Initiative-Aware Self-Supervised Learning for Knowledge-Grounded Conversations

Chuan Meng
Shandong University
Qingdao, China
mengchuan@mail.sdu.edu.cn

Pengjie Ren*
Shandong University
Qingdao, China
renpengjie@sdu.edu.cn

Zhumin Chen
Shandong University
Qingdao, China
chenzhumin@sdu.edu.cn

Zhaochun Ren
Shandong University
Qingdao, China
zhaochun.ren@sdu.edu.cn

Tengxiao Xi
Shandong University
Qingdao, China
tengxiao@mail.sdu.edu.cn

Maarten de Rijke
University of Amsterdam
& Ahold Delhaize Research
Amsterdam, The Netherlands
m.derijke@uva.nl

ABSTRACT

In the knowledge-grounded conversation (KGC) task systems aim to produce more informative responses by leveraging external knowledge. KGC includes a vital part, *knowledge selection*, where conversational agents select the appropriate knowledge to be incorporated in the next response. Mixed initiative is an intrinsic feature of conversations where the user and the system can both take the initiative in suggesting new conversational directions. Knowledge selection can be driven by the user's initiative or by the system's initiative. For the former, the system usually selects knowledge according to the current user utterance that contains new topics or questions posed by the user; for the latter, the system usually selects knowledge according to the previously selected knowledge. No previous study has considered the mixed-initiative characteristics of knowledge selection to improve its performance.

In this paper, we propose a *mixed-initiative knowledge selection method* (MIKe) for KGC, which explicitly distinguishes between user-initiative and system-initiative knowledge selection. Specifically, we introduce two knowledge selectors to model both of them separately, and design a novel *initiative discriminator* to discriminate the initiative type of knowledge selection at each conversational turn. A challenge for training MIKe is that we usually have no labels for indicating initiative. To tackle this challenge, we devise an initiative-aware self-supervised learning scheme that helps MIKe to learn to discriminate the initiative type via a self-supervised task. Experimental results on two datasets show that MIKe significantly outperforms state-of-the-art methods in terms of both automatic and human evaluations, indicating that it can select more appropriate knowledge and generate more informative and engaging responses.

*Corresponding author.

SIGIR '21, July 11–15, 2021, Virtual Event, Canada

ACM ISBN 978-1-4503-8037-9/21/07...$15.00
https://doi.org/10.1145/3404835.3462824

CCS CONCEPTS

• **Computing methodologies** → *Discourse, dialogue and pragmatics*; Natural language generation; • **Information systems** → Question answering; Users and interactive retrieval.

KEYWORDS

Knowledge-grounded conversations; Knowledge selection; Mixed initiative; Self-supervised learning

ACM Reference Format:

Chuan Meng, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Tengxiao Xi, and Maarten de Rijke. 2021. Initiative-Aware Self-Supervised Learning for Knowledge-Grounded Conversations. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '21), July 11–15, 2021, Virtual Event, Canada.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3404835.3462824

1 INTRODUCTION

Open-domain conversational agents [12, 13, 17, 18] are mainly based on the seq2seq framework [43], and they usually condition the response generation only on the conversation context, leading them to generate uninformative responses [20]. Knowledge-grounded conversations (KGCs) mitigate this problem by conditioning the response generation on external knowledge [8, 9]. Because external knowledge contains information that may be redundant or irrelevant to current conversation, *knowledge selection* (KS), that is, choosing the appropriate knowledge to be incorporated in the next response, is a vital part in KGC [14, 29].

In a conversation, *initiative* is the ability to drive the direction of the conversation [36]. Mixed initiative is an intrinsic feature of human-machine conversations [36, 45, 47], where the user and system can both take the initiative in suggesting new conversational directions by introducing new topics, asking questions, and so on. KS also has the potential for mixed-initiative, i.e., the direction of KS can be driven by the user (user-initiative KS) or by the system itself (system-initiative KS). For user-initiative KS, the system usually selects knowledge according to the current user utterance that contains new topics or questions posed by the user. As depicted in Fig. 1, at the first turn in the conversation, the current user utterance "*I love Coca-Cola. How about you?*" contains a question posed by the user, based on which the system chooses a piece of knowledge

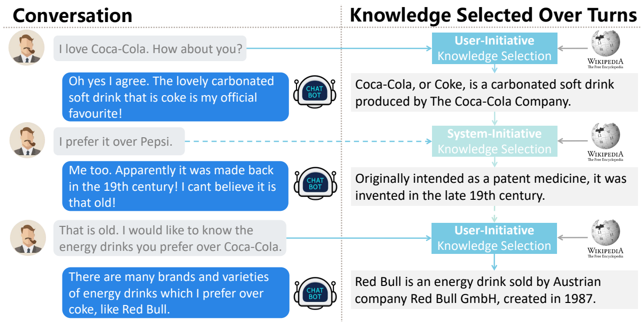

Figure 1: An example of mixed-initiative knowledge selection from the Wizard of Wikipedia (WoW) dataset [8].

about Coca-Cola. A similar case can be found at the third turn. For system-initiative KS, without user's push, the system usually selects knowledge according to the previously selected knowledge. As shown in Fig. 1, at the second turn, the current user utterance "*I prefer it over Pepsi.*" does not contain information suggesting a new conversational direction. In this case, the system selects a new piece of knowledge about the history of Coca-Cola based on the previously selected knowledge about the basic information about Coca-Cola.

No previous studies have considered the mixed-initiative nature of KS. As a result, previous studies treat the roles of the current user utterance and the previously selected knowledge equally, regardless of the different roles they play for different types of KS. We hypothesize that this omission may introduce redundant information and lead to inferior performance of KS. However, modeling mixed-initiative KS is challenging since there is no manual label indicating initiative in KGC training sets. Heuristic methods are far from enough to discriminate initiative effectively. As depicted in Fig. 1, at the third turn in the conversation, the current KS is user-initiative but the current user utterance only contains an implicit information need, which is hard to capture using heuristics.

To tackle the above issues, we propose a *mixed-initiative knowledge selection method* (MIKe) for KGC, that explicitly distinguishes between user-initiative and system-initiative KS. Specifically, we not only introduce two knowledge selectors to model user-initiative and system-initiative, but also design an *initiative discriminator* to discriminate the initiative type of KS at each turn in a conversation. To overcome the challenge of absent manual labels for indicating initiative, we devise an initiative-aware self-supervised learning (ISLe) scheme to make MIKe learn to discriminate the initiative types of KS, which is based on two insights found in data:

(1) If there is an unsmooth knowledge shift at the current conversation turn (i.e., the knowledge previously selected and the knowledge currently selected cannot be directly connected naturally), the current KS tends to be user-initiative. Conversely, the KS tends to be system-initiative. As depicted in Fig. 1, the knowledge shift at the third turn is unsmooth (i.e., knowledge about the history of Coca-Cola and Red Bull cannot be directly connected naturally), and the current KS is user-initiative. The opposite situation can be found in the second turn.
(2) If a piece of knowledge selected at one turn is deleted, the knowledge shift between the knowledge closely before and after the missing knowledge tends to be unsmooth. As depicted in Fig. 1, if the knowledge selected at the second turn is deleted, the shift between the knowledge selected at the first and third turn is also unsmooth.

Building on these intuitions, we hypothesize that learning to locate missing knowledge (detecting the knowledge closely before and after the missing knowledge) is approximately equivalent to learning to detect unsmooth knowledge shifts and detect the user-initiative KS. Thus, ISLe supervises MIKe via a self-supervised task, *locating missing knowledge*, i.e., given all pieces of ground-truth chosen knowledge in a conversation, a piece of knowledge at one of turns is randomly deleted and then the initiative discriminator is required to detect the knowledge closely after the missing knowledge (the knowledge closely before the missing knowledge could be known accordingly). Through the learning, at each turn, given the previously and currently chosen knowledge, if the initiative discriminator identifies the currently chosen knowledge as the knowledge closely after the missing knowledge, the knowledge shift between the previously and currently chosen knowledge would be unsmooth and thus the current KS would be user-initiative. Otherwise, the current KS would be system-initiative.

At each conversation turn, the knowledge currently selected is the target to predict and cannot be fetched during inference. Therefore, we further distinguish the initiative discriminator as a *teacher initiative discriminator* and a *student initiative discriminator* and upgrade ISLe to two tasks: (1) *locating missing knowledge*; the teacher is required to learn to locate the missing knowledge; (2) *learning with pseudo initiative labels*; at each turn, the teacher is first executed to generate the pseudo-label indicating the initiative type of KS, and then the student is required to learn the pseudo initiative label estimated by the teacher. During inference, we only execute the student discriminator.

Experiments on the WoW [8] and Holl-E [30] datasets indicate that MIKe can choose more appropriate knowledge, and generate more informative and engaging responses. ISLe helps MIKe to effectively discriminate the initiative type for KS.

