Knowledge-Based Systems 234 (2021) 107499

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Topic-level knowledge sub-graphs for multi-turn dialogue generation

Jing Li [a], Qingbao Huang [a,c,e,*], Yi Cai [b,c], Yongkang Liu [a], Mingyi Fu [a], Qing Li [d]

[a] School of Electrical Engineering, Guangxi University, Nanning, Guangxi, China
[b] School of Software Engineering, South China University of Technology, Guangzhou, China
[c] Key Laboratory of Big Data and Intelligent Robot (SCUT), MOE of China, Guangzhou, China
[d] Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
[e] Guangxi Key Laboratory of Intelligent Control and Maintenance of Power Equipment, Guangxi, China

ARTICLE INFO

Article history:
Received 9 May 2021
Received in revised form 2 August 2021
Accepted 14 September 2021
Available online 2 October 2021

Keywords:
Knowledge-based dialogue system
Multi-turn dialogue generation
Knowledge graph
Topic-level

ABSTRACT

Previous multi-turn dialogue approaches based on global Knowledge Graphs (KGs) still suffer from generic, uncontrollable, and incoherent responses generation. Most of them neither consider the local topic-level semantic information of KGs nor effectively merge the information of long dialogue contexts and KGs into the dialogue generation. To tackle these issues, we propose a Topic-level Knowledge-aware Dialogue Generation model to capture context-aware topic-level knowledge information. Our method thus accounts for topic-coherence, fluency, and diversity of generated responses. Specifically, we first decompose the given KG into a set of topic-level sub-graphs, with each sub-graph capturing a semantic component of the input KG. Furthermore, we design a Topic-level Sub-graphs Attention Network to calculate the comprehensive representation of both sub-graphs and previous turns of dialogue utterances, which then decoded with the current turn into a response. By using sub-graphs, our model is able to attend to different topical components of the KG and enhance the topic-coherence. We perform extensive experiments on two datasets of DuRecDial and KdConv to demonstrate the effectiveness of our model. The experimental results demonstrate that our model outperforms existing strong baselines.

1. Introduction

Multi-turn open domain dialogue generation system, aiming at endowing a machine the ability to chat with people naturally and continuously, which has been viewed as a crucial technique of the human–computer interaction and attracted much more attention recently. Compared with a single-turn dialogue generation system, the multi-turn dialogue generation not only needs to pay attention to the diversity of the generation response, but also needs to ensure the coherence between dialogue history and the response [1–3]. To effectively handle the multi-turn dialogue information, some researchers process the dialogue history turn by turn to reduce the impact from the noise of history dialogue [4–6]. Although such models have enhanced the coherence of responses, they still struggle to understand the semantic information of the context and tend to generate generic and meaningless responses.

To facilitate language understanding and generation, some researchers integrate structured knowledge Graphs (KGs) into the dialogue generation [7,8]. With external knowledge, these models could chat with human on a more commonsense and informative level. However, due to a massive amount of noise hidden in KGs, an over reliance on external commonsense knowledge beyond context could lead to deviating from the main theme or even resulting in logical conflict with dialogue history.

To effectively distill the context-related information from KGs, some studies propose N-hop [9–11] or PageRank [12] to create sub-graphs for guiding generation. To some extent, these models have indeed got a balance between enhancing informativeness and limiting noise for generation. Nevertheless, the topic generation of dialogue remains uncontrollable and incoherent. A graph about the topic of the movie star *Cecilia Cheung* is shown in Fig. 1(a). The node of *Cecilia Cheung* contains eight child nodes, which belong to different sub-topics as the *Achievement*, the *Opinion*, and the *Actor*, semantically. The aforementioned sub-graph-based models switch themes within two-hop domain of the initial node. However, these multi-hop methods ignore the different semantics of sub-topics, so that models cannot determine which sub-topic to generate for the next response. We consider that the fine-grained topic information is capable of according with the direction of global topic flow throughout the dialogue history more correctly. Different from above methods, we propose to decompose the KG into a set of topic-level sub-graphs, with each sub-graph containing a integrate sub-topic which represents a direction of the conversation, cf. Fig. 1(b).

* Corresponding author.
E-mail addresses: laizirise@gmail.com (J. Li), qbhuang@gxu.edu.cn (Q. Huang), zhkkun@gmail.com (Y. Liu).

https://doi.org/10.1016/j.knosys.2021.107499
0950-7051/

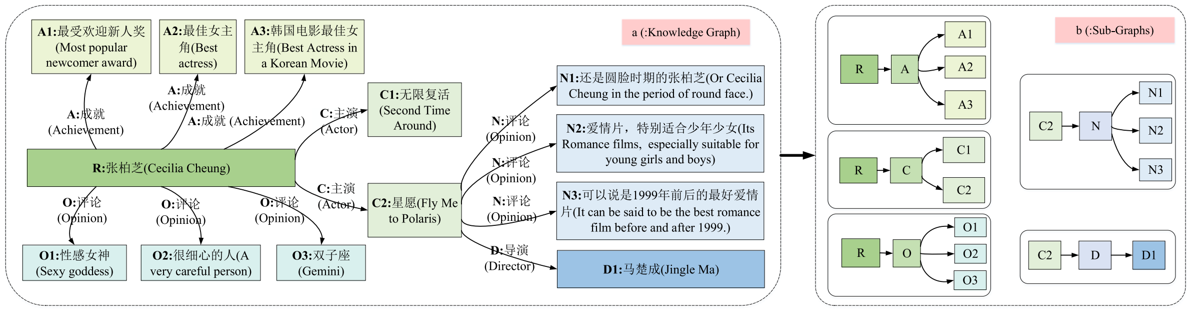

Fig. 1. An example of a knowledge graph (a) and a set of topic-level sub-graphs (b). *Cecilia Cheung* have eight child nodes, which belong to different topic components, and each component represents a sub-topic direction, i.e., *Achievement* (A), *Opinion* (O) and *Actor* (C). These fine-grained topic information is capable of according with the direction of global topic flow of contexts.

Table 1
An example of the difference between previous turns and current turn.

Turn 1 (User)	Who is the actor of Fly Me to Polaris?
Turn 2 (Bot)	It's Cecilia Cheung.
Turn 3 (User)	I like her, Cecilia Cheung is the best actress.
Turn 4 (Bot)	Yes, she have won the Most Popular Newcomer Award.

We also model previous turns and the current turn of dialogue separately to further fuse the KGs and multi-turn dialogue. For multi-turn dialogue, the long dialogue utterances of current turn and previous turns have different effects on the dialogue generation. Specifically, previous turns usually contain main topics information of the whole conversation, such as topics of the movie star *Cecilia Cheung* and the movie *Fly me to Polaris*. While the current turn may be a continuation, transfer or reversal of previous turns. As shown in Table 1, the response of *Turn 3* is the continuation of *Turn 2* instead of *Turn 1*, and the *Turn 4* is the continuation of the topic of *best actress* in *Turn 3*. Whereas, existing methods of merging the KGs with jointed two kinds of utterances, lack the grasp of the key context-related topical information in KGs. These methods may generate disfluent and incoherent response.

Consequently, we propose a **T**opic-level **K**nowledge-aware Dialogue **G**eneration model (TKG). Our model obtains information of sub-graphs more relevant to previous utterances to enable the fluency, and decodes the corresponding response between current utterance and related sub-graphs to facilitate diversity and topic-coherence of generated responses. A major advantage of this design is that, by modelling multiple distinct topic-level sub-graphs and two types of utterances, diversity and coherence are naturally enabled.

