between input sources and the fixed vocabulary, where $p_t^{gen} = \sigma\left(W_{gen}\left[y_{t-1}; h_t; c_t\right]\right)$, and $W_{gen} \in \mathbb{R}^{1\times(d_{emb}+d_h+d_e)}$. The overall distribution is obtained by:

$$P(w_t) = p_t^{gen} P_{vocab}(w_t) + (1 - p_t^{gen}) \cdot \sum_{\{\phi:\phi\in\{U,K,G\}\}} \alpha_t^{(\phi)} P_\phi(w_t) \tag{13}$$

Training

Apart from the KLDiv Loss and BOW Loss, we also use the NLL Loss to capture the word order information. More precisely, given a model that produces a probability $P(w_t | y_{1:t-1})$, we train the whole model end-to-end with the negative log-likelihood loss function defined as:

$$L(\theta) = -\frac{1}{|Y|}\sum_{t=1}^{|Y|} \log\left(P\left(y_t | y_{1:t-1}, U, K, G\right)\right) \tag{14}$$

Where θ denotes all the trainable model parameters. In summary, the final loss is defined by:

$$L(\theta) = L_{NLL}(\theta) + L_{BOW}(\theta) + L_{KL}(\theta) \tag{15}$$

Experiments

Datasets

We conduct our experiments on two goal-oriented knowledge-driven datasets. One is the DuConv (Wu et al. 2019), and the other is DuRecDial (Liu et al. 2020).

DuConv. A proactive conversation dataset, which consists of about 30k dialogs and 270k utterances. Each dialog contains 9.1 utterances and each utterance contains 10.6 words on average. There are 96.2 words per dialog, and 17.1 knowledge facts appeared in each dialog. For each turn of the conversation, the machine needs proactively initiate the dialog with the explicit conversation goal and the related knowledge triplets, where the conversation goal is extracted from the knowledge triplets. Similar to (Wu et al. 2019), the data is normalized by replacing the specific two topics in the knowledge triplets. Besides, we also select some relationships from knowledge triplets and substitute the object entity that appeared in these triplets with the specific tokens according to the relationships. Then, the model is required to generate responses that are closer to the conversation goal until the end of the dialog.

DuRecDial. A goal-oriented knowledge-driven conversation recommendation dataset, which contains multi-type dialogs. This dataset contains about 10k dialogs and 156k utterances. In each dialog session, there are 15.32 utterances and 21.93 knowledge facts on average. The average number of words in each utterance and knowledge fact is 11.53 and 12.73, respectively. At each dialog turn, the machine needs figure as a recommender to leads a multi-type dialog to approach recommendation targets with full consideration of dialog goal. The dialog goal is a sequence-like string obtained from the knowledge triplets. Here, we extract each goal by templets described in (Liu et al. 2020). For proving the influence of the dialog goal, we assume the complete goal is explicitly given at the beginning of the conversation, which is different from (Liu et al. 2020), where the goal needs to be completely planned before the response generation.

Comparison Models

We implement our model on both DuConv dataset and DuRecDial datasets, and compare our model with several competitive models. On the DuConv, we compared our model with: **Pointer Generator Network (PGN)**[3]: The model proposed by (See, Liu, and Manning 2017), which can copy tokens from input sources. This model has exhibited impressive performance on many natural language generation tasks. **KIC**: The model exhibits state-of-the-art performance on DuConv data which is reported in (Lin et al. 2020). In addition, we also compare our model with the baseline models mentioned in (Wu et al. 2019), which are retrieval-based models and generation-based models. On the DuRecDial, we compared our model with: **Seq2Seq**[4]: The vanilla sequence-to-sequence models proposed by (Sutskever, Vinyals, and Le 2014). **PostKS**: A knowledge-grounded response generation model proposed by (Lian et al. 2019), which utilizes a posterior knowledge selection mechanism to select an appropriate knowledge for response generation. **MGCG_R/G**: The retrieval-based models and generative-based models proposed by (Liu et al. 2020), which shows the state-of-the-art results on DuRecDial. We remove the goal planning module of these two models since the dialog goal is supposed to be known. We re-implemented these models by ourselves with the default settings described in (Liu et al. 2020).

Inplementation Details

Our model is implemented by the Pytorch Framework[5]. In our model, all of encoder and decoder have two-layer structures, each layer has 800 hidden units with the dropout rate 0.3 and the gradient clipping threshold is set to 5. The vocabulary size we used is 15k. We set the word embedding size to be 300, and initialize the embedding vectors randomly instead of using pre-trained word embeddings. We used the Adam optimizer (Kingma and Ba 2014), to minimize loss,

[3]We implement PGN using the code shared by https://opennmt.net/OpenNMT-py/.