The contributions of this paper can be summarized as follows:

- We propose *mixed-initiative knowledge selection method* (MIKe) for KGC, which explicitly distinguishes between user-initiative and system-initiative KS at each conversation turn so as to improve the performance of KS.
- We devise *initiative-aware self-supervised learning* (ISLe), which helps MIKe discriminate KS initiative types via an approximately equivalent self-supervised task, *locating missing knowledge*.
- We conduct automatic and human evaluations on two benchmark datasets, which reveals that MIKe can choose more appropriate knowledge and generate more informative and engaging responses, significantly outperforming state-of-the-art methods in terms of both automatic and human evaluation.

2 RELATED WORK

We survey two categories of related work: knowledge-grounded conversations (KGCs) and self-supervised learning.

2.1 Knowledge-grounded conversation

Existing work on KGC can be categorized into two groups. Methods in the first group aims to leverage *structured knowledge* (given

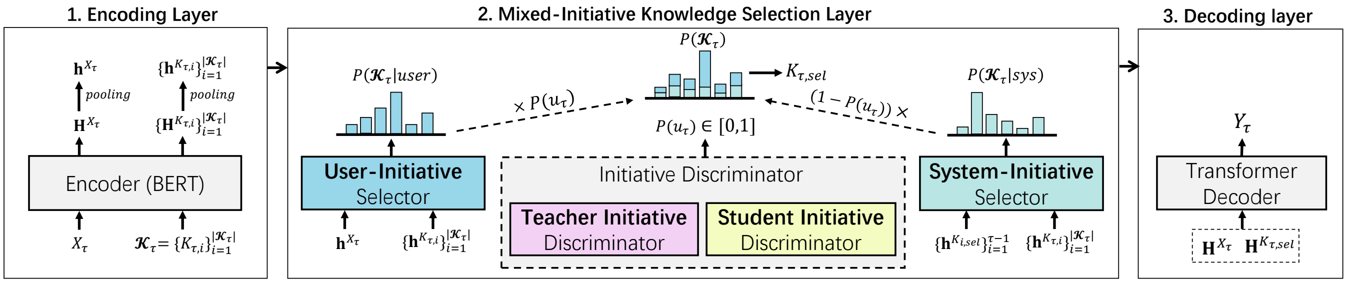

Figure 2: An overview of MIKe. Section 3 contains a walkthrough of the model.

knowledge graphs) [31, 53, 54, 64, 65]. Methods in the second group focus on leveraging *unstructured knowledge* such as *document-based unstructured knowledge* (given a whole document, e.g., Wikipedia article) [22, 28, 30, 34, 37, 44, 66] or *"piece-based" unstructured knowledge* (given some separate pieces of knowledge, e.g., Foursquare tips) [8, 9, 29]. For both groups, there are key research directions: (1) improving knowledge selection (KS) [29]; (2) improving knowledge-aware response generation [59] or response selection [11]; (3) leveraging multiple knowledge modalities [27, 57]; and (4) overcoming data scarcity [21, 58]. (5) leveraging cross-lingual knowledge [42]. We leverage *piece-based unstructured knowledge* and focus on improving KS.

Several previous publications focus on *implicit KS* to calculate a weight for each piece of knowledge and get the weighted sum of their representations [9, 25, 63]. Other studies focus on *explicit KS* [1, 62], i.e., calculating a weight on each piece of knowledge and then directly sample a piece of knowledge with the highest weight. Our work also focuses on *explicit KS*. Dinan et al. [8] propose transformer MemNet (TMemNet), which uses context to predict a distribution over pieces of knowledge and then samples one of them for the decoder. They also introduce a KS loss to supervise KS during training. Lian et al. [23] propose posterior knowledge selection (PostKS), which uses a context to predict a prior distribution over pieces of knowledge. During training, the prior distribution is supervised by a posterior distribution that is predicted by the context and the corresponding response. Kim et al. [14] propose sequential knowledge transformer (SKT), which makes use of the previously selected knowledge (tracked by the model) and context to jointly facilitate KS. Chen et al. [3] upgrade SKT by adding posterior information prediction module (PIPM) and proposing knowledge distillation based training strategy (KDBTS) to further improve KS. Zheng et al. [61] propose a difference-aware knowledge selection (DiffKS) method, which introduces the difference between the previously selected knowledge and the current pieces of candidate knowledge to further facilitate KS. Meng et al. [29] propose dual knowledge interaction network (DukeNet), which regards tracking the previously selected knowledge and selecting the current knowledge as dual tasks, supervised by dual learning [35] to teach each other. Zhao et al. [59] propose a reinforcement learning enhanced knowledge selection (RLKS) method, where the selected knowledge is sent to a decoder to generate a response that would be compared with the ground truth response to give feedback used to further supervise KS.

Unlike the work listed above, MIKe considers the mixed-initiative characteristic of KS to improve the performance of KS.

2.2 Self-supervised learning

Self-supervised learning aims at supervising a network via an objective where the ground-truth labels are automatically obtained from the raw data itself [26] It benefits a range of tasks, such as pre-trained models [7, 19], recommender systems [40, 41, 50, 52, 67], summarization [48] and open-domain conversational agents [51, 55, 56, 60]. The application in the last task is closest to our work. Specifically, Wu et al. [51] devise a self-supervised inconsistent order detection task, which aims to guide a detection model to predict whether the utterances sampled in a conversation is in sequential order. The trained detection model further provides signals used to optimize the conversational systems via adversarial training. Zhang et al. [56] devise self-supervised topic and persona feature extractors. The extracted features are sent to a decoder to help generate more consistent responses. Xu et al. [55] and Zhao et al. [60] devise a group of self-supervised tasks that help their models to produce better features for their primary task (response generation and response selection, respectively), and jointly train their primary tasks with their self-supervised tasks in a multi-task manner.

To the best of our knowledge, self-supervised learning has not been applied in KGC. Unlike the work listed above, ISLe contains a new self-supervised task to discriminate the initiative type of KS.

3 METHODOLOGY

3.1 Task formulation

Suppose that we have a conversation $C = \{(X_\tau, Y_\tau)\}_{\tau=1}^{|C|}$ with $|C|$ turns, where X_τ and Y_τ are the utterances produced by a user and a system at turn τ, separately. Each turn is accompanied with a knowledge pool $\mathcal{K}_\tau = \{K_{\tau,1}, \ldots, K_{\tau,i}, \ldots, K_{\tau,|\mathcal{K}_\tau|}\}$ (See §3.4.1 to know the source of knowledge), with $|\mathcal{K}_\tau|$ pieces of knowledge. At turn τ, given the current user utterance X_τ, the previously selected knowledge $\{K_{1,sel}, \ldots, K_{\tau-1,sel}\}$ (also written as $\{K_{i,sel}\}_{i=1}^{\tau-1}$) and the knowledge pool $\mathcal{K}_\tau$, our target is to select a piece of knowledge $K_{\tau,sel}$ from $\mathcal{K}_\tau$ and then leverage $K_{\tau,sel}$ to generate the response $Y_\tau = (y_{\tau,1}, y_{\tau,2}, \ldots, y_{\tau,|Y_\tau|})$ with $|Y_\tau|$ tokens.

3.2 Overview of MIKe

As depicted in Fig. 2, MIKe consists of three layers: (1) an *encoding layer*, (2) a *mixed-initiative knowledge selection layer*, and (3) a *decoding layer*. The encoding layer uses a BERT encoder to encode the current user utterance X_τ and the knowledge pool $\mathcal{K}_\tau$ into latent representations. The mixed-initiative knowledge selection layer contains a *user-initiative selector*, a *system-initiative selector* and an *initiative discriminator*. The user-initiative selector predicts the distribution $P(\mathcal{K}_\tau|user)$ over the knowledge pool $\mathcal{K}_\tau$ given the current user utterance X_τ, while the system-initiative selector predicts the distribution $P(\mathcal{K}_\tau|sys)$ over the knowledge pool K_τ given the previously selected knowledge $\{K_{i,sel}\}_{i=1}^{\tau-1}$. The initiative discriminator discriminates the current initiative type of KS by estimating the probability of user-initiative KS $P(u_\tau) \in [0,1]$ given the current user utterance X_τ and the previously selected knowledge $\{K_{i,sel}\}_{i=1}^{\tau-1}$. Based on $P(u_\tau)$, the distributions estimated by both selectors are combined to obtain the distribution $P(\mathcal{K}_\tau) = P(u_\tau)P(\mathcal{K}_\tau|user) + (1 - P(u_\tau))P(\mathcal{K}_\tau|sys)$, from which we select the piece of knowledge $K_{\tau,sel}$. The *decoding layer* contains a transformer decoder to generate the response Y_τ given $K_{\tau,sel}$ and X_τ.

We refer to the initiative discriminator defined above as the *student initiative discriminator* and also introduce a *teacher initiative discriminator*. During training, ISLe supervises MIKe to discriminate the initiative type of KS via two tasks: (1) *locating missing knowledge*; given all pieces of ground-truth chosen knowledge in a conversation, a piece of knowledge at one of the turns is randomly deleted and then the teacher is required to locate missing knowledge (detecting the knowledge closely after the missing knowledge); and (2) *learning with pseudo initiative labels*; at each turn, the teacher generates a pseudo-label indicating the initiative type of KS; then, the student is required to learn the pseudo-label estimated by the teacher via mean squared error (MSE) loss. During inference, only the student initiative discriminator is run.