Specifically, the relation in a triple contains essential semantic information, we first construct a Node-level Graph (NG) by regarding all the relation edges as nodes. Next, we decompose the NG into a set of Topic-level sub-Graphs (TGs) to attend to copious information of a original KG.

Afterwords, we develop a Topic-level sub-Graph Attention Network (TGAN) to calculate representations of previous utterances with NG and TGs, respectively. With TGAN, our model captures major topic-level components and nodes in the KG that are related with previous utterances. Finally, we feed the representation of the current utterance and fused representations of two levels to the decoder to generate a response. In conclusion, our method utilizes the KG in both the node level and the topic level, and exploits the previous utterances and the current utterance in distinct approaches to generate diverse, controllable, and topic-coherent responses.

We conduct the dialogue generation evaluation of our model on the DuRecDial [13] and the KdConv [14] datasets. The experimental results show that our model improves the performance of diversity and topic-coherence. The contributions of our work are summarized as follows:

- To the best of our knowledge, this is the first attempt to decompose the KGs into the topic-level sub-graphs, which helps the model capture more relevant fine-grained topic-level semantic information in KGs, and generate more informativeness and topic-coherent responses.
- We propose a Topic-level Knowledge-aware dialogue Generation model (TKG) to construct topic-level sub-graphs and effectively process contexts and graphs. By taking advantage of the model, the fluency, topic-coherence, and diversity of dialogue generation have been enhanced.
- We design a Topic-level sub-Graph Attention Network (TGAN) to better merge the contexts and sub-graphs. With TGAN, our model obtains the information of topic-level sub-graphs more relevant to previous utterances, and decodes the corresponding response via the current utterance to facilitate informativeness and topic-coherence.
- The automatic and human evaluation results on two datasets of DuRecDial and KdConv demonstrate the superiority of our proposed model. Particularly, BELU-2 and F1 are improved by 25% and 12.7% on DuRecDial, respectively, and DIST-2 by 7.5% on KdConv with multi domains.

2. Related work

Recently, dialogue generation systems have gained more attention and achieved promising results. To generate a coherent and fluent response, early works employ a simple encoder–decoder network to generate [3,15]. However, these models tend to generate repeated words with less semantic information. To tackle this problem, [1,4,16] use a memory network to storage the long-term context, [2] uses an attention network to get the important information of context, and [17] uses a two-stage encoder–decoder model to generate informativeness response. Furthermore, due to the lengthy and redundant context, the above models still suffer to generating the incoherent response. To filter irrelevant contexts information, [6] proposes to encode dialogue contexts turn by turn and tracking the contexts dynamically. To further discover the contextual relationship, topic models [18–20] are proposed to find the topics in context. Additionally, [5] utilizes the topic model to improve topical relevance for multi-turn dialogue generation. However, because of the difficulty of semantic understanding, these models are prone to generate generic and incorrect responses.

Thus, many researchers begin to introduce external knowledge [21] to enhance the ability of semantic understanding tasks [22–24] and knowledge representation methods is then proposed to facilitate the ability of understanding [25–27]. While for the dialogue generation systems, unstructured knowledge information models [28,29] and structured KGs-enhanced models [30–32] are aiming at facilitating the understanding of dialogue utterances. To integrate the proper knowledge into models, [7,33,34] propose to use knowledge grounded attention network, incorporating knowledge information which is related to utterances.

Although utilizing KGs increases the diversity of response, the problem of generating incoherent conversations remains. Some researches introduce goal-oriented methods to generate response in line with expectations [35–37]. [38] introduces a topic fact-based knowledge-aware approach to control the direction of generation. Nevertheless, these systems are limited by the massive KGs, which cannot make full use of KGs. The introduced multi-hop relations concepts enhance the ability of incorporating more KGs information into models [8–11]. For example, [11] uses N-hop and PageRank [12] approaches to get different sub-graphs for the task of recommendation.

To better utilize the KGs, different from the methods which use KGs triples directly, we propose to decompose the KGs into a set of topic-level sub-graphs to pay attention to fine-grained topic-level semantic information. Then we design a model, focusing on different topic-level semantic components of the KG and processing the dialogue utterance of current turn and previous turns separately, to generate more fluent, diverse, and topic-coherent responses.

3. Methodology

This section presents our Topic-level Knowledge-aware Dialogue Generation model (TKG) in details.

3.1. Overview

Given a current utterance $X_n = \{x_1, \ldots, x_m\}$ with m words, a sequence of previous utterances $C = \{c_1, \ldots, c_n\}$, with n words, a set of KG triples $G = \{g_1, g_2, \ldots, g_t\}$, where each triple is composed of entity pairs with their relations as [*entity*$_1$, *relation*, *entity*$_2$], our aim is to generate a topic-coherent and diverse response $Y = \{y_1, \ldots, y_l\}$.

To cope with two problems described in introduction, we propose the TKG model (cf. Fig. 2), which consists of three modules: (1) a Sub-Graphs Construction and Decomposition module, (2) a Topic-level sub-Graphs Attention Network, and (3) a Response Decoder. We first construct the node-level knowledge graph (NG) by changing relation edges of the KG to nodes. Then, we decompose the NG into a set of Topic-level sub-Graphs (TGs), with each sub-graph representing a topic-level semantic component of the KG. Next, we employ a Topic-level sub-Graph Attention Network (TGAN) to learn the node-level and topic-level semantic representations of both graphs and previous utterances. Finally, we aggregate two levels of representations and decode the current utterance to generate a topic-coherent and informative response.

3.2. Sub-graphs construction and decomposition

The input KG composed of triples, which have several sub-topic semantic components. Besides, each relation edge in a KG contains the semantic information. There are three forms of connection between entity nodes and edges in KG, i.e., 1-1, 1-n, and n-1.

To obtain above three forms explicitly, and distill the topic-level information from KGs triples, we first convert the implicit relation edges to explicit nodes, and make up to a Node-level Graph $NG = \{V, E\}$. This process is described detailedly in Algorithm 1.

To further distinguish the topic of each node in the NG and get topic-level information, we decompose the NG into a set of Topic-level sub-Graphs $TGs = [T_0, \ldots, T_t]$, each T_i in TGs represents a sub-Graph. In detail, there are two kinds of nodes should be decomposed in NG. First, if a node has more than one child (1-n), we define this node as a sub-topic node (n_t) of a NG.

Algorithm 1 *Node-Level Graph Construction*

Input:
A knowledge graph G;
Output:
A Node-level graph $NG = \{V, E\}$;
Node set V is initialized with a node dict of {"*word*" : "*root*", "*index*" : 0, "*child*" : 0}, edge set E is initialized with a set of list;
for each triple $t = [e1, r, e2]$ in G **do**
  Initialize a new node dict n for each element in t;
  Record the child number and index on the n;
  update $V = V \cup n$
  update E
end for
return NG;

Second, if a node just has one child (1-1), we define this node as a sub node (n_d) of the n_t. Following the above two rules, the children of n_t, the child of n_d, and the related edges constitute to a TG, $T_i = \{V_i^t, E_i^t\}$, where $V_i^t \subseteq V_i$ and $E_i^t \subseteq E_i$. Each node $V_i^t = \{s_j\}$ is a word or a sentence. Fig. 1(b) is an example of *TGs*. This process is stated detailedly in Algorithm 2.

