Intention Reasoning Network for Multi-Domain End-to-end Task-Oriented Dialogue

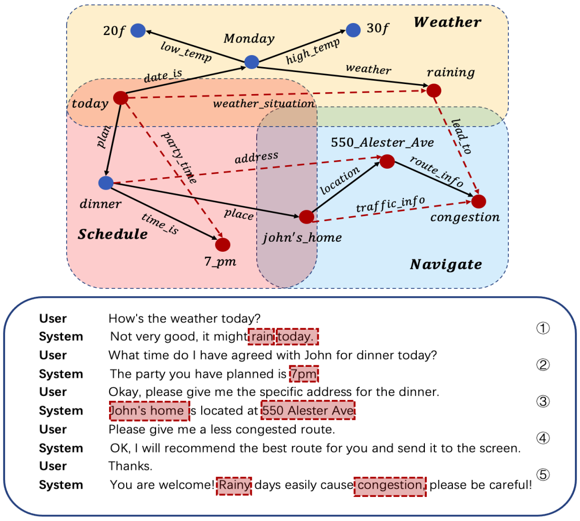
Zhiyuan Ma[1], **Jianjun Li**[1]*, **Zezheng Zhang**[1], **Guohui Li**[1], **Yongjing Cheng**[2]
[1] Huazhong University of Science and Technology (HUST), China
[2] National Unveristy of Defense Technology (NUDT), China
{zhiyuanma,jianjunli,zezhengzhang,guohuili}@hust.edu.cn
davidcheng1001@163.com

Abstract

Recent years has witnessed the remarkable success in end-to-end task-oriented dialog system, especially when incorporating external knowledge information. However, the quality of most existing models' generated response is still limited, mainly due to their lack of fine-grained reasoning on deterministic knowledge (*w.r.t.* conceptual tokens), which makes them difficult to capture the concept shifts and identify user's real intention in cross-task scenarios. To address these issues, we propose a novel intention mechanism to better model deterministic entity knowledge. Based on such a mechanism, we further propose an intention reasoning network (IR-Net), which consists of joint and multi-hop reasoning, to obtain intention-aware representations of conceptual tokens that can be used to capture the concept shifts involved in task-oriented conversations, so as to effectively identify user's intention and generate more accurate responses. Experimental results verify the effectiveness of IR-Net, showing that it achieves the state-of-the-art performance on two representative multi-domain dialog datasets.

1 Introduction

Task-oriented dialogue systems are designed to help users to achieve specific goals such as schedule arrangement or weather inquiry via natural language. Compared with traditional pipeline dialogue systems (Young et al., 2013) that include multiple modules each requiring a huge amount of human effort to design, end-to-end approaches (Gülçehre et al., 2016; Wen et al., 2017; Eric and Manning, 2017; Eric et al., 2017; Zhao et al., 2017; Quan et al., 2019; Moon et al., 2019; Jung et al., 2020; Dai et al., 2020) that can directly output system responses with plain text as input have recently gained much attention. In recent years, sequence to sequence (Seq2Seq) models have dominated the study of end-to-end task-oriented dialog systems, and many memory augmented Seq2Seq models have been proposed (Bordes et al., 2017; Madotto et al., 2018; Wu et al., 2019; Wen et al., 2018; Qin et al., 2019; Reddy et al., 2019; Wang et al., 2020), which exploit both dialog history and domain-specific knowledge base (KB) to incorporate KB information and perform knowledge-based reasoning for better performance.



User How's the weather today?
System Not very good, it might rain today. ①
User What time do I have agreed with John for dinner today?
System The party you have planned is 7pm ②
User Okay, please give me the specific address for the dinner.
System John's home is located at 550 Alester Ave ③
User Please give me a less congested route.
System OK, I will recommend the best route for you and send it to the screen. ④
User Thanks.
System You are welcome! Rainy days easily cause congestion, please be careful! ⑤

Figure 1: Example of task-oriented dialog including concept shifts from the SMD dataset (Eric et al., 2017). The solid arrows indicate the existed relationships between entities, and the dotted arrows indicate the latent entity relationships (captured by IR-Net). The colored dotted boxes in the dialog indicate generated entities.

Though achieving remarkable progress, existing memory augmented Seq2Seq models still suffer from the following two limitations. **First**, prior models rely heavily on the soft attention mechanism (Vaswani et al., 2017) to generate responses by adopting a weighted sum over the embeddings of memory triples (from both dialog history and external KB) as the output representation. Since the representation acquired in this way is scattered by the context, it is difficult to model deterministic knowledge *w.r.t.* specific conceptual tokens. Take

[*] Corresponding author.

Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pages 2273–2285
November 7–11, 2021.

the dialog in Figure 1 as an example, when answering the user's query on today's (Monday's) weather, the response generated by existing Seq2Seq models may be ambiguous or even incorrect due to the impact of contextual triples such as *(Tuesday, weather, sunny)* and *(Wednesday,weather, cloudy)*. **Second**, the soft attention mechanism is inherently not suitable for performing fine-grained (token-level) multi-hop reasoning, which makes it hard to capture user's real intention to generate accurate responses, especially in complex cross-task scenarios where concept shifts (Zhang et al., 2020) may occur. For example, in Figure 1, when the system is asked *"please give me the specific address for the dinner"*, it is expected to explore the pivot "*john's_home*" that connects the start token "*diner*" (in Schedule domain) with the target token "*550_Alester_Ave*" (in Navigate domain), and finally return the answer "*550_Alester_Ave*". Existing attention-based models generally fail to perform such a token-level multi-hop reasoning, which hampers them from obtaining accurate responses.

To address the aforementioned limitations, we propose a novel Intention Reasoning Network (IR-Net), which is a memory-augmented Seq2Seq model equipped with an intention reasoning module that is responsible for obtaining an intention-aware representation, with the goal of generating more accurate responses. Specifically, to address the first limitation, we propose a novel intention mechanism (Sec. 2.3.1), which directly incorporates the tail-token of a knowledge triple by comparing the similarity between the query vector and the triple's head-token to model deterministic knowledge. Based on the intention mechanism, we further address the second limitation by proposing an intention reasoning module that consists of token-level joint reasoning and multi-hop reasoning (Sec. 2.3.2), which are responsible for capturing specific target information from breadth and depth respectively to generate intention-aware representations, so as to improve the integrality and accuracy of the generated responses.

We conduct experiments on two publicly available multi-domain datasets, namely SMD (Eric et al., 2017) and Multi-WOZ 2.1 (Budzianowski et al., 2018). The experimental results show that IR-Net consistently outperforms the current state-of-the-art models in both automatic and human evaluation. To our best knowledge, we are the first to effectively explore fine-grained token-level intention reasoning in multi-domain end-to-end task-oriented dialog.

2 Model Description

Our proposed model is based on a Seq2Seq dialog generation model (Sec. 2.1), which encodes dialogue history X and knowledge base B and ultimately obtains a response sequence Y. An external memory module $M = [X; B]$ is set up for knowledge query (Sec. 2.2). Moreover, to capture potential concept shift and user's intention to generate a fluid and intention-aware response, an intention reasoning module based on a novel intention mechanism is proposed (Sec. 2.3). The workflow of our proposed model is depicted in Figure 2.

2.1 Seq2Seq Dialogue Generation

We define the Seq2Seq dialogue generation task as generating the most likely response sequence $Y = \{y_1, y_2, \cdots, y_n\}$, giving the input with multiple rounds of dialogue history X and knowledge base B. The probability of a response can be formally defined as,

$$p(Y|X, B) = \prod_{t=1}^{n} p\left(y_t | y_1, \ldots, y_{t-1}, X, B\right) \quad (1)$$

where y_t represents the current output token. Different from the vanilla Seq2Seq dialogue generation model (Eric and Manning, 2017), we use $p_C\left(y_t\right)$ to denote the probability that the generated token y_t is a conceptual token within M, and $p_{\overline{C}}\left(y_t\right)$ to denote the probability that y_t is a general token. Finally, we choose the highest probability to generate the token y_t at time t.