[4]The model is implemented using the code shared by https://opennmt.net/OpenNMT-py/ with the default settings.

[5]An open-source deep learning platform (https://pytorch.org/).

the mini-batch size is 32 and the learning rate is 0.0001. We trained our model on a GPU-V100 machine. The whole training process is split into two stages. In the first stage, we train the model for 5 epochs to minimize the BOW loss only for pre-training the knowledge discernment module. In the second stage, we train the model at most 25 epochs to minimize overall loss.

Metrics

We use both human evaluation and automatic evaluation to evaluate each model, the evaluation metrics are adapted from (Wu et al. 2019; Liu et al. 2020). The automatic evaluation metrics include Bleu-1/2, F1 score, Distinct-1/2, and Perplexity. The F1 score measures the precision and recall of the generated response at character level; The Bleu score estimates the fluency of the response over n-grams level; The Distinct-1/2 are used to evaluate the diversity of response; The Perplexity is a widely used approach to estimate how likely the model to generate the ground truth response. Our model discerns the knowledge in a soft way instead of selecting knowledge facts explicitly, thus we evaluate the performance of the whole dialog but not measure the results of knowledge selection independently.

As for human evaluation, we adopted the strategy suggested by (Liu et al. 2020). We randomly selected 100 turns of dialogs from DuRecDial and employed seven well-educated annotators to evaluate the experimental models. The annotators are required to evaluate the models on 4 aspects, which are fluency, informativeness, appropriateness and proactivity. The scores are settled from $\{0, 1, 2\}$ to estimate fluency, informativeness, and appropriateness, while the proactivity scores are assigned from $\{-1, 0, 1\}$. For a fair comparison, the model name is masked during evaluation process. The agreement among the annotators is measured by the Fleiss' kappa (Fleiss 1971).

Results

Automatic Evaluation The automatic evaluation results on datasets DuConv and DuRecDial are shown in Table 1. Our approach outperforms all the comparison models, and obtain a significant improvement over most of the evaluation metrics. Specifically, On DuConv dataset, our model achieves about 1.0%, 8.7%, and 3.8% on F1, BLEU-1, and BLEU-2 compared to the best results of KIC model, which indicates that our model prefers to capture more useful information in n-gram's level, thus generate more readable and fluent responses. Besides, our model achieves about 4.2% reduction on metrics PPL compared to KIC. The metrics PPL reflects the perplexity of generated response, whose reduction indicating that the model is more likely to generate the ground truth responses. On DuRecDial dataset, our GOKC model achieves about 12.5% improvements on F1 compared to the best performance of MGCG_G, and obtains 9.0% and 19.5% improvements on BLEU-1 and BLEU-2 compared to the most competitive model MGCG_R, which is attributed to that the knowledge discernment mechanism endows the GOKC with the ability to discern the coherent knowledge facts that are close to the target response, and the context manager helps the model to accurately restate the facts appeared in discerned knowledge as well as the dialog goal and dialog context.

Model	F1	BLEU-1/2	DIST-1/2	PPL
DuConv				
norm rtr	34.73	0.291 / 0.156	0.118 / **0.373**	-
norm s2s	39.94	0.283 / 0.186	0.093 / 0.222	10.96
norm gen	41.84	0.347 / 0.198	0.057 / 0.155	24.3
PGN	42.13	0.336 / 0.211	0.087 / 0.201	10.62
KIC	44.61	0.377 / 0.262	**0.123** / 0.308	10.36
GOKC	**45.09**	**0.410 / 0.272**	0.105 / 0.272	**9.92**
DuRecDial				
seq2seq	26.08	0.188 / 0.102	0.006 / 0.013	22.82
PGN	33.95	0.243 / 0.161	0.010 / 0.039	24.28
PostKS	39.87	0.343 / 0.244	0.015 / 0.056	15.32
MGCG_R	38.82	0.379 / 0.266	**0.069 / 0.187**	-
MGCG_G	42.04	0.362 / 0.252	0.023 / 0.081	14.89
GOKC	**47.28**	**0.413 / 0.318**	0.025 / 0.084	**11.38**

Table 1: Automatic Evaluation on datasets DuConv and DuRecDial. The results of baselines on DuConv are taken from (Wu et al. 2019), norm generation is the PostKS under DuRecDial.

