Dual-space Hierarchical Learning for Goal-guided Conversational Recommendation

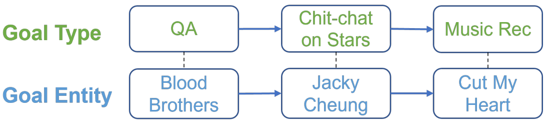
Anonymous ACL submission

Abstract

Proactively and naturally guiding the dialog from the non-recommendation context (e.g., Chit-chat) to the recommendation scenario is crucial for the Conversational Recommender System (CRS). Prior studies mainly focus on planning the next dialog goal (e.g., chat on a movie star) conditioned on the previous dialog. However, we find the dialog goals can be simultaneously observed at different levels, which can be utilized to improve CRS. In this paper, we propose the ***Dual-space Hierarchical Learning*** (**DHL**) to leverage multi-level goal sequences and their hierarchical relationships for conversational recommendation. Specifically, we exploit multi-level goal sequences from both the representation space and the optimization space. In the representation space, we propose the hierarchical representation learning where a cross attention module derives mutually enhanced multi-level goal representations. In the optimization space, we devise the hierarchical weight learning to reweight lower-level goal sequences, and introduce bi-level optimization for stable update. Additionally, we propose a soft labeling strategy to gradually guide the optimization direction. Experiments on two real-world datasets verify the effectiveness of our approach.

1 Introduction

Recent years have witnessed the fast development of the Conversational Recommender System (CRS) (Sun and Zhang, 2018; Konstantina et al., 2018; Zhang et al., 2018; Yu et al., 2019; Lei et al., 2020a,b), which aims to recommend proper items through human-machine natural language interactions. Compared with traditional recommender systems which rely on historical logs, CRS captures dynamic user interests by interacting with users in a more free-form way (i.e. asking questions or recommending items). Therefore, CRS has been widely adopted for various recommendation scenarios, including e-commerce, search engine, and virtual assistant.



Goal Type	Role	Hierarchical Goal Guided Conversation
QA	User	Hi, do you know the leading actor of Blood Brothers?
	Bot	Yeah, I know, Jacky Cheung.
	User	Really, you are awesome!
Chit-chat on Stars	Bot	Besides, Jacky Cheung is the best male singer in the Golden Melody Awards.
	User	He is the God of the singer!
	Bot	Yeah, he is a perfect man.
	User	Sure, anyone who watches his concert says he is the best in both movie and music.
Music Rec	Bot	Do you want to listen to his "Cut My Heart"? It is a classic of lyrical pop songs.
	User	Does it sound good?
	Bot	It sounds great and all of his songs sound good.
	User	Wow, then I'd like to listen to it.

Figure 1: An illustrative example of muti-level goal sequences in a human-machine conversation.

Rather than assuming users always bear in mind what they want, one emerging direction (Liu et al., 2020; Zhou et al., 2020) in CRS is to explore proactively discovering users' interests and naturally leading the conversation from the non-recommendation context to the recommendation scenario. As illustrated in Figure 1, the CRS can intelligently lead the goal type from *QA* to *Chit-chat on Stars* and finally reach the goal of *Music Rec*. To achieve this transition, the study in (Liu et al., 2020) explicitly constructs dialog goal sequences and proposes a CNN based model to plan the next goal. Besides, the study in (Zhou et al., 2020) combines the sequential recommendation model and the pre-trained language model to guide topic transitions by leveraging various conversation signals.

In summary, such methods learn a single representation of previous goals to guide the conversation towards the final recommendation goal.

However, after exploring large-scale recommendation dialogs, we observe the sequence of dialog goals co-exists in multiple levels. Take the dialog in Figure 1 again for example, we can observe the sequence of dialog goals in at least two levels: the goal type level and the goal entity level. To be specific, the goal type of the dialog first transits from *QA* to *Chit-chat on Stars* and then reaches *Music Rec*. The corresponding goal entity sequence can be summarized as *Blood Brothers* $\rightarrow$ *Jacky Cheung* $\rightarrow$ *Cut My Heart*. Such multi-level dialog goal sequences are correlated and contain an intrinsic dependency hierarchy, which can be leveraged to enhance the next goal planning capability and improve the effectiveness of CRS.