Contextual Dialog History Encoder In order to overcome the challenge of modeling long dialogue text, we encode dialog history utterances round by round. We first encode every sentence pair $(Q^p, Y^p) \in X$ as a semantic representation, where Q^p and Y^p respectively represent the p-th round question sequence (with m tokens) and response sequence (with n tokens). To better encode the contextual information of the dialogue, we send (Q^p, Y^p) into an effective pre-trained language representation model BERT (Devlin et al., 2019) to get the representation for the p-th round dialog sequence,

$$\mathbf{H}^p_{1:m+n} = \text{BERT}([CLS]Q^p[SEP]Y^p) \quad (2)$$

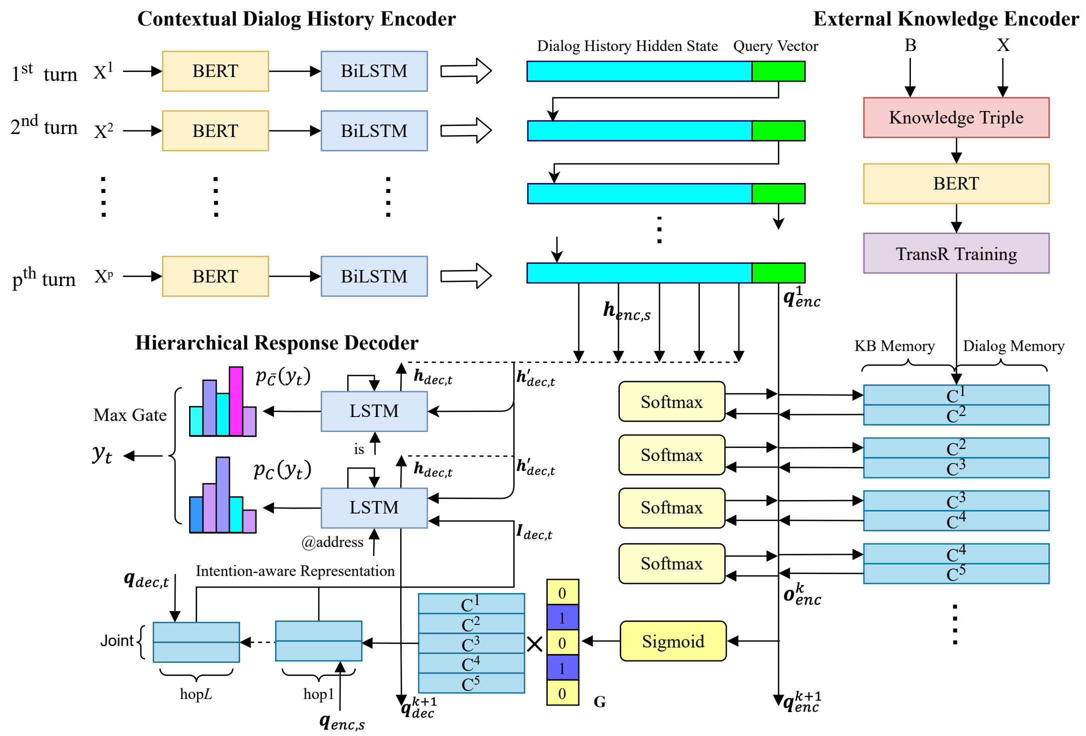


Figure 2: Workflow of the proposed model.

where $\mathbf{H}^p_{1:m}$ denotes the representation for the question sequence, and $\mathbf{H}^p_{m+1:m+n}$ for the response sequence. Afterward, we fed $\mathbf{H}^p_{1:m+n}$ into a Bidirectional Long Short-Term Memory network (BiLSTM) (Hochreiter and Schmidhuber, 1997) to produce contextual hidden states $\mathbf{h}_{enc} = (\mathbf{h}_{enc,1}, \mathbf{h}_{enc,2}, \ldots, \mathbf{h}_{enc,m+n})$, where,

$$\mathbf{h}_{enc,i} = \mathrm{BiLSTM}\left(\mathbf{H}^p_i, \mathbf{h}_{enc,i-1}, \mathbf{h}_{enc,i+1}\right) \quad (3)$$

Note that the first hidden state will be initialized with the last hidden state of the previous round, i.e., $\mathbf{h}^p_{enc,0} = \mathbf{h}^{p-1}_{enc,m+n}$ (the superscript p is omitted if no confusion occurs in the following text).

Hierarchical Response Decoder We exploit a hierarchy mechanism to decode the response sequence. Specifically, when decoding y_t, we use a coarse-grained LSTM decoder and a fine-grained LSTM decoder to compute the probability simultaneously.

We first use a coarse-grained decoder. Given $(\mathbf{h}_{enc,1}, \mathbf{h}_{enc,2}, \ldots, \mathbf{h}_{enc,m+n})$, an LSTM is used to repeatedly predict outputs $(y_1, y_2, \ldots, y_{t-1})$ by the decoder hidden states $(\mathbf{h}_{dec,1}, \mathbf{h}_{dec,2}, \ldots, \mathbf{h}_{dec,t})$. For the generation of y_t, we first calculate an attentive representation $\mathbf{h}'_{dec,t}$ of the dialogue history over the hidden state $\mathbf{h}_{enc}$, and then concatenate it with $\mathbf{h}_{dec,t}$ to get the context-aware output representation,

$$\mathbf{o}_{\overline{\mathcal{C}},t} = \mathbf{W}_1 \left[\mathbf{h}_{dec,t}, \mathbf{h}'_{dec,t}\right] \quad (4)$$

where $\mathbf{o}_{\overline{\mathcal{C}},t}$ is the score (logit) for the next token generation, and $\mathbf{W}_1$ is a trainable parameter. The probability of the next word y_t being regarded as a general token is then calculated as follows,

$$p_{\overline{\mathcal{C}}}\left(y_t\right) = \mathrm{Softmax}(\mathbf{o}_{\overline{\mathcal{C}},t}) \quad (5)$$

Next, we use a fine-grained decoder. In addition to incorporating $\mathbf{h}'_{dec,t}$ to ensure the relevance between the generated response and the question, we further derive an intention-aware representation $\mathbf{I}_{dec,t}$ (which will be detailed in Sec. 2.3) to enhance the representation of the target entity for generating more accurate response. By concatenating $\mathbf{h}_{dec,t}$ with $\mathbf{h}'_{dec,t}$ and $\mathbf{I}_{dec,t}$, we can get the output representation as,

$$\mathbf{o}_{\mathcal{C},t} = \mathbf{W}_2 \left[\mathbf{h}_{dec,t}, \mathbf{h}'_{dec,t}, \mathbf{I}_{dec,t}\right] \quad (6)$$

The probability of y_t being regarded as a conceptual token is then calculated as follows:

$$p_{\mathcal{C}}\left(y_t\right) = \mathrm{Softmax}(\mathbf{o}_{\mathcal{C},t}) \quad (7)$$

Finally, we can get the probability of y_t as,

$$p(y_t) = \max\{p_{\overline{\mathcal{C}}}\left(y_t\right), p_{\mathcal{C}}\left(y_t\right)\} \quad (8)$$

2.2 External Knowledge Memory

As well known, the successful conversations for task-oriented dialogue system heavily depend on accurate knowledge queries. We build our external knowledge memory M based on two parts: dialogue history X and multi-domain knowledge base B, i.e., $M = [X; B] = (m_1, m_2, \dots, m_l)$. Each entity in M is represented in a triple format, i.e., $m_i = (h, r, t)$. To better encode the external knowledge to make it more suitable for multi-hop reasoning and vector calculation, we embed the knowledge triples into a word vector space rich in strong entity relationships and semantic shift information. Specifically, for each triple $m_i = (h, r, t) \in M$, we use the TransR model (Lin et al., 2015) to perform fine-grained representation learning and obtain $(\mathbf{e}_h, \mathbf{e}_r, \mathbf{e}_t)$ as the memory embeddings. More details about TransR learning can be found in Appendix A.2.