Algorithm 2 *Topic-level Graph Construction*

Input:
A Node-level Graph $NG = \{V, E\}$;
Output:
A set of Topic-level sub Graphs $TGs = [T_0, \ldots, T_t]$
for each node n_n in V **do**
  if n_n have more than one child **then**
    Initialize a new sub-graph T_i with the node dict n_n and the edge list e_n
    for each node dict n_t in V **do**
      if $e[n_n, n_t]$ in NG have an edge **then**
        update e_n and e_t
        update $T_i \cup n_t$
        update $T_i \cup n_e$
      end if
    end for
    update TGs = $TG \cup T_i$
  end if
end for
return TGs

3.3. Topic-level sub-graphs attention network

A NG contains all node-level information, while TGs contain topic-level information. In consideration of decomposing the KG with logical integrity into fragmentary sub-graphs risks introducing semantic ambiguity, we separately encode the NG and the TGs in two levels: global-node-level, local-topic-level. In order to obtain the node and topic information related to previous utterances, we first encode previous utterances via an RNN-based passage encoder, and then initialize the representations of the KG and TGs to get utterances-to-node representation and utterances-to-topic representation. Next, we add the above two representations together. Finally, we use an Attention-based Gated Graph Neural Network mechanism (Attn-GGNN) [39] to encode the representation. With TGAN, we can obtain more utterances-related information and filter the redundant information in the KG.

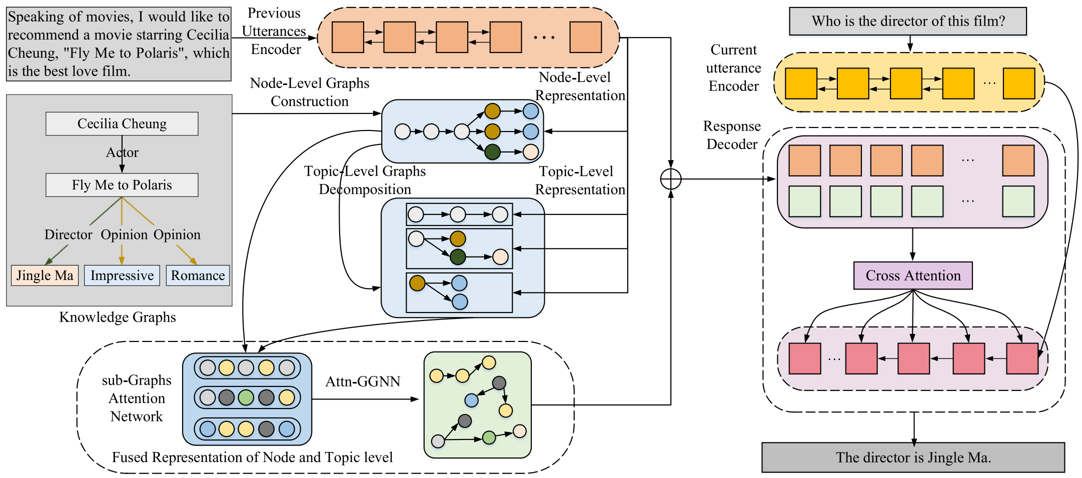

Fig. 2. An overview of our proposed TKG model. Our model takes a node-level Graph (NG) reconstructed from a full KG, and decomposes the NG into a set of Topic-level sub-Graphs (TGs). We design a Topic-level sub-Graph Attention Network (TGAN) to calculate and encode the representation by two levels of the node and the topic. Furthermore, we decode a response by fusing the current utterance and encoded representations. By leveraging sub-graphs, our model enables diverse, controllable and topic-coherent dialogue generation.

Previous utterances encoding. Given a sequence of previous utterances, we combine them into one sentence $C = [c_1, \ldots, c_n]$ with n words. And we employ the Bi-directional Gated Recurrent Unit (BiGRU) [40] to encode previous utterances C:

$$H_C = [h_1, \ldots, h_n], h_j = BiGRU(h_{j-1}, e(c_j)), \tag{1}$$

where $h_j = \{\overrightarrow{h}_j; \overleftarrow{h}_j\}$ is the concatenated bi-directional hidden states of the word c_j, and $e(c_j)$ is the word embedding of c_j. Then we concatenate the last hidden states of the utterance encoder in both directions as the final representation $R_C = [\overrightarrow{h}_n, \overleftarrow{h}_1]$ of previous utterances.

Two levels of graphical representation. To capture not only the semantic information of each node, but also the topic information of both the previous utterance and graphs. We propose to calculate the similarity between the C and KG in two levels. (1) Node-level: the model calculates the attention scores between each word in the C and each node in the NG, to get the node-level representation globally. (2) Topic-level: the model also calculates the attention scores between each word in the C and each sub-graph in the TGs, to get the topic-level representation locally.

Given a node-level Graph $NG = \{V_i, E_i\}_{i=0:N_k}$ and a set of sub-Graphs $TGs = [T_t]_{t=0:N_t}$, we first calculate the embedding of each word W_j in the node V_i:

$$e(V_i) = [e(W_j)]_{j=0:N_{V_i}}, \tag{2}$$

where N_{V_i} is the number of words. Then, we use $e(V_i)$ to initialize the embedding of TGs as e_{TGs} and NG as e_{NG}, respectively:

$$e_{NG} = \{e(V_i), E_i\}_{i=0:N_k}, \tag{3}$$

where N_k is the number of nodes in the NG.

$$e_{TGs} = [T_t]_{t=0:N_t}, T_t = \{e(V_i), E_i\}_{i=0:N_{T_t}}, \tag{4}$$

where N_t is the number of sub-graphs in the TGs and N_{T_t} is the number of nodes in a sub-graph T_t. Then, to better utilize the TG structure into the model, we obtain two-level representations by computing the attention scores of context-to-node and context-to-topic.

For the node-level, given a initialized NG, we calculate the attention distribution of R_C over all the words in the node V_i in NG as follows:

$$\alpha_l^{V_i} = \frac{exp(Attn(R_C, W_l))}{\sum_{k=N_{V_i}}^{0} exp(Attn(R_C, W_k))}, \tag{5}$$

where $\alpha_l^{V_i}$ is the attention coefficient of previous utterances encoding R_C over a word W_l in the node V_i. The representation of node V_i is then calculated by attention-weighted sum of the embedding of its constitute words:

$$h_p^{V_i} = \sum_{p=N_{V_i}}^{0} \alpha_p^{V_i} W_p^{V_i}, \tag{6}$$

which constitute to the representation of the NG $H_{node} = [h_q]_{q=0:N_k}$.

For topic-level, we calculated the attention distribution of R_C over all the nodes $[V_i]_{i=0:N_{T_j}}$ in the T_j:

$$\alpha_l^{T_j} = \frac{exp(Attn(R_C, v_l))}{\sum_{k=N_{T_j}}^{0} exp(Attn(R_C, v_k))}, \tag{7}$$

where $\alpha_l^{T_j}$ is the attention coefficient of the R_C over a node v_l in the sub-graph T_j. The representation of node T_j is then calculated by attention-weighted sum of the embedding of its constitute nodes:

$$h_p^{T_j} = \sum_{p=N_{T_j}}^{0} \alpha_p^{T_j} v_p^{T_j}, \tag{8}$$

which constitute to the representation of TGs $H_{topic} = [h_q]_{q=0:N_t}$.

Finally, we add two kinds of levels representations together to get the node-enhanced representation of the topic-level:

$$H = H_{node} + H_{topic}. \tag{9}$$

3.4. Graphs encoding

The fused topic-level representation $H = [H_j]_{j=0:N_t}$ ensures the model to capture not only the meaning of each word in nodes, but also the key information of each topical component. Moreover, it is important to pay different attention to neighbour nodes of each node and aggregate information from their neighbour nodes. Thus, we utilize the Attn-GGNN to update the representations by aggregating information from neighbour nodes.