To this end, in this paper, we propose the ***D**ual-space **H**ierarchical **L**earning* (**DHL**) to exploit multi-level goal sequences for proactive and natural conversational recommendation. Specifically, DHL models the hierarchical and sequential structure of dialog goals from two spaces. In the representation space, we propose the hierarchical representation learning where a cross attention module captures the hierarchical dependency between multi-level goal sequences. In particular, the cross attention module consists of two symmetric components, *type2entity* and *entity2type*, to derive mutually enhanced representations of multi-level goal sequences. In the optimization space, a hierarchical weight learning module is introduced to reweight goal sequences based on intermediate prediction results for better information use. More specifically, higher accuracy of high-level prediction indicates more useful information for low-level goal planning task, which is assigned with a larger hierarchical weight. To avoid trivial solutions of joint model parameter and hierarchical weight optimization, we introduce bi-level optimization (Luca et al., 2018) to distill weak supervision signals from the training data for the stable update of hierarchical weights. Additionally, we propose a soft labeling strategy to guide the dialog to the final recommendation goal gradually. By assigning a small parameter of the final recommendation goal to each one-hot encoded current goal, the soft label can incorporate the global optimization direction information in the model training phase. In summary, we make the following three major contributions.

- To model hierarchical relationships between multi-level goal sequences in the representation space, we propose the hierarchical representation learning in which a cross attention module derives multi-grained goal representations in a mutual reinforcement way.
- To leverage the hierarchical structure of goal sequences in the optimization space, we develop the hierarchical weight learning to adaptively reweight multi-level goal planning tasks. The bi-level optimization is introduced to stabilize the update of hierarchical weights.
- To guide the conversation to the final recommendation goal, we propose a novel soft labeling strategy to adjust the global optimization direction information.

We have conducted extensive experiments on two real-world conversational recommendation datasets, and the results demonstrate the effectiveness of our approach.

2 Preliminaries

We first introduce some important definitions and then formalize the problem we aim to investigate.

Let $D = \{d_k\}_{k=1}^{N_d}$ denote a set of dialogs, where N_d is the total number of dialogs in the dataset. Each dialog $d_k \in D$ consists of multiple utterances between user and machine.

Definition 1. ***Goal****. A goal g is defined as the topic or knowledge (e.g., an event, a movie star, etc.) the utterances focus on to keep the conversation natural and engaging.*

Depend on the granularity, the goal can be defined at different levels. In this paper, we define dialog goals in three levels, (1) goal type g_p (e.g., QA, Chit-chat and recommendation), (2) goal entity g_e (e.g., movie star, music), and (3) goal attribute g_r (e.g., descriptions about the movie star). We use N_p, N_e and N_r to denote the number of goal type, goal entity, and goal attribute in D. Note that some consecutive utterances may share the same goal.

Definition 2. ***Goal sequence****. Given a dialog d_i, the goal sequence $\mathbf{g}_i = [g_1, g_2, \ldots, g_f]$ is defined as a knowledge path that describes the semantic transition of topics in d_i.*

Note that g_f denotes the final recommendation goal, which is given before the dialog. To ensure consistency, we constrain all goals in a goal sequence at the same level. We denote $\mathbf{g}^p$ as the goal type sequence, $\mathbf{g}^e$ as the goal entity sequence, and $\mathbf{g}^r$ as the goal attribute sequence. Consider a goal sequence $[g_1, g_2, \dots, g_t]$ in a particular level, denote the fixed-length initialization for each goal as $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t]$.

Definition 3. ***Adjacency matrix***. *The adjacency matrix $\mathbf{C}^{pe} \in \mathbb{R}^{N_p \times N_e}$ captures the co-occurrence relationship between the goal type and the goal entity, where $c_{ij}^{pe} \in C^{pe}$ between goal type g_i^p and goal entity g_j^e is defined as*

$$c_{ij}^{pe} = \begin{cases} 1.0 & \textit{if } g_i^p \textit{ and } g_j^e \textit{ co-occurred in } D \\ \epsilon & \textit{otherwise} \end{cases}, \tag{1}$$

where ϵ is a small number. Similarly, the adjacency matrix $\mathbf{C}^{er}$ is constructed by considering the co-occurrence relationship between goal entities and goal attributes.

Problem statement. Given a dialog context $\mathbf{X}$ of previous utterances and the predefined final recommendation goal g_f, we aim to simultaneously plan multi-level goal sequences $\mathbf{g}^p$, $\mathbf{g}^e$, and $\mathbf{g}^r$ to proactively and naturally lead the conversation to reach the final recommendation goal. Based on the item recommendation requirement of CRS, we regard the entity level goal planning as the main task, the type level and attribute level goal planning as auxiliary tasks.

