HiTKG: Towards Goal-Oriented Conversations via Multi-Hierarchy Learning

Jinjie Ni, Vlad Pandelea, Tom Young, Haicang Zhou, Erik Cambria

Nanyang Technological University, Singapore

{jinjie001, yang0552, haicang001}@e.ntu.edu.sg, {vlad.pandelea, cambria}@ntu.edu.sg

Abstract

Human conversations are guided by short-term and long-term goals. We study how to plan short-term goal sequences as coherently as humans do and naturally direct them to an assigned long-term goal in open-domain conversations. Goal sequences are a series of knowledge graph (KG) entity-relation connections generated by KG walkers that traverse through the KG. The existing recurrent and graph attention based KG walkers either insufficiently utilize the conversation states or lack global guidance. In our work, a hierarchical model learns goal planning in a hierarchical learning framework. We present **HiTKG**, a hierarchical transformer-based graph walker that leverages multiscale inputs to make precise and flexible predictions on KG paths. Furthermore, we propose a two-hierarchy learning framework that employs two stages to learn both turn-level (short-term) and global-level (long-term) conversation goals. Specifically, at the first stage, HiTKG is trained in a supervised fashion to learn how to plan turn-level goal sequences; at the second stage, HiTKG tries to naturally approach the assigned global goal via reinforcement learning. In addition, we propose MetaPath as the backbone method for KG path representation to exploit the entity and relation information concurrently. We further propose Multi-source Decoding Inputs and Output-level Length Head to improve the decoding controllability. Our experiments show that HiTKG achieves a significant improvement in the performance of turn-level goal learning compared with state-of-the-art baselines. Additionally, both automatic and human evaluation prove the effectiveness of the two-hierarchy learning framework for both short-term and long-term goal planning.

Introduction

Building a human-like dialogue system has been a long-lasting goal in the community of conversational AI (Ni et al. 2021; Ma et al. 2020). In the pursuit of this goal, multiple research topics have emerged: context awareness (Qiu et al. 2020), response coherence (Liu et al. 2019a) and diversity (Su et al. 2020), speaker consistency (Madotto et al. 2019), empathetic response (Song et al. 2019), conversation topic (Wu et al. 2019), knowledge-grounded system (Chen et al. 2020), etc. Conversation goal is one of the most representative elements that reflect human intelligence.

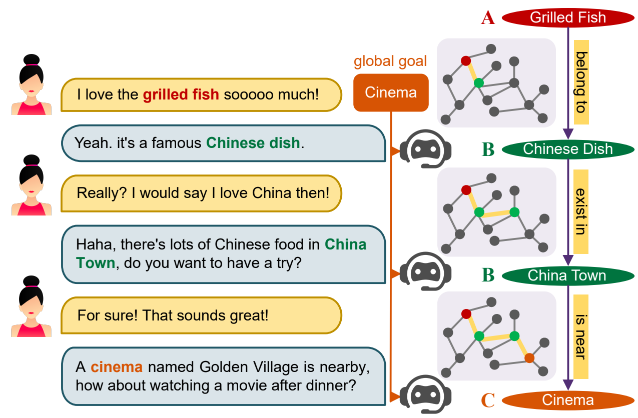

Figure 1: A goal-driven dialogue sample. Starting from an initial entity (A), the chatbot plans turn-level conversation goals (B) based on dialogue content and history goal trajectory, also trying to naturally direct B to a global goal (C).

Human conversations are usually guided by several small goals or a global goal. As shown in Fig. 1, *Grilled Fish*, *Chinese Dish*, *China Town*, and *Cinema* are turn-level goals, while the *Cinema* is also the global goal at the same time. During the conversation, the agent intends to approach the global goal by naturally transitioning between turn-level goals. However, most dialogue systems passively respond to the user without explicit goals, causing incoherent or illogical responses. In recent years, some researchers attempt to ground dialogue systems on knowledge graphs (KGs) to actively guide conversation topics/goals. KG is a structured knowledge network that consists of vertices, or entities, being connected by edges, or relations (Ji et al. 2022).

KGs contain commonsense relationships between real-world entities that can also be seen as conversation goals. Generating or retrieving responses according to the walking trajectory in KGs is effective in generating goal-oriented responses. The current graph walkers can generally be divided into recurrent walkers (Young et al. 2018; Zhang et al. 2019; Moon et al. 2019; Li et al. 2022) and graph attention based walkers (Jung, Son, and Lyu 2020). Recurrent walkers decode KG paths depending on a fixed-length vector, which creates a bottleneck for the performance.

Graph attention based KG walkers are good at achieving an overall optimal since they reserve all potential paths, but such mechanism is too high in computation complexity to be scalable to multi-hop reasoning. In addition, these walkers neglect the hierarchical structure of the input source and make separate predictions for entity and relation paths, which affects their performance. Besides, these walkers only plan turn-level goals based on the dialogue history, which means that their reasoning is local and undirected. However, many of the conversations between humans, especially adults, are guided with an ultimate goal.

To this end, we propose **Hi**erarchical **T**ransformer based **K**nowledge **G**raph Walker (**HiTKG**), a graph walker that leverages multiscale inputs to make precise and flexible predictions on KG paths. We learn this hierarchical model in a two-hierarchy learning framework which employs two stages to learn goal planning. In the first stage, we train HiTKG in a supervised fashion, where it learns how to plan turn-level conversational goal sequences naturally based on the dialogue content; in the second stage, we manually assign a global goal for HiTKG and it learns to approach the global goal via reinforcement learning, where a user simulator is trained to provide user messages for the conversation. In other words, in such two-stage learning, the model learns to approach the target goal without losing the naturalness of the goal sequence. Specifically, HiTKG has a transformer-based structure, as shown in Fig. 2. The graph decoder computes hierarchical attention with different-level memories to obtain a better representation of current states, based on which it reasons over the multihop neighbors and decodes a KG path consisting of the related entities and relations. Note that our model is scalable to KG paths of flexible lengths. To conclude, our contributions are as follows:

- We propose the first transformer-based KG walker that attentively reads multiscale inputs for graph decoding. We also propose Multi-source Decoding Inputs (**MDI**) and Output-level Length Head (**OLH**) to strengthen controllability of the Hierarchical Attention based Graph Decoder (**HAGD**).
- We propose a two-hierarchy learning framework to train the proposed hierarchical KG walker, in order to learn both turn-level and global-level conversation goals. This is the first attempt to learn models to make natural transitions towards the global goal in KG, where we propose a distance embedding to incorporate distance information.
- We propose MetaPath (**MP**) to concurrently exploit entity and relation information when reasoning, which is proved essential as the backbone method for KG path representation, providing a paradigm for KG reasoning.