Model	Fluen.	Appro.	Info.	Proact.
seq2seq	1.24	1.21	0.93	0.30
PGN	1.45	1.35	1.24	0.21
PostKS	1.94	1.62	1.49	0.62
MGCG_R	**1.98**	1.46	1.57	0.49
MGCG_G	1.93	1.74	1.60	0.65
GOKC	1.96	**1.77**	**1.63**	**0.69**
kappa	0.73	0.55	0.61	0.57

Table 2: Human Evaluation on DuRecDial. "Fluen.", "Appro.", "Infor.", "Proact." respectively denote fluency, appropriateness, informativeness and proactivity.

Human Evaluation The results are shown in Table 2, we can conclude: (1) Our model achieves the highest score compared with other comparison models in terms of appropriateness, informativeness and proactivity, which demonstrates the superiority of the proposed GOKC model. (2) Our model can generate more fluent and informative responses. It is highly like that the context manager allows the model to directly copy tokens appeared in input sources, which helps the model restates the facts accurately. (3) The highest score of *Appro.* and *Proact.* indicate that our model is prone to generate more goal-relevant responses, which is attributed to that the goal-oriented knowledge discernment mechanism allows the model to choose more coherent knowledge facts, these facts facilitate the model to generate more appropriate responses and proactively lead the conversation to complete the conversation goal.

Ablation Study We take an ablation study on both DuConv and DuRecDial datasets. The key components of GOKC are removed respectively for further dissection, Ta-

Dialog Goal
[START]→ QA about <**公转自转 (Revolution)**> → Chitchat about <**王力宏 (Leehom Wang)**> → Recommending the music <**流泪手心 (Tears Palm)**>

Knowledge Facts		
王力宏 (Leehom Wang)	出生地 (Birthplace)	美国纽约(New York in America)
	简介 (Introduction)	著名歌手(A famous singer)
	获奖 (Award)	亚洲音乐最受欢迎男歌手 (The most popular male singer in Asian music)
	演唱 (Song)	公转自转(Revolution)
	演唱 (Song)	流泪手心(Tears Palm)
流泪手心 (Tears Palm)	评论 (Comment)	属于力宏早期比较有影响力的歌曲，和现在的演唱风格迥然不同。 (It belongs to one of the Leehom's early influential songs, which is disparate to his current singing style.)
公转自转 (Revolution)	评论 (Comment)	毕竟是拿了金曲奖双料大奖的一张专辑。 (After all, it was an album that won the Golden Melody Award.)

Dialog	
Dialog Context	**User:** 你知道<**公转自转**>的音乐主唱是谁吗？ (Do you know who is the singer of < **Revolution** >?) **Bot:** 是**王力宏**！ (It's **Leehom Wang**!) **User:** 知道啦，谢谢你。 (Okay, thank you.) **Bot:** 我知道他是**亚洲音乐最受欢迎男歌手**呢！ (I know that he is **the most popular male singer in Asian music**!) **User:** 他可是我最喜欢的歌手之一呢！ (He is one of my most favorite singers!)
MGCG_R	是的呢，它可是**亚洲音乐最受欢迎男歌手呢**！ (Yes, he is **the most popular male singer in Asian music**!)
MGCG_G	那你听过他的<**公转自转**>吗？**属于<unk>早期比较有影响力的歌曲，和现在的演唱风格<unk>。** Have you heard his < **Revolution** >? **It belongs to one of the <unk> early influential songs, which is <unk> to his current singing style.**
GOKC	给你推荐一首他演唱的<**流泪手心**>,**属于力宏早期比较有影响力的歌曲，和现在的演唱风格迥然不同。** (Let me recommend you <**Tears Palm**>. **It belongs to one of the Leehom's early influential songs, which is disparate to his current singing style.**)

Figure 3: Case of example dialog. We use different colors to indicate different goals and use boldface to denote knowledge facts, the red boldface means the wrong usage of knowledge facts.

