



# A Unified Multi-task Learning Framework for Multi-goal Conversational Recommender Systems

YANG DENG, The Chinese University of Hong Kong

WENXUAN ZHANG, DAMO Academy, Alibaba Group, Singapore

WEIWEN XU, The Chinese University of Hong Kong

WENQIANG LEI, Sichuan University, China

TAT-SENG CHUA, Sea-NExT Joint Lab, National University of Singapore, Singapore

WAI LAM, The Chinese University of Hong Kong

Recent years witnessed several advances in developing multi-goal conversational recommender systems (MG-CRS) that can proactively attract users' interests and naturally lead user-engaged dialogues with multiple conversational goals and diverse topics. Four tasks are often involved in MG-CRS, including Goal Planning, Topic Prediction, Item Recommendation, and Response Generation. Most existing studies address only some of these tasks. To handle the whole problem of MG-CRS, modularized frameworks are adopted where each task is tackled independently without considering their interdependencies. In this work, we propose a novel Unified Multi-goal conversational recommender system (UniMIND). Specifically, we unify these four tasks with different formulations into the same sequence-to-sequence paradigm. Prompt-based learning strategies are investigated to endow the unified model with the capability of multi-task learning. Finally, the overall learning and inference procedure consists of three stages, including multi-task learning, prompt-based tuning, and inference. Experimental results on two MG-CRS benchmarks (DuRecDial and TG-ReDial) show that UniMIND achieves state-of-the-art performance on all tasks with a unified model. Extensive analyses and discussions are provided for shedding some new perspectives for MG-CRS.

CCS Concepts: • **Information systems** → Users and interactive retrieval; Recommender systems; • **Computing methodologies** → Discourse, dialogue and pragmatics; Natural language generation;

Additional Key Words and Phrases: Conversational recommender system, dialogue generation, paradigm shift, prompt-based learning

## ACM Reference format:

Yang Deng, Wenxuan Zhang, Weiwen Xu, Wenqiang Lei, Tat-Seng Chua, and Wai Lam. 2023. A Unified Multi-task Learning Framework for Multi-goal Conversational Recommender Systems. *ACM Trans. Inf. Syst.* 41, 3, Article 77 (February 2023), 25 pages.  
<https://doi.org/10.1145/3570640>

The work described in this article is substantially supported by a grant from the Research Grant Council of the Hong Kong Special Administrative Region, China (Project Code No. 14200719) and the Sea-NExT Joint Lab. Work done when Yang Deng was a visiting research scholar in the Sea-NExT Joint Lab.

Authors' addresses: Y. Deng, W. Xu, and W. Lam, The Chinese University of Hong Kong, Hong Kong, 999077; emails: {ydeng, wxu, wlam}@se.cuhk.edu.hk; W. Zhang, DAMO Academy, Alibaba Group, Singapore; email: isakzhang@gmail.com; W. Lei (corresponding author), Sichuan University, China; email: wenqianglei@gmail.com; T.-S. Chua, Sea-NExT Joint Lab, National University of Singapore, Singapore; email: chuats@comp.nus.edu.sg.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2023 Association for Computing Machinery.

1046-8188/2023/02-ART77 \$15.00

<https://doi.org/10.1145/3570640>

## 1 INTRODUCTION

**Conversational Recommender Systems (CRS)** aim to make recommendations by learning users' preferences through interactive conversations [23, 27, 60]. CRS has become one of the trending research topics for recommender systems and is gaining increasing attention, due to its natural advantage of explicitly acquiring users' real-time preferences and providing a user-engaged recommendation procedure. Based on different scenarios, various CRS have been proposed, either from the perspective of recommender systems, being an enhanced interactive recommender system [7, 23, 48], or from the perspective of dialogue systems, being a variation of goal-oriented conversational systems [19, 24, 27].

Most of these CRS assume that users always know what they want and the system passively and solely targets at making the successful recommendation on users' desired items. Recently, several efforts have been made on developing **multi-goal conversational recommender systems (MG-CRS)** [33, 65] that can proactively discover users' interests and naturally lead user-engaged dialogues with multiple conversation goals or topics, not just recommendations. As the examples illustrated in Figure 1, MG-CRS is expected to dynamically adjust the conversational goals (e.g., from QA, Chit-chat, to Recommendation in Figure 1(a)) and adaptively lead the dialogue topics (e.g., from Alien, Hero, to Animation in Figure 1(b)). In specific, the class of goal is defined as the dialogue type that is related to the function of the dialogue, while the class of topic is defined as the dialogue topic that is related to the content of the dialogue. Accordingly, the problem of MG-CRS can be decomposed into four tasks:

- *Goal Planning* aims to plan the goal sequence to guide the conversation to the final recommendation.
- *Topic Prediction* predicts the appropriate topics<sup>1</sup> for the current conversational goal.
- *Item Recommendation* provides suitable items that meet the user's need.
- *Response Generation* produces appropriate natural language responses to users.

Early works typically adopt modularized frameworks that address different tasks in MG-CRS with independent models. For example, Liu et al. [33] propose a CNN-based goal planning model to predict the next goal and a multi-GRU-based response model to jointly select topical knowledge and generate responses with the guidance of the predicted goal. Zhou et al. [65] propose three separated modules based on pre-trained language models for topic prediction, item recommendation, and response generation, while the conversational goals (Chit-chat or Recommend) at each turn are pre-defined. Due to the complexity of the whole problem of MG-CRS, some recent studies tend to simplify the MG-CRS problem by either (i) assuming some information (e.g., the goal sequence) is priorly known [1, 56] or (ii) only performing joint learning on some of the tasks (e.g., topic prediction and response generation) [52], instead of solving the whole problem of MG-CRS.

Despite their effectiveness, there remain some limitations in the existing systems: (i) The four tasks in MG-CRS are supposed to be closely related, while existing systems often learn each task individually without considering inter-task interactions for mutual enhancement. (ii) In reality, it is impossible to always have the pre-defined goal sequences or topic threads for guiding the conversation, which is also the challenge that characterizes MG-CRS from other CRS. Therefore, all tasks are expected to be learned and inferred by the system automatically. (iii) There are some substantial differences among the four tasks, including various task paradigms (from multi-label classification, ranking, to text generation) and diverse required information (e.g., dialogue context, user profile, knowledge base, etc.). It would be time-consuming and labor-intensive to design and train several independent models for these distinct tasks.

<sup>1</sup>According to different applications, the topics can be specific topic classes (e.g., TG-ReDial), topical knowledge entities (e.g., DuRecDial), and so on.

Conversation	Goal	Topic	Item
1. User: Hi~ Good evening.	Greeting	N/A	
2. System: Good evening. What can I help you?			
3. User: Do you know who starred in <Bruce Lee: Warrior>?	QA	Bruce Lee: Warrior	N/A
4. System: It's Jackie Chan.			
5. User: I like him very much. He's really great.	Chit-chat about Star		
6. System: You have a good taste. He has won the best actor twice in the Golden Horse Award.			
7. User: He's very good at acting.		Jackie Chan	
8. System: Do you want to see her movie <Jackie Chan: My Stunts>? It can be regarded as the most classic movie of Jackie Chan.	Movie Rec.		Jackie Chan: My Stunts
9. User: Is this movie with a high rating?			
10. System: The rating is 7.8, which is very high!	QA	Jackie Chan: My Stunts	N/A
11. User: But I prefer Nicholas Tse's movies now.			
12. System: Nicholas Tse's <The Witness> is also very good and worth seeing. The director is good at portraying emotions.	Movie Rec.	Nicholas Tse	The Witness
13. User: I will watch it!			
14. System: Sure, you will like it!			
15. User: Thank you! See you next time!	Goodbye	N/A	N/A
16. System: Great! Bye!			

(a) An example dialogue in the DuRecDial dataset

Conversation	Goal	Topic	Item
1. User: Hey, are you here?	Greeting	N/A	N/A
2. System: Hi~ What are you up to?	Chit-chat	Alien	
3. User: I'm watching a video about <u>aliens</u> .	Rec. Request	Hero	
4. System: I haven't seen aliens yet. Do they look like monsters?	Movie Rec.	Hero / Marvel	The Black Widow
5. User: Like the <u>hero</u> after transforming.		Animation	Doctor Strange
6. System: Is it the kind of hero in the movie?			
7. User: Yes, exactly.	Feedback	N/A	N/A
8. System: Do you want to see a movie about hero? I can recommend to you.			
9. User: Sure, Are there any <u>Marvel hero</u> movies? I like Marvel very much.			
10. System: <The Black Widow> is a great movie. The action shots are very cool.			
11. User: Great. Could you recommend some other <u>animated</u> movies?			
12. System: Have you watched Marvel Animation <Doctor Strange>? It is more original than the live-action movie.			
13. User: Thanks for your recommendation! I will watch it later.			
14. System: Great! Bye!			

(b) An example dialogue in the TG-ReDial dataset

Fig. 1. Two illustrative examples of Multi-goal CRS from DuRecDial [33] and TG-ReDial [65] datasets.

To tackle the aforementioned challenges, we propose a novel **Unified Multi-goal conversational recommender system (UniMIND)**, which unifies four tasks in MG-CRS into the same **sequence-to-sequence (Seq2Seq)** paradigm and utilizes prompt-based learning strategies to endow the model with the capability of multi-task learning. In specific, motivated by recent successes of paradigm shifts [47] on many NLP tasks, we reformulate each task in MG-CRS as a Seq2Seq problem. Seq2Seq is a general and flexible paradigm that can handle any task whose input and output can be recast as a sequence of tokens. In the context of MG-CRS, there are several merits of the unified Seq2Seq formulation: (i) **The Seq2Seq formulation can better leverage the semantic relationships between input and output. Topic prediction is a multi-label classification problem, where the Seq2Seq paradigm can fully utilize the label semantics [55] among multiple class labels.**

Item recommendation requires to rank a list of candidate items, where we expand the original vocabulary of the Seq2Seq model with an extra item vocabulary to capture the relationship between context words and candidate items [6]. (ii) **Pre-trained language models (PLMs)**, e.g., GPT2 [39], T5 [40], which can largely improve the performance, have become the de-facto methods for text generation tasks. The reformulation into Seq2Seq can make better use of the knowledge from PLMs to enhance the learning of each task in MG-CRS. (iii) The unified Seq2Seq formulation to perform the multi-task learning of all tasks in MG-CRS. To adapt PLMs to each task of MG-CRS, we investigate prompt-based learning strategies [30] to manipulate the model behavior so that the PLM itself can be used to predict the desired output and facilitate multi-task learning.

The contributions are summarized as follows:

- We propose a novel method, namely, UniMIND, that tackle all tasks in MG-CRS with a unified model. To the best of our knowledge, it is the first attempt toward a unified framework for MG-CRS.
- We reformulate all tasks in MG-CRS with diverse modeling paradigms into the same Seq2Seq paradigm to seamlessly unify the complex MG-CRS problem.
- We investigate prompt-based learning strategies to enable the multi-task learning of all tasks in MG-CRS, and develop a special token prompting strategy that bridges the relationships within each type of information.
- Experimental results on two benchmark MG-CRS datasets show that UniMIND achieves state-of-the-art performance on all tasks. Extensive analyses provide some new insights of the features in different types of dialogues, and some takeaways for future MG-CRS studies.