3 Method

3.1 Framework Overview

Figure 2 shows an overview of DHL that includes the following three tasks, (1) learning enhanced multi-level goal representations in a shared latent space, (2) robustly optimizing goal representations by exploiting cross-level supervision signals, and (3) guiding the optimization direction of goal sequences toward the final recommendation goal. For the first task, we construct a two-layer goal sequence hierarchy between goal type and goal entity, and propose the *hierarchical representation learning* to obtain goal representations by capturing cross-level dependencies in a mutual reinforcement way. For the second task, we propose the *hierarchical weight learning* to achieve robust goal representation optimization by reweighting multi-level goal sequences via bi-level optimization. For the third task, we propose a *soft labeling* strategy to gradually enforce the global optimization direction by attaching information of the final recommendation target to the goal sequence.

3.2 Base Model

We adopt LSTM (Sepp and Jürgen, 1997) as the basic building block for proactive goal planning, with the consideration of sequential and hierarchical relationships in multi-level dialog goal sequences.

Sequential representation learning. We first devise multiple LSTMs for goal sequence dependency learning. We derive the hidden representation of the goal sequence by $\mathbf{h}_i = \mathrm{LSTM}(\mathbf{x}_\mathrm{i}, \mathbf{h}_{\mathrm{i}-1})$ where $\mathrm{LSTM}(\cdot, \cdot)$ is the LSTM function and i ranges from 1 to t. Then we can derive the logit of the next goal by

$$\mathbf{l}_{t+1} = \mathrm{MLP}(\mathbf{h}_\mathrm{t}), \tag{2}$$

where $\mathbf{l}_{t+1}$ is a logit vector for all goals, and $\mathrm{MLP}(\cdot)$ is a fully connected neural network.

Hierarchical information exchange. The multi-level goal sequences describe the dialog topic in multiple granularities, which can be exploited to improve the goal prediction. For instance, if the logit score of goal type *Music Rec* is high and it ever co-occurred with a goal entity *Cut My Heart* in historical dialogs, the likelihood of the goal entity *Cut My Heart* should be relatively larger than others. To incorporate such hierarchical knowledge, we integrate high-level information into low-level tasks based on $\mathbf{C}^{pe}$ and $\mathbf{C}^{er}$:

$$\begin{aligned} \mathbf{l}^e &= \mathbf{l}^e + \mathrm{softmax}(\mathbf{l}^\mathrm{p}) \cdot \mathbf{C}^{\mathrm{pe}}, \\ \mathbf{l}^r &= \mathbf{l}^r + \mathrm{softmax}(\mathbf{l}^\mathrm{e}) \cdot \mathbf{C}^{\mathrm{er}}, \end{aligned} \tag{3}$$

where $\mathbf{l}^p \in \mathbb{R}^{N_p}$, $\mathbf{l}^e \in \mathbb{R}^{N_e}$ and $\mathbf{l}^r \in \mathbb{R}^{N_r}$ are the logit vectors of goal type, the goal entity and the goal attribute derived from Eq. (2), and $\cdot$ is the matrix multiplication operation. The next goal type, the next goal entity and the next goal attribute can be calculated as the one with the largest logit in $\mathbf{l}^p$, $\mathbf{l}^e$ and $\mathbf{l}^r$, respectively.

Optimization. We optimize multi-level goal planning tasks jointly. Take goal entity for instance, we optimize the cross-entropy loss $\mathcal{L}_e(\theta)$ between the one-hot encoded next goal entity $\mathbf{g}^e_{next}$ and the goal entity probability vector $\mathrm{softmax}(\mathbf{l}^\mathrm{e})$, where θ are learnable model parameters. The goal type loss $\mathcal{L}_p(\theta)$ and the goal attribute loss $\mathcal{L}_r(\theta)$ can be calculated similarly. The overall learning objective can be written as $\mathcal{L} = \mathcal{L}_p(\theta) + \mathcal{L}_e(\theta) + \mathcal{L}_r(\theta)$.