Given a fused topic-level sub-graph representation $H_i = [h_l]_{l=0:N_{t_i}}$ in TGs, an aggregation function is applied to each node v_p in H_i to collect messages from nodes directly connected to v_p.

We first calculate the attention coefficient of v_p to each node v_q:

$$\alpha_{pq}^{k} = \frac{exp(Attn(h_p^{(k)}, h_q^{(k)}))}{\sum_{r=N_{t_i}}^{0} exp(Attn(h_p^{(k)}), h_r^{(k)})}, \tag{10}$$

where k is the layer of state transitions, h is the node representation, and N_{t_i} is the number of nodes in a TG T_i. Each node is then distinguished by their incoming and out-coming edges, derived as follows:

$$h_{\overrightarrow{N}_p}^{(k)} = \sum_{v_p \in \overrightarrow{N}_p} \alpha_{pq}^{(k)} h_q^{(k)}, \tag{11}$$

$$h_{\overleftarrow{N}_p}^{(k)} = \sum_{v_p \in \overleftarrow{N}_p} \alpha_{pq}^{(k)} h_q^{(k)}, \tag{12}$$

where $\overrightarrow{N}_p$ and $\overleftarrow{N}_p$ denote the sets of incoming and outgoing edges of v_p, respectively. Next we employ a GRU to update the node state by incorporating information of two directions of edges $h_p^{(k+1)} = GRU(h_p^{(k)}, [h_{\overrightarrow{N}_p}^{(k)}; h_{\overleftarrow{N}_p}^{(k)}])$, which constitute to the updated representation of TGs $H^{(k)}$. Finally, we concatenate the $H^{(k)}$ with previous utterances R_C to obtain the two-level-enhanced context representation T_D, which contains weighted topic and node information.

3.5. Response decoder

Based on the fused representation of the node level and the topic level, the model obtains the main topical information related with previous utterances. Furthermore, we decode the current utterance to generate a topic-coherent response. In detail, we employ another attention-based RNN and coverage mechanisms [41] as the decoder to generate the response Y. First, to make the decoder aware of the current utterance, we employ the BiGRU to encode the current utterance X to initialize the decoder hidden states H_X. Then, the decoder takes the two-level-enhanced context representation T_D from the encoder as the attention memory, and utilizes the utterance representation H_X to generate the output sequence one word at a time. Finally, we incorporate the coverage mechanisms [41] to encourage the decoder to apply diverse words. Specifically, at each step, we calculate the sum of attention distributions over all previous decoder steps. And a coverage loss is then computed to penalize repeatedly attending to the same locations of the input dialogue.

4. Experiments

This section describes the experimental details, including datasets, training settings, and performance on both automatic and human evaluation.

4.1. Settings

Datasets. To verify the validity of our model, we evaluate our model on the task of knowledge-based multi-turn dialogue, using recently released datasets of DuRecDial [13] and KdConv [14]. Although DuRecDial is the dataset of multi-turn dialogue recommendation task, the quality of dialogue utterances are higher, and both of two datasets reflect the process of topic transfer. Thus, we utilize the both of above datasets to better training and evaluation our model.

In detail, a DuRecDial sample consists of 4 parts, a user profile, a user profile related knowledge graph, a set of goals, and a set of dialogue utterances. Considering that the information of goals may affect the fluency of the dialogue, we abandon the goals to make batter use of DuRecDial. Besides, we follow the original data splits of DuRecDial, which contains 10190 dialogues and 155477 utterances. And for KdConv, which contains three domains of film, music, and travel. We first train and evaluate our model in three domains separately, with each domain contains 1200 dialogues. Then, we fuse three domains into one dataset to train and evaluate our model. The fused dataset contains 3600 dialogues totally. We follow the original data splits of KdConv as 8/1/1 for train/test/validation.

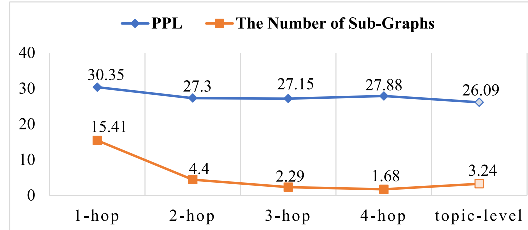

Fig. 3. Comparison between different segmentation methods for KGs on the dataset of KdConv with multi domains. The ordinate represents the average number of sub-graphs and PPL scores of each segmentation method.

Metrics. We conduct automatic evaluation with the following metrics: **F1**, **PPL**, **ROUGE**, and **BLEU1/2/3/4**. BLEU [42] measures the average n-gram overlap the generated response and reference sentence. Both BLEU and ROUGE evaluate the fluency of generated response. Moreover, we report the distinct DIST [43] results to reveal the diversity in all generated words. Critically, we also conduct human evaluation to prove the superiority of our model. Annotators of human evaluation evaluate the quality of generated dialogue from four important aspects: fluency, topic-coherent, informativeness, and knowledge-coherence.

Implementation details. Our model is implemented over PyTorch framework. And our implementation of Graph Encoder is inspired by the public code provided by DQG [44]. For fair comparison, we use the original implement settings of the DQG to apply them on DurecDial and KdConv datasets. All words in contexts and KGs share a sentence encoder with hidden units of 300 dimensions. Word embeddings are initialized with 300 dimensional pre-trained FastText [45] for all the baselines. For the graph encoder, the number of layers k of state transitions of knowledge encoder is set to 1 and hidden state size is 300. The batch size is set to 8 and learning rate is set to le-5. We train all the models for 25 epochs on a GPU-2080Ti machine.

Baselines. We compare our proposed model with the following strong baselines:

- **Seq2Seq** [46]: A standard encoder–decoder model without attention over the input context which is widely used in open-domain dialogue.
- **Seq2Seq (Attn/GGNN)** [44]: An encoder–decoder model augmented with an attention mechanism (Attn) and a GGNN mechanism (GGNN) to better incorporate the knowledge into the model.
- **MGCG (G/R)** [13]: The SOTA generation (G) and retrieval (R) model of datasets DuRecDial, which present a multi-goal driven conversation generation framework to handle multi-type dialogues. Simultaneously, We also implement a model without strategy encoder (MGCG) to fair competition.
- **PostKS (fusion/concat)** [47]: A knowledge grounded model with incorporated knowledge by proposed Hierarchical Gated Fusion Unit (HGFU) mechanism (fusion), and by a GRU decoder where knowledge is concatenated (concat).
- **HRED** [3]: A hierarchical recurrent encoder–decoder model, incorporating context with a specific context RNN encoder.

Table 2
Automatic Evaluation on DuRecDial and multi-domain of KdConv, separately. Best performance is highlighted in **boldface**, the second performance is highlighted in underline.