Method

Overview

We define the knowledge graph $\mathbf{G}_{KG} = \mathbf{V}_{KG} \times \mathbf{E}_{KG}$, where the KG is composed of the commonsense vertices (entities) $\mathbf{V}_{KG}$ and edges (relations) $\mathbf{E}_{KG}$ that connect the vertices. We define $\mathbf{V}_{E,n}(v)$ as the set of nodes that are connected to v with n-hop edge connections E.

There are two learning stages for the model to learn different levels of goal planning: turn-level goal learning (stage 1) and global-level goal learning (stage 2). Specifically, at stage 1 (supervised learning), the model takes the multiscale state $\mathbf{x}$ as input, and predicts an n-hop KG path $\mathbf{Y}_p$ which represents the transition of conversation goals. The KG path is made up of the entities and relations in $\mathbf{G}_{KG}$. $\mathbf{x}$ has two scales: $\mathbf{x} = \{\mathbf{x}_p; \mathbf{x}_d\}$, where $\mathbf{x}_p = \{\mathbf{x}_p^{(1)}, \mathbf{x}_p^{(2)}, ..., \mathbf{x}_p^{(i)}\}$ denotes KG path history with a fixed window size i that is generated in previous turns; the $\mathbf{x}_d = \{\mathbf{x}_d^{(1)}, \mathbf{x}_d^{(2)}, ..., \mathbf{x}_d^{(j)}\}$ denotes dialogue history utterances with a fixed window size j that is produced by both speakers. The KG path inference at the t-th dialogue turn of **stage 1** is formulated as:

$$\mathbf{Y}_p^t = \underset{\mathbf{Y}}{\operatorname{argmax}} \prod_{k=1}^{T} P(\mathbf{v}_k^t | \mathbf{x}_p^t, \mathbf{x}_d^t, \mathbf{V}_{E,1}(v_{k-1}^t)) \qquad (1)$$

Where $\mathbf{Y}_p^t = \{\mathbf{v}_1^t, \mathbf{v}_2^t, ..., \mathbf{v}_T^t\}$ denotes the predicted T-hop KG path. $\mathbf{v}_k^t = [v_{k-1}^t, e_k^t, v_k^t]$ denotes the k-th one-hop KG path of $\mathbf{Y}_p^t$. $v_{k-1}^t \in \mathbf{V}_{KG}$ denotes the starting KG vertex of $\mathbf{v}_k^t$, and $e_k^t \in \mathbf{E}_{KG}$ denotes the edge that connects v_{k-1}^t and v_k^t. At stage 2 (supervised and reinforcement learning), $\mathbf{x}$ is composed of three scales: $\mathbf{x} = \{\mathbf{x}_p; \mathbf{x}_d; \mathbf{g}\}$, where $\mathbf{g}$ denotes global goal. The model takes $\mathbf{x}$ as input to predict turn-level KG path $\mathbf{v}_{y,k}^t$, and tries to make the entity of $\mathbf{v}_{y,k}^t$ closer to $\mathbf{g}$. Since no n-hop KG path annotation is available at this stage, we predict one-hop paths here to analyze the problem. The KG path inference at **stage 2** is formulated as:

$$\mathbf{v}_{y,k}^t = \underset{\mathbf{v}_k^t}{\operatorname{argmax}} P(\mathbf{v}_k^t | \mathbf{x}_p^t, \mathbf{x}_d^t, \mathbf{g}, \mathbf{V}_{E,1}(v_{k-1}^t)) \qquad (2)$$

Multiscale Source Representation

The KG path is a series of conversation goals based on which the utterances are organized. We argue that dialogue history is the surface-level representation of a dialogue and KG path history is a higher level one that can be interpreted as the outline of a conversation. At stage 2, the global goal is the top-level input source that decides the topic flow. As shown in Fig. 2, HiTKG encodes multiscale dialogue sources with separate transformer encoders, where MetaPath is the cornerstone of KG path representation and reasoning. The global goal is represented as a part of the decoder input.

MetaPath (MP) In previous works such as (Moon et al. 2019) and (Jung, Son, and Lyu 2020), entities and relations are represented separately in both KG path encoding and decoding. At the decoding stage, KG paths are predicted by scoring entity paths and relation paths respectively, and then rerank. Thus, the model only considers one distribution of the entities or relations at a time, while a KG triple is composed of both. The prediction quality is decided by entity path reasoning, relation path reasoning, and reranking algorithm jointly, which makes it harder to achieve optimum. We propose MetaPath, an effective method to represent and score the KG paths by concurrently considering the entity and relation information. A MetaPath is a flexible combination of embeddings.