To integrate knowledge information into the end-to-end dialogue system, the memory network (MN) (Sukhbaatar et al., 2015) is adopted to store global cross-domain knowledge, which is shared between the encoder and the decoder. For a k-hop MN, the external knowledge is composed of a set of trainable embedding matrices $\mathbf{C} = (\mathbf{C}^1, \dots, \mathbf{C}^{k+1})$.

Query Knowledge in Encoder We use the last hidden state as the initial query vector:

$$\mathbf{q}_{enc}^1 = \mathbf{h}_{enc,m+n} \tag{9}$$

It can loop over k hops and compute the attention weights at each hop k using

$$p_i^k = \text{Softmax}((\mathbf{q}_{enc}^k)^\mathsf{T} \mathbf{c}_i^k) \tag{10}$$

where $\mathbf{c}_i^k$ is the embedding in i-th memory position using the embedding matrix $\mathbf{C}^k$, and $\mathbf{q}_{enc}^k$ is the query vector for hop k. Finally, the model reads out the memory $\mathbf{o}_{enc}^k$ by the weighted sum over $\mathbf{c}_i^{k+1}$ and updates the query vector $\mathbf{q}_{enc}^{k+1}$. Formally,

$$\mathbf{o}_{enc}^k = \sum\nolimits_i p_i^k \mathbf{c}_i^{k+1}, \quad \mathbf{q}_{enc}^{k+1} = \mathbf{q}_{enc}^k + \mathbf{o}_{enc}^k \tag{11}$$

where $\mathbf{q}_{enc}^{k+1}$ is a coarse-grained representation containing KB information, and can be used to initialize the coarse-grained LSTM decoder.

By the above steps, we can obtain a global memory pointer $G = (g_1, \dots, g_l)$ to filter out worthless external knowledge for further decoding, where,

$$g_i^k = \text{Sigmoid}((\mathbf{q}_{enc}^k)^\mathsf{T} \mathbf{c}_i^k) \tag{12}$$

Note that G is finally trained as a n-dimensional $0/1$ prediction vector, and its training details are shown in Appendices A.3 and A.4.

Query Knowledge in Decoder Recall that we adopt two LSTMs as the decoder. For the coarse-grained LSTM decoder, following Wu et al. (2019) and Qin et al. (2020), we use the concatenation of $\mathbf{h}_{dec,t}$ (initialized by $\mathbf{q}_{enc}^{k+1}$) with the attentive representation $\mathbf{h}'_{dec,t}$ to query knowledge.

For the fine-grained LSTM decoder, we use the concatenation of the hidden states $\mathbf{h}_{dec,t}$ (initialized by $\mathbf{q}_{dec}^{k+1}$ obtained when generating the previous conceptual word), the attentive representation $\mathbf{h}'_{dec,t}$ and the intention-aware representation $\mathbf{I}_{dec,t}$, to query knowledge. Formally,

$$\mathbf{q}_{dec}^1 = [\mathbf{h}_{dec,t}, \mathbf{h}'_{dec,t}, \mathbf{I}_{dec,t}] \tag{13}$$

$$p_i^k = \text{Softmax}((\mathbf{q}_{dec}^k)^\mathsf{T} \mathbf{c}_i^k g_i^k) \tag{14}$$

Instead of selecting the maximum p_i^k to generate y_t, we read out the memory $\mathbf{o}_{dec}^k$ by the weighted sum over $\mathbf{c}^{k+1}$ and update the query vector $\mathbf{q}_{dec}^{k+1}$,

$$\mathbf{o}_{dec}^k = \sum\nolimits_i p_i^k \mathbf{c}_i^{k+1}, \quad \mathbf{q}_{dec}^{k+1} = \mathbf{q}_{dec}^k + \mathbf{o}_{dec}^k \tag{15}$$

Note that $\mathbf{q}_{dec}^{k+1}$ is a fine-grained representation containing user intention, and can be fed to the fine-grained LSTM decoder for the next conceptual word generation.

2.3 Intention Reasoning Module

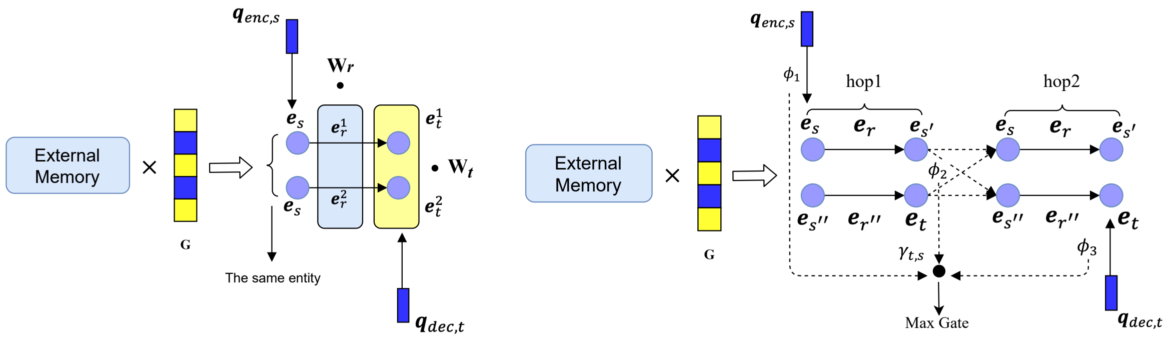
To obtain intention-aware presentation $\mathbf{I}_{dec,t}$, we first propose a novel intention mechanism (Sec. 2.3.1), based on which we further propose a fine-grained intention reasoning module (Sec. 2.3.2) that includes joint reasoning and multi-hop reasoning.

2.3.1 Intention Mechanism

Previous works usually use soft attention mechanisms (Vaswani et al., 2017) to calculate a weighted sum of all the knowledge based on the whole vector of each triple, which may not be conducive to generating accurate task-oriented responses. To address this issue, we propose a new intention mechanism to directly incorporate tail-entity information by comparing the similarity between query vector and the head-entity, which is formally defined as,

$$\text{Intention}(\mathbf{q}, (\mathbf{e}_h, \mathbf{e}_r, \mathbf{e}_t)) = \phi(\mathbf{q}, \mathbf{e}_h) \cdot \mathbf{e}_t \tag{16}$$

where $\mathbf{q}$ is query vector, and $(\mathbf{e}_h, \mathbf{e}_r, \mathbf{e}_t)$ represents the representation of the selected knowledge triple.



(a) Joint reasoning

(b) Multi-hop reasoning

Figure 3: The architecture of intention reasoning module.

Note here ϕ denotes for similarity score function, such as $cos(\cdot)$, dot product and scaled dot-product. We have tried these three functions and finally chose $cos(\cdot)$ based on their performance.

2.3.2 Fine-grained Intention Reasoning

Based on the intention mechanism, we further perform fine-grained intention reasoning to obtain an intention-aware representation $\mathbf{I}_{dec,t}$, which can be used to capture the concept shift information for final response generation. Specifically, giving the encoder query vector $\mathbf{q}_{enc,s}$ (i.e., $\mathbf{h}_{enc,s}$), the decoder query vector $\mathbf{q}_{dec,t}$ (i.e., $\mathbf{h}_{dec,t}$) and the global memory pointer G, $\mathbf{I}_{dec,t}$ is obtained by performing joint reasoning and multi-hop reasoning sequentially. Note that before conducting intention reasoning, we first use G to filter the external knowledge to obtain the target triples.