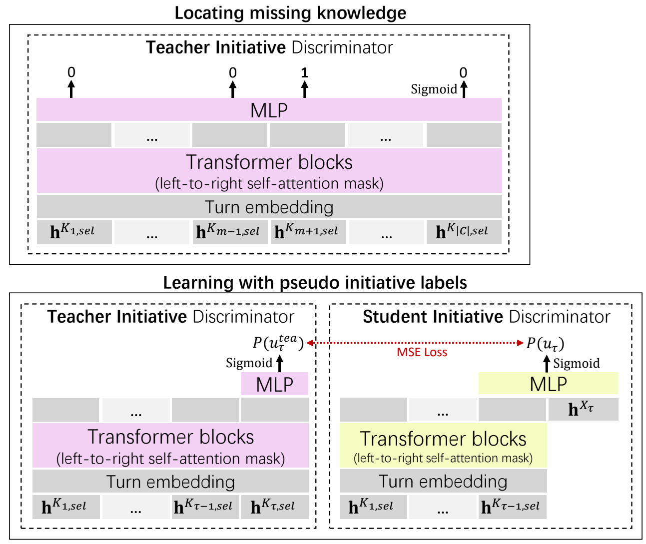

Figure 3: Two tasks of ISLe

3.3 Encoding layer

Given the current user utterance X_τ, we encode it into latent representation $\mathbf{H}^{X_\tau}$ via BERT [7], and then convert it into the condensed representation $\mathbf{h}^{X_\tau}$ via an average pooling operation [2]:

$$\mathbf{H}^{X_\tau} = \text{BERT}(X_\tau) \in \mathbb{R}^{|X_\tau| \times d},\ \mathbf{h}^{X_\tau} = \text{pooling}(\mathbf{H}^{X_\tau}) \in \mathbb{R}^{1\times d}, \quad (1)$$

where d denotes the hidden size. Likewise, given the knowledge pool $\mathcal{K}_\tau = \{K_{\tau,1}, \ldots, K_{\tau,i}, \ldots, K_{\tau,|\mathcal{K}_\tau|}\}$, we get the representations of each piece of knowledge, $\mathbf{H}^{K_{\tau,i}} \in \mathbb{R}^{|K_{\tau,i}|\times d}$ and $\mathbf{h}^{K_{\tau,i}} \in \mathbb{R}^{1\times d}$. We also track the previously selected knowledge representations $\{\mathbf{h}^{K_{i,sel}}\}_{i=1}^{\tau-1} \in \mathbb{R}^{(\tau-1)\times d}$.

3.4 Mixed-initiative knowledge selection layer

3.4.1 User-initiative selector. Given the current user utterance representation $\mathbf{h}^{X_\tau}$ and the knowledge pool representation $\{\mathbf{h}^{K_{\tau,1}}, \ldots, \mathbf{h}^{K_{\tau,|\mathcal{K}_\tau|}}\}$, the user-initiative selector predicts the probability distribution $P(\mathcal{K}_\tau|user)$ over the knowledge pool $\mathcal{K}_\tau$, which is estimated as follows:

$$\begin{aligned} P(\mathcal{K}_\tau|user) &= \text{Softmax}(\mathbf{Q}_{user}\mathbf{K}_{user}^\top) \in \mathbb{R}^{1\times|\mathcal{K}_\tau|} \\ \mathbf{Q}_{user} &= \text{MLP}(\mathbf{h}^{X_\tau}) \in \mathbb{R}^{1\times d} \\ \mathbf{K}_{user} &= \text{MLP}([\mathbf{h}^{K_{\tau,1}};\ldots;\mathbf{h}^{K_{\tau,|\mathcal{K}_\tau|}}]) \in \mathbb{R}^{|\mathcal{K}_\tau|\times d} \end{aligned} \quad (2)$$

where $\text{MLP}(\cdot) = \cdot\mathbf{W} + \mathbf{b}$ is a multilayer perceptron (MLP), and $[\cdot;\cdot]$ denotes the vector concatenation operation.

3.4.2 System-initiative selector. Given the previously selected knowledge representation $\{\mathbf{h}^{K_{i,sel}}\}_{i=1}^{\tau-1}$ and the knowledge pool representation $\{\mathbf{h}^{K_{\tau,1}}, \ldots, \mathbf{h}^{K_{\tau,|\mathcal{K}_\tau|}}\}$, the system-initiative selector predicts the probability distribution $P(\mathcal{K}_\tau|sys)$ over the knowledge pool $\mathcal{K}_\tau$, which is estimated as follow:

$$\begin{aligned} P(\mathcal{K}_\tau|sys) &= \text{Softmax}(\mathbf{Q}_{sys}\mathbf{K}_{sys}^\top) \in \mathbb{R}^{1\times|\mathcal{K}_\tau|} \\ \mathbf{Q}_{sys} &= \text{MLP}(\mathbf{h}_{trans}^{K_{\tau-1,sel}}) \in \mathbb{R}^{1\times d} \\ \mathbf{K}_{sys} &= \text{MLP}([\mathbf{h}^{K_{\tau,1}};\ldots;\mathbf{h}^{K_{\tau,|\mathcal{K}_\tau|}}]) \in \mathbb{R}^{|\mathcal{K}_\tau|\times d} \\ [\mathbf{h}_{trans}^{K_{1,sel}};&\ldots;\mathbf{h}_{trans}^{K_{\tau-1,sel}}] = \\ &\text{TransformerE}([\mathbf{h}^{K_{1,sel}};\ldots;\mathbf{h}^{K_{\tau-1,sel}}]) \in \mathbb{R}^{(\tau-1)\times d}, \end{aligned} \quad (3)$$

where TransformerE is a stack of transformer encoder blocks [46]; it first adds positional embeddings to inputs, where each turn has its own distinct positional embedding. We use a left-to-right self-attention mask in these blocks, where every position can only attend to previous positions, e.g. $K_{\tau-2,sel}$ cannot attend to $K_{\tau-1,sel}$.

3.4.3 Initiative discriminator. Given the current user utterance representation $\mathbf{h}^{X_\tau}$ and the previously selected knowledge representations $\{\mathbf{h}^{K_{i,sel}}\}_{i=1}^{\tau-1}$, the initiative discriminator predicts the probability of user-initiative KS $P(u_\tau)$ at turn τ, which is estimated as follow:

$$\begin{aligned} P(u_\tau) &= \text{Sigmoid}(\psi(u_\tau)) \in \mathbb{R}^{1\times 1} \\ \psi(u_\tau) &= \text{MLP}([\mathbf{h}_{trans}^{K_{\tau-1,sel}};\mathbf{h}^{X_\tau}]) \in \mathbb{R}^{1\times 1} \\ [\mathbf{h}_{trans}^{K_{1,sel}};&\ldots;\mathbf{h}_{trans}^{K_{\tau-1,sel}}] = \\ &\text{TransformerE}([\mathbf{h}^{K_{1,sel}};\ldots;\mathbf{h}^{K_{\tau-1,sel}}]) \in \mathbb{R}^{(\tau-1)\times d}, \end{aligned} \quad (4)$$

where TransformerE here has the same setting as the one in Eq. 3, but they do not share parameters. Accordingly, $(1 - P(u_\tau))$ is regarded as the probability of system-initiative KS. Given above results, we get the final distribution $P(\mathcal{K}_\tau) = P(u_\tau)P(\mathcal{K}_\tau|user) + (1 - P(u_\tau))P(\mathcal{K}_\tau|sys)$, from which we select the piece of knowledge $K_{\tau,sel}$ with the highest probability.

3.5 Decoding layer

The concatenated representations of the current user utterance and the selected knowledge $\mathbf{H}^{XK_\tau} = [\mathbf{H}^{X_\tau};\mathbf{H}^{K_{\tau,sel}}] \in \mathbb{R}^{|XK_\tau|\times d}$ are fed into a transformer decoder [46] with a copying mechanism [10, 39] to generate Y_τ. Note that during training, $\mathbf{H}^{K_{\tau,sel}}$ would be $\mathbf{H}^{K_{\tau,*}}$ ($K_{\tau,*}$ is the ground-truth selected knowledge). Concretely, the probability of generating $y_{\tau,t}$ at the timestamp t is modeled as:

$$P(y_{\tau,t}) = (1 - P(c))P_{vocab}(y_{\tau,t}) + P(c) \sum_{i:xk_{\tau,i}=y_{\tau,t}} \alpha_{\tau,t,i}, \tag{5}$$

where $P_{vocab}(y_{\tau,t})$ is the probability of generating $y_{\tau,t}$ from a pre-defined vocabulary V:

$$\begin{aligned} P_{vocab}(y_{\tau,t}) &= \text{Softmax}(\text{MLP}(\mathbf{h}^{dec_{\tau,t}}_{trans})) \in \mathbb{R}^{1\times|V|} \\ \mathbf{h}^{dec_{\tau,t}}_{trans} &= \text{TransformerD}(\mathbf{emb}(y_{\tau,<t}), \mathbf{H}^{XK_\tau}) \in \mathbb{R}^{1\times d}, \end{aligned} \tag{6}$$

where TransformerD is a stack of transformer decoder blocks [46] and $\mathbf{emb}(y_{\tau,<t})$ denotes the embedding of $y_{\tau,<t}$.