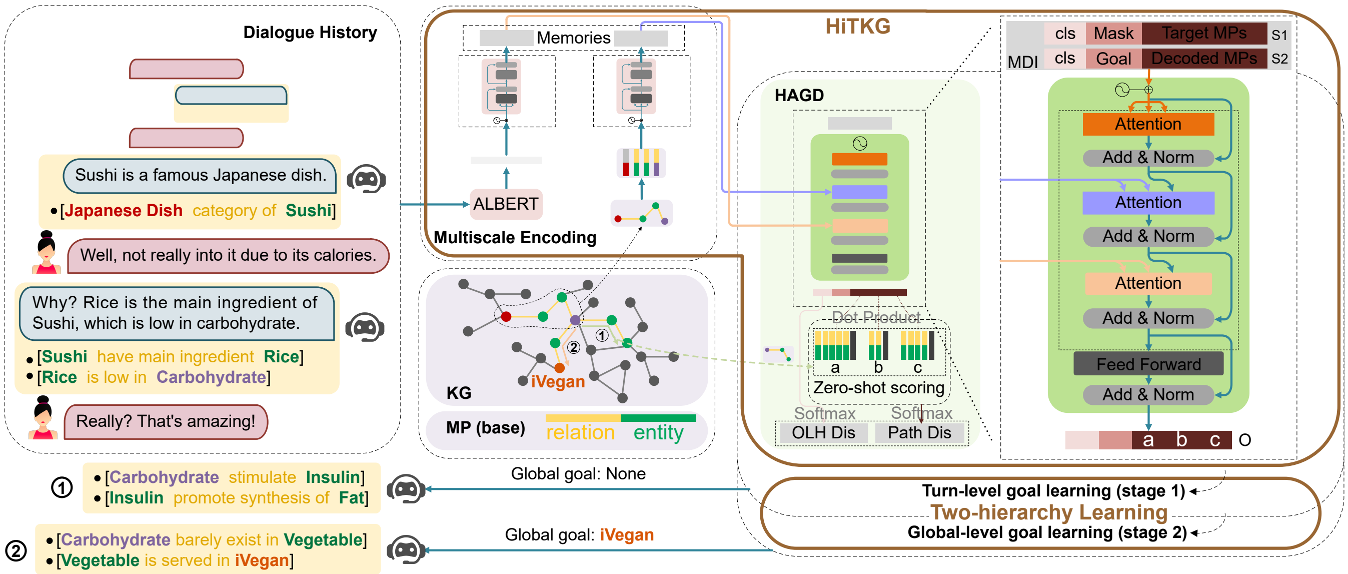

Figure 2: The overall architecture. HiTKG is composed of multiscale encoders and the Hierarchical Attention based Graph Decoder (HAGD). It first employs two separate transformers to learn dialogue history and KG path history representations, and then HAGD leverages the multiscale memories to plan KG paths. HiTKG has different reasoning strategies when trained with stage 1 only and with both stages. We optimize the whole HiTKG during training. Note that our task **only** predicts KG paths.

Given a KG triple $(\mathbf{v}_1, \mathbf{e}, \mathbf{v}_2)$, a base MP contains the concatenated embeddings of $\mathbf{e}$ and $\mathbf{v}_2$: MP=$[\vec{\mathbf{e}}_e; \vec{\mathbf{e}}_{v_2}]$. Where $\vec{\mathbf{e}}_e, \vec{\mathbf{e}}_{v_2} \in \mathbb{R}^{d_{kg}}$, d_{kg} denoting the KG embedding dimension. Although $\mathbf{v}_1$ is not included, the MP still expresses the full triple, benefiting from the base graph embedding. MP=$[\vec{\mathbf{e}}_e; \vec{\mathbf{e}}_{v_2}]$ instead of $[\vec{\mathbf{e}}_{v_1}; \vec{\mathbf{e}}_e; \vec{\mathbf{e}}_{v_2}]$ because the later form causes information redundancy when encoding KG path history with multiple MPs (each entity appears twice), causing worse convergence during training. As the cornerstone of KG path reasoning, MP also decides the scoring logic, bringing high efficiency, accuracy, and flexibility.

KG Vertex & Edge Following previous KG walkers, we use TransE (Bordes et al. 2013) to represent the vertices and edges of $\mathbf{G}_{KG}$. TransE is a simple but powerful graph embedding method that is scalable to large KGs and represents multi-scale relationships. The main idea is that given a ground KG triple $(\mathbf{v}_1, \mathbf{e}, \mathbf{v}_2)$, it satisfies $\vec{\mathbf{e}}_{v_1} + \vec{\mathbf{e}}_e \approx \vec{\mathbf{e}}_{v_2}$. The MetaPath and TransE form an ideal combo because according to the principle $\vec{\mathbf{e}}_{v_1} + \vec{\mathbf{e}}_e \approx \vec{\mathbf{e}}_{v_2}$, the embedding of $\mathbf{v}_1$ can be inferred by the model given $\mathbf{e}$ and $\mathbf{v}_2$, which makes up the information loss for a single MP (no loss for MP sequences).

Dialogue History Dialogue history is composed of the conversations from past turns and has a fixed window size. The conversation sentences are first encoded with an ALBERT[1] (Lan et al. 2019) layer ϕ^{al}, which is frozen during training, to obtain contextual representation $\mathbf{e}_d = \phi^{al}(\mathbf{x}_d) = \{\vec{\mathbf{e}}_d^{(1)}, \vec{\mathbf{e}}_d^{(2)}, ..., \vec{\mathbf{e}}_d^{(n_d)}\}$. Memory $\mathbf{h}_d = \{\vec{\mathbf{h}}_d^{(1)}, \vec{\mathbf{h}}_d^{(2)}, ..., \vec{\mathbf{h}}_d^{(n_d)}\}$ is obtained by parallelly applying $\mathbf{h}_d = E_{dial}(\mathbf{e}_d)$. Where E_{dial} is composed of learnable positional embedding (Wang and Chen 2020), multi-head attention α_n (Vaswani et al. 2017), and feedforward network (FFN) (we simplify multihead attention and omit bias terms, and illustrate only one transformer layer for conciseness):

$$\begin{aligned} E_{dial}(\mathbf{e}_d) &= \psi(\alpha_{self}(\mathbf{e}_d^{pos}) + \mathbf{W}_2\sigma(\mathbf{W}_1\alpha_{self}(\mathbf{e}_d^{pos}))) \\ \alpha_{self}(\mathbf{e}_d^{pos}) &= \alpha_n(\tilde{\mathbf{e}}_d^{pos}|\bar{\mathbf{e}}_d^{pos}) \end{aligned} \tag{3}$$

Where $\mathbf{e}_d^{pos}$ denote the position-embedded inputs. $\bar{\mathbf{e}}_d^{pos}$ denotes attention query and $\tilde{\mathbf{e}}_d^{pos}$ denotes attention key/value. ψ denotes layer normalization (Ba, Kiros, and Hinton 2016). $\mathbf{W}_1 \in \mathbb{R}^{d_m \times d_f}$ and $\mathbf{W}_2 \in \mathbb{R}^{d_f \times d_m}$ are the weight matrices of FFN, where d_m and d_f denotes the model and FFN dimension, respectively. We add $[usr1]$ and $[usr2]$ in front of respective messages to identify the utterances from different speakers, which has been proved effective in dialogue tasks (Quan and Xiong 2020).