Joint Reasoning This operation is used to improve the integrality of the generated responses. Specifically, for multiple knowledge triples with the same head entity (or same tail entity), we fuse them into a single triple. Take the triples $\left(\mathbf{e}_s, \mathbf{e}_r^1, \mathbf{e}_t^1\right)$ and $\left(\mathbf{e}_s, \mathbf{e}_r^2, \mathbf{e}_t^2\right)$ in Figure 3(a) as an example, the joint reasoning is conducted as,

$$\mathbf{e}_t = \mathbf{W}_t(\mathbf{e}_t^1, \mathbf{e}_t^2), \quad \mathbf{e}_r = \mathbf{W}_r(\mathbf{e}_r^1, \mathbf{e}_r^2) \tag{17}$$

where $\mathbf{W}_t$ and $\mathbf{W}_r$ are trainable weight matrices. Then, $(\mathbf{e}_s, \mathbf{e}_r, \mathbf{e}_t)$ can be regarded as a new triple for multi-hop reasoning below.

Multi-hop Reasoning This operation aims at improving the accuracy of the generated responses. Specifically, an intention weight $\gamma_{t,s}$ is calculated to evaluate the probability that a set of ordered triples can generate the optimal reasoning chain. As shown in Figure 3(b), after filtering by G, there are two triples: $(\mathbf{e}_s, \mathbf{e}_r, \mathbf{e}_{s'})$ and $(\mathbf{e}_{s''}, \mathbf{e}_{r''}, \mathbf{e}_t)$. Suppose we perform 2-hop reasoning here, then there are totally 2^2 possible chains, and their intention weights can be calculated as follows:

$$\begin{cases} \gamma_{t,s}^1 = \phi\left(\mathbf{q}_{enc,s}, \mathbf{e}_s\right) \cdot \phi\left(\mathbf{e}_{s'}, \mathbf{e}_s\right) \cdot \phi\left(\mathbf{e}_{s'}, \mathbf{q}_{dec,t}\right) \\ \gamma_{t,s}^2 = \phi\left(\mathbf{q}_{enc,s}, \mathbf{e}_s\right) \cdot \phi\left(\mathbf{e}_{s'}, \mathbf{e}_{s''}\right) \cdot \phi\left(\mathbf{e}_t, \mathbf{q}_{dec,t}\right) \\ \gamma_{t,s}^3 = \phi\left(\mathbf{q}_{enc,s}, \mathbf{e}_{s''}\right) \cdot \phi\left(\mathbf{e}_t, \mathbf{e}_{s''}\right) \cdot \phi\left(\mathbf{e}_t, \mathbf{q}_{dec,t}\right) \\ \gamma_{t,s}^4 = \phi\left(\mathbf{q}_{enc,s}, \mathbf{e}_{s''}\right) \cdot \phi\left(\mathbf{e}_t, \mathbf{e}_s\right) \cdot \phi\left(\mathbf{e}_{\mathbf{e}_{s'}}, \mathbf{q}_{dec,t}\right) \end{cases} \tag{18}$$

Finally, we choose $\max\left\{\gamma_{t,s}^i\right\}$ as the final $\gamma_{t,s}$. Note that the above procedures can be generalized to L hops, where L is a model hyper-parameter.

After performing L-hop reasoning, we can get $\gamma_{t,s}$, and the corresponding optimal reasoning chain that contains L ordered triples, denoted by $\{(\mathbf{e}_h^1, \mathbf{e}_r^1, \mathbf{e}_t^1), \ldots, (\mathbf{e}_h^L, \mathbf{e}_r^L, \mathbf{e}_t^L)\}$ (note duplicate may occur when the number of target triples is less than L). Finally, we can obtain the intention-aware representation as,

$$\begin{aligned} \mathbf{I}_{dec,t} = & \mathbf{W}^{(1)}\text{Intention}\left(\mathbf{q}_{enc,s}, (\mathbf{e}_h^1, \mathbf{e}_r^1, \mathbf{e}_t^1)\right) \\ & + \sum\nolimits_{i=2}^{L} \mathbf{W}^{(i)}\text{Intention}(\mathbf{e}_t^{i-1}, (\mathbf{e}_h^i, \mathbf{e}_r^i, \mathbf{e}_t^i)) \end{aligned} \tag{19}$$

where $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(i)}$ are trainable parameters that are used to weigh the tail-token information obtained from the reasoning chain.

2.3.3 Degeneration

Note that when the encoded and decoded word is a general word, our model will no longer perform joint and multi-hop reasoning. Accordingly, the intention weight is reduced to the attentive weight:

$$\gamma_{t,s} = \phi\left(\mathbf{q}_{enc,s}, \mathbf{q}_{dec,t}\right) \propto \alpha_{t,s} \tag{20}$$

where $\alpha_{t,s}$ represents the attentive weight. This means the intention mechanism actually degenerates to the attention mechanism, which proves the robustness of our model.

3 Experimental Setup

3.1 Datasets and Metrics

Two publicly available datasets: SMD (Eric et al., 2017) and an extended version of Multi-WOZ 2.1 (Qin et al., 2020), are used to evaluate the performance of our model. We follow Eric et al. (2017), Madotto et al. (2018) and Wu et al. (2019) to partition SMD, and follow Budzianowski et al. (2018) and Qin et al. (2020) to partition Multi-WOZ 2.1. The statistics of the datasets after partition are presented in Table 1.

Follow several previous work (Eric et al., 2017; Madotto et al., 2018; Wu et al., 2019; Qin et al., 2019, 2020), we use BLEU and F1 (including both macro-F1 and micro-F1) to evaluate our model versus existing models. Moreover, to evaluate the performance in a more fine-grained level, we also choose Rouge-1 and Rouge-2 as metrics.

3.2 Baselines

We compare our model with the following state-of-the-art baselines.

- **Mem2Seq** (Madotto et al., 2018)[1]: the model takes dialog history and KB entities as input and utilizes a pointer gate to control either generating a vocabulary word or copying an entity word.
- **KB-retriever** (Qin et al., 2019)[2]: the model adopts a retriever module to extract the most relevant knowledge items and filter irrelevant information for response generation.
- **GLMP** (Wu et al., 2019)[3]: the model adopts a global-to-local pointer to query knowledge, where the global memory pointer is used to filter the external KB information, and the local memory pointer is used to instantiate a slot value generated by a sketch RNN.
- **DF-Net** (Qin et al., 2020)[4]: the framework uses a dynamic fusion network to dynamically exploit the correlation between all domains for fine-grained knowledge transfer and achieves state-of-the-art performance.

Dataset	Domains	Train	Dev	Test
SMD	Navigate, Weather, Schedule	2,425	302	304
Multi-WOZ 2.1	Restaurant, Attraction, Hotel	1,839	117	141

Table 1: Statistics of two datasets.

For BLEU and micro-F1 scores of the above baselines, we adopt the reported results from Wu et al. (2019) and Qin et al. (2020). For macro-F1 and Rouge scores, we rerun their public code to obtain results on same datasets.

3.3 Implementation Details

We train our model end-to-end by using Adam optimizer (Kingma and Ba, 2015) and choose the learning rate between $[1e^{-3}, 1e^{-4}]$. The loss functions are described in Appendix A.4. The dropout ratio is selected from $\{0.1, 0.15, 0.2, 0.25, 0.3\}$ and the batch size from $\{8, 16, 32\}$. The hyper-parameters such as hidden size, dropout, batch size, and embedding dimensionality are all tuned with grid-search over the development set. All experiments are conducted with PyTorch and our adopted BERT inherits huggingface's implementation[5]. Appendix A.1 presents more details about hyper-parameters.