3.3 Hierarchical Representation Learning

Then we introduce the hierarchical representation learning to further exploit relationships between

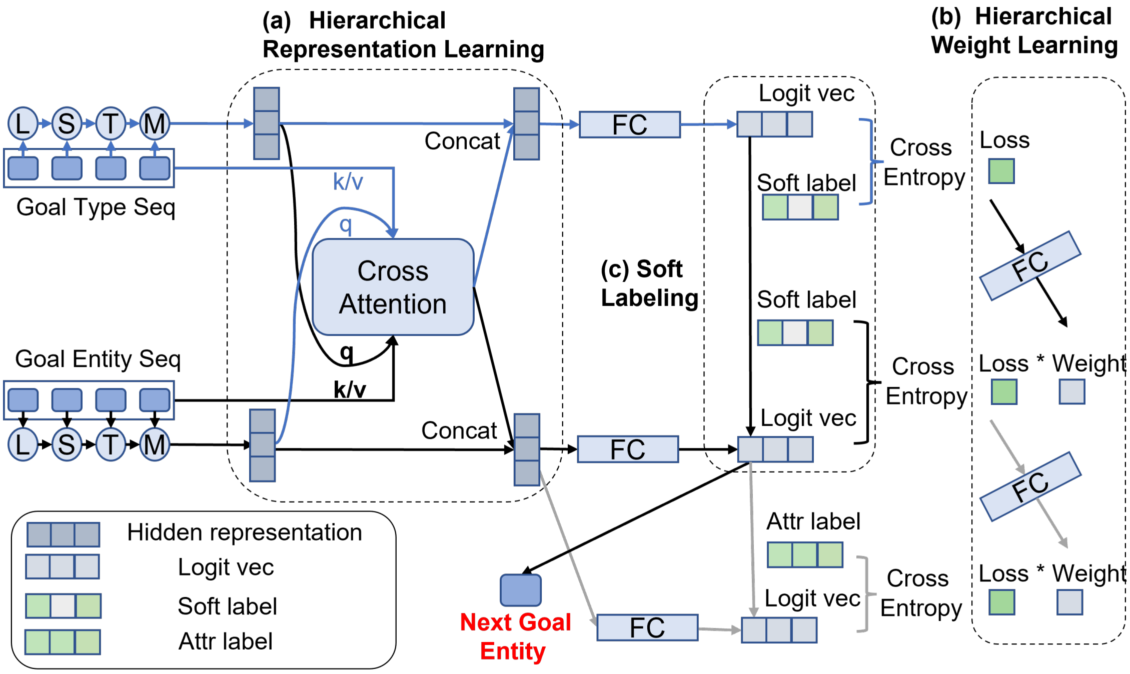


Figure 2: Overview of DHL. Taking the goal type sequence and the goal entity sequence as input, DHL outputs the next goal entity to lead the conversation via dual-space hierarchical learning and soft labeling guidance.

multi-level goal sequences. In particular, we propose the *cross attention* module to learn enhanced goal representations, which including two symmetric components: the type2entity attention component and the entity2type attention component. Specifically, the type2entity attention component adaptively absorbs the goal type knowledge during the goal entity sequence learning, and the entity2type attention component dynamically integrates the goal entity knowledge during the goal type sequence learning. Due to page limit, we use the type2entity attention component for illustration, and the entity2type attention component works in a similar way.

Formally, consider the sequence of goal entity embeddings $[\mathbf{x}_1^e, \mathbf{x}_2^e, \ldots, \mathbf{x}_t^e]$ and the hidden state of the goal type sequence $\mathbf{h}_t^p$ derived from the LSTM encoder. We transform the sequence of goal entity embeddings as the sequence of key embeddings $[\mathbf{x}_1^k, \mathbf{x}_2^k, \ldots, \mathbf{x}_t^k]$ and value embeddings $[\mathbf{x}_1^v, \mathbf{x}_2^v, \ldots, \mathbf{x}_t^v]$ by MLPs and treat $\mathbf{h}_t^p$ as the query embedding. The distilled knowledge representation by the type2entity component is defined by

$$\mathbf{h}^{pe} = \sum_{j=1}^{t} \frac{f(\mathbf{h}_t^p, \mathbf{x}_j^k)}{\sum_{l=1}^{t} f(\mathbf{h}_t^p, \mathbf{x}_l^k)} \mathbf{x}_j^v, \qquad (4)$$

where $f(\mathbf{h}_t^p, \mathbf{x}_j^k)$ represents the exponential kernel $exp(\frac{\mathbf{h}_t^p(\mathbf{x}_j^k)^T}{\sqrt{d}})$ and d denotes the embedding size. Then we simply concatenate $\mathbf{h}^{pe}$ with $\mathbf{h}_t^e$ to integrate the distilled knowledge into the next goal entity prediction.