## 2 RELATED WORK

### 2.1 Conversational Recommendation

CRS [14, 18] generally consists of two main components: a dialogue component to interact with the user and a recommender component to select items for recommendations based on user preference. According to the form of conversation, existing mainstream CRS can be divided into two groups: attribute-based CRS and open-ended CRS [53]. Attribute-based CRS [11, 12, 41, 54, 60] asks clarification questions about the item attributes to acquire user preferences for making better recommendation. For these CRS, the system usually asks questions about the user's preferences or makes recommendations multiple times, with the goal of achieving engaging and successful recommendations with fewer turns of conversations [8, 25, 50]. Open-ended CRS [6, 19, 27] focuses on how to understand users' preferences and intentions from their utterances and interacts with user through natural language conversations. Some latest studies explore the utilization of knowledge graph [36, 57, 62, 64] and textual reviews [34, 66] for enhancing the performance of open-ended CRS. Zou et al. [67] propose a Transformer-based sequential conversational recommendation method to models the sequential dependencies in the conversations for improving the performance of item recommendation in CRS.

All aforementioned studies on CRS typically target at a single goal, i.e., making successful recommendations. To be specific, attribute-based CRS studies [8, 23, 25] mainly focus on the policy learning of when to ask a question or recommend items as well as which items to recommend. While open-ended CRS studies [6, 19, 27] mainly investigate the problem of generating appropriate responses to make successful recommendations. Some latest studies [16, 33, 65] aim at solving the problem of **multi-goal conversational recommender systems (MG-CRS)**, which involves a sequence of goals to lead a user-engaged conversation, such as recommendation, chit-chat, QA, topic-focused dialogue, and so on. Therefore, the problem of MG-CRS becomes more complicated than conventional CRS, since it involves four different tasks, including goal planning, topic

Table 1. Comparison of the Addressed Tasks between Conventional CRS and MG-CRS Methods

Method	Goal Planning	Topic Prediction	Item Recommendation	Response Generation
Attribute-based [8, 23, 25]	✓*	×	✓	×
Open-ended [6, 19, 27]	×	×	✓	✓
MGCG [33]	✓	✓	×	✓
GOKC [1]	○	✓	×	✓
KERS [56]	○	✓	×	✓
Union [65]	○	✓	✓	✓
TopicRef. [52]	○	✓	✓	✓
UniMIND	✓	✓	✓	✓

\*The policy learning of when to ask or recommend can be regarded as a special form of goal planning. ○ denotes that the information is pre-defined without learning.

prediction, item recommendation, and response generation. Existing studies mainly focus on some of the tasks in MG-CRS. For example, Bai et al. [1] and Zhang et al. [56] assume that the complete goal sequence is given and study the task as a knowledge-grounded response selection problem. Wang et al. [52] jointly predict the topics and generate responses without the consideration of item recommendation. As presented in Table 1, we summarize the addressed tasks in the existing works of conventional CRS as well as MG-CRS. To tackle the entire problem of MG-CRS, we investigate a unified framework for all the tasks.

## 2.2 Pre-trained Seq2Seq Models for Dialogue

Conventional dialogue systems [5] can be generally categorized into chitchat-based dialogue systems, which aim at conversing with users on open-domain topics, and task-oriented dialogue systems, which target at assisting users to accomplish certain goals. Recently, both kinds of dialogue systems benefit from the advances in pre-trained Seq2Seq models, e.g., GPT-2 [39] and T5 [40]. For example, DialoGPT [61] and Plato [2] extend GPT-2 [39] and BERT [9], respectively, to pre-train on open-domain conversational data for dialogue generation. In task-oriented dialogue systems, several attempts have been made on leveraging pre-trained Seq2Seq models for generative dialogue state tracking [17, 22], or further applying multi-task dialogue pre-training over external dialogue corpora [38, 46]. In the scope of CRS, pre-trained Seq2Seq models are typically adopted as the response generation module [52, 65]. In this work, we aim to maximize the utility of pre-trained Seq2Seq models on all tasks in MG-CRS.

## 2.3 Paradigm Shift and Prompt-based Learning

In the past years, modeling for most NLP tasks have converged to several mainstream paradigms [47], including Classification, Matching/Ranking, Sequence Labeling, Seq2Seq, and so on. Recent work has shown that models under some paradigms also generalize well on tasks with other paradigms. For example, multi-label or multi-task classification may be challenging for conventional classification modeling, where Yang et al. [55] adopt the Seq2Seq paradigm to better capture interactions between the labels and Chai et al. [4] adopt the Matching paradigm to predict whether the text is matched with the label with descriptions. To better utilize powerful **pre-trained language models (PLMs)** with diverse pre-training paradigms, prompt-based learning [30] has been widely studied for shifting the target task to adaptive modeling paradigms with the PLM by using appropriate prompts. Text classification tasks can also be solved by PLMs with the **Masked Language Modeling (MLM)** paradigm [15, 42, 43] or the Seq2Seq paradigm [40]. In addition, several attempts have been made on adapting pre-trained Seq2Seq models to the task of document ranking [10, 37]. Some latest studies investigate paradigm shift or prompt-based learning on more diverse tasks, such as information extraction [35], sentiment analysis [58, 59], and so on.

In the field of dialogue systems, such paradigm shift techniques also bring inspiring progresses on task-oriented dialogue systems. As mentioned in Section 2.2, recent state-of-the-art performance [17, 22] on dialogue state tracking is achieved by the reformulation of such a structure prediction task as a text generation task. In addition, Byrne et al. [3] investigate the unified Seq2Seq paradigm for action prediction in transaction-based dialogue systems. However, there are more complicated subtasks with diverse paradigms in MG-CRS, varying from multi-label classification, ranking/recommendation, to text generation. In this work, we propose to unify all the subtasks with different paradigms in MG-CRS into Seq2Seq paradigm, and design prompt-based learning approaches for utilizing pre-trained Seq2Seq models for MG-CRS.

### 3 PROBLEM DEFINITION

Let  $C_t = \{c_1, \dots, c_{t-1}\}$  denote the dialogue context at the current conversation turn  $t$ . Correspondingly, let  $\mathcal{G}_t = \{g_1, \dots, g_{t-1}\}$  and  $\mathcal{K}_t = \{k_1, \dots, k_{t-1}\}$  denote the historical goal sequence and topic thread, respectively. The CRS maintains a pre-defined set of goals  $\mathbb{G}$ , topics  $\mathbb{K}$  to be predicted, and a large set of items  $\mathbb{V}$  to be recommended during the conversation. In some applications, there also exist the user profiles  $\mathcal{P}_u$  for each user  $u$ , which can be historical item interaction data or certain personal knowledge. Overall, MG-CRS aims to (1) plan the next goal  $g_t \in \mathbb{G}$ , (2) predict the next topic  $k_t \in \mathbb{K}$ , (3) recommend appropriate items  $v_t \in \mathbb{V}$ , and (4) produce a proper response  $c_t$  for the current turn  $t$ . Specifically, the problem of MG-CRS can be decomposed into the following four tasks:

- **Goal Planning.** At each turn  $t$ , given the dialogue context  $C_t$  and the goal history  $\mathcal{G}_t$ , MG-CRS first selects the appropriate goal  $g_t \in \mathbb{G}$  to determine where the conversation goes.
- **Topic Prediction.** The second task is to predict the next conversational topics  $k_t \in \mathbb{K}$  for completing the planned goal  $g_t$ , with respect to the dialogue context  $C_t$ , the historical topic thread  $\mathcal{K}_t$ , and the user profile  $\mathcal{P}_u$  (if exists).
- **Item Recommendation.** If the selected goal  $g_t$  is to make recommendations, then the CRS should recommend an item  $v_t \in \mathbb{V}$ , based on the dialogue context  $C_t$  and the user profile  $\mathcal{P}_u$  (if exists). In general, the recommended item  $v_t$  is supposed to be related to the predicted topics  $k_t$ .
- **Response Generation.** The end task is to generate a proper response  $c_t$  concerning the predicted topics  $k_t$  for completing the selected goal  $g_t$ . When the goal is to make recommendation, the generated response is also expected to provide persuasive reasons for the recommended item  $v_t$ .

### 4 METHOD

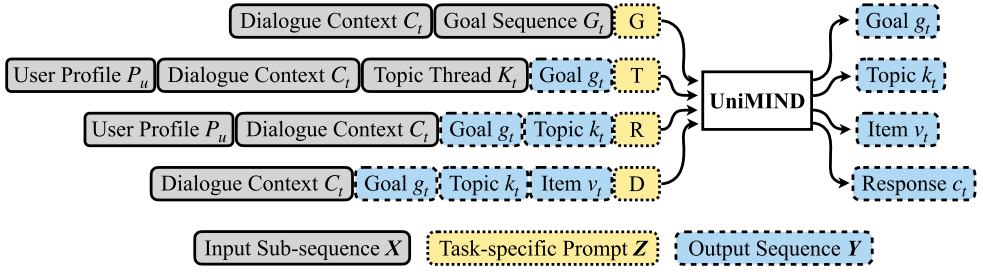
In this section, we first describe the paradigm shift that reformulates each task in MG-CRS into the unified Seq2Seq paradigm, and then introduce the prompt-based learning strategies for the multi-task learning of all tasks. The overview and examples of the input and output sequences for UniMIND are illustrated in Figure 2. Overall, the learning and inference procedure consists of three stages, including multi-task learning, prompt-based learning, and inference.

#### 4.1 Unified Paradigm Shifts

**4.1.1 Goal Planning.** To dynamically adjust the conversation goal in the multi-goal conversational recommendation systems, Liu et al. [33] divide the task of goal planning into two sub-tasks, **goal completion estimation and goal prediction**. Goal completion estimation aims at estimating the probability of goal completion by performing a *binary classification*:

$$\mathbf{y}_{GC} = \text{CLS}(\text{Enc}(C_t, g_{t-1})) \in \{0, 1\}, \quad (1)$$





(a) Overview of the input and output sequences.

“[user] Who is the star of the movie <stolen life>? [goal] QA [system] It is Xun Zhou. [user] She is my goddess. [goal] Chit-chat about Star [system] You have a good taste. She is the most popular actress in the Golden Eagle Award. [user] I like her very much. [goal]”	“ <b>Movie Recommendation</b> ”
“[profile] I often miss childhood [user] I’m looking at the photo album. I miss the old days. [topic] Miss [system] There must be many good memories. [user] I miss my childhood toys. Do you have any movie to recommend? [goal] <u>Movie Recommendation</u> [topic]”	“ <b>Childhood &lt;/k&gt; Toys</b> ”
“[profile] favorite singer: Leehom Wang [user] Do you know who is the singer of <Revolution>? [system] It’s Leehom Wang! [user] Yes! He is one of my favorite singer! [goal] <u>Music Recommendation</u> [topic] <u>Leehom Wang</u> [item]”	“ <b>_20803_ Tears Palm</b> ”
“[user] I like superhero movies. [system] Do you want to see a movie about superhero? [user] Sure. Are there any Marvel hero movies? [goal] <u>Movie Recommendation</u> [topic] <u>Marvel &lt;/k&gt; Hero</u> [item] <u>The Black Widow</u> [system]”	“ <b>&lt;The Black Widow&gt; is a great movie. The action shots are very cool.</b> ”

(b) Examples of the input and output sequences.

Fig. 2. Overview and examples of the input and output sequences for UniMIND.

where  $\text{CLS}(\cdot)$  is a binary classifier, and  $\text{Enc}(\cdot)$  is the encoder that encodes the input information. Goal prediction aims to predict the next goal when the previous goal is completed by performing *multi-task classification*:

$$\mathbf{y}_{\text{GP}} = \text{MT-CLS}(\text{Enc}(C_t, \mathcal{G}_t)) \in \{0, 1\}^{|\mathbb{G}|}, \quad (2)$$

where  $\text{MT-CLS}(\cdot)$  is a multi-task classifier.  $|\mathbb{G}|$  is the total number of all possible goal classes.