4 Evaluation Results

4.1 Response Quality Evaluation

Automatic Evaluation Follow the prior work (Madotto et al., 2018; Wu et al., 2019; Qin et al., 2020; Zhang et al., 2020), we evaluate model performance automatically from two aspects: relevancy and novelty, where the corresponding results are presented in Tables 2 and 3, respectively.

From Table 2, we can observe that our model IR-Net achieves the state-of-the-art performance on two multi-domain datasets SMD and Multi-WOZ 2.1. Specifically, On SMD dataset, IR-Net exhibits the highest BLEU compared with other baselines, indicating that our model can generate responses closer to the golden ones. Moreover, our model outperforms DF-Net, a recent model that can capture the correlation between domains for fine-grained knowledge transfer, by 2.6% and 0.5% on macro-F1 and micro-F1 respectively, which verifies the effectiveness of our intention reasoning model in capturing the concept shifts across multiple domains to generate more accurate and appropriate responses. On Multi-WOZ 2.1, a trend for a similar performance improvement can be observed, which further demonstrates the effectiveness of our model.

[1] https://github.com/HLTCHKUST/Mem2Seq.

[2] We reproduce KB-retriever as no open-source code is available. Moreover, since Multi-WOZ 2.1 cannot be processed by KB-retriever, we only report its results on SMD.

[3] https://github.com/jasonwu0731/GLMP.

[4] https://github.com/LooperXX/DF-Net.

[5] https://github.com/huggingface/pytorch-transformers.

	SMD					Multi-WOZ 2.1				
Model	BLEU	Macro-F1	Micro-F1	Rouge-1	Rouge-2	BLEU	Macro-F1	Micro-F1	Rouge-1	Rouge-2
Mem2Seq (Madotto et al., 2018)	12.6	31.2	33.4	64.0	38.2	6.6	19.8	21.6	58.2	25.4
KB-retriever (Qin et al., 2019)	13.9	51.2	53.7	70.5	45.9	-	-	-	-	-
GLMP (Wu et al., 2019)	13.9	52.0	60.7	68.8	43.8	6.9	28.4	32.4	62.5	27.6
DF-Net (Qin et al., 2020)	14.4	59.4	62.7	70.6	46.1	9.4	32.2	35.1	65.1	31.9
IR-Net (ours)	**16.3**	**62.0**	**63.2**	**71.3**	**48.0**	**10.9**	**35.3**	**37.5**	**66.1**	**33.6**

Table 2: Main results. Relevance (higher better) between generated responses and golden responses. Note all our results are statistically significant with $p < 0.05$ under t-test.

	SMD			Multi-WOZ 2.1		
Model	BLEU	Rouge-1	Rouge-2	BLEU	Rouge-1	Rouge-2
Mem2Seq	2.21	26.5	9.8	3.09	33.0	13.5
KB-retriever	1.90	19.2	3.5	-	-	-
GLMP	0.12	9.8	1.2	0.17	18.4	3.1
DF-Net	0.06	11.7	1.3	0.11	19.2	3.3
IR-Net	**0.01**	**9.2**	**0.8**	**0.02**	**12.4**	**2.1**

Table 3: Repetitiveness (lower better) between generated responses and user's questions. All our results are statistically significant with $p < 0.05$ under t-test.

Model	Hel.	App.	Cor.	Flu.	Fri.	Hum.	Overall Average	Relative Ratio
Mem2Seq	1.50	2.64	3.15	3.62	2.20	1.80	2.49	52.8%
GLMP	2.45	2.86	3.24	3.84	3.95	3.90	3.37	71.4%
DF-Net	3.24	3.95	3.68	4.15	4.20	4.00	3.87	82.0%
IR-Net	3.90	4.00	3.80	4.20	4.35	4.15	4.07	86.2%
Golden	4.60	4.48	4.82	4.85	4.60	4.98	4.72	100%

Table 4: Human evaluation of responses on helpfulness (Hel.), appropriateness (App.), correctness (Cor.), fluency (Flu.), friendliness (Fri.), and human-likeness (Hum.) on randomly selected dialogs.

Model	**Entity F1 (%)**	
	Test	Δ
Complete model	63.2	-
w/o IR module & Fine-grained Decoder	60.9	2.3
w/o Coarse-grained Decoder	58.2	5.0
w/o Bert Embedding	61.4	1.8
w/o TransR Training	62.2	1.0

Table 5: Ablation study on SMD dataset.

From Table 3, we can see that compared with other baselines, IR-Net achieves consistently lower BLEU and Rouge scores, which demonstrates its capability in generating more innovative responses, possibly due to the following two reasons: 1) The integration of cross-domain knowledge in multi-hop reasoning makes the generated responses more diverse; 2) The hierarchical LSTM decoder in IR-Net can learn more forms of expressions.

Human Evaluation The human evaluation mainly focuses on six aspects: helpfulness, appropriateness, correctness, fluency, friendliness, and human-likeness, which are all important for task-oriented dialogue systems (Zhou et al., 2018; Zhang et al., 2020; Qin et al., 2020). We first randomly selected 100 dialogs 1:1 from the SMD and Multi-WOZ 2.1 datasets, and used different models to generate responses, including Mem2Seq, GLMP, DF-Net and IR-Net. Then, we hired human experts to score the responses and golden responses on a scale from 1 to 5, which simulated a real-life task-oriented conversation scenario. By calculating the average score of the above metrics, we obtained the final manual evaluation result, as shown in Table 4. It can be seen that IR-Net outperforms the other three models on all metrics, which is consistent with the results of automatic evaluation.

4.2 Ablation Study

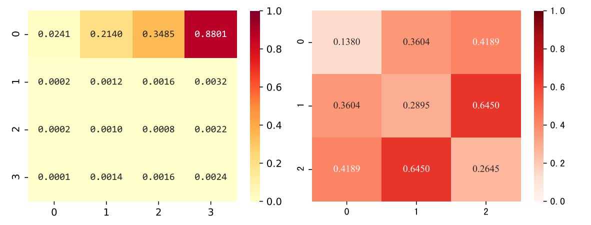
In this part, we perform ablation experiments to evaluate the effectiveness of each component. We focused on four crucial components and set them accordingly: 1) w/o IR module and Fine-grained Decoder denotes that we remove the intention reasoning module and the fine-grained decoder, and just adopt the "coarse-grained decoder" with querying external KB attentively; 2) w/o Coarse-grained Decoder denotes that we only use attentive KB to return answer; 3) w/o Bert Embedding denotes that we simply feed randomly initialized embeddings into the contextual dialog encoder; 4) w/o TransR Training denotes that we discard the TransR-based knowledge triple embedding learning. From the results in Table 5, we can observe that removing each component will result in a performance degradation. In particular, w/o Intention Reasoning and Fine-grained Decoder causes 2.3% drops in entity F1 score, which further verifies the effectiveness of our model.

4.3 Case Study and Visualization

We take the dialog in Table 6 as an example. To better illustrate the advantage of our model and understand what the intention reasoning module has learned, we visualize the intention weights, as well as the attention weights of an attention-based Seq2Seq model for this dialog, as depicted in Figure 4. It can be observe that our intention-based

Dialog	Content	Reasoning Chain
User	please check the temperature for me today.	(today, temperature, ?)
IR-Net	toady's temperature is 20f-30f.	0: (today, date, Monday) 1: (Monday, low_temp, 20f) 2: (Monday, high_temp, 30f) 3: (Monday, temperature, 20f-30f) 0→3: **(today, temperature, 20f-30f)**

Table 6: A dialog example illustrating joint reasoning and 2-hop reasoning.