Model	DuRecDial					KdConv-multi domains				
	BLEU-2	BLEU-1	PPL	DIST-2	F1	BLEU-2	BLEU-1	PPL	DIST-2	F1
S2S	0.088	0.166	29.20	0.029	19.01	0.064	0.141	33.27	0.043	27.56
S2S + Attn	0.114	0.202	13.93	0.051	20.38	0.069	0.162	32.12	0.106	31.79
S2S + GGNN	0.221	<u>0.336</u>	12.99	0.045	33.94	0.076	0.169	<u>31.68</u>	0.096	32.68
PostKS (fusion)	0.225	0.330	16.54	0.070	27.77	0.177	0.303	–	<u>0.107</u>	<u>43.18</u>
PostKS (concat)	0.220	0.320	<u>12.15</u>	0.033	26.87	<u>0.201</u>	<u>0.319</u>	–	0.073	41.77
MGCG	0.180	0.240	24.40	0.017	22.83	–	–	–	–	–
MGCG (G) [13]	0.219	–	17.69	0.052	<u>36.81</u>	–	–	–	–	–
MGCG (R) [13]	<u>0.232</u>	–	–	**0.187**	33.93	–	–	–	–	–
TKG (Our)	**0.290**	**0.414**	**11.26**	<u>0.145</u>	**41.49**	**0.219**	**0.333**	**31.35**	**0.115**	**47.45**

Table 3
Automatic Evaluation on single-domain of KdConv. The Best performance is highlighted in **boldface**, the second performance is highlighted in underline.

Model	Metrics					
	BLEU-4	BLEU-3	BLEU-2	BLEU-1	DIST-2	DIST-1
KdConv-Film						
S2S [14]	0.053	0.087	0.145	<u>0.275</u>	0.080	0.029
HRED [14]	0.054	0.087	0.147	**0.280**	0.081	0.029
PostKS (fusion)	<u>0.071</u>	<u>0.098</u>	<u>0.148</u>	0.263	<u>0.093</u>	<u>0.041</u>
TKG (Our)	**0.079**	**0.115**	**0.156**	0.253	**0.112**	**0.044**
KdConv-Music						
S2S [14]	0.078	0.114	0.173	0.296	0.124	0.039
HRED [14]	0.080	0.116	0.175	0.297	0.117	0.038
PostKS (fusion)	<u>0.094</u>	<u>0.128</u>	<u>0.190</u>	<u>0.337</u>	<u>0.120</u>	<u>0.050</u>
TKG (Our)	**0.118**	**0.156**	**0.203**	**0.358**	**0.126**	**0.056**
KdConv-Travel						
S2S [14]	**0.189**	**0.222**	<u>0.273</u>	0.370	<u>0.136</u>	<u>0.043</u>
HRED [14]	<u>0.180</u>	0.213	0.267	0.369	0.133	0.040
PostKS (fusion)	0.129	0.168	0.221	<u>0.395</u>	0.093	0.040
TKG (Our)	0.170	<u>0.218</u>	**0.273**	**0.395**	**0.200**	**0.087**

Table 4
Automatic Evaluation of ablation study of TKG model. Best performance is highlighted in **boldface**.

Model	DuRecDial						
	BLEU-4	BLEU-2	BLEU-1	DIST-2	DIST-1	F1	ROUGE-L
TKG w/o TGs,Cu	0.151	0.264	0.382	0.100	0.038	38.59	0.347
TKG w/o TGs	0.162	0.265	0.383	0.126	0.045	39.10	0.376
TKG w/o NG	0.176	0.285	0.406	0.129	0.044	40.18	0.391
TKG	**0.182**	**0.290**	**0.414**	**0.145**	**0.047**	**41.49**	**0.394**

4.2. Evaluation

Automatic evaluation results. Tables 2 and 3 show the overall evaluation results of the TKG model and baselines on two datasets of KdConv and DuRecDial. From the table, we can observe that our model achieves the state-of-the-art performance in each metric of DuRecDial dataset, and overall BLEU and DIST on KdConv. On DuRecDial dataset of Table 2, S2S + GGNN model obtains better BLEU and DIST than S2S model, which is due to the GGNN mechanism facilitate the understanding of semantic information of dialogue and knowledge. PostKs model achieves a much higher score on BLEU, which is mainly benefited from its knowledge selecting process. However, while the dialogue turns increased, the BLEU and DIST scores of PostKs decreased and the generated responses are incoherent, which illustrate that PostKs have difficulty in dealing with the long dialogue context. MGCG (R) model has achieved a strong performance on BLEU and DIST, on account of taking BERT [48] as the encoder to select proper knowledge. Our model performs even better than both of MGCG and PostKs, which verifies the effectiveness of our model in generating fluent and appropriate responses. Concretely, On the DuRecDial, the BLEU and F1 yield considerable increments. Compared with Seq2Seq + GGNN model, the BLEU-2 of our model improves 0.069 points, the DIST-2 improves about 320%. In particular, our model achieves significant higher F1 score, which improves 7.56 points compared with MGCG(R) model. Furthermore, We can observe the similar trend on KdConv.

On the multi-domains dataset of KdConv in Table 2, our model achieves significant improvement with highest BLEU and DIST. Specifically, compared with PostKs(fusion) model, the BLEU-2 of our model improves 0.042 points, the DIST-2 improves about 107%. On the single-domain dataset of film, music, and travel in Table 3, our model still performs better on diversity of generated responses. The DIST and BLEU-2 scores of our model are the highest compared with each baselines on single-domain KdConv dataset.

In order to verify the influence of different segmentation methods on our proposed model, we use the N-hop method to segment the KG on the KdConv with multi domains. We compare our proposed topic-level segmentation methods with the following N-hop segmentation methods. A N-hop sub-graph includes all the edges and nodes reachable from the root of KG by routes of length N. We consider four cases of N = 1, 2, 3 and 4. As shown in Fig. 3, each segmentation methods obtains different numbers of sub-graphs. Compared with the 2-hop method, the number of sub-graphs of our topic-level segmentation method decreased by 1.16 graphs, but the PPL increased by 1.21 points. Compared with the 4-hop segmentation method, the number of sub-graphs of our model increased by 0.95 graphs, but increased by 1.06 points on PPL. By analysing the responses generated by these N-hop methods, we reveal that these methods tend to generate topic-incoherent responses, which mainly due to the fact that the N-hop segmentation method ignore the integrity of the subtopics on KGs. Compared with all other N-hop segmentation methods, the PPL result of our topic-level segmentation achieves best performance, which demonstrates that our proposed topic-level segmentation method is more effective.

The superior performance on two datasets illustrates responses generated by our proposed model TKG are more diverse, fluent, and appropriate.

Ablation study. To verify the effectiveness of each module on TKG, we perform ablation studies on DuRecDial dataset from three aspects, as shown in Table 4. Firstly, we remove the TGs constraint from TKG model and jointly encode previous utterances (Pu) and current utterance (Cu) (w/o TGs, Cu), leaving only NG module, which resulting in significant decline on all metrics. This proves that our proposed model of TKG and TGAN enhance the fluency, diversity, and quality of the dialogue generation. Secondly, to verify the effectiveness of separately encoding Pu and the Cu, we only remove the TGs constraint (w/o TGs) from

Table 5
Human evaluation results of our models and baselines. *Flu.*, *Topic-Coh.*, *Inf.*, and *Know-Rel* denotes the *Fluency*, *Topic-Coherency*, *Informative*, and *Knowledge-Related*, respectively. Each metric is rated on a 0-5 scale (5 for best).

Model	Turn-Level				Dialogue-Level			
	Flu	Topic-Coh	Inf	Know-Rel	Flu	Topic-Coh	Inf	Know-Rel
Ground-truth	4.47	4.44	4.06	0.54	3.11	3.86	2.81	0.38
MGCG (G)	3.20	1.55	2.12	0.35	2.16	1.08	1.47	0.24
PostKS (fusion)	3.49	2.94	2.69	0.19	2.24	2.05	1.87	0.13
TKG w/o TGs (Our)	3.61	3.29	3.21	0.26	2.32	2.29	2.23	0.18
TKG (Our)	**3.90**	**3.73**	**3.23**	**0.37**	**2.71**	**2.56**	**2.25**	**0.25**

our model. Compared with the first experiment (w/o TGs,Cu), the BLEU-4 of the second experiment (w/o TGs) promotes 0.1 points and DIST-2 promotes 0.026 points, which demonstrates that it is necessary to decode Pu and the Cu separately. Furthermore, to explore the influence of NG on the generated response, we also remove the NG from encoder of our model (w/o NG). Compared with the second experiment (w/o TGs), the results show that TGs contribute more to the performance of TKG than NG, since KGs information cannot be attended to efficiently without TGs. And compared with the TKG, BLEU and DIST scores of both TGs constraint (w/o TGs) and NG constraint (w/o NG) drop significantly, which suggests that both the NG and TGs facilitate the diversity and fluency.