KG Path History TransE-based MetaPath transformation is denoted as φ^{mp}. We stack all MPs corresponding to their respective one-hop KG paths to represent KG path history $\mathbf{x}_p$. A starting MetaPath MP$_{bop}$=$[\vec{\mathbf{e}}_{bop}; \vec{\mathbf{e}}_{1st}]$ is placed at the beginning. $\vec{\mathbf{e}}_{bop}$ is a special token embedding indicating the beginning of path and $\vec{\mathbf{e}}_{1st}$ is the starting entity embedding of $\mathbf{x}_p$. The MP$_{bop}$ ensures that all entities of the original KG path history are presented and indicates the starting point. E_{kg} further aggregates the MP representations of path history $\mathbf{e}_p = \varphi^{mp}(\mathbf{x}_p) = \{\vec{\mathbf{e}}_p^{(1)}, \vec{\mathbf{e}}_p^{(2)}, ..., \vec{\mathbf{e}}_p^{(n_p)}\}$ to get memory $\mathbf{h}_p = \{\vec{\mathbf{h}}_p^{(1)}, \vec{\mathbf{h}}_p^{(2)}, ..., \vec{\mathbf{h}}_p^{(n_p)}\}$ by applying $\mathbf{h}_p = E_{kg}(\mathbf{e}_p)$, where E_{kg} has the same architecture as E_{dial}.

[1]A state-of-the-art contextual representation model. It achieves better performance than BERT-large with fewer parameters.

Global Goal The global goal $\mathbf{g}$ is an entity of $\mathbf{G}_{KG}$ that is manually or randomly selected. At the stage of global-level goal learning, the global goal is a significant source of input, while it is not required when learning the turn-level goals. Considering that stage 1 and stage 2 share the same model, they need to keep the same input forms. Thus, we put the global goal in front of the target sequence, jointly as an input of the graph decoder. At stage 1, we use a target mask to mask out the global goal embedding when computing self-attentions for decoder inputs.

Turn-level Goal Learning

We propose **Hierarchical Attention based Graph Decoder (HAGD)** to predict turn-level KG paths and train it in a supervised fashion. Given KG environment $\mathbf{V}_e^{tar}$ of the target sequence, graph decoder ξ decodes KG paths:

$$\mathbf{y}_p = \xi(S_a, \mathbf{h}_d, \mathbf{h}_p || \mathbf{V}_e^{tar}) \tag{4}$$

$S_a = [\vec{\mathbf{c}}; \vec{\mathbf{e}}_{gm}; \mathbf{t}]$ denotes turn-level **Multi-source Decoding Inputs (MDI)**, which aggregates static and dynamic states for ξ, being the top level of multi-scale sources. $\vec{\mathbf{c}} \in \mathbb{R}^{d_m}$ is the corresponding cls embedding of OLH. $\vec{\mathbf{e}}_{gm} \in \mathbb{R}^{d_m}$ denotes masked global goal embedding (padded). $\mathbf{t} \in \mathbb{R}^{n_s \times d_m}$ is the shifted right target sequence starting with MP_{bop}.

Multi-hierarchy Attention Block ξ has three scales of sources: MDI, dialogue history and KG path history memories, being denoted as S_a, $\mathbf{h}_d$, and $\mathbf{h}_p$, respectively. We build a multi-hierarchy attention block to aggregate the multiscale information. Specifically, the proposed HAGD has three attention layers α_{self}, α_{kg}, and α_{dial} that align with MDI, KG path history, and dialogue history, respectively:

$$\begin{aligned} \kappa_s^{top} &= \tau(\alpha_{self}(S_a^{pos})) &&= \tau(\alpha_n(\tilde{S}_a^{pos} | \bar{S}_a^{pos})) \\ \kappa_p^{mid} &= \tau(\alpha_{kg}(\mathbf{h}_p | \kappa_s^{top})) &&= \tau(\alpha_n(\tilde{\mathbf{h}}_p | \bar{\kappa}_s^{top})) \\ \kappa_d^{bot} &= \tau(\alpha_{dial}(\mathbf{h}_d | \kappa_p^{mid})) &&= \tau(\alpha_n(\tilde{\mathbf{h}}_d | \bar{\kappa}_p^{mid})) \end{aligned} \tag{5}$$

τ denotes the residual operation $\tau(y(x)) = x + y(x)$. A self-attention layer α_{self} computes attention over the top-level source MDI; the resulting context vectors $\kappa_s^{top} = \{\vec{\kappa}_{s,1}^{top}, \vec{\kappa}_{s,2}^{top}, ..., \vec{\kappa}_{s,n_s}^{top}\}$ interact with $\mathbf{h}_p$ at the middle layer α_{kg}; then taking the resulting residual state $\kappa_p^{mid} = \{\vec{\kappa}_{p,1}^{mid}, \vec{\kappa}_{p,2}^{mid}, ..., \vec{\kappa}_{p,n_s}^{mid}\}$ from α_{kg}, α_{dial} leverages $\mathbf{h}_d$ and obtain $\kappa_d^{bot} = \{\vec{\kappa}_{d,1}^{bot}, \vec{\kappa}_{d,2}^{bot}, ..., \vec{\kappa}_{d,n_s}^{bot}\}$. $\vec{\kappa}_{s,k}^{top}$, $\vec{\kappa}_{p,k}^{mid}$, and $\vec{\kappa}_{d,k}^{bot} \in \mathbb{R}^{d_m}$.

Output-level Length Head (OLH) Humans have a general estimation about how long should the goal trajectory be when they plan the conversation goals, though this could be subconscious. To this end, we place a cls token at the beginning of decoder input and attach the corresponding OLH to the output layer for path length prediction. The output $O = [\vec{\mathbf{c}}_\iota; \vec{\mathbf{g}}_\iota; \mathbf{t}_\iota]$ has the corresponding shape of S. $\vec{\mathbf{c}}_\iota$ is the OLH state representation for KG path length prediction; $\vec{\mathbf{g}}_\iota$ is a placeholder; $\mathbf{t}_\iota$ is the KG path state representation. We compute path length distribution $\mathbf{y}_l$ as follows:

$$\mathbf{y}_l^i = \frac{exp(\mathbf{c}_\iota^i)}{\sum_{k=1}^{N} exp(\mathbf{c}_\iota^k)} \tag{6}$$

Where $\mathbf{c}_\iota^k \in \mathbb{R}$. This multi-task learning framework predicts path length and KG paths concurrently and achieves improvement in the performance of KG path prediction.