(a) Intention weights. (b) Attention weights.

Figure 4: Visualization of intention weights of IR-Net and attention weights of an attentiton-based Seq2Seq model for the dialog in Table 6.

model is more adept at mining potential reasoning chains, while previous attention-based model, which is limited by scattered attention weights, is hard to capture explicit reasoning relations. Specifically, for this dialog, IR-Net first performs joint reasoning to derive triple 3 by triples 1 and 2. Then, it performs a 2-hop reasoning to obtain a set of intention weights, as shown in Figure 4 (a). Unlike scattered attention weights in Figure 4 (b), it is clear to see that chain 0 → 3 achieves the highest intention weight (0.8801) in Figure 4 (a), indicating that *(today, date, Monday)→(Monday temperature, 20f-30f)* has been mined by IR-Net to be the optimal 2-hop reasoning chain. Finally, IR-Net can generate a relatively accurate response "today's temperature is 20f-30f". More analyses and experimental details regarding the visualization of intention and attention weights can be found in Appendix B.2.

5 Related Work

Sequence to sequence approaches, which use an encoder-decoder structure to capture the contextual dialog semantics and generate responses directly, have recently gained much attention in task-oriented dialogue systems (Zhao et al., 2017). These models have effective language modeling ability, but cannot work well in KB retrieval, even with sophisticated attention-based mechanism. To alleviate this problem, copy augmented Seq2Seq models (Gülçehre et al., 2016; Eric and Manning, 2017) have been adopted, but still suffer from the challenge of performing reasoning over KB triples.

To address this problem, memory augmented Seq2Seq models, such as end-to-end Memory Network (Bordes et al., 2017) and DQMN (Wu et al., 2018), have been proposed and shown promising results. Later, Mem2Seq (Madotto et al., 2018) and GLMP (Wu et al., 2019) further augmented memory based methods by incorporating the copy mechanism (Gülçehre et al., 2016), which enables copying words from both dialog history and KB. DSR (Wen et al., 2018) proposed to leverage dialogue state representation to retrieve the KB implicitly. Multi-level memory model (Reddy et al., 2019) represented the KB results with a multi-level memory instead of the form of triples. KB-retriever (Qin et al., 2019) adopted a KB retriever module to extract the most relevant knowledge items and improve the consistency of generated entities. DDMN (Wang et al., 2020) adopted a dual dynamic memory network to track the dialog context and KB triples respectively. DF-Net (Qin et al., 2020) introduced a dynamic fusion model to capture the correlation between domains for fine-grained knowledge transfer.

Different from existing models that rely on the soft attention mechanism to perform coarse-grained reasoning, our IR-Net can model more deterministic knowledge and capture the entity (or concept) shift by performing fine-grained token-level reasoning based on the intention mechanism. To our best knowledge, we are the first to effectively explore fine-grained token-level reasoning in multi-domain task-oriented dialog generation.

6 Conclusion

In this paper, we propose a novel intention mechanism to directly incorporate the tail-token information of a knowledge triple to better model deterministic knowledge for multi-domain task-oriented dialog. Moreover, based on the intention mechanism, we further propose an intention reasoning module that consists of token-level joint reasoning and multi-hop reasoning to obtain an intention-aware representation, aiming at improving the integrality and accuracy of the generated response. Experiments on two publicly available multi-domain datasets demonstrate the effectiveness and superior performance of our model in both automatic and human evaluation.

Acknowledgements

We would like to thank all anonymous reviewers for their valuable comments. The work was partially supported by the National Natural Science Foundation of China under Grant No. 61672252.

References

Antoine Bordes, Y-Lan Boureau, and Jason Weston. 2017. Learning end-to-end goal-oriented dialog. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net.

Pawel Budzianowski, Tsung-Hsien Wen, Bo-Hsiang Tseng, Iñigo Casanueva, Stefan Ultes, Osman Ramadan, and Milica Gasic. 2018. Multiwoz - A large-scale multi-domain wizard-of-oz dataset for task-oriented dialogue modelling. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018*, pages 5016–5026. Association for Computational Linguistics.

Yinpei Dai, Hangyu Li, Chengguang Tang, Yongbin Li, Jian Sun, and Xiaodan Zhu. 2020. Learning low-resource end-to-end goal-oriented dialog for fast and reliable system deployment. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 609–618. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186. Association for Computational Linguistics.

Mihail Eric, Lakshmi Krishnan, Francois Charette, and Christopher D. Manning. 2017. Key-value retrieval networks for task-oriented dialogue. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue, Saarbrücken, Germany, August 15-17, 2017*, pages 37–49. Association for Computational Linguistics.

Mihail Eric and Christopher D. Manning. 2017. A copy-augmented sequence-to-sequence architecture gives good performance on task-oriented dialogue. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017, Valencia, Spain, April 3-7, 2017, Volume 2: Short Papers*, pages 468–473. Association for Computational Linguistics.

Çaglar Gülçehre, Sungjin Ahn, Ramesh Nallapati, Bowen Zhou, and Yoshua Bengio. 2016. Pointing the unknown words. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016, August 7-12, 2016, Berlin, Germany, Volume 1: Long Papers*. The Association for Computer Linguistics.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Comput.*, 9(8):1735–1780.

Jaehun Jung, Bokyung Son, and Sungwon Lyu. 2020. Attnio: Knowledge graph exploration with in-and-out attention flow for knowledge-grounded dialogue. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, pages 3484–3497. Association for Computational Linguistics.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. 2015. Learning entity and relation embeddings for knowledge graph completion. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA*, pages 2181–2187. AAAI Press.

Andrea Madotto, Chien-Sheng Wu, and Pascale Fung. 2018. Mem2seq: Effectively incorporating knowledge bases into end-to-end task-oriented dialog systems. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers*, pages 1468–1478. Association for Computational Linguistics.

Alexander H. Miller, Adam Fisch, Jesse Dodge, Amir-Hossein Karimi, Antoine Bordes, and Jason Weston. 2016. Key-value memory networks for directly reading documents. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, EMNLP 2016, Austin, Texas, USA, November 1-4, 2016*, pages 1400–1409. The Association for Computational Linguistics.

Seungwhan Moon, Pararth Shah, Anuj Kumar, and Rajen Subba. 2019. Opendialkg: Explainable conversational reasoning with attention-based walks over knowledge graphs. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pages 845–854. Association for Computational Linguistics.

Libo Qin, Yijia Liu, Wanxiang Che, Haoyang Wen, Yangming Li, and Ting Liu. 2019. Entity-consistent end-to-end task-oriented dialogue system with KB retriever. 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 2019, Hong Kong, China, November 3-7, 2019*, pages 133–142. Association for Computational Linguistics.

Libo Qin, Xiao Xu, Wanxiang Che, Yue Zhang, and Ting Liu. 2020. Dynamic fusion network for multi-domain end-to-end task-oriented dialog. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 6344–6354. Association for Computational Linguistics.

Jun Quan, Deyi Xiong, Bonnie Webber, and Changjian Hu. 2019. GECOR: an end-to-end generative ellipsis and co-reference resolution model for task-oriented dialogue. 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 2019, Hong Kong, China, November 3-7, 2019*, pages 4546–4556. Association for Computational Linguistics.

Revanth Reddy, Danish Contractor, Dinesh Raghu, and Sachindra Joshi. 2019. Multi-level memory for task oriented dialogs. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 3744–3754. Association for Computational Linguistics.

Sainbayar Sukhbaatar, Arthur Szlam, Jason Weston, and Rob Fergus. 2015. End-to-end memory networks. In *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada*, pages 2440–2448.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 5998–6008.