$\sum_{i:xk_{\tau,i}=y_{\tau,t}} \alpha_{\tau,t,i}$ is the probability of copying $y_{\tau,t}$ from the concatenated sequence of the current user utterance and the selected knowledge $XK_\tau = [X_\tau; K_{\tau,sel}]$. $xk_{\tau,i}$ is the i-th token in XK_τ and $\alpha_{\tau,t}$ is the attention distribution over XK_τ with $\mathbf{h}^{dec_{\tau,t}}_{trans}$ attending to $\mathbf{H}^{XK_\tau}$ (see Eq. 8).

$P(c)$ is used as a soft switch to choose between generating from the vocabulary V and copying from XK_τ, which is estimated as follows:

$$P(c) = \text{Sigmoid}(\text{MLP}([\mathbf{h}^{dec_{\tau,t}}_{trans}; \mathbf{c}_{\tau,t}]) \in \mathbb{R}^{1\times1}, \tag{7}$$

where $\mathbf{c}_{\tau,t}$ is the attention vector derived from $\mathbf{h}^{dec_{\tau,t}}_{trans}$ attending to $\mathbf{H}^{XK_\tau}$, which is calculated as follows:

$$\begin{aligned} \mathbf{c}_{\tau,t} &= \alpha_{\tau,t}\mathbf{H}^{XK_\tau} \in \mathbb{R}^{1\times d} \\ \alpha_{\tau,t} &= \text{Softmax}(\mathbf{Q}_c\mathbf{K}_c^\top) \in \mathbb{R}^{1\times|XK_\tau|} \\ \mathbf{Q}_c &= \text{MLP}(\mathbf{h}^{dec_{\tau,t}}_{trans}) \in \mathbb{R}^{1\times d},\ \mathbf{K}_c = \text{MLP}(\mathbf{H}^{XK_\tau}) \in \mathbb{R}^{|XK_\tau|\times d}. \end{aligned} \tag{8}$$

3.6 Initiative-Aware Self-Supervised Learning

As depicted in Fig. 3, ISLe is devised to supervise MIKe to discriminate the initiative type of KS via two tasks, *locating missing knowledge* and *learning with pseudo initiative labels*.

3.6.1 Locating missing knowledge. Given all pieces of ground-truth chosen knowledge in a conversation $\{K_{1,sel}, \ldots, K_{|C|,sel}\}$, we first corrupt it via randomly deleting one piece of knowledge at one of conversation turns (e.g., deleting the piece of knowledge $K_{\tau=m,sel}$ at the m-th turn, $1 < m < |C|$), and then we introduce a *teacher initiative discriminator* to learn to locate the missing knowledge $K_{m,sel}$ (detecting the knowledge $K_{m+1,sel}$ that is closely after the missing knowledge $K_{m,sel}$), which is estimated as follows:

$$\begin{aligned} &[\ldots; P(u^{tea}_{m-1}); P(u^{tea}_{m+1}); \ldots] = \\ &\quad \text{Sigmoid}([\ldots; \psi(u^{tea}_{m-1}); \psi(u^{tea}_{m+1}); \ldots]) \in \mathbb{R}^{1\times(|C|-1)} \\ &[\ldots; \psi(u^{tea}_{m-1}); \psi(u^{tea}_{m+1}); \ldots] = \\ &\quad \text{MLP}([\ldots; \mathbf{h}^{K_{m-1,sel}}_{trans}; \mathbf{h}^{K_{m+1,sel}}_{trans}; \ldots]) \in \mathbb{R}^{1\times(|C|-1)} \\ &[\ldots; \mathbf{h}^{K_{m-1,sel}}_{trans}; \mathbf{h}^{K_{m+1,sel}}_{trans}; \ldots] = \\ &\quad \text{TransformerE}([\ldots; \mathbf{h}^{K_{m-1,sel}}; \mathbf{h}^{K_{m+1,sel}}; \ldots]) \in \mathbb{R}^{1\times(|C|-1)}, \end{aligned} \tag{9}$$

where TransformerE here has the same settings as in Eq. 3, but they do not share parameters. $[\ldots; P(u^{tea}_{m-1}); P(u^{tea}_{m+1}); \ldots]$ are the probabilities of being the knowledge closely after the missing knowledge, where we want that $P(u^{tea}_{m+1})$ equals 1 and others equal 0, meaning that the knowledge at the m-th turn is missing and the knowledge shift between $K_{m-1,sel}$ and $K_{m+1,sel}$ is unsmooth. Therefore, the objective function is defined as:

$$\begin{aligned} &\mathcal{L}_{loc}(\theta) = \\ &\quad -\frac{1}{|C|}\sum_{\tau=1}^{|C|} I(\tau)\log P(u^{tea}_\tau) + (1-I(\tau))\log(1 - P(u^{tea}_\tau)), \end{aligned} \tag{10}$$

where $\tau \neq m$ and $I(\tau)$ is an indicator function that equals 1 if $\tau = m + 1$ and 0 otherwise.

3.6.2 Learning with pseudo initiative labels. We regard the initiative discriminator defined in §3.4.3 as *student initiative discriminator*. At turn τ, given the previously and currently selected knowledge $\{K_{1,sel}, \ldots, K_{\tau,sel}\}$, the teacher initiative discriminator predicts $P(u^{tea}_\tau)$, the probability that $K_{\tau,sel}$ is the knowledge closely after the missing knowledge, which is approximately equivalent to the probability of user-initiative KS at turn τ based on our insights:

$$\begin{aligned} P(u^{tea}_\tau) &= \text{Sigmoid}(\psi(u^{tea}_\tau)) \in \mathbb{R}^{1\times1} \\ \psi(u^{tea}_\tau) &= \text{MLP}(\mathbf{h}^{K_{\tau,sel}}_{trans}) \in \mathbb{R}^{1\times1} \\ [\mathbf{h}^{K_{1,sel}}_{trans}; \ldots; \mathbf{h}^{K_{\tau,sel}}_{trans}] &= \\ &\text{TransformerE}([\mathbf{h}^{K_{1,sel}}; \ldots; \mathbf{h}^{K_{\tau,sel}}]) \in \mathbb{R}^{\tau\times d}, \end{aligned} \tag{11}$$

where $P(u^{tea}_\tau)$ is also regarded as the pseudo initiative label. The student initiative discriminator learns with it via an MSE loss:

$$\mathcal{L}_{mse}(\theta) = -\frac{1}{|C|}\sum_{\tau=1}^{|C|}(P(u_\tau) - P(u^{tea}_\tau))^2, \tag{12}$$

where $P(u_\tau)$ is the user-initiative KS probability predicted by the student initiative discriminator (see Eq. 4). Note that only the student initiative discriminator would execute during inference.

3.7 Final learning objective

Given a conversation $C = \{(X_\tau, Y_\tau)\}_{\tau=1}^{|C|}$ with $|C|$ turns, MIKe is optimised in a multi-task learning manner and the final objective function is defined as:

$$\begin{aligned} \mathcal{L}(\theta) &= \mathcal{L}_{primary}(\theta) + \lambda L_{ISLe}(\theta) \\ \mathcal{L}_{ISLe}(\theta) &= \mathcal{L}_{loc}(\theta) + \mathcal{L}_{mse}(\theta) \\ \mathcal{L}_{primary}(\theta) &= \mathcal{L}_{ks}(\theta) + \mathcal{L}_g(\theta), \end{aligned} \tag{13}$$

where θ are all the parameters of MIKe and λ is a hyper-parameter as a trade-off between the objectives of learning primary tasks (KS and response generation) and ISLe. And $\mathcal{L}_{ks}(\theta)$ and $\mathcal{L}_g(\theta)$ are the learning objective functions for KS and response generation, separately, which are defined as:

$$\begin{aligned} \mathcal{L}_{ks}(\theta) &= -\frac{1}{|C|}\sum_{\tau=1}^{|C|}\log P(K_{\tau,*}) \\ \mathcal{L}_g(\theta) &= -\frac{1}{|C|}\sum_{\tau=1}^{|C|}\sum_{t=1}^{|Y_\tau|}\log P(y_{\tau,t} \mid y_{\tau,<t}, X_\tau, K_{\tau,*}), \end{aligned} \tag{14}$$

where $*$ refers to the index of the ground-truth selected knowledge in the knowledge pool $\mathcal{K}_\tau$.

4 EXPERIMENTAL SETUP

To assess the performance of MIKe we compare it against a number of state-of-the-art baselines. We also analyze the contribution of some of the ingredients that make up MIKe, in particular, the initiative discriminator and ISLe.