Zero-shot Scoring Note that different from traditional language models, the total number of candidates M is not fixed at different path positions n. Thus, given the path state representation $\mathbf{t}_\iota$, we compute the zero-shot embedding similarity to score the paths:

$$\mathbf{y}_p^{n,j} = \frac{exp(\vec{\mathbf{t}}_\iota^n \cdot \vec{\pi}_n^j)}{\sum_{k=1}^{M} exp(\vec{\mathbf{t}}_\iota^n \cdot \vec{\pi}_n^k)} \tag{7}$$

Where $\Pi_n = \{\vec{\pi}_n^1, \vec{\pi}_n^2, ..., \vec{\pi}_n^m\}$ denotes a set containing neighbor MP candidates of the n-th node in a KG path and the *end of path* embedding MP_{eop}. $\vec{\pi}_n^k \in \mathbb{R}^{d_{mp}}$ denotes the k-th MP candidate of Π_n. We compute Cross Entropy loss of length and KG path prediction to optimize HiTKG:

$$\mathcal{L}_{sup} = \gamma \mathcal{L}_{CE}(\mathbf{y}_l, \hat{\mathbf{y}}_l) + \lambda \mathcal{L}_{CE}(\mathbf{y}_p, \hat{\mathbf{y}}_p) + \epsilon \sum_w w^2 \tag{8}$$

Where $\hat{\mathbf{y}}_l$ and $\hat{\mathbf{y}}_p$ denote the ground truth path length and KG path, respectively. γ and λ are weight coefficients. $\epsilon \sum_w w^2$ is the weight decay term.

Global-level Goal Learning

We propose a reinforcement KG walker HiTKG-RL, which has the same architecture as HiTKG, to walk on the $\mathbf{G}_{KG}$ under the guidance of a global goal. This learning stage can be viewed as the combination of a pretrain stage (generally the same as stage 1) and a fine-tune stage where we apply a reinforcement framework to teach the pretrained KG walker how to approach the global goal without losing naturalness.

User Simulation We train a user simulator to generate user responses as dialogue history when interacting. The user simulator has the same architecture as the KG walker, which takes the dialogue history and KG path history as input sources. The decoder input sequence $S^{usr} = [\vec{\mathbf{v}}_c; \mathbf{t}]$ is different from that of the KG walker, where $\vec{\mathbf{v}}_c$ denotes the KG vertex of current turn. Instead of predicting the KG paths, the decoder output is modified to predict the probability distribution over a fixed vocabulary set to generate human responses. The dialogue history is solely composed of the simulated responses. Omission of the second speaker's response barely influences the overall performance of KG path reasoning, which is indicated by an ablation experiment.

Distance Embedding To measure how close the current node is from the global goal entity, a distance metric is required. We directly use the graph distance[2] as the distance metric since the dot product of TransE embeddings does not provide good estimations of distances when the vertices are far away from each other. We traverse the graph to obtain a distance matrix $\mathbf{D}$ between all vertex pairs and then perform matrix factorization to get two low-dimensional matrices. Given a vertex, we retrieve the vector at the corresponding position as its distance embedding $\vec{\mathbf{e}}_d$.

[2]The minimum length of the paths connecting two vertices.

Policy The policy model HiTKG-RL has the same architecture as HiTKG, while the MetaPath is modified: MP_{rl}=$[\vec{\mathbf{e}}_e; \vec{\mathbf{e}}_v; \vec{\mathbf{e}}_d]$, in order to incorporate the distance information. To maintain the same MetaPath structure between stage 1 and 2, we conduct the turn-level goal learning at stage 2 with MP_{rl} instead of MP (HiTKG performs best with MP at stage 1). The encoder and decoder inputs constitute the observable states. HiTKG-RL predicts a one-hop KG path at each step and tries to approach the global goal. We employ A2C[3] (Mnih et al. 2016) to optimize the model.

Reward At the t-th turn, we directly obtain the distance of two vertices $d_t(\mathbf{v}_1, \mathbf{v}_2)$ from $\mathbf{D}$ and estimate the reward based on this. If $d_t(\mathbf{v}_{eop}, \mathbf{g}) < d_{t-1}(\mathbf{v}_{eop}, \mathbf{g})$, then the reward is set to 1, otherwise the reward is 0. $\mathbf{v}_{eop}$ denotes the ending entity of the path predicted. Currently, due to the lack of automatic KG path evaluation metrics, we use the distance as the only criteria. The future work will introduce more evaluation criteria of KG paths such as the naturalness.

Experiments and Results

Dataset

We conduct our evaluation on OpenDialKG (Moon et al. 2019). It is a dialogue - KG dataset where each utterance of a dialogue is annotated with a KG path, which enables learning graph walkers to reason over the KG based on the conversations. It consists of 15K dialogues and 91K turns. Each dialogue is produced by two crowd-workers and grounds in a given topic. We follow the baselines and split it into train (70%), dev (15%), and test set (15%).

Experimental Settings

Baselines To evaluate the stage 1 learning, we compare our results with six baseline models. However, we do not benchmark against previous work to evaluate the stage 2 learning, since we can hardly find any similar work, or related codes are not available.

Generally, the six baseline models can be divided into breadth-centric and depth-centric models. Tri-LSTM (Young et al. 2018) is a breadth-centric model that augments its dialogue inputs with wide-ranging shallow KG facts to retrieve short KG paths. The other five baselines and HiTKG are depth-centric models which focus on a small set of KG entity-relation connections and perform deep reasoning over the KG. Among them Seq2Seq and DialKG Walker were proposed in (Moon et al. 2019), while Seq2Path, AttnFlow and AttnIO were proposed in (Jung, Son, and Lyu 2020).

Implementation Details The MetaPath is the basic component of KG path representations, while we perform moderate modifications under different situations. When encoding an n-hop KG path history wherein the one-hop KG path components are in series connections, a starting MetaPath MP_{bop}=$[\vec{\mathbf{e}}_{bop}; \vec{\mathbf{e}}_{1st}]$ is added to the beginning to indicate the starting point in KG.

[3] A2C replaces the Q value of Actor-critic's gradient with the expected advantage and the learning process is more stable compared with policy gradient methods.