Jian Wang, Junhao Liu, Wei Bi, Xiaojiang Liu, Kejing He, Ruifeng Xu, and Min Yang. 2020. Dual dynamic memory network for end-to-end multi-turn task-oriented dialog systems. In *Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8-13, 2020*, pages 4100–4110. International Committee on Computational Linguistics.

Haoyang Wen, Yijia Liu, Wanxiang Che, Libo Qin, and Ting Liu. 2018. Sequence-to-sequence learning for task-oriented dialogue with dialogue state representation. In *Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, New Mexico, USA, August 20-26, 2018*, pages 3781–3792. Association for Computational Linguistics.

Tsung-Hsien Wen, David Vandyke, Nikola Mrksic, Milica Gasic, Lina Maria Rojas-Barahona, Pei-Hao Su, Stefan Ultes, and Steve J. Young. 2017. A network-based end-to-end trainable task-oriented dialogue system. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017, Valencia, Spain, April 3-7, 2017, Volume 1: Long Papers*, pages 438–449. Association for Computational Linguistics.

Chien-Sheng Wu, Andrea Madotto, Genta Indra Winata, and Pascale Fung. 2018. End-to-end dynamic query memory network for entity-value independent task-oriented dialog. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2018, Calgary, AB, Canada, April 15-20, 2018*, pages 6154–6158. IEEE.

Chien-Sheng Wu, Richard Socher, and Caiming Xiong. 2019. Global-to-local memory pointer networks for task-oriented dialogue. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net.

Steve J. Young, Milica Gasic, Blaise Thomson, and Jason D. Williams. 2013. Pomdp-based statistical spoken dialog systems: A review. *Proc. IEEE*, 101(5):1160–1179.

Houyu Zhang, Zhenghao Liu, Chenyan Xiong, and Zhiyuan Liu. 2020. Grounded conversation generation as guided traverses in commonsense knowledge graphs. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 2031–2043. Association for Computational Linguistics.

Tiancheng Zhao, Allen Lu, Kyusong Lee, and Maxine Eskénazi. 2017. Generative encoder-decoder models for task-oriented spoken dialog systems with chatting capability. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue, Saarbrücken, Germany, August 15-17, 2017*, pages 27–36. Association for Computational Linguistics.

Hao Zhou, Tom Young, Minlie Huang, Haizhou Zhao, Jingfang Xu, and Xiaoyan Zhu. 2018. Commonsense knowledge aware conversation generation with graph attention. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden*, pages 4623–4629. ijcai.org.

A Model Details

A.1 Hyperparameters Setting

Hyperparameter Name	SMD	Multi-WOZ 2.1
Batch Size	32	16
Hidden Size	128	128
Bert Embedding Size	768	768
Learning Rate	0.001	0.001
Dropout Ratio	0.15	0.15
Teacher Forcing Ratio	0.9	0.9
Memory Network's Hop	3	3
Intention Reasoning's Hop	3	3

Table 7: Hyperparameters we used for SMD and Multi-WOZ 2.1.

A.2 Knowledge Embedding Training

In the KB memory module, each element m_i is represented in the triple format as (head, relation, tail), e.g., (*dinner, time_is,* 7_pm), which is a commonly used format to represent a knowledge item (Miller et al., 2016; Eric et al., 2017). On the other hand, the dialog history X is stored in the dialogue memory, where the user and temporal encoding are included as in (Bordes et al., 2017) like a triple format, e.g., the first utterance from the user in Figure 1 will be denoted as *{($user, turn1, How's), ($user, turn1, the), ($user, turn1, weather), ($user, turn1, today)}*.

For each triple $m_i \in M$, we use the TransR model (Lin et al., 2015) to perform representation learning and obtain $(\mathbf{e}_h, \mathbf{e}_r, \mathbf{e}_t)$ as the memory embeddings. Specifically, for triple $m_i = (h, r, t)$, where $h, t \in \mathcal{E}$ and $r \in \mathcal{R}$ ($\mathcal{E}$ and $\mathcal{R}$ represent the entity space and relation space, respectively), we first use BERT to pre-train them:

$$(\mathbf{e}_h, \mathbf{e}_r, \mathbf{e}_t) = \text{BERT}(h, r, t) \tag{21}$$

Then, we embed $\mathbf{e}_h$ and $\mathbf{e}_t$ into the relation space through a trainable projection matrix $\mathbf{M}_r$, where the evaluation function is described as follows:

$$f_r(h, t) = \|\mathbf{M}_r\mathbf{e}_h + \mathbf{e}_r - \mathbf{M}_r\mathbf{e}_t\|_{L_2} \tag{22}$$

Finally, we minimize the following loss function to get the optimal knowledge triple embedding,

$$\mathcal{L}_{emb} = \sum_{(h,r,t)\in S} \sum_{(h',r,t')\in S'} \max(\lambda + f_r(h,t) - f_r(h',t'), 0) \tag{23}$$

where S and S' are positive triple set and negative triple set (1:3 selected in both SMD and Multi-WOZ 2.1 datasets) respectively, and λ is the distance between the scores of positive and negative triples.

A.3 Description on the Global Pointer

Follow prior work GLMP (Wu et al., 2019), we employ a global memory pointer to select knowledge and regard it as a multi-label classification problem, that is, selecting k target knowledge triples from n candidate triples. For the training of the global memory pointer G, we first use the sigmoid function to activate the dot product of the query vector and the memory representation, and then convert the multi-label classification problem into n binary classification problems (each predicted value $1/0$ represents whether the triplet is selected), and finally, we use the sum of the cross-entropy as the loss function. Therefore, G is regarded as a final n-dimensional $1/0$ prediction vector to filter worthless knowledge triples, and its training details are shown in Appendix A.4.

A.4 Loss Function

The loss $\mathcal{L}$ used in IR-Net is similar to that of GLMP. We first define $G^{label} = (\hat{g}_1, \dots, \hat{g}_l)$ by checking whether the object words in the memory exist in the expected system response Y,

$$\hat{g}_i = \begin{cases} 1 & \text{if } Object\,(m_i) \in Y \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

where m_i is one triple in the external knowledge $M = [X; B] = (m_1, m_2, \dots, m_l)$ and $Object(\cdot)$ is denoted as getting the object word from a triple. Then, the cross-entropy loss $\mathcal{L}_g$ between G and G^{lable} can be written as,

$$\mathcal{L}_g = -\sum_{i=1}^{l} \left(\hat{g}_i \cdot \log g_i + (1 - \hat{g}_i) \cdot \log\left(1 - g_i\right)\right) \tag{25}$$

We exploit a hierarchy mechanism to decode the response sequence. Specifically, when decoding y_t, we use a coarse-grained LSTM decoder and a fine-grained LSTM decoder to generate a rough response $Y^c_{\overline{\mathcal{C}}} = (y^c_1, \dots, y^c_n)$ and a fine-grained response $Y^f_{\mathcal{C}} = (y^f_1, \dots, y^f_n)$, respectively. Their output probabilities are calculated as follows,

$$p_{\overline{\mathcal{C}}}\left(y_t\right) = \text{Softmax}(\mathbf{W}_1\left[\mathbf{h}_{dec,t}, \mathbf{h}'_{dec,t}\right]) \tag{26}$$

$$p_{\mathcal{C}}\left(y_t\right) = \text{Softmax}(\mathbf{W}_2\left[\mathbf{h}_{dec,t}, \mathbf{h}'_{dec,t}, \mathbf{I}_{dec,t}\right]) \tag{27}$$

Then, we calculate standard cross-entropy losses $\mathcal{L}_{\overline{\mathcal{C}}}$ and $\mathcal{L}_{\mathcal{C}}$ as follows:

$$\mathcal{L}_{\overline{\mathcal{C}}} = \sum_{t=1}^{n} -\log\left(p_{\overline{\mathcal{C}}}(y_t) \cdot (y^c_t)\right) \tag{28}$$