3.4 Hierarchical Weight Learning

Besides capturing the hierarchical relationship in the representation space, we also propose the hierarchical weight learning to leverage the goal sequence hierarchy in the optimization space.

Hierarchical weight. Intuitively, an accurate prediction of the high-level goal (e.g., goal type) can provide much useful information to guide the optimization of the low-level prediction task (e.g., goal entity prediction), which can be leveraged for model training. Therefore, we introduce the hierarchical weight for the low-level prediction task, where a more accurate high-level prediction indicates a larger hierarchical weight for the low-level prediction task.

However, assigning a scalar weight for each goal sequence is not scalable due to the large number of goal sequences in real-world datasets. Therefore, we approximate the weight for each goal sequence via a neural network. Inspired by (Shu et al., 2019), we adopt an MLP network followed by a sigmoid function to output the hierarchical weight. Specifically, the MLP network takes the loss of goal type as input and outputs the hierarchical weight for the goal entity loss. And the same weight assignment operation can be applied for goal attributes based on goal entities. Denote the MLP parameters as α, the hierarchical weighted loss function can be written as:

$$\mathcal{L}(\theta,\alpha)=\mathcal{L}_p(\theta)+\omega_e^\alpha(\mathcal{L}_p)\mathcal{L}_e(\theta)+\omega_r^\alpha(\mathcal{L}_e)\mathcal{L}_r(\theta) \tag{5}$$

where $\omega_e^\alpha(\mathcal{L}_p)$ and $\omega_r^\alpha(\mathcal{L}_e)$ denote the hierarchical weights for the goal entity loss and the goal attribute loss, respectively. In the following, we denote $\omega_e^\alpha(\mathcal{L}_p)$ as ω_e^α and $\omega_r^\alpha(\mathcal{L}_e)$ as ω_r^α for brevity.

Bi-level optimization for weight learning. Directly optimizing Eq. (5) may lead to a trivial solution where all hierarchical weights reduce to zeros. In this work, we propose to formulate hierarchical weight learning as a bi-level optimization problem (Luca et al., 2018) where the hierarchical weight is decided by an outer level learning task:

$$\begin{aligned}&\min_{\alpha}\mathcal{L}_{out}=\mathcal{L}_p(\theta^*(\alpha))+\mathcal{L}_e(\theta^*(\alpha))+\mathcal{L}_r(\theta^*(\alpha))\\ &\text{s.t. } \theta^*(\alpha)=\arg\min_{\theta}\mathcal{L}_p(\theta)+\omega_e^\alpha L_e(\theta)+\omega_r^\alpha\mathcal{L}_r(\theta)\end{aligned} \tag{6}$$

In this formulation, the inner variable is the model parameters θ and the outer variable is the MLP network parameters α. We build the connection between θ and α in the inner loop via a gradient descent step, and optimize α in the outer loop. Denote $\mathcal{L}_{in}=\mathcal{L}_p(\theta)+\omega_e^\alpha L_e(\theta)+\omega_r^\alpha\mathcal{L}_r(\theta)$. For the the inner level loop, we have:

$$\theta^*(\alpha)\approx\theta-\eta\frac{\partial\mathcal{L}_{in}(\theta,\alpha)}{\partial\theta}. \tag{7}$$

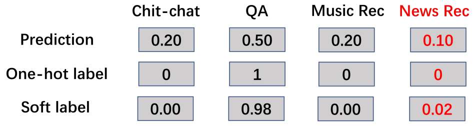
In the outer level, we update α by minimizing $\mathcal{L}_{out}(\theta^*(\alpha))$ via the gradient descent method with the learning rate $\eta^{'}$:

$$\alpha^*\approx\alpha-\eta^{'}\frac{\partial\mathcal{L}_{out}(\theta^*(\alpha))}{\partial\alpha}. \tag{8}$$

In this way, we can leverage the weak supervision signals derived from the outer level task to update hierarchical weights stably. We calculate the weighted loss in Eq. (5) by using the newly updated hierarchical weights ω_e^α and ω_r^α.

3.5 Soft Labeling

As the number of goals is finite, the next goal prediction task can be formulated as a multi-class classification task. Traditional classification tasks optimize the cross-entropy loss between the probability score and the one-hot label. Different from such formulation, one unique characteristic of the goal guided conversational recommender system is to proactively and naturally lead the user to the final recommendation goal when planning the short-term goal in each step.