We unify these two sub-tasks as one Seq2Seq task, which aims at directly generating the label of the next goal as natural language with a text generation model, i.e.,  $\text{UniMIND}(\cdot)$ :

$$g_t = \text{UniMIND}(C_t, \mathcal{G}_t), \quad (3)$$

where the generated  $g_t$  will be assigned to a specific goal class in  $|\mathbb{G}|$  by lexical matching with the label name. The objective of the Goal Planning task  $G$  is to maximize the sequential log-likelihood:

$$\mathcal{L}_G = \log p(g_t | C_t, \mathcal{G}_t) = \sum_{l=1}^{L_g} \log p_{\theta}(g_t, l | g_{t, <l>}; C_t, \mathcal{G}_t), \quad (4)$$

where  $L_g$  denotes the length of the generated goal label. Such a paradigm shift can alleviate the error propagation in the two-stage classification method.

**4.1.2 Topic Prediction.** According to different applications, the topic can be knowledge entities or specific topic classes, and there are also different corresponding solutions. For instance, Liu et al.

[33] implicitly select the relevant knowledge triples by assigning attention weights to all candidate knowledge and then fusing them into a single vector. Zhou et al. [65] predict the next topics by performing a *text matching* task:

$$y_{KS_i} = \text{CLS}(\text{Enc}(k_i), \text{Enc}(C_t, \mathcal{K}_t, \mathcal{P}_u)), \quad (5)$$

where the predicted topic is ranked with the highest relevance score, i.e.,  $\arg \max_i y_{KS_i}$ . However, these approaches may fail to handle those scenarios with a large set of topics to be predicted. Besides, the fixed number of predicted topics may also affect the cases where there is no topic needed or multiple topics are involved.

We regard the topic prediction as an multi-label classification problem and reformulate it into the Seq2Seq paradigm, where the predicted labels are concatenated into a single sequence as the target sequence to be generated. A special token, e.g., “</k>”, is used to separate each individual label:

$$k_t = \text{UniMIND}(C_t, \mathcal{K}_t, \mathcal{P}_u, g_t), \quad (6)$$

$$\begin{aligned} \mathcal{L}_K &= \log p(k_t | C_t, \mathcal{K}_t, \mathcal{P}_u, g_t) \\ &= \sum_{l=1}^{L_k} \log p_\theta(k_{t,l} | k_{t,<l}; C_t, \mathcal{K}_t, \mathcal{P}_u, g_t), \end{aligned} \quad (7)$$

where  $L_k$  denotes the length of generated sequence of topic labels, and  $\mathcal{L}_K$  is the objective function of the Topic Prediction task  $T$ . The final prediction results can be recovered by splitting the sequence using “</k>” and then match to the topic classes in  $\mathbb{K}$ . By doing so, the MG-CRS will be more scalable and flexible to different situations, even when the number of possible topics is extremely large and variable. In addition, it can be observed from Figure 1 that the next topic is highly related to the dialogue context. And the Seq2Seq paradigm can make full use of the semantic relationships between the dialogue context and the predicted topic labels as well as among different topic labels.

**4.1.3 Item Recommendation.** There are two main-stream solutions for item recommendation in CRS: (i) Following traditional recommendation systems, the recommendation module [64, 65] aims to rank all the items by computing the probability that recommend an item  $v_i$  to the user  $u$ :

$$y_{\text{Rec}_i} = \text{CLS}(e_i, \text{Enc}(C_t, \mathcal{P}_u)), \quad (8)$$

where  $e_i$  is the trainable item embedding, and the recommended item is ranked with the highest recommendation probability, i.e.,  $\arg \max_i y_{\text{Rec}_i}$ . (ii) Some studies [6, 53] perform the item recommendation within an end-to-end CRS, by regarding the item set as an additional vocabulary for generation, i.e., each item index as a word in the expanded vocabulary. For example, if the item index of the target item is 100, then this item will be represented as “\_100\_” in the expanded vocabulary.

Motivated by the second type of approach, we extend the idea of vocabulary expansion with the candidate item set to further take into account the semantic information of the item. Specifically, the target sequence of the item recommendation task is composed of the item expressions in both the original vocabulary and the expanded item vocabulary. For example, if the target item in a sample is “The Witness” with the item index as 100, then the target sequence for this sample will be “\_100\_ The Witness.” By doing so, not only can the model capture the relationship between



context words and candidate items but also exploit the semantic information of items:

$$v_t = \text{UniMIND}(C_t, \mathcal{P}_u, g_t, k_t), \quad (9)$$

$$\begin{aligned} \mathcal{L}_R &= \log p(v_t | C_t, \mathcal{P}_u, g_t, k_t) \\ &= \sum_{l=1}^{L_r} \log p_\theta(v_{t,l} | v_{t,<l}; C_t, \mathcal{P}_u, g_t, k_t), \end{aligned} \quad (10)$$

where  $L_r$  denotes the length of the generated sequence, and  $\mathcal{L}_R$  is the objective function of the Item Recommendation task  $R$ . In the inference, we just conduct decoding for one step and rank the largest probability of the item in the expanded item vocabulary to recommend:

$$v_t = \arg \max_{v^{(i)} \in \mathbb{V}} p_\theta(v^{(i)} | [\text{sos}]; C_t, \mathcal{P}_u, g_t, k_t), \quad (11)$$

where  $[\text{sos}]$  denote the start-of-sentence token.

**4.1.4 Response Generation.** Since response generation is a standard text generation problem in the Seq2Seq paradigm, it does not require any paradigm shift:

$$c_t = \text{UniMIND}(C_t, g_t, k_t, v_t), \quad (12)$$

$$\begin{aligned} \mathcal{L}_D &= \log p(c_t | C_t, g_t, k_t, v_t) \\ &= \sum_{l=1}^{L_d} \log p_\theta(c_{t,l} | c_{t,<l}; C_t, g_t, k_t, v_t), \end{aligned} \quad (13)$$

where  $L_d$  denotes the length of generated responses, and  $\mathcal{L}_D$  is the objective function of the Response Generation task  $D$ .

## 4.2 Prompt-based Learning

After unifying all tasks in MG-CRS into the Seq2Seq paradigm, each input and output sequence pair in each task forms an annotated data instance for training a unified encoder-decoder model. Inspired by prompt-based learning [30, 40], we add a task-specific prompt to the original input sequence to specify which task the model should perform as well as enrich the input with task-specific information.

As discussed in Section 4.1, the combination of required information in the input sequence varies in different tasks. For example, the dialogue context is necessary for all tasks, while the goal history and the topic thread are only required in goal planning task and topic prediction task, respectively. To indicate each information segment with a specific type, all sub-sequences are concatenated with special segment tokens, such as  $[\text{user}]$ ,  $[\text{system}]$ ,  $[\text{goal}]$ ,  $[\text{topic}]$ ,  $[\text{item}]$ ,  $[\text{profile}]$ , and so on. These special tokens are regarded as additional vocabulary with randomly initialized embeddings to be learned. For instance, the original input  $X$  can be represented as follows:

$$\begin{aligned} X_G &= [\text{goal}]g_1[\text{user}]c_1[\text{goal}]g_2[\text{system}] \dots [\text{user}]c_{t-1}; \\ X_T &= [\text{profile}]\mathcal{P}_u[\text{topic}]k_1[\text{user}]c_1[\text{topic}]k_2[\text{system}] \dots [\text{user}]c_{t-1}[\text{goal}]g_t; \\ X_R &= [\text{profile}]\mathcal{P}_u[\text{user}]c_1[\text{system}] \dots [\text{user}]c_{t-1}[\text{goal}]g_t[\text{topic}]k_t; \\ X_D &= [\text{user}]c_1[\text{system}] \dots [\text{user}]c_{t-1}[\text{goal}]g_t[\text{topic}]k_t[\text{item}]v_t. \end{aligned}$$

Specific examples are presented in Figure 2(b). The input for prompt-based learning is composed of the original input sequence  $X$  and a task-specific prompt  $Z$ . We investigate two types of task-specific prompts, namely, *Natural Language Prompt* and *Special Token Prompt*.

**Natural Language Prompt (UniMIND<sub>N</sub>).** Similar to T5 [40], the natural language prompt employs a guidance sentence to indicate each task as follows:

$$\begin{aligned} Z_G &= \text{"Plan the next goal:"}; \\ Z_T &= \text{"Predict the next topic:"}; \\ Z_R &= \text{"Recommend an item:"}; \\ Z_D &= \text{"Generate the response:"}. \end{aligned}$$

**Special Token Prompt (UniMIND<sub>S</sub>).** Another realization of the task-specific prompt is based on the special segment tokens as follows:

$$\begin{aligned} Z_G &= [\text{goal}]; & Z_T &= [\text{topic}]; \\ Z_R &= [\text{item}]; & Z_D &= [\text{system}], \end{aligned}$$

which can be categorized into the family of *continuous prompts* [30]. Originally, the special segment tokens are designed to indicate the beginning of each type of information in the input sequence. Therefore, the learned representations are supposed to preserve the knowledge of the different forms or patterns of each type of information. For example, [goal] and [topic] can separate two groups of classification labels, while [system] can also distinguish the speaker features of the system responses from the user utterances.

### 4.3 Multi-task Training and Inference

The proposed UniMIND model is initialized with weights from a pre-trained LM in an encoder-decoder fashion, e.g., BART [26] or T5 [40]. The overall learning and inference procedure of UniMIND is presented in Algorithm 1, which consists of three stages, including multi-task training, prompt-based learning, and inference.

In the multi-task training stage, the model is trained to perform all tasks in MG-CRS with all training data. Since all tasks in MG-CRS have been reformulated as the Seq2Seq problem, they can be collectively optimized by the same objective as presented in Equations (3), (6), (11), and (13). Given the training sample  $(X, Y, Z)$ , the objective  $\mathcal{L}_\theta$  is to maximize the log-likelihood:

$$\mathcal{L}_\theta = \mathbb{E}_{(X, Y, Z) \sim \mathcal{D}_1} \sum_{l=1}^L \log p_\theta(y_l | y_{<l}; X, Z), \quad (14)$$

where  $\mathcal{D}_1 = \{\mathcal{D}_G, \mathcal{D}_T, \mathcal{D}_R, \mathcal{D}_D\} = \{(X, Y, Z)_i\}_{i=1}^{|\mathcal{D}_1|}$  denotes the whole train-set, and  $L$  denotes the length of the generated sequence.

When applying the trained UniMIND to a specific task, we use the same objective function as in the multi-task training stage to perform task-specific prompt-based learning. In specific, for each task, we fine-tune the trained UniMIND with the corresponding training data:

$$\mathcal{L}_{\theta_t} = \mathbb{E}_{(X, Y, Z) \sim \mathcal{D}_t} \sum_{l=1}^L \log p_\theta(y_l | y_{<l}; X, Z), \quad (15)$$

where  $\theta_t$  and  $\mathcal{D}_t$  denote the tuned parameters of the model and the train-set for each task  $t \in \{G, T, R, D\}$ , respectively.