4.1 Datasets

Following [3, 14, 29, 61], we evaluate our model on two KGC datasets, Wizard of Wikipedia (WoW) [8] and Holl-E [30]. Both contain the ground-truth labels for KS (only one piece of knowledge is true for KS per turn). We split the data into training, validation and test as per the original papers.

WoW is a KGC dataset based on *piece-based unstructured knowledge*, i.e., each conversation is given some separate pieces of knowledge. In this dataset, a piece of knowledge is defined as a knowledge sentence. Each conversation is conducted between a wizard who can retrieve knowledge sentences from Wikipedia and then choose one to produce a response and an apprentice who is active in talking with the wizard. It contains 18,430/1,948/1,933 conversations for training/validation/test. The test set is split into two subsets, Test Seen (in-domain) and Test Unseen (out-of-domain): the former contains conversations on topics appearing in the training set, while the latter contains conversations on new topics. There are around 67 pieces of knowledge on average in a knowledge pool.

Holl-E is originally a KGC dataset based on *document-based unstructured knowledge*. Kim et al. [14] have changed it to a version having the same format as WoW by splitting the document into separate sentences, and recreated the ground-truth labels for KS, so we use the version released by them. It contains 7,228/930/913 conversations for training/validation/test. There are also two versions of the test set: one with a single reference and the other with multiple references (more than one ground-truth pieces of knowledge and responses for each given conversation context). There are nearly 60 pieces of knowledge on average in a knowledge pool.

4.2 Baselines

We compare MIKe with state-of-the-art KGC methods focusing on *explicit KS* and leveraging *piece-based unstructured knowledge*.

- **TMemNet** [8] extends a transformer model with a memory network storing knowledge in an end-to-end manner. Following [3, 14, 29], we replace the original transformer encoder with a BERT encoder, naming it TMemNet+BERT.
- **PostKS** [23] uses context and response to jointly predict a posterior knowledge distribution and regards it as pseudo-labels to supervise KS. Following [3, 14, 29], we use a BERT encoder for this baseline, naming it PostKS+BERT.
- **SKT** [14] makes use of the previously selected knowledge and context to jointly facilitate KS. It uses a BERT encoder and incorporates a copying mechanism in decoder [10, 39].
- **SKT+PIPM+KDBTS** [3] upgrades SKT by adding a posterior information prediction module (PIPM) and proposing knowledge distillation based training strategy (KDBTS) to improve KS.
- **DukeNet** [29] regards tracking the previously selected knowledge and selecting the current knowledge as dual tasks and use dual learning [35] to supervise them. It uses a BERT encoder and incorporates a copying mechanism in the decoder.
- **DiffKS** [61] utilizes the difference in information between the previously selected knowledge and the current candidate knowledge to improve KS. It uses a GRU [4] encoder, pre-trained GloVe embedding[33] and a copying mechanism in decoder. For a fair comparison, we replace the GRU encoder and GloVe embedding with a BERT encoder, naming it DiffKS+BERT.

4.3 Evaluation metrics

For automatic evaluation, we evaluate KS with Recall@1 (R@1) and evaluate response generation with sentence-level BLEU-4 [32], METEOR [6], ROUGE-1/2/L [24], which are widely-used in previous studies [3, 14, 29, 61, 63]. For the evaluation of multiple references in the Holl-E dataset, we follow Meng et al. [29] who evaluate KS by regarding the knowledge chosen by the model as correct once it matches any of the ground-truth knowledge, and evaluate response generation by taking the max score between responses generated by models and the multiple ground-truth responses.

We conduct human evaluation on Amazon Mechanical Turk.[1] We first randomly sample 300 examples from each test set, and each of them is annotated by three annotators. Concretely, each annotator is shown an example containing a context, the knowledge pool (at most 10 pieces of knowledge are shown to reduce the workload of the annotators), the pieces of knowledge chosen by MIKe, and a baseline (their names are masked out during annotation), as well as the responses generated by the both. Each annotator then needs to give a preference (ties are allowed) between MIKe and a baseline based on three aspects [29]: (1) *appropriateness*, i.e., which chosen knowledge is more appropriate according to the given context; (2) *informativeness*, i.e., which response looks more informative; and (3) *engagingness*, i.e., which response is better in general.

4.4 Implementation details

For all models, we apply BERT-Base-Uncased (110M) as the encoder[2] (hidden size 768), use the BERT vocabulary (the size is 30,522), set the learning rate to 0.00002, use the Adam optimizer [15] to optimize parameters, use gradient clipping with a maximum gradient norm of 0.4, train up to 10 epochs, and select the best checkpoints based on performance on the validation set. For MIKe, we set λ in Eq. 13 to 0.5, batch whole conversations rather than individual turns, train our model on one NVIDIA TITAN RTX GPU.

5 EXPERIMENTAL RESULTS

5.1 Automatic evaluation

Tables 1 and 2 show the results of all approaches on the WoW and Holl-E datasets, respectively. Overall, MIKe achieves new state-of-the-art performance on all metrics on both datasets. There are two main observations from the results.

First, MIKe distinctly outperforms other baselines in terms of KS (see R@1) on both datasets. Concretely, the R@1 percentage of MIKe exceeds the percentage of the strongest baseline SKT+PIPM+KDBTS by 1.01%–1.27% on the WoW dataset and by 1.06%–1.08% on the Holl-E dataset. The gains indicate that explicitly distinguishing the initiative type of KS makes MIKe focus on the more important one

[1] https://www.mturk.com/

[2] https://github.com/huggingface/transformers

Table 1: Automatic evaluation results on the WoW dataset. Bold face indicates the best result in terms of the corresponding metric. Significant improvements over the best baseline results are marked with * (t-test, $p < 0.05$).

Methods	Test Seen (%)						Test Unseen (%)					
	BLEU-4	METEOR	ROUGE-1	ROUGE-2	ROUGE-L	R@1	BLEU-4	METEOR	ROUGE-1	ROUGE-2	ROUGE-L	R@1
PostKS + BERT	0.77	14.16	22.68	4.27	16.59	4.83	0.39	12.59	20.82	2.73	15.25	4.39
TMemNet + BERT	1.61	15.47	24.12	4.98	17.00	23.86	0.60	13.05	21.74	3.63	15.60	16.33
SKT	1.76	16.04	24.61	5.24	17.61	25.36	1.05	13.74	22.84	4.40	16.05	18.19
DiffKS + BERT	2.22	16.82	24.75	6.27	17.90	25.62	1.69	14.69	23.62	5.05	16.82	20.11
DukeNet	2.43	17.09	25.17	6.81	18.52	26.38	1.68	15.06	23.34	5.29	17.06	19.57
SKT+PIPM+KDBTS	2.47	17.14	25.19	7.01	18.47	27.40	1.71	14.83	23.56	5.46	17.14	20.20
MIKe (ours)	**2.78***	**17.76***	**25.40**	**7.11**	**18.78***	**28.41***	**2.00***	**15.64***	**23.78***	**5.61**	**17.41***	**21.47***

Table 2: Automatic evaluation results on the Holl-E dataset. Same conventions as in Table 1.

Methods	Single golden reference (%)						Multiple golden references (%)					
	BLEU-4	METEOR	ROUGE-1	ROUGE-2	ROUGE-L	R@1	BLEU-4	METEOR	ROUGE-1	ROUGE-2	ROUGE-L	R@1
PostKS + BERT	6.54	19.30	28.94	9.89	22.15	3.95	8.49	23.97	32.85	13.10	26.17	6.40
TMemNet + BERT	8.99	24.48	31.65	13.24	25.90	28.44	12.36	28.61	35.29	16.14	29.51	37.30
SKT	17.81	29.41	35.28	21.74	30.06	28.99	24.69	35.78	41.68	28.30	36.24	39.05
DiffKS + BERT	19.08	30.87	36.37	22.88	31.30	29.39	26.20	37.32	42.77	29.57	37.53	38.99
DukeNet	19.15	30.93	36.53	23.02	31.46	30.03	26.83	37.73	43.18	30.13	38.03	40.33
SKT+PIPM+KDBTS	20.07	31.07	36.78	24.29	31.70	30.80	27.49	37.34	43.07	30.91	37.82	40.70
MIKe (ours)	**21.14***	**32.28***	**37.78**	**25.31***	**32.82***	**31.86***	**28.52***	**38.55***	**44.06**	**31.92***	**38.91***	**41.78***

for the current KS from two parts (the current user utterance or previously selected knowledge), leading to better performance of KS. It is worth noting that MIKe has a stronger ability of generalization than the baselines. Specifically, the R@1 gap between MIKe and the strongest baseline SKT+PIPM+KDBTS is 1.01% on Test Seen (in-domain) and 1.27% on Test Unseen (out-of-domain). The self-supervised task in ISLe exploits more natural and universal patterns contained in KGC data compared to other baselines, making MIKe perform well when fed with out-of-domain data.

Second, MIKe significantly outperforms other baselines in terms of response generation (see BLEU-4, METEOR, ROUGE-1/2/L) on both datasets. Note that MIKe has almost the same decoder as these strong baselines, indicating that the better KS performance of MIKe further improves the quality of generated responses.