	Chit-chat	QA	Music Rec	News Rec
Prediction	0.20	0.50	0.20	0.10
One-hot label	0	1	0	0
Soft label	0.00	0.98	0.00	0.02

Figure 3: An illustrative example of the soft labeling strategy. A soft parameter 0.02 is attached to the one-hot goal label, which guides the prediction to approach the final recommendation goal *News Rec*.

In this work, we achieve gradual guidance by introducing the *soft labeling* strategy, which enforces the prediction closer to the final goal as the conversation went on. As shown in Figure 3, we attach the one-hot goal label with a soft parameter in the final goal position. As the goal sequence unrolls, we gradually increase the soft parameter. The soft parameter s_p is calculated as

$$s_p=s_0\min\left(\frac{L}{10},1\right), \tag{9}$$

where L is the goal sequence length and $s_0=0.02$ is a hyper-parameter that controls the strength of soft labeling. From Eq. (9), We can observe that s_p increases as L increases and it remains unchanged when L is larger than 10. Denote the soft label of goal entity as $\mathbf{g}^e_{soft}$, the goal entity loss can be computed by $\mathcal{L}_e(\theta)=\mathrm{CE}(\mathbf{g}^{\mathrm{e}}_{\mathrm{soft}},\mathrm{softmax}(\mathbf{l}^{\mathrm{e}}))$ where $\mathrm{CE}(\cdot)$ is the cross entropy loss function. The loss function for other levels of goals can be computed similarly. By leveraging the soft labeling strategy, the goal sequence transits to the final recommendation goal gradually.

4 Experiments

4.1 Datasets

We perform experiments on two datasets: DuRecDial(Liu et al., 2020) and TG-ReDial(Zhou et al., 2020). Note that there are type level, entity level and attribute level goals in DuRecDial while there are no explicit goal attributes in TG-ReDial. Therefore, we only consider goal type prediction as the auxiliary task in the experiments on TG-ReDial. See Appendix 7.1 for the details of datasets.

4.2 Baselines

We have chosen five baselines for comparison:

Table 1: Overall Results on DuRecDial Dataset.

Metrics	Next	LSTM	CNN	TG	MGNN	DHL
Acc(%)	68.75	78.24	81.85	79.48	67.31	**83.51**
Rec(%)	59.34	46.56	55.84	51.21	25.95	**62.62**
Prec(%)	66.34	54.33	61.82	61.79	26.81	**72.52**
F1(%)	61.37	48.29	56.50	54.31	23.91	**65.25**
LS(%)	11.49	13.38	12.87	14.22	**14.46**	13.52

Table 2: Overall Results on TGReDial Dataset.

Metrics	Next	LSTM	CNN	TG	MGNN	DHL
Acc(%)	33.59	37.98	38.78	37.06	36.98	**39.99**
Rec(%)	15.17	18.19	17.60	14.10	12.56	**20.78**
Prec(%)	29.18	24.26	15.63	8.98	7.76	**46.32**
F1(%)	12.76	17.68	14.45	10.23	8.74	**20.43**
LS(%)	0.00	9.26	20.16	19.29	2.79	**24.82**

Table 3: Ablation Study on DuRecDial Dataset.

Metrics	base	w/o att	w/o weight	w/o soft	DHL
Acc(%)	82.39	82.54	82.95	82.80	**83.51**
Rec(%)	58.72	60.82	61.34	60.27	**62.62**
Prec(%)	67.34	69.84	69.20	67.81	**72.52**
F1(%)	60.40	62.98	62.90	61.73	**65.25**
LS(%)	12.84	13.46	13.24	12.95	**13.52**

- **Next** simply chooses the last goal entity as the prediction of the next goal entity;
- **LSTM** leverages an LSTM to model the goal entity sequence dependency;
- **CNN** (Liu et al., 2020) proposes a CNN for the goal entity sequence learning;
- **TG** (Zhou et al., 2020) uses the dot product between the learned representation and every goal entity embedding for ranking;
- **MGNN** (Xu et al., 2020a) is the state-of-the-art graph neural network based model for dialog generation policy learning. Compared with the original paper, we construct graphs for each level based on the goal co-occurrence relationship and use these graphs to enhance the goal representation learning.