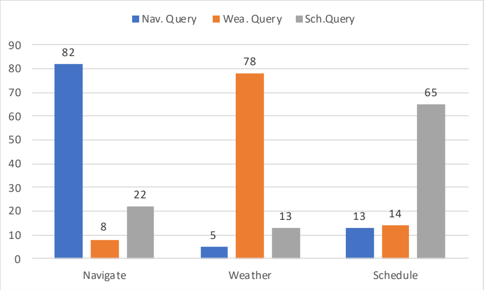


Figure 5: Knowledge distribution of cross-domain query for randomly selected 100 examples in each domain on the SMD dataset.

$$\mathcal{L}_{\mathcal{C}} = \sum_{t=1}^{n} - \log(p_{\mathcal{C}}(y_t) \cdot (y_t^f)) \tag{29}$$

Finally, $\mathcal{L}$ is the weighted-sum of three losses:

$$\mathcal{L} = \beta_g \mathcal{L}_g + \beta_{\mathcal{C}} \mathcal{L}_{\mathcal{C}} + \beta_{\overline{\mathcal{C}}} \mathcal{L}_{\overline{\mathcal{C}}} \tag{30}$$

where β_g, $\beta_{\mathcal{C}}$ and$\beta_{\overline{\mathcal{C}}}$ are hyperparameters. Note that these three weights are initialized equally, i.e., 0.33, 0.33 and 0.33. Then we tune them on the verification set to obtain a better weight setting of 0.39, 0.36 and 0.25.

B Experimental Details

B.1 Additional experiments

Experiments on Domain-shift In this experiment, we randomly selected 100 examples of knowledge queries in each domain on the SMD test set. By parsing the global memory pointer G, we obtain the distribution of the selected knowledge, as shown in Figure 5. We can find that: (1) A small fraction of knowledge query successfully implements cross-domain knowledge-selection through the attention mechanism, while the majority of knowledge is selected within the domain. It means that cross-domain knowledge query occurs in the task-oriented dialogue. (2) Navigation-related query selects more knowledge in the schedule domain than in the weather domain. Similarly, schedule-related query also selects more knowledge in the navigation domain than in the weather domain. This indicates that the navigation domain and the schedule domain are more closely related.

Analysis on L-hops To analyze the impact of the hop number L in intention reasoning, we keep other hyper-parameters unchanged, and vary L in the range of $[1, 2, 3, 4, 5, 6, 7]$. From Figure 6, we can observe that with the increase of L, the entity (micro) F1 score first increases and then decreases, and reaches the best result at $L = 3$, less or more hops would decrease the performance. It is straightforward that less hops are insufficient to capture user's real intention, while too more hops may also lead to more noisy, which is harmful to the expressiveness of the obtained intention-aware representations. Hence, it is necessary to choose appropriate hops for intention reasoning.

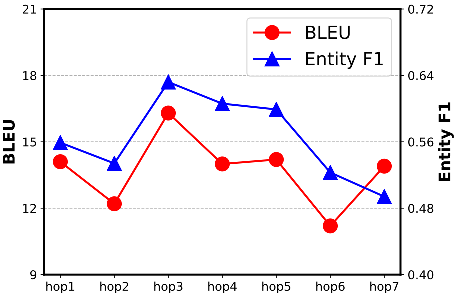


Figure 6: BLUE and Entity F1 under different reasoning hops on SMD dataset.

B.2 Visualization of Attention and Intention Weights

To further illustrate what our intention reasoning module has learned, we visualize the attention and intention weights (denoted by α and γ respectively) of the dialog generation process in dialog #1 and #2, as shown in Figure 7 (note that only parts of knowledge triples are presented). Darker colors represent higher attention or intention weights. G represents for $(0, 1)$ distribution vector generated by α. From Figure 7, we can observe that: 1) There is a joint reasoning guided by γ, i.e., *"today temperature is 20f-30f"* by combining *(monday, low_temp,* $20f$*)* with *(monday, high_temp,* $30f$*)*; 2) There are two 2-hop reasoning guided by γ, one is *(today, date, monday)→(monday, temperature,* $20f$*-*$30f$*)*, and the other is *(friends_house, poi, jills_house)→(jills_house, address, 347_alta_mesa_ave)*. The above two observations illustrate that our intentional reasoning module can: 1) effectively perform cross-domain knowledge selection (by the attention mechanism); 2) effectively perform fine-grained knowledge reasoning (by the intention mechanism).

B.3 Error Analysis

To better understand the limitations of our model, we conduct an error analysis on IR-Net. We randomly select 100 responses generated by IR-Net that achieve low human evaluation scores in the test

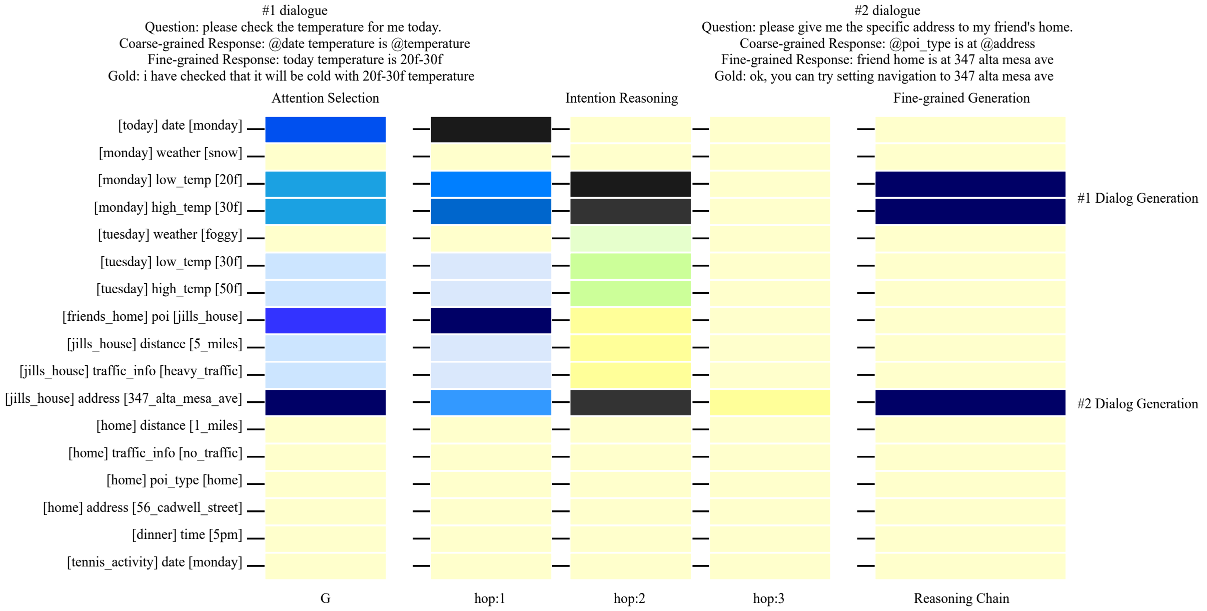


Figure 7: Visualization of attention weight α and intention weight γ. The leftmost column denotes the attention wights selected by the global memory pointer G; The three columns in the middle represent the intention weights selected by the intention reasoning module; The rightmost column denotes the derived reasoning chain for #1 and #2 dialog generation.

set of SMD. We report several reasons for the low scores, which can roughly be classified into four categories. (1) KB information in the generated responses is incorrect (35%), especially when the corresponding equipped knowledge base is large and complex. (2) The sentence structure of the generated responses is incorrect and there are serious grammatical and semantic errors (26%). (3) The model makes incomplete response when there are multiple options corresponding to the user intention (24%). (4) The conceptual tokens generated by the fine-grained decoder cannot be well matched with the golden entities (15%).