In the inference stage, given a data sample  $d = (C, \mathcal{G}, \mathcal{K}, \mathcal{P}_u)$ , we perform the four tasks in order using corresponding tuned models. Finally, the predicted output contains the next goal  $g$ , the next topic  $k$ , the recommended item  $v$ , and the generated response  $c$ .

**ALGORITHM 1:** Learning and Inference Procedure

---

**Input:** Train-set  $\mathcal{D}_1 = \{\mathcal{D}_G, \mathcal{D}_T, \mathcal{D}_R, \mathcal{D}_D\} = \{(X, Y, Z)_i\}_{i=1}^{|\mathcal{D}_1|}$ ; Test-set  $\mathcal{D}_2 = \{(C, \mathcal{G}, \mathcal{K}, \mathcal{P}_u)_i\}_{i=1}^{|\mathcal{D}_2|}$ ;  
 Pre-trained parameters  $\theta$ ; Training epoch number  $e_1$ ; Fine-tuning epoch number  $e_2$ ; Loss function  $\mathcal{L}_\theta$ ;

```

1 // Multi-task Training;
2 for  $epoch = 1, 2, \dots, e_1$  do
3   for Batch  $B$  in  $\text{Shuffle}(\mathcal{D}_1)$  do
4     | Use  $B = \{(X, Y, Z)_i\}_{i=1}^{|B|}$  to optimize  $\mathcal{L}_\theta$  in Equation (14);
5   end
6 end
7 // Prompt-based Learning;
8 for Task  $t$  in  $[G, T, R, D]$  do
9    $\theta_t = \theta$ ;
10  for  $epoch = 1, 2, \dots, e_2$  do
11    for Batch  $B$  in  $\text{Shuffle}(\mathcal{D}_t)$  do
12      | Use  $B = \{(X, Y, Z)_i\}_{i=1}^{|B|}$  to optimize  $\mathcal{L}_{\theta_t}$  in Equation (15);
13    end
14  end
15 end
16 // Inference;
17 for Sample  $d_i$  in  $\mathcal{D}_2$  do
18    $d_i = (C, \mathcal{G}, \mathcal{K}, \mathcal{P}_u)_i$ ;
19   Obtain  $g_i, k_i, v_i, c_i$  in order via Equations (3), (6), (11), and (13) using tuned model  $\theta_G, \theta_T, \theta_R, \theta_D$ , respectively;
20 end

```

**Output:** Models  $\theta_G, \theta_T, \theta_R, \theta_D$ ; Prediction results  $\{g, k, v, c\}_i^{|\mathcal{D}_2|}$ .

---

## 5 EXPERIMENTAL SETUPS

### 5.1 Research Questions

The empirical analysis targets the following research questions:

- **RQ1:** How is the performance of the proposed method on the end task of MG-CRS, i.e., Response Generation, compared to existing methods?
- **RQ2:** How is the performance of the proposed method on each sub-task of MG-CRS, including Goal Planning, Topic Prediction, and Item Recommendation, compared to existing methods?
- **RQ3:** How does the three-stage training and inference procedure as well as each sub-task contribute to the overall performance?
- **RQ4:** What is the difference between multi-goal conversational recommender systems and traditional conversational recommender systems?

### 5.2 Datasets

We conduct the experiments on two multi-goal conversational recommendation datasets, namely, DuRecDial<sup>2</sup> [33] and TG-ReDial<sup>3</sup> [65]. The dataset statistics are presented in Table 2. We adopt the same train/dev/test split in the original datasets.

<sup>2</sup><https://github.com/PaddlePaddle/Research/tree/48408392e152ffb2a09ce0a52334453e9c08b082/NLP/ACL2020-DuRecDial>.

<sup>3</sup><https://github.com/RUCAIBox/TG-ReDial>.

Table 2. Statistics of Datasets

Dataset	TG-ReDial	DuRecDial
#Dialogues	10,000	10,190
train/dev/test	8,495/757/748	6,618/946/2,626
#Utterances	129,392	155,477
#Goals	8	21
#Topics/Entities	2,571	701
#items	33,834	11,162

- **DuRecDial** is a goal-oriented knowledge-driven conversational recommendation dataset, which contains dialogues across multiple domains, including movie, music, restaurant, and so on. The conversational goals and knowledge entities at each conversation turn are given. There are 21 types of conversational goals. We treat the knowledge entities as conversational topics, as the example in Figure 1(a). The dataset provides the user profiles with historical item interactions and the knowledge base related to each dialogue.
- **TG-ReDial** is a topic-guided conversational recommendation dataset in the movie domain. We regard the labeled actions in the TG-ReDial dataset as the conversational goals at each turn. There are eight types of conversational goals in TG-ReDial. An example is shown in Figure 1(b). The user profiles with historical item interactions are also given.

### 5.3 Baselines and Evaluation Metrics

*5.3.1 Response Generation.* We first include three general text generation methods for comparisons:

- **Trans.** Reference [51] applies a self-attention-based encoder-decoder framework to generate responses.
- **GPT-2** [39] and **BART** [26] are two pre-trained language models for text generation, which are fine-tuned on the target dataset for response generation.

Since the original settings of TG-ReDial and DuRecDial have some differences, we further compare to different groups of response generation baselines. For TG-ReDial, we consider the following baselines for comparisons:

- **ReDial** [27] and **KBRD** [6] apply the hierarchical RNN and the Transformer architecture, respectively, for response generation in conversational recommendation.
- **Union** [65] and **TopRef.** [52] employ GPT-2 to generate the response conditioned on the predicted topic or the recommended item.

For DuRecDial, we compare to the following baselines:

- **seq2seq** [49] and **PGN** [44] are two classic text generation methods. For seq2seq and Transformer, we also report the knowledge-grounded performance, i.e., seq2seq+kg and Trans.+kg.
- **PostKS** [28] is a knowledge-grounded response generation method.
- **MGCG**<sup>4</sup> [33] adopts multi-type GRUs to encode the dialogue context, the goal sequence, and the topical knowledge and uses another GRU to generate responses.

<sup>4</sup>Following previous studies [1, 56], we only compare our model with MGCG\_G, since MGCG\_R is a retrieval-based dialogue model.

- **GOKC** [1] and **KERS** [56] both assume that the goal sequence is given and focus on the knowledge-grounded response generation problem. Therefore, we also report their performance without the given goal sequence (“-w/o goal”).

Following previous studies [1, 33, 56, 65], we adopt word-level F1, BLEU, **Distinct scores (Dist)**, and **Perplexity (PPL)** as automatic evaluation metrics.

**5.3.2 Goal Planning.** We compare to three text classification baselines:

- **MGCG** [33] employs two CNN-based classifiers to perform two sub-tasks, goal completion estimation and goal prediction.
- **BERT** uses the dialogue context as the input for performing a classification task to predict the next goal.
- **BERT+CNN** combines the obtained representations from MGCG and BERT to predict the next goal.

We adopt Macro-averaged **Precision (P)**, **Recall (R)**, and **F1** as the evaluation metrics.

**5.3.3 Topic Prediction.** We compare to five methods:

- **MGCG** [33] adopts multi-type GRUs to encode the dialogue context, the historical topic sequence, and the user profile for performing a text matching task to rank the candidate topics.
- **Conv/Topic/Profile/-BERT** [65] utilizes BERT to encode historical utterances/topics/user profiles to rank the candidate topics, respectively.
- **Union** [65] combines the obtained representations from Conv/Topic/Profile-BERT.

Following Zhou et al. [65], we adopt Hit@ $k$  as the evaluation metrics for ranking all the possible topics. Besides, we also report Micro-average P, R, and F1 of all the test instances. Since there are some responses that have no topic, if the prediction on a response with no gold labels is also empty, it means the model performs well and we set the P, R, F1 to 1, otherwise 0.

**5.3.4 Item Recommendation.** We compare to five recommendation baselines:

- **GRU4Rec** [31] and **SASRec** [20] apply GRU and Transformer, respectively, to encode user interaction history without using conversation data.
- **TextCNN** [21] and **BERT** [9] adopt a CNN-based model and a BERT-based model, respectively, to encode the dialogue context without using historical user interaction data.
- **Union** [65] combines the learned representations from SASRec and BERT to make recommendations.

Following Zhou et al. [65], we adopt NDCG@ $k$  and MRR@ $k$  as the evaluation metrics.

## 5.4 Implementation Details

Most results of the baselines are reported in previous works. For reproducing some additional results, we implement those baselines with the open-source CRS toolkit, CRSLab<sup>5</sup> [63]. To make a fair comparison with other baselines, we choose BART<sub>base</sub> as the pre-trained Seq2Seq model for UniMIND, which shares a similar number of model parameters with BERT<sub>base</sub> and GPT-2. The pre-trained weights of BART<sub>base</sub> are initialized using the Chinese BART<sup>6</sup> [45]. We use the same hyper-parameter settings for these two datasets. The learning rate and the weight decay rate are set to be

<sup>5</sup><https://github.com/RUCAIBox/CRSLab>.

<sup>6</sup><https://huggingface.co/fnlp/bart-base-chinese>.

Table 3. End-to-end Evaluation of Response Generation on TG-ReDial

Model	Ground-truth		Evaluation Metrics			
	Goal	Topic	F1	BLEU-1/2	Dist-2	PPL
ReDial* [27]	×	×	—	0.177/0.028	0.025	81.61
KBRD* [6]	×	×	—	0.221/0.028	0.025	28.02
Trans.* [51]	×	×	—	0.287/0.071	0.083	32.86
GPT-2* [39]	×	×	—	0.279/0.066	0.094	13.38
Union [65]	○	✓	—	0.280/0.065	0.094	<u>7.22</u>
TopicRef. [52]	○	✓	—	0.294/0.086	—	—
BART [26]	×	×	<u>32.80</u>	0.291/0.070	<u>0.097</u>	7.59
<b>UniMIND<sub>N</sub></b>	✓	✓	35.40 <sup>†</sup>	0.310/0.089 <sup>†</sup>	<b>0.200<sup>†</sup></b>	6.81 <sup>†</sup>
<b>UniMIND<sub>S</sub></b>	✓	✓	<b>35.62<sup>†</sup></b>	<b>0.314/0.090<sup>†</sup></b>	0.198 <sup>†</sup>	<b>5.22<sup>†</sup></b>

\*Results reported from Zhou et al. [65].

The mark ○ denotes that the ground-truth labels are given on both training and testing.

✓ denotes that the ground-truth labels are given for training only, and × for none. <sup>†</sup> indicates statistically significant improvement (t-test with  $p < 0.05$ ) over the best baseline.

5e-5 and 0.01, respectively. The max source sequence length and the max target sequence length are 512 and 100, respectively. For DuRecDial dataset, we extract the knowledge triples concerning the predicted topical entities from the given knowledge base, and regard these knowledge triples as the complete topic context for the response generation task, whose maximum length is set to be 256. We train all the baselines up to 20 epochs. For the proposed method, UniMIND, we conduct multi-task training for 15 epochs and prompt-based learning for 5 epochs.<sup>7</sup>

## 6 EXPERIMENTAL RESULTS

We evaluate the proposed method on both the end-to-end response generation and each sub-task of MG-CRS. Note that **UniMIND<sub>N</sub>** and **UniMIND<sub>S</sub>** denote the **UniMIND** using Natural Language Prompt and Special Token Prompt, respectively, during training and inference.

### 6.1 Evaluation on Response Generation (RQ1)

We conduct both automatic and human evaluation on the response generation task.