5.2 Human evaluation

Table 3 shows the results of a comparison between MIKe and the three most competitive baselines (SKT+PIPM+KDBTS, DukeNet, DiffKS+BERT) on the more challenging WoW dataset; qualitatively similar results were observed on the Holl-E dataset.

Overall, MIKe achieves the best performance on all metrics on Test Seen and Test Unseen. MIKe outperforms the baselines in terms of Appropriateness (evaluating KS), especially having more obvious advantages over the baselines on Test Unseen. For example, the win ratio of MIKe versus the most competitive baseline SKT+PIPM+KDBTS is 25% on Test Seen and 29% on Test Unseen, which is consistent with the observations of our automatic evaluation. Surprisingly, in spite of having almost the same decoder as these baselines, MIKe still has a clear advantage over these baselines in terms of Informativeness and Engagingness (evaluating response generation). The knowledge selected by MIKe is more appropriate and thus the corresponding generated responses contain more coherent and useful information.

6 ANALYSIS

6.1 Ablation study

To analyze where the improvements of MIKe come from, we conduct an ablation study. Table 4 shows the results on the WoW dataset; qualitatively similar results were observed also for the Holl-E dataset. We consider four settings: (1) No ISLe (MIKe-ISLe in Table 4); (2) No initiative discriminator (MIKe-ISLe-ID in Table 4); (3) No user-initiative selector (MIKe-ISLe-ID-UIS in Table 4); (4) No system-initiative selector (MIKe-ISLe-ID-SIS in Table 4).

The results show that all components are beneficial for MIKe because removing any of them will decrease the results. Without ISLe, the performance of MIKe falls greatly in terms of all metrics. Concretely, it drops 0.89% and 1.03% in terms of R@1 on Test seen and Test Unseen, separately, indicating that the training signals only provided by KS and generation losses are not sufficient to discriminate the initiative type of KS, and thus it's necessary to design ISLe. Without initiative discriminator, the performance of MIKe further goes down a lot in terms of all metrics compared to the case without ISLe, dropping by 0.94% and 1.09% in terms of R@1 on Test Seen and Test Unseen, respectively, which means that explicitly distinguishing the initiative type of KS is effective. Without either user-initiative or system-initiative selector, the performance drops dramatically to the level of TMemNet+BERT, indicating that user-initiative and system-initiative KS have their own roles and should be coordinated to work together.

Table 3: Human evaluation results on the WoW dataset.

Methods	Test Seen (%)									Test Unseen (%)								
	Appropriateness			Informativeness			Engagingness			Appropriateness			Informativeness			Engagingness		
	Win	Tie	Lose	Win	Tie	Lose	Win	Tie	Lose	Win	Tie	Lose	Win	Tie	Lose	Win	Tie	Lose
MIKe vs DiffKS + BERT	32	59	9	18	76	6	26	62	12	27	67	6	19	77	4	24	64	12
MIKe vs DukeNet	27	64	9	18	75	7	22	65	13	30	66	4	18	74	8	24	61	15
MIKe vs SKT+PIPM+KDBTS	25	67	8	17	78	5	20	69	11	29	66	5	19	76	5	25	62	13

Table 4: Ablation study on the WoW dataset. -ISLe denotes removing initiative-aware self-supervised learning. -ID denotes removing initiative discriminator. -SIS and -UIS denote removing system-initiative selector and user-initiative selector.

Methods	Test Seen (%)						Test Unseen (%)					
	BLEU-4	METEOR	ROUGE-1	ROUGE-2	ROUGE-L	R@1	BLEU-4	METEOR	ROUGE-1	ROUGE-2	ROUGE-L	R@1
MIKe (ours)	**2.78**	**17.76**	**25.40**	**7.11**	**18.78**	**28.41**	**2.00**	**15.64**	**23.78**	**5.61**	**17.41**	**21.47**
MIKe-ISLe	2.63	17.22	25.15	6.97	18.67	27.52	1.67	15.38	23.42	5.28	17.04	20.44
MIKe-ISLe-ID	2.48	17.28	24.90	6.64	18.24	26.58	1.46	14.70	22.87	5.16	16.36	19.35
MIKe-ISLe-ID-UIS	1.70	15.88	24.37	5.17	17.33	23.95	0.89	13.68	22.17	4.09	15.98	16.67
MIKe-ISLe-ID-SIS	1.68	15.76	24.33	5.08	17.21	23.88	0.87	13.44	22.01	3.88	15.79	15.99

Table 5: Statistics about the manual annotation of initiative type of KS on the WoW dataset.

Initiative type	Test Seen (%)	Test Unseen (%)
User-initiative KS	47.80	48.70
System-initiative KS	52.20	51.30

6.2 Initiative discrimination evaluation

To verify whether ISLe helps MIKe to discriminate the initiative type of KS effectively, we again hire annotators on Amazon Mechanical Turk to manually annotate the initiative type of KS, and then regard the manual annotation as ground truth to evaluate the performance of the initiative discrimination of MIKe.

As for the collection of annotation, we randomly sample 1,000 examples from the two test sets of the WoW dataset, respectively (1,000 examples is almost one fifth of the two test sets), and each example is annotated by an annotator. Given an example containing a context, the ground-truth chosen knowledge and the corresponding response, the annotator needs to distinguish whether the KS in the given example is user-initiative or system-initiative. Table 5 shows the statistics about the manual annotation on the two test sets, where we found the initiative type of KS is skewed towards system-initiative KS.

We evaluate the initiative discrimination of MIKe, MIKe-ISLe, and a heuristic method with Macro-F1 score and F1 scores for these two initiative types of KS. For the first two methods, a current KS is classified into user-initiative KS if $P(u_\tau)$ in Eq. 4 is greater than 0.5 and system-initiative KS otherwise. For the heuristic method, a current KS is classified as user-initiative KS if the current user utterance contains a question mark or begins with a question word, such as “how”, “why”, “who”, “where”, “what” or “when”, and as system-initiative KS otherwise.

As shown in Table 6, MIKe markedly outperforms others in terms of all metrics on both. Specifically, MIKe exceeds MIKe-ISLe by 11.83% and 14.50% on Test Seen and Test Unseen in terms of Macro-F1, respectively, indicating that ISLe is very effective in helping MIKe to distinguish between the two initiative types of KS. Though the heuristic method outperforms MIKe-ISLe in some cases, we found there are two common situations that the heuristic method cannot handle but MIKe can. First, the current user utterance usually contains implicit information needs without any question words or question mark, such as “Let’s talk about…” and “I would like to know…”, and thus the current KS is user-initiative. Second, the current user utterance can also contains a simple question that does not need a knowledge-grounded respons, such as “really?” and “did you enjoy it?”. In this case, the current KS is still system-initiative, and the corresponding response should directly answer the simple question at first and then incorporates the chosen knowledge, suggesting a new conversational direction.

Table 6: The evaluation results of initiative discrimination on the WoW dataset. M-F1 denotes Macro F1 percentage. U-F1 and S-F1 denote the F1 percentages for user-initiative and system-initiative KS, respectively.

Methods	Test Seen (%)			Test Unseen (%)		
	M-F1	U-F1	S-F1	M-F1	U-F1	S-F1
MIKe	**62.87**	**61.79**	**63.95**	**61.79**	**61.10**	**62.48**
MIKe-ISLe	51.04	60.59	41.49	47.29	60.89	33.69
Heuristic	51.74	48.16	55.31	52.69	49.52	55.86

6.3 Case study

We randomly select two examples from the WoW test set to compare the performance of MIKe, SKT+PIPM+KDBTS, DukeNet and DiffKS+BERT in Table 7. We see that MIKe chooses more appropriate knowledge and hence generates more engaging responses with the help of its distinctions between the two initiative types of KS. For instance, in Example 1, given the current user utterance implicitly suggesting a new topic about “bulls/dream team”, MIKe identifies the current KS as user-initiative KS and then select a piece of knowledge about “Jordan”. In contrast, the baselines all ignore the implicit topic drive by the user and continue to push the current

Table 7: Case study. We marked the chosen piece of knowledge in parentheses before each response and the system utterance in the context. Note that ✓ and ✗ denote that the chosen knowledge is true or false, respectively.