4.3 Implementation Details

For all methods, we set the embedding size of goal type, goal entity and goal attribute as 256 and the batchsize as 128. We use the dev dataset to tune other hyper-parameters such as the learning rate. For the evaluation metrics, we adopt the accuracy/recall/precision/F1 scores to evaluate the performance of the goal entity prediction on all methods. Besides, we use the dialog-leading success rate metric(LS) to measure how well a model can lead the dialog to approach the final recommendation goal. More specifically, LS is the ratio of the number of achieved final recommendation goals over the number of all goal predictions. Please refer to Appendix 7.2 for more training details.

4.4 Overall Results

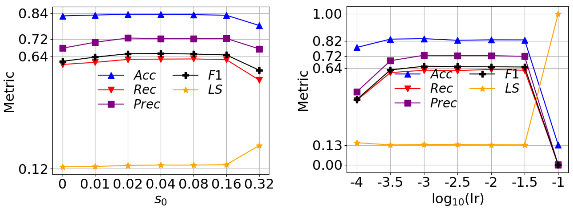
Table 1 and Table 2 summarize the results for all methods where the best results are in bold and the second-best results are marked by underlines. Firstly, we can observe that DHL outperforms all comparison methods in all metrics except for the LS metric on DuRecDial . TG and MGNN perform slightly better in terms of the LS metric but get relatively low performance in other metrics. We increase s_0 in Eq. (9) from 0.02 to 0.20 in DHL and get better performance than TG and MGNN in terms of all metrics on DuRecDial (Accuracy: 82.65%; Recall: 61.56%; Precision: 65.45%; LS: 17.00%). Furthermore, we notice that the Next method gets much better performance in DuRecDial compared with the experimental results in TG-ReDial. The reason is that the goal entity sequence in DuRecDial contains more consecutive goal entities than that in TG-ReDial, and thus setting the last goal entity as the prediction of the next goal entity prediction performs better in DuRecDial . Last but not least, CNN serves as a good baseline in both datasets. One possible explanation is that the next goal entity prediction highly relies on the last several goal entities and CNN can model this relation well via a small sliding window. We also analyze the computational efficiency of all comparison methods in Appendix 7.3.

4.5 Ablation Study

We have also conducted additional experiments on both DuRecDial and TG-ReDial with ablation consideration. More specifically, we remove the hierarchical representation learning, the hierarchical

Table 4: Ablation Study on TGReDial Dataset.

Metrics	base	w/o att	w/o weight	w/o soft	DHL
Acc(%)	38.92	39.19	38.88	39.09	**39.99**
Rec(%)	18.56	20.00	20.17	20.72	**20.78**
Prec(%)	29.78	30.55	24.10	33.96	**46.32**
F1(%)	17.26	19.31	18.80	20.19	**20.43**
LS(%)	11.15	16.12	20.82	22.35	**24.82**



(a) Soft parameter (b) Learning rate

Figure 4: Parameter sensitivity of DHL.

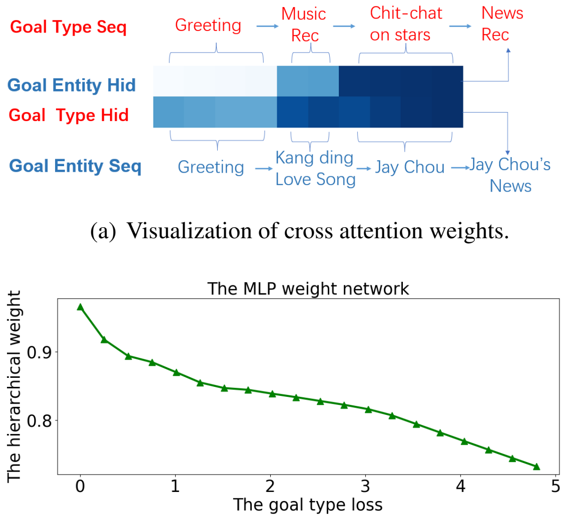
weight learning, and the soft labeling in DHL, and denote these variants as *w/o att*, *w/o weight*, *w/o soft* respectively. Besides, we denote DHL without all the three designs as *base*. The results are reported in Table 3 and Table 4.