The ablation experimental results illustrate that the topic-level decomposition of KGs enhances the understanding of knowledge and boost the fluency of generated responses. Separately processing previous utterances and the current utterance significantly

Human evaluation results. impacts the informativeness and diversity of the generation model.

We conduct human evaluation at turn-level and dialogue-level. For turn-level, the model generates a response conditioned on a given context and related knowledge, while the given context is from given datasets. For dialogue-level, given the relevant KG, the model generates the response by conversing with a human, while the context is obtained iteratively by chatting with a human. For both of levels, the annotators evaluate the quality of responses according to the dialogue history and the related knowledge, in terms of fluency, relevance, and informativeness. We conduct human evaluation on 200 dialogues of the DuRecDial dataset, generated by our annotators (Ground Truth), the TKG model, the TKG model without TGs, MGCG (G) model, the PostKS (fusion) model, and the S2S+GGNN model. Each dialogue of random test samples contains at least 7 turns. We ask three annotators to rate the 200 dialogues for each model between 0 (worst) and 5 (best) on four metrics: (1) *Fluency*, indicates whether the grammar of response with the correct logic. (2) *Topic-Coherence*, indicates whether the topic of response is coherent with the current turn. (3) *Informative*, indicates whether the words of response is diverse and informativeness. (4) *Knowledge-Related*, indicates whether entities in the utterance exist in the given KG, which rated between 0 (in) and 1 (out).

As shown in Table 5, on both turn and dialogue levels, our TKG model outperforms the strong baselines on all metrics. Specifically, compared with model without TGs module, the topic-coherence of our model absolute improves 0.53 points on turn level and 0.27 points on dialogue level, which demonstrates that TGs play a crucial role in promoting topic-coherence. Similarly, our model also achieves the best performance in Know-Rel, which indicates that the response generated by our model obtain more relevant knowledge. Human evaluations indicate that the responses generated by our TKG model preform better informativeness and topic-coherence in human view, which are closest to ground-truth.

Case study. We present the responses generated by our proposed model TKG and some baselines in Fig. 4. We observe that the MGCG model tends to generate incoherent and short response with less information. While our model without TGs or Cu module tends to generate knowledge-irrelevant or context-incoherent responses. PostKs performs much better than TKG without TGs module, generating more related response with more information from given KGs. However, with the dialogue turn increased, PostKs tends to generate repetitive responses that are irrelevant to given KGs, e.g., PostKs fails to extract correct the key graph information in the third turn. Compared with the above four models, our TKG model without NG module is able to utilizes the related knowledge into generation and generates topic-coherent response even in the last few turns during the conversation. While TKG model can integrate more related knowledge into response and generate more informative and natural responses. For example, responses in the third turn generated by TKG and TKG without NG module are more related to the topic of *Jacky Cheung*, which is the main topic in previous turns. However, the other four responses generated by baselines are irrelevant to *Jacky Cheung*.

In addition, we can observe that some of the responses of TKG are almost the same as the ground truth, the main reason is that, by using topic-level sub-graphs and separately modelling the previous utterances and current utterance, TKG obtains context-relevant nodes from KGs, which is then decoded to generate a topic-coherent response. This further shows that our model facilitate the ability to capture the key knowledge and information, generating more diverse and coherent responses.

5. Conclusion and future work

In this paper, we propose a TKG model by exploring the topic-level semantic information of KGs. Our model calculates the representation of incorporated previous utterances into sub-graphs, and decodes response from the fused representation and the current utterance. Experimental results on automatic evaluation and human evaluation illustrate that our TKG is superior to all the strong baselines, and the topic-level decomposition method based on KGs can facilitate the diversity and topic-coherence of the dialogue generation.

In future work, we intend to utilize our approach to both contexts and KGs to further control the dialogue generation. There are at least two potential future directions. First, integrating KGs into dialogue models to generate a more appropriate response will become an urgent problem to be solved. The method of this paper is helpful for researchers to explore the interrelationship between KGs and dialogues. Secondly, it is beneficial to the future development to improve our model in terms of implementation method to reduce the training cost of the model. We hope our work can provide insights into the design of explainable and controllable models for the dialogue generation task.