In contrast, at the graph decoding stage, when predicting probability distribution over the one-hop neighbors of the current entity, MP_{bop} is not required, since all of the paths start from the current entity and they are in parallel relationships. In addition, stage 1 and 2 use MetaPaths of different structures, as stated in Section . We use Pytorch (Paszke et al. 2019) to implement our model, which is trained on two RTX 8000 GPUs. We tune the hyperparameters by grid searching the hyperparameter space and choose the following settings that perform best: number of encoder/decoder layers: 2/6; dimension of the KG walker: 768; dimension of the KG embedding: 384 (stage 1), 256 (stage 2); loss coefficients γ/λ: 0.1/0.9; number of attention heads: 12; learning rate: 10^{-3}; dropout rate: 0.1; L2 regularization parameter ϵ: 10^{-5}; batch size: 10. We use learning rate scheduler to tune the learning rate manually and patient & early stopping to avoid overfitting. In addition, we use gradient clipping to avoid gradient explosions.

Evaluation

Results The turn-level goal planning performance of baseline models and HiTKGs are presented in Table 1. Following the baselines, we use recall@k as the evaluation metric of path-level ($path@k$) and target entity-level ($tgt@k$) correctness. HiTKG outperforms all baselines we benchmark against in both $path@k$ and $tgt@k$, with two metrics worse than the ablation models. The performance gain is significant, especially in recalls with larger k: there is a 13% relative improvement in $path@25$ and 9% in $tgt@25$. As illustrated in Section , at the second learning stage, we use a different MetaPath structure to represent the KG paths and KG neighbors for both the supervised and reinforcement learning. Thus, we also report the performance of turn-level goal planning at stage 2. HiTKG-RL is designed for reinforcement learning, while it shows comparable performance with HiTKG when trained in a supervised fashion, even outperforming in $tgt@3$. This result indicates that the introduced distance embedding does not significantly influence the performance of turn-level goal planning at stage 2.

Tri-LSTM, Seq2Seq, Seq2Path, and DialKG Walker are recurrent graph walkers, which deliver history information with a fixed-length vector. The use of a fixed-length vector creates a performance bottleneck in KG reasoning and these recurrent baseline models show at least 42.59% lower performance than HiTKG in $path@1$. In addition, recurrent units suffer from short memories, restricting the performance in long KG path predictions. Trajectory is a good form to represent dynamic information and we leverage dialogue history and KG path history as two trajectory sources for goal planning. Most of the baselines omit the KG path history and only learn utterance patterns for KG walking. However, KG path history records the KG trajectory up to the previous turn and is an important guide to a KG walker, e.g., the model knows which paths have been walked in previous turns, which avoids or reduces repeated attempts. DialKG Walker and AttnIO are two state-of-the-art KG walkers. The recurrent architecture of DialKG Walker limits itself in feature representation, which causes its comparatively low performance, especially in recalls with small k.

Model	Recall@k									
	path@1	path@3	path@5	path@10	path@25	tgt@1	tgt@3	tgt@5	tgt@10	tgt@25
Tri-LSTM	3.2	14.2	22.6	36.3	56.2	-	-	-	-	-
Seq2Seq	3.1	18.3	29.7	44.1	60.2	-	-	-	-	-
DialKG Walker	13.2	26.1	35.3	47.9	62.2	-	-	-	-	-
Seq2Path	14.92	24.95	31.1	38.68	48.15	15.65	27.04	33.86	42.52	53.28
AttnFlow	17.37	24.84	30.68	39.48	51.4	18.97	36.23	45.48	58.84	71.35
AttnIO	23.72	37.53	43.57	52.17	62.86	24.98	43.78	53.49	65.48	78.79
HiTKG - 2EMP	24.16	37.01	46.78	58.12	67.83	30.01	43.22	55.19	66.73	85.36
HiTKG - DK	24.55	38.01	48.02	58.17	70.39	30.89	43.92	53.68	71.01	85.26
HiTKG - W2	25.82	38.56	48.63	57.93	71.25	30.99	46.29	**55.91**	70.88	85.15
HiTKG - WLH	24.49	37.31	49.10	58.22	70.12	30.13	43.29	54.76	70.16	85.52
HiTKG - SP	23.98	35.21	39.29	53.25	66.81	25.02	45.15	50.52	63.77	80.31
HiTKG-RL	25.12	36.55	45.67	56.37	69.18	30.99	**46.67**	54.16	67.20	85.19
HiTKG	**25.99**	**38.67**	**49.18**	**59.32**	**71.27**	**31.11**	46.29	55.59	**71.61**	**86.09**

Table 1: Path-level ($path@k$) and target-level ($tgt@k$) performance of supervised KG path reasoning at stage 1 (metric: recall@k). HiTKG is benchmarked against several state-of-the-art baselines and ablation models on the OpenDialKG dataset.

Besides, when computing the context vector at the decoding stage, it needs to compute attention scores over the whole relation space, which can be computationally expensive and may affect the quality of the resulting context vector. AttnIO computes an incoming attention flow to represent entities and an outgoing attention flow to select KG paths. This design ensures an optimum path at the decoding stage because it reserves all potential paths at each step. However, to predict a T-hop path where each node has N neighbors on average, the computation of outgoing attention flow has a complexity of $O(N^T)$, which is not scalable to long KG path predictions. In addition, all baseline models separately deal with the entity and relation paths, which breaks the semantic structure of KG paths.

Ablation Study We conduct five ablation studies as reported in Table 1. **(1)** First, we experiment with the 2-entity MetaPaths (2EMP) where MP=$[\vec{\mathbf{e}}_{v_1}; \vec{\mathbf{e}}_e; \vec{\mathbf{e}}_{v_2}]$. The performance degradation suggests that the redundancy of entity information harms the training. **(2)** Next, the encoder-decoder attention layers α_{kg} and α_{dial} are swapped (DK). Placing the layer α_{kg} in front of α_{dial} outperforms the reversed condition, which implies that it is more reasonable to select low-level information (dialogue history) with a higher one (path history), demonstrating a better way to compute hierarchical attention. **(3)** We test the performance of supervised path learning without the utterances from speaker 2 (W2). We find that, although performance is slightly degraded, results are still comparable, even higher than the standard setting in $tgt@5$. We infer that this is because given speaker 1 (user) and speaker 2 (agent), speaker 2 will pay more attention to the utterances from speaker 1 instead of his own for goal planning. In addition, the KG path history contains most of the essential information in the utterances from speaker 2. **(4)** To investigate the contribution of OLH, we train the KG walker without it (WLH) and this causes performance to drop 1-2%. **(5)** The fifth ablation model separately predicts entity and relation paths (SP), using both distributions for one-hop KG path reranking at each decoding step. A drop in performance suggests the contribution of MetaPath, which concurrently considers entity and relation information.