**6.1.1 Automatic Evaluation.** Tables 3 and 4 summarize the experimental results on the end task of MG-CRS, i.e., Response Generation, with different conversational recommender systems. Most of the baseline systems simplify the whole MG-CRS problem by assuming the conversational goals are pre-defined (○) at each turn or ignoring some tasks (×), including the current state-of-the-art methods on both datasets, i.e., TopicRef. and KERS. Given pre-defined goals, the systems perform much better than their original counterparts, e.g., GOKC and KERS, indicating the importance of the conversational goals in MG-CRS. Besides, simply adapting pre-trained language models to response generation in MG-CRS fails to achieve a promising performance improvement, e.g., GPT-2 and BART. Finally, UniMIND not only achieves the state-of-the-art performance on the content preservation metrics (F1, BLEU) but also has a promising performance on diversity (Dist) and fluency (PPL) on both datasets.

Overall, the experimental results provide the answer to **RQ1**: *UniMIND substantially and consistently outperforms existing strong baselines, including those baselines with pre-defined goals, with a noticeable margin on Response Generation, which is the end task of MG-CRS.*

<sup>7</sup>The code will be publicly released via <https://github.com/dengyang17/UniMIND>.



Table 4. End-to-end Evaluation of Response Generation on DuRecDial

Model	Ground-truth		Evaluation Metrics			
	Goal	Topic	F1	BLEU-1/2	Dist-2	PPL
seq2seq* [49]	×	×	26.08	0.188/0.102	0.013	22.82
seq2seq+kg** [49]	×	✓	24.52	0.165/0.079	0.013	24.75
PGN* [44]	×	×	33.95	0.243/0.161	0.039	24.28
PostKS* [28]	×	✓	39.87	0.343/0.244	0.056	15.32
MGCG [33]	✓	✓	42.04	0.362/0.252	0.081	14.89
Trans.** [51]	×	×	41.79	0.393/0.288	0.050	9.78
Trans.+kg** [51]	×	✓	44.73	0.419/0.318	0.055	9.40
GPT-2 [39]	×	×	47.01	0.392/0.295	0.165	15.56
GOKC [1]	○	✓	47.28	0.413/0.318	0.084	11.38
- w/o goal	×	✓	45.59	0.401/0.303	0.081	12.45
KERS [56]	○	✓	50.47	0.463/0.362	0.079	8.34
- w/o goal	×	✓	48.95	0.450/0.351	0.082	8.76
BART [26]	×	×	48.41	0.418/0.328	0.049	8.72
UniMIND <sub>N</sub>	✓	✓	52.19 <sup>†</sup>	0.479/0.398 <sup>†</sup>	0.079	6.63 <sup>†</sup>
UniMIND <sub>S</sub>	✓	✓	51.87 <sup>†</sup>	0.477/0.397 <sup>†</sup>	0.086	6.69 <sup>†</sup>

\*Results reported from Bai et al. [1]. \*\*Results reported from Zhang et al. [56].

The mark ○ denotes that the ground-truth labels are given on both training and testing.

✓ denotes that the ground-truth labels are given for training only, and × for none. <sup>†</sup> indicates statistically significant improvement (t-test with  $p < 0.05$ ) over the best baseline.

Table 5. Human Evaluation of Response Generation

Model	Fluency	Informativeness	Appropriateness	Proactivity
GPT-2	1.36	1.39	1.25	1.71
Union	1.31	1.24	1.58	1.80
BART	1.81	1.31	1.40	1.77
UniMind <sub>N</sub>	1.93	1.52	1.70	1.96
UniMind <sub>S</sub>	1.94	1.62	1.72	1.96
Human	1.98	1.90	1.99	1.98

**6.1.2 Human Evaluation.** We conduct human evaluation to evaluate the generated response from four aspects:

- **Fluency:** How fluent and coherent is the generated response?
- **Informativeness:** How rich is the generated response in information?
- **Appropriateness:** Is the generated response appropriate for the current topic?
- **Proactivity:** How well does the generated response proactively complete the current goal?

We randomly sample 100 dialogues from TG-ReDial and compare their responses produced by four methods (GPT-2, Union, BART, and UniMIND).<sup>8</sup> Three annotators are asked to score each generated response with {0: bad, 1: ok, 2: good}. These annotators are all well-educated research assistants. As the results presented in Table 5, UniMIND consistently outperforms these strong baselines from different aspects of human evaluation. It is noteworthy that the scores on *Appropriateness* and *Proactivity* are substantially improved by UniMIND, which demonstrates that

<sup>8</sup>The generated responses of GPT-2 and Union on the TG-ReDial dataset are provided by [https://github.com/RUCAIBox/TG\\_CRS\\_Code](https://github.com/RUCAIBox/TG_CRS_Code).

Table 6. Evaluation of Goal Planning

Model	TG-ReDial			DuRecDial		
	P	R	F1	P	R	F1
MGCG	0.7546	0.8093	0.7794	0.5739	0.6324	0.5787
BERT	0.8742	0.9004	0.8858	0.9174	0.9337	0.9187
BERT+CNN	0.8777	0.9182	0.8971	0.9248	0.9357	0.9229
UniMIND <sub>N</sub>	0.8879 <sup>†</sup>	0.9403 <sup>†</sup>	0.9122 <sup>†</sup>	<b>0.9327<sup>†</sup></b>	<b>0.9466<sup>†</sup></b>	<b>0.9357<sup>†</sup></b>
UniMIND <sub>S</sub>	<b>0.8887<sup>†</sup></b>	<b>0.9425<sup>†</sup></b>	<b>0.9137<sup>†</sup></b>	0.9326 <sup>†</sup>	0.9369 <sup>†</sup>	0.9335 <sup>†</sup>

<sup>†</sup>indicates statistically significant improvement (t-test with  $p < 0.05$ ) over the best baseline.

Table 7. Evaluation of Topic Prediction on TG-ReDial

Model	TG-ReDial					
	Hit@1	Hit@3	Hit@5	P	R	F1
MGCG	0.3635	0.5173	0.6009	0.5211	0.5013	0.5079
Topic-BERT	0.4381	0.5823	0.6246	0.6221	0.5548	0.5772
Profile-BERT	0.0907	0.1597	0.2248	0.3945	0.3740	0.3808
Conv-BERT	0.4348	0.5873	0.6329	0.6264	0.5589	0.5814
Union	0.4420	0.5923	0.6374	0.6301	0.5586	0.5824
UniMIND <sub>N</sub>	0.7317 <sup>†</sup>	0.7319 <sup>†</sup>	0.7319 <sup>†</sup>	0.6876 <sup>†</sup>	0.6915 <sup>†</sup>	0.6889 <sup>†</sup>
UniMIND <sub>S</sub>	<b>0.7351<sup>†</sup></b>	<b>0.7351<sup>†</sup></b>	<b>0.7351<sup>†</sup></b>	<b>0.6912<sup>†</sup></b>	<b>0.6951<sup>†</sup></b>	<b>0.6925<sup>†</sup></b>

We report P@1, R@1, and F1@1 scores for the matching-based baselines, which achieve the highest F1@ $k$  scores. We regard the order of generated topics as the rank of the predicted topic to compute Hit@ $k$  scores for UniMIND, e.g., the first generated topic as the top-ranked topic to compute the Hit@1 score. <sup>†</sup> indicates statistically significant improvement (t-test with  $p < 0.05$ ) over the best baseline.

UniMIND can effectively lead a proactive conversation with appropriate content. However, compared with the reference responses (Human), there is still much room for improvement on *Informativeness* and *Appropriateness* of the responses generated by UniMIND.

The human judgements further support the above answer to the **RQ1**: *The responses generated by the proposed method preserve a higher degree of fluency, informativeness as well as explicitly reflect the target conversational topics and lead a proactive conversation, which contributes to a higher overall quality.*

## 6.2 Evaluation on Each Task (RQ2)

**6.2.1 Evaluation on Goal Planning.** Table 6 presents the experimental results on the Goal Planning task. Among the baselines, BERT and BERT+CNN perform much better than MGCG by fine-tuning a BERT-based model to encode the contextual information from the dialogue history, which attaches great importance in determining where the conversation should go at the next turn. Two variants of UniMIND achieve similar performance and both of them significantly outperform these baselines on two datasets.

**6.2.2 Evaluation on Topic Prediction.** Tables 7 and 8 present the experimental results on the Topic Prediction task. Note that the results on TG-ReDial reported in Reference [65] are based on the assumption that the target recommendation topic is given, so that all the baselines achieve a similar result, due to the strong supervision signal of the target topic. In fact, without the guidance of the target topic, the task of topic prediction becomes more difficult, since the system is required to predict the next topic based on the coherency and relevancy to the dialogue context, the

Table 8. Evaluation of Topic Prediction on DuRecDial

Model	DuRecDial					
	Hit@1	Hit@3	Hit@5	P	R	F1
MGCG	0.6639	0.7510	0.7810	0.6697	0.5419	0.5880
Topic-BERT	0.6337	0.7460	0.7785	0.6957	0.6015	0.6289
Profile-BERT	0.4908	0.6844	0.7698	0.4576	0.4036	0.4181
Conv-BERT	0.7791	0.8226	0.8425	0.8122	0.7065	0.7377
Union	0.7877	0.8462	0.8696	0.8327	0.7270	0.7582
<b>UniMIND<sub>N</sub></b>	<b>0.9024<sup>†</sup></b>	<b>0.9055<sup>†</sup></b>	<b>0.9056<sup>†</sup></b>	<b>0.8981<sup>†</sup></b>	<b>0.8994<sup>†</sup></b>	<b>0.8978<sup>†</sup></b>
<b>UniMIND<sub>S</sub></b>	<b>0.8999<sup>†</sup></b>	<b>0.9021<sup>†</sup></b>	<b>0.9023<sup>†</sup></b>	<b>0.8957<sup>†</sup></b>	<b>0.8964<sup>†</sup></b>	<b>0.8952<sup>†</sup></b>

We report P@1, R@1, and F1@1 scores for the matching-based baselines, which achieve the highest F1@ $k$  scores. We regard the order of generated topics as the rank of the predicted topic to compute Hit@ $k$  scores for UniMIND, e.g., the first generated topic as the top-ranked topic to compute the Hit@1 score. <sup>†</sup> indicates statistically significant improvement (t-test with  $p < 0.05$ ) over the best baseline.

historical topic thread, and the user profile. Therefore, we can observe that Profile-BERT barely works on the topic prediction task. Overall, UniMIND significantly outperforms these strong baselines on both Hit scores and F1 scores, where the Hit@1 scores of UniMIND are even higher than the Hit@5 scores of Union. Since it is difficult to determine the number of topic classes with these matching-based baselines, UniMIND owns remarkable flexibility and scalability to this multi-label classification problem, which is proven by the F1 scores.

**6.2.3 Evaluation on Item Recommendation.** Tables 9 and 10 summarize the experimental results on the Item Recommendation task. Among these baselines, due to the sparsity of the historical user-item interaction data, traditional recommendation methods (e.g., GRU4Rec and SASRec) fall short of handling the item recommendation task in MG-CRS, while text-based recommendation methods (e.g., TextCNN and BERT) show more promising performance. Union further improves the performance by combining the advantages of BERT and SASRec. Without using the historical interaction data, UniMIND achieves competitive performance with Union.

In specific, UniMIND<sub>S</sub> significantly and consistently outperforms Union on both datasets and UniMIND<sub>N</sub> outperforms Union on DuRecDial. Different from the observations on other tasks, UniMIND<sub>S</sub> performs much better than UniMIND<sub>N</sub> on item recommendation, showing that the natural language prompts may not fully utilize the relationships with the expanded item vocabulary through prompt-based learning.