	Example 1 (Test seen)	Example 2 (Test unseen)
Knowledge pool	K_1: no knowledge used . K_2: while basketball is most often played as a team sport with five players on each side , two-on-two , and one-on-one competitions are also common . K_3: basketball is a limited contact sport played on a rectangular court . K_4: jordan played 15 seasons in the nba for the chicago bulls and washington wizards	K_1: no knowledge used . K_2: instagram is a mobile , desktop , and internet - based photo - sharing application and service that allows users to share pictures and videos either publicly , or privately to pre-approved followers . K_3: google photos is a photo sharing and storage service developed by google . K_4: instagram is owned by facebook . ...
Context	User: are you a basketball fan ? System: (K_2) yes , i am a fan of the five player sport . are you ? User: not as much as i used to be . i watched the bulls in the 90s when they were the dream team . (the current user utterance)	User: i hate to admit it but i spend way too much time on instagram ! System: (K_2) i use it for sharing photos and videos User: do you have a lot of followers ? (the current user utterance)
Initiative type	User-initiative KS	System-initiative KS
Response	DiffKS + BERT: (K_3 ✗) i do know that basketball is a limited contact sport played on a rectangular court . DukeNet: (K_2 ✗) i agree . i like to play basketball . i like the sport with five players on each side . SKT+PIPM+KDBTS: (K_2 ✗) i ' m not sure but i know that while basketball is most played as a team sport with five players . MIKe: (K_4 ✓) i know that jordan played 15 seasons in the nba for the chicago bulls and washington wizards .	DiffKS + BERT: (K_3 ✗) i have a google DukeNet: (K_1 ✗) i have a lot of followers . SKT+PIPM+KDBTS: (K_1 ✗) i have not i have not . MIKe: (K_4 ✓) i have a lot of followers and i do know that it is owned by facebook .

KS based on the previously selected knowledge. In Example 2, given the current user utterance containing a question "do you have a lot of followers?" that cannot be answered with knowledge, MIKe identifies the current KS as system-initiative KS and then selects a piece of knowledge about the owner of Instagram, based on the previously selected knowledge about the definition of Instagram. The part "i have a lot of followers" in MIKe's generated response answers the simple question at first and the part "i do know that it is owned by facebook" incorporates the chosen knowledge. No baseline handles this case well. Specifically, SKT+PIPM+KDBTS ignores the previously selected knowledge, but cannot find an appropriate piece of knowledge to answer the question, generating uninformative responses. These baselines cannot distinguish between the two initiative types of KS, and so they cannot know which (the current user utterance or previously selected knowledge) is the more important feature for the current KS; misled by unimportant features, their performance on KS suffers.

7 CONCLUSION AND DISCUSSION

In this paper, we propose a mixed-initiative knowledge selection method (MIKe), which explicitly distinguishes between user-initiative and system-initiative knowledge selection (KS) to enhance the performance of KS. We also devise an ISLe scheme that allows MIKe to learn to discriminate the initiative type of KS without manually labeling. Extensive experiments on two benchmark datasets demonstrate that MIKe achieves state-of-the-art performance, indicating it can select more appropriate knowledge and generate more informative and engaging responses.

Next, we discuss limitations of MIKe and future work. First, as shown in Table 6, there is still a large room for MIKe to improve the performance of discriminating the initiative type of KS. This shows a limitation of MIKe, that is, the learning scheme ISLe refers to two assumptions, which lead to inconsistencies in some situations, e.g., user-initiative KS may go hand-in-hand with a smooth knowledge shift, and a smooth knowledge shift may exist between knowledge before and after a missing knowledge gap. In future work, we plan to develop a semi-supervised scheme that leverages the self-supervised learning signals and manual annotation together to improve the performance of initiative discrimination of KS. Second, compared to previous methods, the introduced initiative discriminator increases the computational burden of our model during training and inference [49]. We plan to design a simple but effective initiative discriminator to improve the efficiency. Finally, KGC has seen its first applications to the task of *conversational search* [5, 38] during the past few years. Thus, we also plan to extend our method to model the mixed initiative [16] in this scenario.

REPRODUCIBILITY

The code used to produce the results in this paper is available online at https://github.com/ChuanMeng/MIKe.

ACKNOWLEDGMENTS

This research was supported by the Natural Science Foundation of China (61972234, 61902219, 62072279), the National Key R&D Program of China with grant No. 2020YFB1406704, the Key Scientific and Technological Innovation Program of Shandong Province (2019JZZY010129), the Tencent WeChat Rhino-Bird Focused Research Program (JR-WXG-2021411), the Fundamental Research Funds of Shandong University, and the Hybrid Intelligence Center, a 10-year programme funded by the Dutch Ministry of Education, Culture and Science through the Netherlands Organisation for Scientific Research, https://hybrid-intelligence-centre.nl. All content represents the opinion of the authors, which is not necessarily shared or endorsed by their respective employers and/or sponsors.

REFERENCES

[1] Yeonchan Ahn, Sang-Goo Lee, and Jaehui Park. 2020. Exploiting Text Matching Techniques for Knowledge-Grounded Conversation. *IEEE Access* 8 (2020), 126201–126214.

[2] Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, et al. 2018. Universal Sentence Encoder. *arXiv preprint arXiv:1803.11175* (2018).

[3] Xiuyi Chen, Fandong Meng, Peng Li, Feilong Chen, Shuang Xu, Bo Xu, and Jie Zhou. 2020. Bridging the Gap between Prior and Posterior Knowledge Selection for Knowledge-Grounded Dialogue Generation. In *EMNLP*. 3426–3437.

[4] Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. In *EMNLP*. 1724–1734.

[5] Jeffrey Dalton, Chenyan Xiong, and Jamie Callan. 2019. TREC CAsT 2019: The Conversational Assistance Track Overview. *TREC* (2019).

[6] Michael Denkowski and Alon Lavie. 2014. Meteor Universal: Language Specific Translation Evaluation for Any Target Language. In *Proceedings of the ninth workshop on statistical machine translation*. 376–380.

[7] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *NAACL*. 4171–4186.

[8] Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. 2019. Wizard of Wikipedia: Knowledge-Powered Conversational Agents. In *ICLR*.

[9] Marjan Ghazvininejad, Chris Brockett, Ming-Wei Chang, Bill Dolan, Jianfeng Gao, Wen-tau Yih, and Michel Galley. 2018. A Knowledge-grounded Neural Conversation Model. In *AAAI*. 5110–5117.

[10] Jiatao Gu, Zhengdong Lu, Hang Li, and Victor OK Li. 2016. Incorporating Copying Mechanism in Sequence-to-Sequence Learning. In *ACL*. 1631–1640.

[11] Kai Hua, Zhiyuan Feng, Chongyang Tao, Rui Yan, and Lu Zhang. 2020. Learning to Detect Relevant Contexts and Knowledge for Response Selection in Retrieval-based Dialogue Systems. In *CIKM*. 525–534.

[12] Minlie Huang, Xiaoyan Zhu, and Jianfeng Gao. 2020. Challenges in Building Intelligent Open-domain Dialog Systems. *TOIS* 38, 3 (2020), 1–32.

[13] Xisen Jin, Wenqiang Lei, Zhaochun Ren, Hongshen Chen, Shangsong Liang, Yihong Zhao, and Dawei Yin. 2018. Explicit State Tracking with Semi-Supervision for Neural Dialogue Generation. In *CIKM*. 1403–1412.

[14] Byeongchang Kim, Jaewoo Ahn, and Gunhee Kim. 2020. Sequential Latent Knowledge Selection for Knowledge-Grounded Dialogue. In *ICLR*.

[15] Diederik P Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. In *ICLR*.

[16] Antonios Minas Krasakis, Mohammad Aliannejadi, Nikos Voskarides, and Evangelos Kanoulas. 2020. Analysing the Effect of Clarifying Questions on Document Ranking in Conversational Search. In *ICTIR*. 129–132.

[17] Wenqiang Lei, Xiangnan He, Yisong Miao, Qingyun Wu, Richang Hong, Min-Yen Kan, and Tat-Seng Chua. 2020. Estimation-action-reflection: Towards Deep Interaction between Conversational and Recommender Systems. In *WSDM*. 304–312.

[18] Wenqiang Lei, Xisen Jin, Min-Yen Kan, Zhaochun Ren, Xiangnan He, and Dawei Yin. 2018. Sequicity: Simplifying Task-oriented Dialogue Systems with Single Sequence-to-sequence Architectures. In *ACL*. 1437–1447.

[19] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In *ACL*. 7871–7880.

[20] Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2016. A Diversity-Promoting Objective Function for Neural Conversation Models. In *NAACL*. 110–119.

[21] Linxiao Li, Can Xu, Wei Wu, Yufan Zhao, Xueliang Zhao, and Chongyang Tao. 2020. Zero-Resource Knowledge-Grounded Dialogue Generation. In *NeurIPS*.

[22] Zekang Li, Cheng Niu, Fandong Meng, Yang Feng, Qian Li, and Jie Zhou. 2019. Incremental Transformer with Deliberation Decoder for Document Grounded Conversations. In *ACL*. 12–21.

[23] Rongzhong Lian, Min Xie, Fan Wang, Jinhua Peng, and Hua Wu. 2019. Learning to Select Knowledge for Response Generation in Dialog Systems. In *IJCAI*.

[24] Chin-Yew Lin. 2004. Rouge: A Package for Automatic Evaluation of Summaries. In *Text Summarization Branches Out*. 74–81.

[25] Xiexiong Lin, Weiyu Jian, Jianshan He, Taifeng Wang, and Wei Chu. 2020. Generating Informative Conversational Response using Recurrent Knowledge-Interaction and Knowledge-Copy. In *ACL*. 41–52.