Model	Distance (%)				
	3	5	7	9	11
HiTKG	2	0	0	0	0
HiTKG-RL	66	27	11	2	0

Table 2: The success rate of reaching the global goal entity.

Model	Naturalness				Semantic Closeness			
	1st	2nd	3rd	4th	1st	2nd	3rd	4th
SP	9	5.8	18.2	67	41	33.8	10.2	15
GT	39.2	27.8	24	9	12.6	17.2	39.8	30.4
Hi	32.6	43.2	18	6.2	7	11.8	41.2	40
Hi-RL	19.2	23.2	39.8	17.8	39.4	37.2	8.8	14.6

Table 3: Ranking results of the semantic closeness between ending node and global goal, and the path naturalness. The results are presented as the number of instances a certain model is ranked as a certain ranking (averaged).

Success Rate Whether the agent can reach the global goal entity is a natural way to evaluate whether stage 2 works. For each case, we randomly select a beginning node $\mathbf{v}_{1st}$ and a target global goal $\mathbf{g}$ which has a graph distance of 3/5/7/9/11 from $\mathbf{v}_{1st}$. We report and compare the success rate of 100 independent attempts by HiTKG and HiTKG-RL, respectively, as shown in Table 2. The HiTKG is only trained at stage 1 while HiTKG-RL undergoes both stages. It is indicated that without global goal guided training, the HiTKG can barely succeed (only 2 cases succeeded by chance). Whereas HiTKG-RL has a 66% success rate at distance 3 and declines as the distance rises. The decline is partially ascribed to the trade-off between naturalness and success.

Human Evaluation We aim to plan natural turn-level goals at stage 1, while at stage 2 we aim to approach the target without losing naturalness. We conduct human evaluation to further evaluate the path naturalness of both stages and the approaching effectiveness of stage 2. We sample 100 2-hop KG paths from Shortest Path (SP), Ground Truth

Conversation	P1: Fiona Stafford wrote Emma. It's a romance novel. Are you into that genre? P2: Any other books that might fall under *comedy*? I'm in the mood for something light. P1: [response]
GT	([BOP], Comedy) → (genre of, One Crazy Summer)
AttnFlow	([BOP], Comedy) → (parent genre, Slapstick)
AttnIO	([BOP], Comedy) → (subject of, The War of the Worlds) → (written by, Arthur. C. Clarke)
HiTKG (WLH)	([BOP], Comedy) → (genre of, Slacker) → (release year, 1991)
HiTKG	([BOP], Comedy) → (genre of, One Crazy Summer)

Table 4: Comparison of KG paths generated from models trained at stage 1 under a context (including ground truth).

(GT), HiTKG (Hi), and HiTKG-RL (Hi-RL), respectively, and combine them into 100 four-tuples where each path tuple corresponds to a certain starting entity $\mathbf{v}_{1st}$ in OpenDialKG and global goal $\mathbf{g}$ which is 3 units away from $\mathbf{v}_{1st}$. Each path starts from $\mathbf{v}_{1st}$ and does not reach $\mathbf{g}$. SP denotes the shortest paths from $\mathbf{v}_{1st}$ to $\mathbf{g}$. Hi is only trained at stage 1 while Hi-RL undergoes both stages. Five human evaluators independently rank the four paths in each tuple regarding the path naturalness and the semantic closeness between the ending node and global goal. As shown in Table 3, GT and Hi paths show dominance in top rankings of naturalness, which indicates the effectiveness of stage 1; Hi-RL paths have significantly more instances than SP in top naturalness rankings, indicating that Hi-RL reserves naturalness when learning to approach global goal (comparing with Hi, it sacrifices some naturalness to approach the target). Hi-RL paths rank the first in 39.4% closeness rankings, only second to the SP. It indicates that Hi-RL effectively approaches the target in semantic space when reasoning over the KG.

Case Study Table 4 presents the example KG path predictions from models trained at stage 1 under the same conversational context in (Jung, Son, and Lyu 2020). With the starting entity *Comedy*, the proposed HiTKG predicts precisely in both content and length. In contrast, the baseline models tend to select paths that are different from the ground truth, which lacks naturalness and may not fit the context. In addition, we also showcase the output of an ablation model which predicts without the OLH. The generated path is not consistent in length compared with the ground truth.

We compare how KG walkers choose one-hop paths given a global goal, as shown in Table 5. At the node *Beatrice and Virgil*, there are three candidate paths to select, which ends with entities that are 1/3/3 units away from the global goal *An American in Hollywood*. In general, the walker trained with both stages tends to choose the entity that is the closest to the target, in both topological and semantic spaces.

Conversation	P1: Do you have books by Yann Martel? P2: He wrote Life of Pi and *Beatrice and Virgil*. Have you read either of those? P1: [response]
Candidates	(in language, English) [d=1]; (release year, 2010) [d=3]; (written by, Yann Martel) [d=3]
Global Goal	*An American in Hollywood*
HiTKG	([BOP], Beatrice and Virgil) → (release year, 2010)
HiTKG-RL	([BOP], Beatrice and Virgil) → (in language, English)

Table 5: Comparison of KG path selections among neighbor candidates under global goal guidance and a given context.

Related Work

KG-grounded dialogue systems. Some work augments dialogue inputs with shallow entity and relation information (Liu et al. 2018; Parthasarathi and Pineau 2018; Young et al. 2020; Wu et al. 2019; Liu et al. 2019b; Xu et al. 2020a). A knowledge retriever works with the utterance decoder to generate responses based on the retrieved shallow commonsense knowledge entities. These models enjoy rich knowledge augmentation since all short KG paths relating to the user message are encoded, but they lack the ability to extend the topics along with the KG connections, which is essential when organizing a natural dialogue. Another set of work focuses on developing a series of goals/topics for each conversation turn by walking on the KG (Zhang et al. 2019; Moon et al. 2019; Jung, Son, and Lyu 2020). These graph walkers are recurrent or graph attention based models that attentively read the dialogues and reason over the KG paths, starting from the entity mentioned in the last user message.