Knowledge Graph Tripts
[张学友,生日 1961 - 7 - 10], [张学友,获奖,**十大中文金曲最高荣誉金针奖**], [张学友, 获奖, 香港电影金像奖 最佳 男配角], [张学友,简介,娱乐圈零绯闻的明星], [张学友,简介,娱乐圈真正的好男人],[张学友,成就,**十大劲歌金曲最受欢迎男歌星**],[张学友,演唱,**她来听我的演唱会**], [她来听我演唱会, 评论, 夜晚的心事], [她来听我的演唱会,评论,**《 她来听我的演唱会 》，整首歌通过一个女人几次爱的绽放为主线， 张学友的歌和演唱会为 暗线， 展现了女人一生受的伤害和坚守的情感经历**，再通过张学友自己角度去演唱这首描述着来自世界 任何角度的一个 关于她的故事， 似乎也是我们许多人 一生爱情的故事，其实 也就是每个人在一生不同阶段共同的情感 经历。] ([**Jacky Cheung**, birthday, **1961-7-10**], [Jacky Cheung, award-winning, **Golden Needle Award for Top Ten Chinese Golden songs**], [Jacky Cheung, award-winning, Hong Kong Film Awards-Best supporting Actor], [Jacky Cheung, introduction, star of zero gossip in the entertainment industry], [Jacky Cheung, introduction, a really good man in the entertainment industry],[Jacky Cheung,Achievement,**The top ten hit songs are the most popular male singers**.] [Jacky Cheung, singing, she came to my concert]. [**she comes to listen to my concert**, comments, thoughts of the night], [she comes to my concert, comments, **"she came to my concert." The whole song blossoms through a woman's love several times. Jacky Cheung's songs and concerts are dark lines, showing the hurt and emotional experience of a woman's life.** Then sing this story about her from any angle of the world through Jacky Cheung's own point of view, which seems to be the love story of many of us all our lives. In fact, it is everyone's common emotional experience at different stages of life.])
Persona Profile
姓名：周玲倩 (Name: Linlin Zhou), 居住地：成都 (Address: Chengdu)，年龄区间：18-25 (Age-Range: 18-25), 职业状态：学生 (Profession: Student), 喜欢的明星：杨幂、张学友 (Favorite star: Mi Yang, Xueyou Zhang) 同意的新闻：杨幂的新闻; 张学友 的新闻 (Favorite news: News of Mi Yang, News of Xueyou Zhang)
Turn 1: Dialogue Context: “在吗？今天几号了？” (Hey, What's the date today?)
Ground-truth: “好的亲，今天是**2018年7月10日**” (Okay, today is June 10, 2018.) **PostKs:** “今天时2018年7月8日” (Today is July 7,2018.) **MGCG (G):** “是巨蟹座” (It's Cancer.) **TKG w/o TGs, Cu (Our):** “你好啊，今天是2018 年5月17日。(Hello, today is May 17, 2018.)” **TKG w/o TGs (Our)**: “今天是2018年5月17日” (Today is May 17, 2018.) **TKG w/o NG (Our):** “今天是2018年11月6日。(Today is December 6, 2018.)” **TKG (Our):** “今天是2018年1月18日” (Today is January 18, 2018.)
Turn 2: Dialogue Context: “在吗？今天几号了→ 好的亲，今天是2018年7月10日” (Hey, What's the date today?.→Okay, today is June 10, 2018)
Ground-truth: “嘿嘿，今天还是明星**张学友**的生日呢，很巧呢” (Hey, today is the birthday of star **Jacky Cheung**. What a coincidence.) **PostKs:** “不客气呢，今天也是张学友的生日哦” (You are welcome, today is also **Jacky Cheung**'s birthday.) **MGCG (G):** “是的呢， 他还获得过世界音乐大奖全球销量最高华人歌手 奖” (Yes, he also won the World Music Awards and Chinese singer award in the world.) **TKG w/o TGs, Cu (Our):** “好的，你知道吗？今天也是周杰伦的生日呢。(Okay, do you know today is also Jay Chou's birthday.)” **TKG w/o TGs (Our):** “不客气哦，提醒你，今天多云，东北风” (You're welcome. I'd like to remind you that it's cloudy and northeasterly today.) **TKG w/o NG (Our):** “ 你知道吗，今天还是张学友的生日呢，你知道吗？(Do you know today is **Jacky Cheung**'s birthday?)” **TKG (Our):** “嘿嘿， 今天 也 是 **张学友** 的 生日 呢” (Today is also **Jacky Cheung**'s birthday.)
Turn 3: Dialogue Context: “在吗？今天几号了→ 好的亲，今天是2018年7月10日→嘿嘿，今天还是明星**张学友**的生日呢，很巧呢 →原来是偶像的生日啊，他还是娱乐圈零绯闻的明星。” (Hey, What's the date today?→Okay, today is June 10, 2018→ Hey, today is the birthday of star **Jacky Cheung**. What a coincidence.→ Its the birthday of an idol, and he was also the star of zero gossip in the entertainment industry.)
Ground-truth: “嗯， 那为 您 推荐 他 唱 的 **《 她来听我的演唱会 》** 整首歌 通过 一个 女人 几次 爱 的 绽放 为主 线， 张学友 的 歌 和 演唱会 为 暗线， 展现 了 女人 一生 受 的 伤害 和 坚守 的 情感 经历 (Well, the whole song **"she comes to my Concert"** recommended to you blossoms through a woman's love several times. **Jacky Cheung's** songs and concerts are dark lines, showing the hurt and emotional experience of a woman's life)” **PostKs:** “是的，他曾多次入围金马奖最佳男演员，曾获台湾金曲奖最佳国语男歌手呢” (Yes, he has been nominated for the Golden Horse Award for best actor many times, and has won the Taiwan Golden Melody Award for best Mandarin male singer.) **MGCG (G):** “是的呢，**他还获得了十大中文金曲最高荣誉金针奖** (Yes, he also won the Golden Needle Award, the highest honor of the top ten Chinese songs.)” **TKG w/o TGs, Cu (Our):** “嗯，那为您推荐他主演的《 暴疯语 》，很不错的一部电影呢 。(Well, I'd like to recommend to you his movie "Crazy talk" , which is very good.)” **TKG w/o TGs (Our):** “是啊，他还获得过十五座金曲奖 (Yeah, he also won the 15th Golden Music Awards.)” **TKG w/o NG (Our):** “是的呢，对了，他的电影**《 她来听我的演唱会 》**很不错呢，推荐给你。(Yes, by the way, his movie "**she comes to my Concert**" is very good. I recommend it to you.)” **TKG (Our):** “我 给 你 推荐 一个 他 的 歌 **《 她来听我的演唱会 》**， 整首歌 通过 一个 女人 几次 爱 的 绽放 为主 线， **张学友** 的 歌 和 演唱会 为 暗线， 展现 了 女人 一生 受 的 伤害 和 坚守 的 情感 经历 。(I recommend to you one of his songs, **"she comes to my concert"**. The whole song blossoms through a woman's love several times. **Jacky Cheung's** songs and concerts are dark lines, showing the hurt and emotional experience of a woman's life.)”

Fig. 4. Case study of DuRecDial.

CRediT authorship contribution statement

Jing Li: Data curation, Methodology, Experiments, Writing – original draft. **Qingbao Huang:** Conceptualization, Writing – review & editing, Funding acquisition, Supervision. **Yi Cai:** Supervision, Writing – review & editing. **Yongkang Liu:** Experiments, Writing – original draft. **Mingyi Fu:** Investigation, Writing – original draft. **Qing Li:** Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by National Natural Science Foundation of China (62076100), and the collaborative research grants from the Fundamental Research Funds for the Central Universities, China, SCUT (D2210010, D2200150, and D2201300), the Science and Technology Planning Project of Guangdong Province, China (2020B0101100002), the Science and Technology Programs of Guangzhou, China (201704030076, 201707010223, 201802010027, 201902010046), the Science and Technology Key Projects of Guangxi Province, China (2020AA21077007), and the Hong Kong Research Grants Council, China (PolyU1121417, C1031-18G), and an internal research grant from the Hong Kong Polytechnic University, China (1.9B0V).

References

[1] B. Zhang, X. Xu, X. Li, Y. Ye, X. Chen, Z. Wang, A memory network based end-to-end personalized task-oriented dialogue generation, Knowl. Based Syst. 207 (2020) 106398.

[2] H. Zhang, Y. Lan, L. Pang, J. Guo, X. Cheng, ReCoSa: Detecting the relevant contexts with self-attention for multi-turn dialogue generation, in: ACL, 2019, pp. 3721–3730.

[3] A. Sordoni, Y. Bengio, H. Vahabi, C. Lioma, J.G. Simonsen, J. Nie, A hierarchical recurrent encoder-decoder for generative context-aware query suggestion, in: CIKM, 2015, pp. 553–562.

[4] J. Wang, J. Liu, W. Bi, X. Liu, K. He, R. Xu, M. Yang, Dual dynamic memory network for end-to-end multi-turn task-oriented dialog systems, in: COLING, 2020, pp. 4100–4110.

[5] H. Zhang, Y. Lan, L. Pang, H. Chen, Z. Ding, D. Yin, Modeling topical relevance for multi-turn dialogue generation, in: IJCAI, 2020, pp. 3737–3743.

[6] J. Xu, Z. Lei, H. Wang, Z. Niu, H. Wu, W. Che, Enhancing dialog coherence with event graph grounded content planning, in: IJCAI, 2020, pp. 3941–3947.

[7] M. Ghazvininejad, C. Brockett, M. Chang, B. Dolan, J. Gao, W. Yih, M. Galley, A knowledge-grounded neural conversation model, in: AAAI, 2018, pp. 5110–5117.

[8] H. Zhou, T. Young, M. Huang, H. Zhao, J. Xu, X. Zhu, Commonsense knowledge aware conversation generation with graph attention, in: IJCAI, 2018, pp. 4623–4629.

[9] H. Zhang, Z. Liu, C. Xiong, Z. Liu, Grounded conversation generation as guided traverses in commonsense knowledge graphs, in: ACL, 2020, pp. 2031–2043.