Overall, the experimental results provide the answer to **RQ2**: *UniMIND significantly outperforms existing strong baselines on each sub-task of MG-CRS. The strong performance can not only contribute to the final quality of the generated responses but also provide useful and convenient adaptation for different sub-task applications.*

### 6.3 Discussions on the Choice of Task-specific Prompts

As shown in the results of the four tasks in MG-CRS, we observe that there is not much difference between the performance of UniMIND<sub>N</sub> and UniMIND<sub>S</sub> on goal planning, topic prediction and response generation tasks. However, using special token prompts (i.e., UniMIND<sub>S</sub>) outperforms using natural language prompts (i.e., UniMIND<sub>N</sub>) with a large margin on the item recommendation task. This is mainly for the reason that the knowledge from the pre-trained language models can not be utilized in the expanded item vocabulary, while the sparsity of the expanded item vocabulary makes it difficult for natural language prompt to learn the relationships between two different

Table 9. Evaluation of Item Recommendation on TG-ReDial

Model	TG-ReDial			
	NDCG@10	NDCG@50	MRR@10	MRR@50
GRU4Rec	0.0028	0.0062	0.0014	0.0020
SASRec	0.0092	0.0179	0.0050	0.0068
TextCNN	0.0144	0.0215	0.0119	0.0133
BERT	0.0246	0.0439	0.0182	0.0211
Union	<u>0.0348</u>	<u>0.0527</u>	<u>0.0240</u>	<u>0.0277</u>
UniMIND <sub>N</sub>	0.0306	0.0499	0.0236	0.0277
UniMIND <sub>S</sub>	<b>0.0386<sup>†</sup></b>	<b>0.0638<sup>†</sup></b>	<b>0.0283<sup>†</sup></b>	<b>0.0319</b>

<sup>†</sup> indicates statistically significant improvement (t-test with  $p < 0.05$ ) over the best baseline.

Table 10. Evaluation of Item Recommendation on DuRecDial

Model	DuRecDial			
	NDCG@10	NDCG@50	MRR@10	MRR@50
GRU4Rec	0.2188	0.2734	0.1713	0.1833
SASRec	0.3686	0.4130	0.3071	0.3174
TextCNN	0.5049	0.5344	0.4516	0.4584
BERT	0.5455	0.5719	0.4983	0.5043
Union	<u>0.5568</u>	<u>0.5831</u>	<u>0.5101</u>	<u>0.5159</u>
UniMIND <sub>N</sub>	0.5986 <sup>†</sup>	0.6099 <sup>†</sup>	0.5922 <sup>†</sup>	0.5944 <sup>†</sup>
UniMIND <sub>S</sub>	<b>0.6343<sup>†</sup></b>	<b>0.6471<sup>†</sup></b>	<b>0.6291<sup>†</sup></b>	<b>0.6318<sup>†</sup></b>

<sup>†</sup> indicates statistically significant improvement (t-test with  $p < 0.05$ ) over the best baseline.

vocabularies with limited samples. However, the special token prompt can not only learn from the output sequence but also capture the information between the special tokens and the vocabularies in the input sequence among different tasks, which leads to a better result on item recommendation. Based on these observations, we would recommend adopting the special token prompts for the prompt-based learning with the proposed UniMIND framework.

## 7 DETAILED ANALYSES AND DISCUSSIONS

In this section, we provide a variety of detailed analyses and discussions to look deeper into the proposed method. Note that the following analyses are all conducted with UniMIND<sub>S</sub>, since this variant has generally better performance on both datasets according to Section 6.

### 7.1 Ablation Study (RQ3)

To investigate the effect of the proposed training procedure and each task, the results of the ablation study are presented in Table 11.

**7.1.1 Effect of Training Procedure.** We first evaluate the effectiveness of the multi-task learning and the prompt-based learning strategies. “-w/o MTL” denotes that we train four independent Seq2Seq models for each task. “-w/o PL” denotes that we only train one unified multi-task learning model for all tasks without task-specific prompt-based learning. *The results show that the best performance can only be achieved by combining multi-task learning and prompt-based learning*, which can answer the first part of RQ3.

Table 11. Ablation Study (Results on Response Generation)

Model	TG-ReDial				DuRecDial			
	F1	BLEU-1/2	Dist-2	PPL	F1	BLEU-1/2	Dist-2	PPL
<b>UniMIND<sub>s</sub></b>	35.62	0.314/0.090	0.198	5.22	51.87	0.477/0.397	0.086	6.69
-w/o MTL	34.73	0.305/0.086	0.192	5.98	50.98	0.465/0.381	0.084	7.02
-w/o PL	33.94	0.299/0.084	0.182	6.54	50.67	0.454/0.373	0.084	7.05
<i>Response Generation</i>								
-OracleGen	39.43	0.350/0.104	0.234	4.12	55.07	0.525/0.444	0.145	5.95
-DirectGen	32.80	0.291/0.070	0.135	7.22	48.41	0.418/0.328	0.052	8.44
<i>Goal Planning</i>								
-Oracle	36.78	0.328/0.097	0.194	5.18	53.28	0.498/0.417	0.126	6.35
-BERT+CNN	35.51	0.314/0.090	0.197	5.56	51.67	0.475/0.396	0.087	6.97
-w/o goal	35.13	0.309/0.086	0.156	5.49	50.97	0.467/0.380	0.082	7.22
<i>Topic Prediction</i>								
-Oracle	38.22	0.336/0.100	0.227	4.67	52.68	0.492/0.413	0.133	6.12
-Union	34.32	0.301/0.084	0.178	5.52	51.09	0.469/0.383	0.084	6.92
-w/o topic	35.01	0.307/0.085	0.156	5.76	48.73	0.426/0.340	0.058	8.11
<i>Item Recommendation</i>								
-Oracle	35.72	0.315/0.091	0.199	5.26	52.42	0.488/0.409	0.110	6.33
-Union	35.60	0.314/0.090	0.198	5.24	51.25	0.471/0.394	0.089	6.79
-w/o item	35.73	0.315/0.090	0.202	5.45	51.10	0.469/0.385	0.079	7.18

**7.1.2 Effect of Each Sub-task.** Besides, we also present the performance of “-OracleGen,” which denotes that the input sequence for response generation is composed of the ground-truth goals, topics, and items, while “-DirectGen” denotes the input sequence only contains the dialogue history. The results show that the performance is improved by 15%–25% (DirectGen → OracleGen), which demonstrates the importance of these kinds of information in MG-CRS. In addition, we further investigate the effect of each task, by replacing them with the ground-truth labels (“-Oracle”) or predicted results from other strong baselines (“-BERT+CNN” and “-Union”), or discarding the information (“-w/o”). The results show that Topic Prediction and Goal Planning largely affect the final performance, where precisely predicting topics can bring the most prominent improvement on the final response generation. However, Item Recommendation has the least effect on the final response generation for both datasets.

Overall, for the answer to the second part of **RQ3**, *the experimental results show that all sub-tasks attach more or less importance to the final Response Generation task, among which Topic Prediction and Goal Planning are more influential than Item Recommendation.*

## 7.2 Performance w.r.t. Goal Type (RQ4)

To further analyze the characteristics of MG-CRS against traditional CRS, we present the performance with respect to different goal types in Table 12. To provide a comprehensive scope, we aggregate the results of several dominant goal types for each dataset. For the TG-ReDial dataset that contains responses with multiple goals, the type-wise scores are averaged from all the samples that contain the goal type. For the DuRecDial dataset that contains multi-domain dialogues, the type-wise scores are averaged from all the samples that contain the goal type across all domains. On both datasets, “*Recommendation*” is still the most important goal type with the largest number of samples in MG-CRS. There are several notable observations as follows:

Table 12. Performance w.r.t. Goal Type

Goal Type	%	Goal	Topic	Response Gen.		
		F1	F1	F1	BLEU-1/2	Dist-2
TG-ReDial						
Recommend.	54.4	<b>0.9629</b>	<b>0.8864</b>	37.6	0.337/0.072	0.218
Chit-chat	39.0	0.9428	0.3886	30.5	0.254/0.071	<b>0.327</b>
Rec. Request	31.9	0.8352	0.6926	<b>45.4</b>	<b>0.404/0.167</b>	0.251
DuRecDial						
Recommend.	37.2	0.9235	0.7933	45.9	0.455/0.376	0.101
Chit-chat	15.5	0.8734	0.9787	41.7	0.396/0.309	<b>0.132</b>
QA	16.7	0.9298	0.9278	62.5	0.587/0.505	0.122
Task	11.3	<b>0.9456</b>	<b>0.9963</b>	<b>68.5</b>	<b>0.701/0.637</b>	0.114

- As for Goal Planning, “*Recommendation*” is the easiest goal type to plan for TG-ReDial, since this type always comes after “*Recommendation Request*.” However, the timing for “*Recommendation Request*” would be more difficult to determine.
- As for Topic Prediction, the topic labels in TG-ReDial are the abstractive conversational topics, while those in DuRecDial are the knowledge entities discussed in the current conversation. The difference in annotations leads to the different observations in the two datasets. For “*Chit-chat*” dialogues, the abstractive topics are hard to predict in TG-ReDial, while the discussed entities can be effectively predicted in DuRecDial. Conversely, for “*Recommendation*” dialogues, the abstractive topics are often the genre of the recommended movies in TG-ReDial, while there might be multiple knowledge entities related to the recommended item in DuRecDial.
- As for Response Generation, “*Chit-chat*” dialogues reach the lowest scores on the content preservation metrics (i.e., F1 and BLEU), while achieving the highest scores on the diversity metrics (i.e., Dist) in both datasets. This phenomenon is prevalent in chit-chat dialogue studies. “*Recommendation*” dialogues reach the lowest scores on Dist, due to the similar expressions when making recommendations.

Therefore, we can derive the answer to **RQ4** from this analysis: *The performance w.r.t. different goal types demonstrate the difficulties of MG-CRS, since there are some great differences among different types of dialogues, not just making recommendations in traditional conversational recommender systems.*

### 7.3 Varying the Number of Training Data

To better investigate the training efficiency of the proposed method, we illustrate the testing performance curves of UniMIND compared to several state-of-the-art methods, by varying the number of training data. Figure 3 shows the learning curves of the F1 score on the end task, Response Generation, during the training process on both TG-ReDial and DuRecDial datasets, respectively.

On DuRecDial, it can be observed that those methods using pre-trained language models (UniMIND, BART) can achieve a stronger performance than those methods that are trained from scratch (GOKC<sup>9</sup> and KERS<sup>10</sup>) in the low-resource situation, e.g., <30% training data. This results indicate the effectiveness of the utilization of pre-trained language models on MG-CRS. However, the vanilla BART becomes less competitive with KERS as the growth in the number of training data, which indicates the importance of other sub-tasks in achieving a promising performance on

<sup>9</sup><https://github.com/jq2276/Learning2Copy>.

<sup>10</sup><https://github.com/z562/KERS>.