Global goal guided dialogue systems. To the best of our knowledge, we are the first that study global goal oriented KG path transitions in dialogue. Existing similar work either omits explicit approaching (Xu et al. 2020b) or grounds in unstructured knowledge (Tang et al. 2019). Xu et al. propose a hierarchical policy model to plan and generate responses of different levels where the high-level policy plans a global topic. However, the low-level policy plans responses that are coherent to this topic instead of approaching it. Tang et al. study guiding the conversation to an assigned target subject, which has a similar purpose as ours, but the transition is grounded in a disordered keyword set which lacks commonsense connections.

Conclusion

We propose HiTKG, a hierarchical transformer based KG walker that leverages multiscale inputs for graph reasoning in dialogues. HiTKG first learns to plan natural turn-level goals and then learns to approach a global goal. Both automatic and human evaluation illustrate the effectiveness of our method. In the future, we will investigate how to improve the embedding, learning framework, and evaluation criteria of stage 2 to further extend this topic.

Acknowledgments

This research is supported by the Agency for Science, Technology and Research (A*STAR) under its AME Programmatic Funding Scheme (Project #A18A2b0046) and China Scholarship Council.

References

Ba, J. L.; Kiros, J. R.; and Hinton, G. E. 2016. Layer normalization. *arXiv preprint arXiv:1607.06450*.

Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; and Yakhnenko, O. 2013. Translating embeddings for modeling multi-relational data. *Advances in neural information processing systems*, 26.

Chen, X.; Meng, F.; Li, P.; Chen, F.; Xu, S.; Xu, B.; and Zhou, J. 2020. Bridging the gap between prior and posterior knowledge selection for knowledge-grounded dialogue generation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 3426–3437.

Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; and Yu, P. S. 2022. A Survey on Knowledge Graphs: Representation, Acquisition and Applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33(10).

Jung, J.; Son, B.; and Lyu, S. 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)*, 3484–3497.

Lan, Z.; Chen, M.; Goodman, S.; Gimpel, K.; Sharma, P.; and Soricut, R. 2019. Albert: A lite bert for self-supervised learning of language representations. *arXiv preprint arXiv:1909.11942*.

Li, W.; Shao, W.; Ji, S.; and Cambria, E. 2022. BiERU: Bidirectional Emotional Recurrent Unit for Conversational Sentiment Analysis. *Neurocomputing*, 467: 73–82.

Liu, C.; He, S.; Liu, K.; and Zhao, J. 2019a. Vocabulary pyramid network: Multi-pass encoding and decoding with multi-level vocabularies for response generation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3774–3783.

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

Liu, Z.; Niu, Z.-Y.; Wu, H.; and Wang, H. 2019b. Knowledge aware conversation generation with explainable reasoning over augmented graphs. *arXiv preprint arXiv:1903.10245*.

Ma, Y.; Nguyen, K. L.; Xing, F.; and Cambria, E. 2020. A Survey on Empathetic Dialogue Systems. *Information Fusion*, 64: 50–70.

Madotto, A.; Lin, Z.; Wu, C.-S.; and Fung, P. 2019. Personalizing dialogue agents via meta-learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 5454–5459.

Mnih, V.; Badia, A. P.; Mirza, M.; Graves, A.; Lillicrap, T.; Harley, T.; Silver, D.; and Kavukcuoglu, K. 2016. Asynchronous methods for deep reinforcement learning. In *International conference on machine learning*, 1928–1937. PMLR.

Moon, S.; Shah, P.; Kumar, A.; and Subba, R. 2019. Opendialkg: Explainable conversational reasoning with attention-based walks over knowledge graphs. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 845–854.

Ni, J.; Young, T.; Pandelea, V.; Xue, F.; Adiga, V.; and Cambria, E. 2021. Recent advances in deep learning based dialogue systems: A systematic survey. *arXiv preprint arXiv:2105.04387*.

Parthasarathi, P.; and Pineau, J. 2018. Extending neural generative conversational model using external knowledge sources. *arXiv preprint arXiv:1809.05524*.

Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. 2019. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32: 8026–8037.

Qiu, L.; Zhao, Y.; Shi, W.; Liang, Y.; Shi, F.; Yuan, T.; Yu, Z.; and Zhu, S.-C. 2020. Structured attention for unsupervised dialogue structure induction. *arXiv preprint arXiv:2009.08552*.

Quan, J.; and Xiong, D. 2020. Modeling long context for task-oriented dialogue state generation. *arXiv preprint arXiv:2004.14080*.

Song, Z.; Zheng, X.; Liu, L.; Xu, M.; and Huang, X.-J. 2019. Generating responses with a specific emotion in dialog. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3685–3695.

Su, H.; Shen, X.; Zhao, S.; Zhou, X.; Hu, P.; Zhong, R.; Niu, C.; and Zhou, J. 2020. Diversifying dialogue generation with non-conversational text. *arXiv preprint arXiv:2005.04346*.

Tang, J.; Zhao, T.; Xiong, C.; Liang, X.; Xing, E. P.; and Hu, Z. 2019. Target-guided open-domain conversation. *arXiv preprint arXiv:1905.11553*.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. In *Advances in neural information processing systems*, 5998–6008.

Wang, Y.-A.; and Chen, Y.-N. 2020. What do position embeddings learn? an empirical study of pre-trained language model positional encoding. *arXiv preprint arXiv:2010.04903*.

Wu, W.; Guo, Z.; Zhou, X.; Wu, H.; Zhang, X.; Lian, R.; and Wang, H. 2019. Proactive human-machine conversation with explicit conversation goals. *arXiv preprint arXiv:1906.05572*.

Xu, H.; Peng, H.; Xie, H.; Cambria, E.; Zhou, L.; and Zheng, W. 2020a. End-to-End latent-variable task-oriented dialogue system with exact log-likelihood optimization. *World Wide Web*, 23: 1989–2002.

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

Young, T.; Cambria, E.; Chaturvedi, I.; Zhou, H.; Biswas, S.; and Huang, M. 2018. Augmenting end-to-end dialogue systems with commonsense knowledge. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 4970–4977.

Young, T.; Pandelea, V.; Poria, S.; and Cambria, E. 2020. Dialogue Systems with Audio Context. *Neurocomputing*, 388: 102–109.

Zhang, H.; Liu, Z.; Xiong, C.; and Liu, Z. 2019. Grounded conversation generation as guided traverses in commonsense knowledge graphs. *arXiv preprint arXiv:1911.02707*.