[10] H. Ji, P. Ke, S. Huang, F. Wei, X. Zhu, M. Huang, Language generation with multi-hop reasoning on commonsense knowledge graph, in: EMNLP, 2020, pp. 725–736.

[11] R. Sarkar, K. Goswami, M. Arcan, J.P. McCrae, Suggest me a movie for tonight: Leveraging knowledge graphs for conversational recommendation, in: COLING, 2020, pp. 4179–4189.

[12] L. Page, S. Brin, R. Motwani, T. Winograd, The PageRank citation ranking: Bringing order to the web, Technical Report, Stanford InfoLab, 1999.

[13] Z. Liu, H. Wang, Z. Niu, H. Wu, W. Che, T. Liu, Towards conversational recommendation over multi-type dialogs, in: ACL, 2020, pp. 1036–1049.

[14] H. Zhou, C. Zheng, K. Huang, M. Huang, X. Zhu, Kdconv: A Chinese multi-domain dialogue dataset towards multi-turn knowledge-driven conversation, in: ACL, 2020, pp. 7098–7108.

[15] I.V. Serban, A. Sordoni, Y. Bengio, A.C. Courville, J. Pineau, Building end-to-end dialogue systems using generative hierarchical neural network models, in: AAAI, 2016, pp. 3776–3784.

[16] Z. Tian, W. Bi, D. Lee, L. Xue, Y. Song, X. Liu, N.L. Zhang, Response-anticipated memory for on-demand knowledge integration in response generation, in: ACL, 2020, pp. 650–659.

[17] J. Li, W. Monroe, T. Shi, S. Jean, A. Ritter, D. Jurafsky, Adversarial learning for neural dialogue generation, in: EMNLP, 2017, pp. 2157–2169.
[18] X. Yan, J. Guo, Y. Lan, X. Cheng, A biterm topic model for short texts, WWW, 2013, pp. 1445–1456.
[19] J. Feng, Y. Rao, H. Xie, F.L. Wang, Q. Li, User group based emotion detection and topic discovery over short text, World Wide Web (2020) 1553–1587.
[20] J. Wu, Y. Rao, Z. Zhang, H. Xie, Q. Li, F.L. Wang, Z. Chen, Neural mixed counting models for dispersed topic discovery, in: ACL, 2020, pp. 6159–6169.
[21] Y. Cai, W. Chen, H. Leung, Q. Li, H. Xie, R.Y. Lau, H. Min, F.L. Wang, Context-aware ontologies generation with basic level concepts from collaborative tags, Neurocomputing (2016) 25–38.
[22] R. Yin, K. Li, G. Zhang, J. Lu, A deeper graph neural network for recommender systems, Knowl. Based Syst. 185 (2019) 105020.
[23] W. He, Y. Sun, M. Yang, F. Ji, C. Li, R. Xu, Multi-goal multi-agent learning for task-oriented dialogue with bidirectional teacher-student learning, Knowl. Based Syst. 213 (2021) 106667.
[24] Y. Cai, H. Leung, A formal model of fuzzy ontology with property hierarchy and object membership, in: ER, in: Lecture Notes in Computer Science, Vol. 5231, 2008, pp. 69–82.
[25] Z. Li, X. Wang, J. Li, Q. Zhang, Deep attributed network representation learning of complex coupling and interaction, Knowl. Based Syst. 212 (2021) 106618.
[26] X. Wang, L. Chai, Q. Xu, Y. Yang, J. Li, J. Wang, Y. Chai, Efficient subgraph matching on large RDF graphs using MapReduce, Data Sci. Eng. 4 (2019) 24–43.
[27] F. Zhang, X. Wang, Z. Li, J. Li, TransRHS: A representation learning method for knowledge graphs with relation hierarchical structure, in: IJCAI, 2020, pp. 2987–2993.
[28] W. Zhang, K. Song, Y. Kang, Z. Wang, C. Sun, X. Liu, S. Li, M. Zhang, L. Si, Multi-turn dialogue generation in E-commerce platform with the context of historical dialogue, in: EMNLP, 2020, pp. 1981–1990.
[29] Y. Song, Z. Liu, W. Bi, R. Yan, M. Zhang, Learning to customize model structures for few-shot dialogue generation tasks, in: ACL, 2020, pp. 5832–5841.
[30] W. Yu, C. Zhu, Z. Li, Z. Hu, Q. Wang, H. Ji, M. Jiang, A survey of knowledge-enhanced text generation, 2020, CoRR http://arxiv.org/abs/2010.04389.
[31] M. Nakano, K. Komatani, A framework for building closed-domain chat dialogue systems, Knowl. Based Syst. 204 (2020) 106212.
[32] B. Jiang, J. Yang, C. Yang, W. Zhou, L. Pang, X. Zhou, Knowledge augmented dialogue generation with divergent facts selection, Knowl. Based Syst. 210 (2020) 106479.
[33] X. Lin, W. Jian, J. He, T. Wang, W. Chu, Generating informative conversational response using recurrent knowledge-interaction and knowledge-copy, in: ACL, 2020, pp. 41–52.
[34] J. Jung, B. Son, S. Lyu, Attnio: Knowledge graph exploration with in-and-out attention flow for knowledge-grounded dialogue, in: EMNLP, 2020, pp. 3484–3497.
[35] J. Xu, H. Wang, Z. Niu, H. Wu, W. Che, Knowledge graph grounded goal planning for open-domain conversation generation, in: AAAI, 2020, pp. 9338–9345.
[36] W. Wu, Z. Guo, X. Zhou, H. Wu, X. Zhang, R. Lian, H. Wang, Proactive human-machine conversation with explicit conversation goal, in: ACL, 2019, pp. 3794–3804.
[37] J. Tang, T. Zhao, C. Xiong, X. Liang, E.P. Xing, Z. Hu, Target-guided open-domain conversation, in: ACL, 2019, pp. 5624–5634.
[38] S. Wu, Y. Li, D. Zhang, Y. Zhou, Z. Wu, Topicka: Generating commonsense knowledge-aware dialogue responses towards the recommended topic fact, in: IJCAI, 2020, pp. 3766–3772.
[39] L. Pan, Y. Xie, Y. Feng, T. Chua, M. Kan, Semantic graphs for generating deep questions, in: ACL, 2020, pp. 1463–1475.
[40] K. Cho, B. van Merrienboer, Ç. Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation, in: EMNLP, 2014, pp. 1724–1734.
[41] Z. Tu, Z. Lu, Y. Liu, X. Liu, H. Li, Modeling coverage for neural machine translation, in: ACL, 2016, pp. 1724–1734.
[42] K. Papineni, S. Roukos, T. Ward, W. Zhu, Bleu: a method for automatic evaluation of machine translation, in: ACL, 2002, pp. 311–318.
[43] J. Li, M. Galley, C. Brockett, J. Gao, B. Dolan, A diversity-promoting objective function for neural conversation models, in: NAACL-HLT, 2016, pp. 110–119.
[44] D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate, in: ICLR, 2015.
[45] E. Grave, P. Bojanowski, P. Gupta, A. Joulin, T. Mikolov, Learning word vectors for 157 languages, in: LREC, 2018.
[46] I. Sutskever, O. Vinyals, Q.V. Le, Sequence to sequence learning with neural networks, in: NIPS, 2014, pp. 3104–3112.
[47] R. Lian, M. Xie, F. Wang, J. Peng, H. Wu, Learning to select knowledge for response generation in dialog systems, in: IJCAI, 2019, pp. 5081–5087.
[48] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: pre-training of deep bidirectional transformers for language understanding, in: NAACL-HLT, 2019, pp. 4171–4186.