Model	F1	BLEU-1/2	DIST-1/2	PPL
		DuConv		
GOKC	**45.09**	**0.410 / 0.272**	**0.105 / 0.272**	**9.92**
- gl enc	44.63	0.398 / 0.266	0.102 / 0.261	11.38
- kg discn	39.72	0.363 / 0.220	0.089 / 0.221	14.10
- cxt mgr	36.23	0.348 / 0.211	0.077 / 0.191	16.60
		DuRecDial		
GOKC	**47.28**	**0.413 / 0.318**	**0.025 / 0.084**	**11.38**
- gl enc	45.59	0.401 / 0.303	0.020 / 0.081	12.45
- kg discn	43.62	0.363 / 0.258	0.016 / 0.068	15.79
- cxt mgr	41.32	0.359 / 0.262	0.018 / 0.060	19.38

Table 3: The ablation study on datasets DuConv and DuRecDial. "-" means remove the module from the GOKC model.

ble 3 presents the results. **First**, we remove the goal encoder from our model, the goal information is concatenated with the dialog context and feed into the utterance encoder. The F1, Bleu-1&2 results decrease slightly on DuConv, but decrease distinctly on DuRecDial, it indicates that the goal encoder has a potential influence on knowledge discernment and response generation, especially when the dialog goal contains much information that is relative to the knowledge and response. **Second**, we remove the knowledge discernment module from our model, thus the computation of knowledge facts distribution depicted in Figure 2 is ignored. We observe that the results are significantly decreased overall metrics on both DuConv and DuRecDial, it is highly like that the contribution of each knowledge fact is not equal for response generation. **Third**, we remove the context manager which is substituted by concatenating the final hidden state of each encoder and feeding it into the decoder to generate a response, the results show that such alternation also decreases the performance of GOKC by a significant margin.

Case Study As shown in Figure 3, we present the responses generated by GOKC, MGCG_G, and MGCG_R. Given background knowledge, dialog goal, and dialog context, the proposed GOKC is superior to generate knowledge-enriched responses, and the knowledge facts appeared in generated responses are not only coherent to the dialog context but relevant to the dialog goal as well. Compared with GOKC, MGCG_G/R show less ability to utilize the knowledge facts that are coherent to the dialog goal, since they lack of a proper mechanism to consider the dialog into the knowledge discernment and response generation. Furthermore, our model also shows an ability to alleviate the oov problem, while the MGCG_G seems to suffer from that since it leaves out the important words appeared in knowledge facts, such as "**Leehom's**" and "**disparate**", despite these words are infrequently appeared in the fixed vocabulary.

Conclusion and Future Work

In this paper, we propose a goal-oriented knowledge copy network that could copy tokens from multiple input sources and discern coherent knowledge for response generation. The experimental results show that our model obtained impressive results on two goal-oriented knowledge-driven datasets. In the future, we intend to incorporate transfer learning into dialog system and enhance the quality of generated response by alleviating knowledge repetition problem.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grant Nos.U1636211, 61672081, 61370126), the Beijing Advanced Innovation Center for Imaging Technology (Grant No.BAICIT2016001), and the Fund of the State Key Laboratory of Software Development Environment (Grant No.SKLSDE2019ZX-17).

References

Bahdanau, D.; Cho, K.; and Bengio, Y. 2015. Neural machine translation by jointly learning to align and translate. In *3rd International Conference on Learning Representations, ICLR 2015*.

Cho, K.; van Merrienboer, B.; Bahdanau, D.; and Bengio, Y. 2014. On the properties of neural machine translation: Encoder-decoder approaches. In *Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST-8), 2014*.

Dinan, E.; Roller, S.; Shuster, K.; Fan, A.; Auli, M.; and Weston, J. 2018. Wizard of Wikipedia: Knowledge-Powered Conversational Agents. In *International Conference on Learning Representations*.

Fleiss, J. L. 1971. Measuring nominal scale agreement among many raters. *Psychological bulletin* 76(5): 378.

Ghazvininejad, M.; Brockett, C.; Chang, M.-W.; Dolan, B.; Gao, J.; Yih, W.-t.; and Galley, M. 2018. A knowledge-grounded neural conversation model. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Gu, J.; Lu, Z.; Li, H.; and Li, V. O. 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)*, 1631–1640.

Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* .

Kullback, S.; and Leibler, R. 2006. On information and sufficiency. the annals of mathematical statistics. *The Annals of Mathematical Statistics* 79–86.

Li, J.; Galley, M.; Brockett, C.; Spithourakis, G.; Gao, J.; and Dolan, B. 2016. A Persona-Based Neural Conversation Model. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 994–1003.

Lian, R.; Xie, M.; Wang, F.; Peng, J.; and Wu, H. 2019. Learning to select knowledge for response generation in dialog systems. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, 5081–5087. AAAI Press.