[26] Xiao Liu, Fanjin Zhang, Zhenyu Hou, Jing Zhang, and Jie Tang. 2020. Self-supervised Learning: Generative or Contrastive. *arXiv preprint arXiv:2006.08218* (2020).

[27] Zhibin Liu, Zheng-Yu Niu, Hua Wu, and Haifeng Wang. 2019. Knowledge Aware Conversation Generation with Reasoning on Augmented Graph. In *EMNLP*. 1782–1792.

[28] Chuan Meng, Pengjie Ren, Zhumin Chen, Christof Monz, Jun Ma, and Maarten de Rijke. 2020. RefNet: A Reference-aware Network for Background Based Conversation. In *AAAI*. 8496–8503.

[29] Chuan Meng, Pengjie Ren, Zhumin Chen, Weiwei Sun, Zhaochun Ren, Zhaopeng Tu, and Maarten de Rijke. 2020. DukeNet: A Dual Knowledge Interaction Network for Knowledge-Grounded Conversation. In *SIGIR*. 1151–1160.

[30] Nikita Moghe, Siddhartha Arora, Suman Banerjee, and Mitesh M Khapra. 2018. Towards Exploiting Background Knowledge for Building Conversation Systems. In *EMNLP*. 2322–2332.

[31] Seungwhan Moon, Pararth Shah, Anuj Kumar, and Rajen Subba. 2019. Opendialkg: Explainable Conversational Reasoning with Attention-based Walks over Knowledge Graphs. In *ACL*. 845–854.

[32] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. BLEU: A Method for Automatic Evaluation of Machine Translation. In *ACL*. 311–318.

[33] Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *EMNLP*. 1532–1543.

[34] Lianhui Qin, Michel Galley, Chris Brockett, and Xiaodon Liu. 2019. Conversing by Reading: Contentful Neural Conversation with On-demand Machine Reading. In *ACL*. 5427–5436.

[35] Tao Qin. 2020. Dual Learning for Machine Translation and Beyond. In *Dual Learning*. Springer, 49–72.

[36] Filip Radlinski and Nick Craswell. 2017. A Theoretical Framework for Conversational Search. In *CHIIR*. 117–126.

[37] Pengjie Ren, Zhumin Chen, Christof Monz, Jun Ma, and Maarten de Rijke. 2020. Thinking Globally, Acting Locally: Distantly Supervised Global-to-Local Knowledge Selection for Background Based Conversation. In *AAAI*.

[38] Pengjie Ren, Zhumin Chen, Zhaochun Ren, Evangelos Kanoulas, Christof Monz, and Maarten de Rijke. 2020. Conversations with Search Engines. *TOIS* (2020).

[39] Abigail See, Peter J Liu, and Christopher D Manning. 2017. Get To The Point: Summarization with Pointer-Generator Networks. In *ACL*. 1073–1083.

[40] Peijie Sun, Le Wu, and Meng Wang. 2018. Attentive Recurrent Social Recommendation. In *SIGIR*. 185–194.

[41] Peijie Sun, Le Wu, Kun Zhang, Yanjie Fu, Richang Hong, and Meng Wang. 2020. Dual Learning for Explainable Recommendation: Towards Unifying User Preference Prediction and Review Generation. In *WWW*. 837–847.

[42] Weiwei Sun, Chuan Meng, Qi Meng, Zhaochun Ren, Pengjie Ren, Zhumin Chen, and Maarten de Rijke. 2021. Conversations Powered by Cross-Lingual Knowledge. In *SIGIR*.

[43] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. 2014. Sequence to Sequence Learning with Neural Networks. In *NeurIPS*. 3104–3112.

[44] Zhiliang Tian, Wei Bi, Dongkyu Lee, Lanqing Xue, Yiping Song, Xiaojiang Liu, and Nevin L. Zhang. 2020. Response-Anticipated Memory for On-Demand Knowledge Integration in Response Generation. In *ACL*. 650–659.

[45] Svitlana Vakulenko, Evangelos Kanoulas, and Maarten de Rijke. 2020. An Analysis of Mixed Initiative and Collaboration in Information-Seeking Dialogues. In *SIGIR*. 2085–2088.

[46] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention Is All You Need. In *NeurIPS*. 5998–6008.

[47] Marilyn Walker and Steve Whittaker. 1990. Mixed Initiative in Dialogue: An Investigation into Discourse Segmentation. In *ACL*. 70–78.

[48] Hong Wang, Xin Wang, Wenhan Xiong, Mo Yu, Xiaoxiao Guo, Shiyu Chang, and William Yang Wang. 2019. Self-Supervised Learning for Contextualized Extractive Summarization. In *ACL*. 2221–2227.

[49] Meng Wang, Weijie Fu, Xiangnan He, Shijie Hao, and Xindong Wu. 2020. A Survey on Large-scale Machine Learning. *TKDE* (2020).

[50] Jiancan Wu, Xiang Wang, Fuli Feng, Xiangnan He, Liang Chen, Jianxun Lian, and Xing Xie. 2021. Self-supervised Graph Learning for Recommendation. *SIGIR*.

[51] Jiawei Wu, Xin Wang, and William Yang Wang. 2019. Self-Supervised Dialogue Learning. In *ACL*. 3857–3867.

[52] Le Wu, Peijie Sun, Yanjie Fu, Richang Hong, Xiting Wang, and Meng Wang. 2019. A Neural Influence Diffusion Model for Social Recommendation. In *SIGIR*. 235–244.

[53] Sixing Wu, Ying Li, Dawei Zhang, Yang Zhou, and Zhonghai Wu. 2020. Diverse and Informative Dialogue Generation with Context-specific Commonsense Knowledge Awareness. In *ACL*. 5811–5820.

[54] Jun Xu, Haifeng Wang, Zheng-Yu Niu, Hua Wu, Wanxiang Che, and Ting Liu. 2020. Conversational Graph Grounded Policy Learning for Open-domain Conversation Generation. In *ACL*. 1835–1845.

[55] Ruijian Xu, Chongyang Tao, Daxin Jiang, Xueliang Zhao, Dongyan Zhao, and Rui Yan. 2021. Learning an Effective Context-Response Matching Model with Self-Supervised Tasks for Retrieval-based Dialogues. In *AAAI*.

[56] Yizhe Zhang, Xiang Gao, Sungjin Lee, Chris Brockett, Michel Galley, Jianfeng Gao, and Bill Dolan. 2019. Consistent Dialogue Generation with Self-supervised Feature Learning. *arXiv preprint arXiv:1903.05759* (2019).

[57] Xiangyu Zhao, Longbiao Wang, Ruifang He, Ting Yang, Jinxin Chang, and Ruifang Wang. 2020. Multiple Knowledge Syncretic Transformer for Natural Dialogue Generation. In *WWW*. 752–762.

[58] Xueliang Zhao, Wei Wu, Chongyang Tao, Can Xu, Dongyan Zhao, and Rui Yan. 2019. Low-Resource Knowledge-Grounded Dialogue Generation. In *ICLR*.

[59] Xueliang Zhao, Wei Wu, Can Xu, Chongyang Tao, Dongyan Zhao, and Rui Yan. 2020. Knowledge-Grounded Dialogue Generation with Pre-trained Language Models. In *EMNLP*. 3377–3390.

[60] Yufan Zhao, Can Xu, and Wei Wu. 2021. Learning a Simple and Effective Model for Multi-turn Response Generation with Auxiliary Tasks. In *EMNLP*. 3472–3483.

[61] Chujie Zheng, Yunbo Cao, Daxin Jiang, and Minlie Huang. 2020. Difference-aware Knowledge Selection for Knowledge-grounded Conversation Generation. In *Findings of EMNLP*. 115–125.

[62] Wen Zheng, Natasa Milic-Frayling, and Ke Zhou. 2020. Approximation of Response Knowledge Retrieval in Knowledge-grounded Dialogue Generation. In *Findings of EMNLP*. 3581–3591.

[63] Wen Zheng and Ke Zhou. 2019. Enhancing Conversational Dialogue Models with Grounded Knowledge. In *CIKM*. 709–718.

[64] Hao Zhou, Tom Young, Minlie Huang, Haizhou Zhao, Jingfang Xu, and Xiaoyan Zhu. 2018. Commonsense Knowledge Aware Conversation Generation with Graph Attention.. In *IJCAI*. 4623–4629.

[65] Hao Zhou, Chujie Zheng, Kaili Huang, Minlie Huang, and Xiaoyan Zhu. 2020. KdConv: A Chinese Multi-domain Dialogue Dataset Towards Multi-turn Knowledge-driven Conversation. In *ACL*. 7098–7108.

[66] Kangyan Zhou, Shrimai Prabhumoye, and Alan W Black. 2018. A Dataset for Document Grounded Conversations. In *EMNLP*. 708–713.

[67] Kun Zhou, Hui Wang, Wayne Xin Zhao, Yutao Zhu, Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, and Ji-Rong Wen. 2020. S3-rec: Self-supervised Learning for Sequential Recommendation with Mutual Information Maximization. In *CIKM*. 1893–1902.