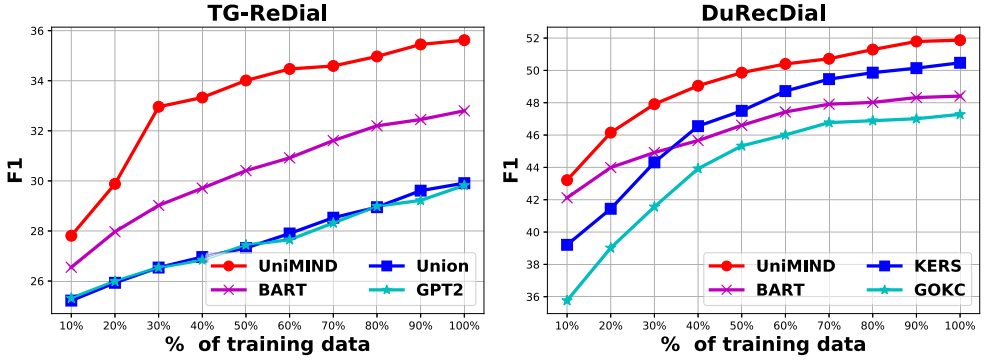


Fig. 3. Performance on Response Generation in terms of different percentage of training data.

MG-CRS. On TG-ReDial, UniMIND further widens the difference with other PLM-based methods (GPT-2 and Union<sup>11</sup>) as the growth in the number of training data. Overall, UniMIND can take advantages of the PLMs for low-resource settings but also make full use of other sub-tasks for training with sufficient data.

#### 7.4 Case Study

To intuitively differentiate UniMIND from other baselines, Figure 4 presents a specific dialogue from TG-ReDial. At the  $(t-1)$ th turn, the ground-truth conversational goal and topics are “*Chit-chat*” and “*Love*,” which means that the system is expected to provide a causal response talking about love. However, Union predicts the topic to be “*Starry Sky*” in advance, due to the strong supervision signals of the target topics of the movie to be recommended. Moreover, without the goal planning task, Union and BART tend to make recommendations frequently, since the majority of the conversational goals in MG-CRS is “*Recommendation*” as shown in Table 12. This is likely to degrade the user experience. UniMIND generates an appropriate response to discuss about love with the user, by making good goal planning and topic prediction. At the  $t$ th turn, the ground-truth conversational goal and topics are “*Movie Recommendation*” and “*Love/Starry Sky*,” which means that the system is expected to recommend a movie about love and starry sky. UniMIND can better capture the multiple topics and provide a more coherent response to the dialogue context.

#### 7.5 Error Analysis and Limitations

Despite the effectiveness of the proposed UniMIND framework for MG-CRS, we would like to better understand the failure modes of UniMIND for further improvement in future studies. After analyzing those cases with low human evaluation scores, we identify the following limitations and discuss the potential solutions:

- **Low Recommendation Success Rate.** All the baselines and UniMIND fail to reach a promising recommendation performance on TG-ReDial as shown in Table 9, due to the sparsity of the user-item interactions. Since the historical interaction data is not utilized in UniMIND, one possible direction is to study how to incorporate this kind of data into the Seq2Seq framework for improving the recommendation performance.
- **Informativeness.** As shown in Table 5, there is still a gap between the generated and the ground-truth response on **Informativeness**. To diversify and enrich the information in dialogue systems, a common practice is to leverage open-domain dialogue corpus to post-train

<sup>11</sup>[https://github.com/RUCAIBox/TG\\_CRS\\_Code](https://github.com/RUCAIBox/TG_CRS_Code).

Reference Conversation		Goal	Topic
Turn 1	System: What are you up to? (你在干啥呢?)	Greeting	N/A
	User: Nothing. Just thinking of a girl whom I have a crush on. (没什么, 在想一个让我心动的女孩。)		
	...		
Turn t-1	User: I've heard about it, a romantic love story that seems like a comedy. (听过的, 看似非常小品的浪漫爱情故事。)	Chit-chat	Love
	System: In fact, having a <b>simple</b> love is the dream of many modern people. (其实拥有一份简简单单的爱情, 是很多现代人的梦想呢。)		
	User: Yes, this dream is as beautiful as the starry sky. (是啊, 这个梦想就像星空一样美丽。)		
Turn t	System: Let me recommend one more movie about love and starry sky. You will definitely enjoy <Solaris>. (我再推荐个关于爱情和星空的电影吧。你会喜欢《飞向太空》的。)	Movie Rec.	Love / Starry Sky
Turn t-1	<b>Union:</b> [topic] Starry Sky [system] Yeah, let me recommend one more movie about starry sky. (是啊, 我再推荐个关于星空的电影吧。)		
	<b>BART:</b> Let me recommend one more movie about romance and love. <_>. What a charming love. (我再推荐一个关于浪漫和爱情的电影吧。《_》, 迷人的爱情。)		
	<b>UniMIND:</b> [goal] Chit-chat [topic] Love [system] Yeah, love is very precious. (是啊, 爱情是非常珍贵的。)		
Turn t	<b>Union:</b> [topic] N/A [system] Let me recommend one more movie about love and love. Have you seen <_>? This is the love. (我再推荐一个关于爱情和爱情的电影吧。《_》看过吗, 就是这样的爱情。)		
	<b>BART:</b> Let me recommend one more movie about starry sky. <_> is about the starry sky. It's very nice. (我再推荐一个关于星空的电影吧。《_》这个电影关于星空的, 很美好。)		
	<b>UniMIND:</b> [goal] Movie Rec. [topic] Love </k> Starry Sky [system] Let me recommend one more movie about love and starry sky. <_> is very suitable for you. Love is <b>simple</b> like this. (我再推荐个关于爱情和星空的电影吧。《_》非常适合你, 爱情就是这样简单。)		

Fig. 4. Case study.

the generation model [61], which can also be easily applied to our unified Seq2Seq framework.

- **Error Propagation.** This is a typical issue of solving multiple tasks in sequential order. Table 13 presents the **Exact Match (EM)** scores between the generated input sequence and the oracle input sequence for each task, which inevitably go down along with the sequential completion of each task. There are some techniques studied to alleviate this issue in cascaded generation methods, such as introducing contrastive objectives [38] or noisy channel models [32].

## 8 CONCLUSIONS

In this work, we propose a novel unified multi-task learning framework for multi-goal conversational recommender systems, namely, UniMIND. Specifically, we unify four tasks in MG-CRS into the same Seq2Seq paradigm and utilize prompt-based learning strategies to endow the model with the capability of multi-task learning. Experimental results on two MG-CRS datasets show that the

Table 13. EM between Generated and Oracle Input Sequences

Dataset	Topic Pred.	Item Rec.	Response Gen.
TG-ReDial	87.20%	80.17%	15.44%
DuRecDial	92.32%	89.16%	84.74%

proposed method achieves state-of-the-art performance on each task with a unified model. Extensive analyses demonstrate the importance of each task and the difficulties of handling different types of dialogues in MG-CRS.

This work is the first attempt toward a unified multi-task learning framework for MG-CRS. There are some limitations and room for further improvement. As discussed in Section 7.5, the error analyses and limitation discussions shed some potential directions for future studies. For example, it can be beneficial to incorporate historical user-item interaction data into the unified framework for making better recommendations or leverage open-domain dialogue corpus to post-train the generation model for generating more informative and diverse responses. It would be also worth investigating approaches to alleviate the error propagation issue in the training procedure. Meanwhile, additional knowledge graphs [36, 64] or extra review data [34, 66] accompanied with items can also be incorporated to further enrich the information for a better performance of MG-CRS. Besides, similar to other prompt-based learning studies, the proposed method can be extended to handle the low-resource scenarios or few-shot learning settings in MG-CRS, which will be more practical in real-world applications. In addition, it is also worth investigating the filter bubble issue [13] and the fairness issue [29] under the setting of MG-CRS.

## REFERENCES

- [1] Jiaqi Bai, Ze Yang, Xinnian Liang, Wei Wang, and Zhoujun Li. 2021. Learning to copy coherent knowledge for response generation. In *Proceedings of AAAI 2021*. 12535–12543.
- [2] Siqi Bao, Huang He, Fan Wang, Hua Wu, and Haifeng Wang. 2020. PLATO: Pre-trained dialogue generation model with discrete latent variable. In *Proceedings of ACL 2020*. 85–96.
- [3] Bill Byrne, Karthik Krishnamoorthi, Saravanan Ganesh, and Mihir Sanjay Kale. 2021. TicketTalk: Toward human-level performance with end-to-end, transaction-based dialog systems. In *Proceedings of ACL/IJCNLP 2021*. 671–680. <https://doi.org/10.18653/v1/2021.acl-long.55>
- [4] Duo Chai, Wei Wu, Qinghong Han, Fei Wu, and Jiwei Li. 2020. Description-based text classification with reinforcement learning. In *Proceedings of ICML 2020*. 1371–1382.
- [5] Hongshen Chen, Xiaorui Liu, Dawei Yin, and Jiliang Tang. 2017. A survey on dialogue systems: Recent advances and new frontiers. *SIGKDD Explor.* 19, 2 (2017), 25–35.
- [6] Qibin Chen, Junyang Lin, Yichang Zhang, Ming Ding, Yukuo Cen, Hongxia Yang, and Jie Tang. 2019. Towards knowledge-based recommender dialog system. In *Proceedings of EMNLP-IJCNLP 2019*. 1803–1813.
- [7] Konstantina Christakopoulou, Alex Beutel, Rui Li, Sagar Jain, and Ed H. Chi. 2018. Q&R: A two-stage approach toward interactive recommendation. In *Proceedings of SIGKDD 2018*. 139–148.
- [8] Yang Deng, Yaliang Li, Fei Sun, Bolin Ding, and Wai Lam. 2021. Unified conversational recommendation policy learning via graph-based reinforcement learning. In *Proceedings of SIGIR 2021*. 1431–1441.
- [9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT 2019*. 4171–4186.
- [10] Cícero Nogueira dos Santos, Xiaofei Ma, Ramesh Nallapati, Zhiheng Huang, and Bing Xiang. 2020. Beyond [CLS] through ranking by generation. In *Proceedings of EMNLP 2020*. 1722–1727.
- [11] Zuohui Fu, Yikun Xian, Shijie Geng, Gerard de Melo, and Yongfeng Zhang. 2021. Popcorn: Human-in-the-loop popularity debiasing in conversational recommender systems. In *Proceedings of CIKM 2021*. 494–503.
- [12] Zuohui Fu, Yikun Xian, Yaxin Zhu, Shuyuan Xu, Zelong Li, Gerard de Melo, and Yongfeng Zhang. 2021. HOOPS: Human-in-the-Loop graph reasoning for conversational recommendation. In *Proceedings of SIGIR 2021*. 2415–2421.
- [13] Chongming Gao, Wenqiang Lei, Jiawei Chen, Shiqi Wang, Xiangnan He, Shijun Li, Biao Li, Yuan Zhang, and Peng Jiang. 2022. CIRS: Bursting filter bubbles by counterfactual interactive recommender system. Retrieved from <https://arxiv.org/abs/2204.01266?context=cs>.
- [14] Chongming Gao, Wenqiang Lei, Xiangnan He, Maarten de Rijke, and Tat-Seng Chua. 2021. Advances and challenges in conversational recommender systems: A survey. Retrieved from <https://arxiv.org/abs/2101.09459>.