Lin, X.; Jian, W.; He, J.; Wang, T.; and Chu, W. 2020. Generating Informative Conversational Response using Recurrent Knowledge-Interaction and Knowledge-Copy. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 41–52.

Liu, S.; Chen, H.; Ren, Z.; Feng, Y.; Liu, Q.; and Yin, D. 2018. Knowledge diffusion for neural dialogue generation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1489–1498.

Liu, Z.; Wang, H.; Niu, Z.-Y.; Wu, H.; Che, W.; and Liu, T. 2020. Towards Conversational Recommendation over Multi-Type Dialogs. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 1036–1049. Online: Association for Computational Linguistics. doi:10.18653/v1/2020.acl-main.98. URL https://www.aclweb.org/anthology/2020.acl-main.98.

Long, Y.; Wang, J.; Xu, Z.; Wang, Z.; Wang, B.; and Wang, Z. 2017. A knowledge enhanced generative conversational service agent. In *Proceedings of the 6th Dialog System Technology Challenges (DSTC6) Workshop*.

Moon, S.; Shah, P.; Kumar, A.; and Subba, R. 2019. OpenDialKG: Explainable Conversational Reasoning with Attention-based Walks over Knowledge Graphs. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 845–854. Florence, Italy: Association for Computational Linguistics. doi:10.18653/v1/P19-1081. URL https://www.aclweb.org/anthology/P19-1081.

See, A.; Liu, P. J.; and Manning, C. D. 2017. Get To The Point: Summarization with Pointer-Generator Networks. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1073–1083.

Serban, I. V.; Sordoni, A.; Lowe, R.; Charlin, L.; Pineau, J.; Courville, A.; and Bengio, Y. 2017. A hierarchical latent variable encoder-decoder model for generating dialogues. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Shang, L.; Lu, Z.; and Li, H. 2015. Neural Responding Machine for Short-Text Conversation. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 1577–1586.

Sun, Y.; Hu, Y.; Xing, L.; Yu, J.; and Xie, Y. 2020. History-Adaption Knowledge Incorporation Mechanism for Multi-Turn Dialogue System. In *AAAI*, 8944–8951.

Sutskever, I.; Vinyals, O.; and Le, Q. V. 2014. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, 3104–3112.

Tian, Z.; Bi, W.; Lee, D.; Xue, L.; Song, Y.; Liu, X.; and Zhang, N. L. 2020. Response-Anticipated Memory for On-Demand Knowledge Integration in Response Generation. *arXiv preprint arXiv:2005.06128* .

Wang, W.; Gao, Y.; Huang, H.-Y.; and Zhou, Y. 2019. Concept pointer network for abstractive summarization. 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)*, 3067–3076.

Wu, W.; Guo, Z.; Zhou, X.; Wu, H.; Zhang, X.; Lian, R.; and Wang, H. 2019. Proactive Human-Machine Conversation with Explicit Conversation Goal. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3794–3804.

Wu, Y.; Wu, W.; Yang, D.; Xu, C.; and Li, Z. 2018. Neural response generation with dynamic vocabularies. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Xing, C.; Wu, W.; Wu, Y.; Liu, J.; Huang, Y.; Zhou, M.; and Ma, W.-Y. 2017. Topic aware neural response generation. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Xu, J.; Wang, H.; Niu, Z.; Wu, H.; and Che, W. 2020. Knowledge graph grounded goal planning for open-domain conversation generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, 9338–9345.

Yao, L.; Zhang, Y.; Feng, Y.; Zhao, D.; and Yan, R. 2017. Towards implicit content-introducing for generative short-text conversation systems. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2190–2199.

Yavuz, S.; Rastogi, A.; Chao, G.-L.; Hakkani-Tür, D.; and AI, A. A. 2019. DEEPCOPY: Grounded Response Generation with Hierarchical Pointer Networks. In *20th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 122.

Zhang, R.; Guo, J.; Fan, Y.; Lan, Y.; Xu, J.; and Cheng, X. 2018. Learning to control the specificity in neural response generation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1108–1117.

Zhang, Y.; Ren, P.; and de Rijke, M. 2019. Improving background based conversation with context-aware knowledge pre-selection. *arXiv preprint arXiv:1906.06685* .

Zhao, T.; Zhao, R.; and Eskenazi, M. 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)*, 654–664.

Zhao, X.; Wu, W.; Tao, C.; Xu, C.; Zhao, D.; and Yan, R. 2019. Low-Resource Knowledge-Grounded Dialogue Generation. In *International Conference on Learning Representations*.