- [15] Tianyu Gao, Adam Fisch, and Danqi Chen. 2021. Making pre-trained language models better few-shot learners. In *Proceedings of ACL/IJCNLP 2021*. 3816–3830.
- [16] Shirley Anugrah Hayati, Dongyeop Kang, Qingxiaoyang Zhu, Weiyan Shi, and Zhou Yu. 2020. INSPIRED: Toward sociable recommendation dialog systems. In *Proceedings of EMNLP 2020*. 8142–8152.
- [17] Ehsan Hosseini-Asl, Bryan McCann, Chien-Sheng Wu, Semih Yavuz, and Richard Socher. 2020. A simple language model for task-oriented dialogue. In *Proceedings of NeurIPS 2020*.
- [18] Dietmar Jannach, Ahtsham Manzoor, Wanling Cai, and Li Chen. 2021. A survey on conversational recommender systems. *ACM Comput. Surv.* 54, 5 (2021), 105:1–105:36.
- [19] Dongyeop Kang, Anusha Balakrishnan, Pararth Shah, Paul A. Crook, Y-Lan Boureau, and Jason Weston. 2019. Recommendation as a communication game: Self-Supervised bot-play for goal-oriented dialogue. In *Proceedings of EMNLP-IJCNLP 2019*. 1951–1961.
- [20] Wang-Cheng Kang and Julian J. McAuley. 2018. Self-Attentive sequential recommendation. In *Proceedings of ICDM 2018*. 197–206.
- [21] Yoon Kim. 2014. Convolutional neural networks for sentence classification. In *Proceedings of EMNLP 2014*. 1746–1751.
- [22] Chia-Hsuan Lee, Hao Cheng, and Mari Ostendorf. 2021. Dialogue state tracking with a language model using schema-driven prompting. In *Proceedings of EMNLP 2021*. 4937–4949.
- [23] 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 *Proceedings of WSDM 2020*. 304–312.
- [24] 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 *Proceedings of ACL 2018*. 1437–1447.
- [25] Wenqiang Lei, Gangyi Zhang, Xiangnan He, Yisong Miao, Xiang Wang, Liang Chen, and Tat-Seng Chua. 2020. Interactive path reasoning on graph for conversational recommendation. In *Proceedings of SIGKDD 2020*. 2073–2083.
- [26] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of ACL 2020*. 7871–7880.
- [27] Raymond Li, Samira Ebrahimi Kahou, Hannes Schulz, Vincent Michalski, Laurent Charlin, and Chris Pal. 2018. Towards deep conversational recommendations. In *Proceedings of NeurIPS 2018*. 9748–9758.
- [28] Rongzhong Lian, Min Xie, Fan Wang, Jinhua Peng, and Hua Wu. 2019. Learning to select knowledge for response generation in dialog systems. In *Proceedings of IJCAI 2019*. 5081–5087.
- [29] Allen Lin, Ziwei Zhu, Jianling Wang, and James Caverlee. 2022. Towards fair conversational recommender systems. Retrieved from <https://arxiv.org/abs/2208.03854>.
- [30] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. 2021. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. Retrieved from <https://arxiv.org/abs/2107.13586>.
- [31] Qiang Liu, Shu Wu, Diyi Wang, Zhaokang Li, and Liang Wang. 2016. Context-Aware sequential recommendation. In *Proceedings of ICDM 2016*. 1053–1058.
- [32] Qi Liu, Lei Yu, Laura Rimell, and Phil Blunsom. 2021. Pretraining the noisy channel model for task-oriented dialogue. *Trans. Assoc. Comput. Linguist.* 9 (2021), 657–674.
- [33] Zeming Liu, Haifeng Wang, Zheng-Yu Niu, Hua Wu, Wanxiang Che, and Ting Liu. 2020. Towards conversational recommendation over multi-type dialogs. In *Proceedings of ACL 2020*. 1036–1049.
- [34] Yu Lu, Junwei Bao, Yan Song, Zichen Ma, Shuguang Cui, Youzheng Wu, and Xiaodong He. 2021. RevCore: Review-Augmented conversational recommendation. In *Proceedings of ACL/IJCNLP 2021 (Findings of ACL)*. 1161–1173.
- [35] Yaojie Lu, Qing Liu, Dai Dai, Xinyan Xiao, Hongyu Lin, Xianpei Han, Le Sun, and Hua Wu. 2022. Unified structure generation for universal information extraction. Retrieved from <https://arxiv.org/abs/2203.12277>.
- [36] Wenchang Ma, Ryuichi Takanobu, and Minlie Huang. 2021. CR-Walker: Tree-Structured graph reasoning and dialog acts for conversational recommendation. In *Proceedings of EMNLP 2021*. 1839–1851.
- [37] Rodrigo Nogueira, Zhiying Jiang, Ronak Pradeep, and Jimmy Lin. 2020. Document ranking with a pretrained sequence-to-sequence model. In *Proceedings of ACL: EMNLP 2020*. 708–718.
- [38] Baolin Peng, Chunyuan Li, Jinchao Li, Shahin Shayanmehr, Lars Liden, and Jianfeng Gao. 2021. SOLOIST: Building task bots at scale with transfer learning and machine teaching. *Trans. Assoc. Comput. Linguist.* 9 (2021), 907–824.
- [39] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multi-task learners. *OpenAI Blog* 1, 8 (2019), 9.
- [40] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.* 21 (2020), 140:1–140:67.
- [41] Xuhui Ren, Hongzhi Yin, Tong Chen, Hao Wang, Zi Huang, and Kai Zheng. 2021. Learning to ask appropriate questions in conversational recommendation. In *Proceedings of SIGIR 2021*. 808–817.

- [42] Timo Schick and Hinrich Schütze. 2021. Exploiting cloze-questions for few-shot text classification and natural language inference. In *Proceedings of EACL 2021*. 255–269.
- [43] Timo Schick and Hinrich Schütze. 2021. It's not just size that matters: Small language models are also few-shot learners. In *Proceedings of NAACL-HLT 2021*. 2339–2352.
- [44] Abigail See, Peter J. Liu, and Christopher D. Manning. 2017. Get to the point: Summarization with pointer-generator networks. In *Proceedings of ACL 2017*. 1073–1083.
- [45] Yunfan Shao, Zhichao Geng, Yitao Liu, Junqi Dai, Fei Yang, Li Zhe, Hujun Bao, and Xipeng Qiu. 2021. CPT: A pre-trained unbalanced transformer for both chinese language understanding and generation. Retrieved from <https://arxiv.org/abs/2109.05729>.
- [46] Yixuan Su, Lei Shu, Elman Mansimov, Arshit Gupta, Deng Cai, Yi-An Lai, and Yi Zhang. 2021. Multi-task pre-training for plug-and-play task-oriented dialogue system. Retrieved from <https://arxiv.org/abs/2109.14739>.
- [47] Tianxiang Sun, Xiangyang Liu, Xipeng Qiu, and Xuanjing Huang. 2021. Paradigm shift in natural language processing. Retrieved from <https://arxiv.org/abs/2109.12575>.
- [48] Yueming Sun and Yi Zhang. 2018. Conversational recommender system. In *Proceedings of SIGIR 2018*. 235–244.
- [49] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. 2014. Sequence to sequence learning with neural networks. In *Proceedings of NeurIPS 2014*. 3104–3112.
- [50] Xintao Tian, Yongjing Hao, Pengpeng Zhao, Deqing Wang, Yanchi Liu, and Victor S. Sheng. 2021. Considering interaction sequence of historical items for conversational recommender system. In *Proceedings of DASFAA 2021*. 115–131.
- [51] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Proceedings of NeurIPS 2017*. 5998–6008.
- [52] Hongru Wang, Mingyu Cui, Zimo Zhou, Gabriel Pui Cheong Fung, and Kam-Fai Wong. 2021. TopicRefine: Joint topic prediction and dialogue response generation for multi-turn end-to-end dialogue system. Retrieved from <https://arxiv.org/abs/2109.05187>.
- [53] Lingzhi Wang, Huang Hu, Lei Sha, Can Xu, Kam-Fai Wong, and Daxin Jiang. 2021. Finetuning large-scale pre-trained language models for conversational recommendation with knowledge graph. Retrieved from <https://arxiv.org/abs/2110.07477>.
- [54] Zhihui Xie, Tong Yu, Canzhe Zhao, and Shuai Li. 2021. Comparison-based conversational recommender system with relative bandit feedback. In *Proceedings of SIGIR 2021*.
- [55] Pengcheng Yang, Xu Sun, Wei Li, Shuming Ma, Wei Wu, and Houfeng Wang. 2018. SGM: Sequence generation model for multi-label classification. In *Proceedings of COLING 2018*. 3915–3926.
- [56] Jun Zhang, Yan Yang, Chencai Chen, Liang He, and Zhou Yu. 2021. KERS: A knowledge-enhanced framework for recommendation dialog systems with multiple subgoals. In *Proceedings of Findings of ACL: EMNLP 2021*. 1092–1101.
- [57] Tong Zhang, Yong Liu, Boyang Li, Peixiang Zhong, Chen Zhang, Hao Wang, and Chunyan Miao. 2022. Toward knowledge-enriched conversational recommendation systems. In *Proceedings of ConvAI@ACL 2022*. 212–217.
- [58] Wenxuan Zhang, Yang Deng, Xin Li, Yifei Yuan, Lidong Bing, and Wai Lam. 2021. Aspect sentiment quad prediction as paraphrase generation. In *Proceedings of EMNLP 2021*. 9209–9219. <https://doi.org/10.18653/v1/2021.emnlp-main.726>
- [59] Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam. 2021. Towards generative aspect-based sentiment analysis. In *Proceedings of ACL/IJCNLP 2021*. 504–510. <https://doi.org/10.18653/v1/2021.acl-short.64>
- [60] Yongfeng Zhang, Xu Chen, Qingyao Ai, Liu Yang, and W. Bruce Croft. 2018. Towards conversational search and recommendation: System ask, user respond. In *Proceedings of CIKM 2018*. 177–186.
- [61] Yizhe Zhang, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, and Bill Dolan. 2020. DIALOGPT: Large-Scale generative pre-training for conversational response generation. In *Proceedings of System Demonstrations, ACL 2020*. 270–278.
- [62] Jinfeng Zhou, Bo Wang, Ruifang He, and Yuxian Hou. 2021. CRFR: Improving conversational recommender systems via flexible fragments reasoning on knowledge graphs. In *Proceedings of EMNLP 2021*. 4324–4334.
- [63] Kun Zhou, Xiaolei Wang, Yuanhang Zhou, Chenzhan Shang, Yuan Cheng, Wayne Xin Zhao, Yaliang Li, and Ji-Rong Wen. 2021. CRSLab: An open-source toolkit for building conversational recommender system. Retrieved from <https://arxiv.org/abs/2101.00939>.
- [64] Kun Zhou, Wayne Xin Zhao, Shuqing Bian, Yuanhang Zhou, Ji-Rong Wen, and Jingsong Yu. 2020. Improving conversational recommender systems via knowledge graph-based semantic fusion. In *Proceedings of SIGKDD 2020*. 1006–1014.
- [65] Kun Zhou, Yuanhang Zhou, Wayne Xin Zhao, Xiaoke Wang, and Ji-Rong Wen. 2020. Towards topic-guided conversational recommender system. In *Proceedings of COLING 2020*. 4128–4139.
- [66] Yuanhang Zhou, Kun Zhou, Wayne Xin Zhao, Cheng Wang, Peng Jiang, and He Hu. 2022. C<sup>2</sup>-CRS: Coarse-to-Fine contrastive learning for conversational recommender system. In *Proceedings of WSDM 2022*. 1488–1496.
- [67] Jie Zou, Evangelos Kanoulas, Pengjie Ren, Zhaochun Ren, Aixin Sun, and Cheng Long. 2022. Improving conversational recommender systems via transformer-based sequential modelling. In *Proceedings of SIGIR 2022*. 2319–2324.

Received 22 April 2022; revised 21 September 2022; accepted 26 October 2022