Firstly, we can observe that DHL outperforms *base* by a large margin (1.12% accuracy, 3.90% recall, 5.18% precision, 4.85% F1 and 0.68% LS) in DuRecDial and (1.07% accuracy, 2.22% recall, 16.54% precision, 3.17% F1 and 13.67% LS) in TG-ReDial. This validates the effectiveness of the proposed DHL. Secondly, we can observe that the hierarchical representation learning and the hierarchical weight learning both play important roles in DHL. In the DuRecDial experiments, the precision score drops 2.68% without the hierarchical representation learning and drops 3.32% without the hierarchical weight learning. A similar trend is observed in TG-ReDial experiments: 15.77% and 22.22% precision score drop after removing the hierarchical representation learning and the hierarchical weight learning respectively. Last not but least, as shown in Table 3 and Table 4, DHL gets lower LS after removing the soft labeling strategy (from 13.52% to 12.95% in the DuRecDial experiments and from 24.82% to 22.35% in the TG-ReDial experiments). This can be explained as the lack of the gradual guidance to the final recommendation goal. We further notice that the other four metrics degrade slightly (accuracy from 83.51% to 82.80% in the DuRecDial experiments and from 39.99% to 39.09% in the TG-ReDial experiments) and the reason may be the soft labeling strategy also acts as an implicit regularization which improves the model training(Christian et al., 2016).

4.6 Parameter Sensitivity

We further analyze the parameter sensitivity of DHL in DuRecDial .

Soft parameter. We first examine the soft parameter and results are shown in Figure 4(a). It can be observed that as s_0 increases from 0 to 0.32, the LS score increases from 12.95% to 22.70%. This verifies that the soft labeling strategy can lead the conversation to the final recommendation goal. Other metrics including the accuracy, the recall, the precision and the F1 score, increase at first and then decrease as s_0 becomes larger than 0.02. The increase may be explained by that the soft labeling serves as an implicit regularization to improve the model training (Christian et al., 2016) while the reason for the decrease is that the soft label gradually becomes a noisy label when s_0 becomes large. We choose $s_0 = 0.02$ in our experiments.



(a) Visualization of cross attention weights.

(b) Visualization of hierarchical weights.

Figure 5: Case study.

Learning rate. Besides, we examine the sensitivity of the learning rate. As shown in Figure 4(b), given a small learning rate of 1e−4, the model gets poor performance. As the learning rate increases from 1e−3.5 to 1e−1.5, the model performs well and is not sensitive to the learning rate during this period. It can be observed that the 1e−3 learning rate performs best for most metrics. When the learning rate becomes large as 1e−1, the model collapses and all goal entity predictions become the final recommendation goal.

4.7 Case Study

To qualitatively analyze the effectiveness of DHL, we visualize the cross attention weights and the hierarchical weights in DuRecDial .

Cross attention weights. We extract the goal type sequence and the goal entity sequence from a dialog. Please refer to Appendix 7.4 for dialog details. As shown in Figure 5(a), we list the goal type sequence and the goal entity sequence, and visualize the cross attention weights between the hidden state and the goal embeddings. We can observe the cross attention weight, between the goal type hidden state and the goal entity embeddings, increases as the goal sequence unrolls. This makes sense, since the last goal entities have more influence on the next goal entity prediction. To be specific, highlighted by the goal type hidden state, the goal entity *Jay Chou* contributes a lot to the next goal entity prediction as *Jay Chou's news*. Similarly, stressed by the goal entity hidden state, the goal type *Chit-chat on Stars* contributes much to the next goal type prediction as *News Rec*. The above observations further validate the effectiveness of the hierarchical representation learning.

Hierarchical weights. As shown in Figure 5(b), we visualize the MLP weight network and the hierarchical weight for the goal entity task decreases as the goal type loss increases. This makes sense: given a small goal type loss, the goal type prediction is accurate and thus provides much useful information for the goal entity prediction. Therefore the goal entity prediction deserves a large task weight for optimization.

Furthermore, in the example of Figure 5(a), we find the goal type loss is 0.012 and this small loss corresponds to a large task weight 0.96 for the goal entity prediction in Figure 5(b). Following Eq. (3), we compute the goal type prediction's contribution score to the goal entity prediction as

$$\mathbf{l}_{p2e} = \text{softmax}(\mathbf{l}_{\text{p}}) \cdot \mathbf{C}^{\text{pe}}. \tag{10}$$

The ground-truth goal entity position of $\text{softmax}(\mathbf{l}_{\text{p2e}})$ is 1.17%, much larger than the average score $\frac{1}{1385} = 0.07\%$ (1385 is the class number of the goal entity in DuRecDial.) The above results further validate the effectiveness of hierarchical weight learning.

5 Related work

This work is related to conversational recommender system and hierarchical structure modeling.

Conversational recommender system. The research on the conversational recommender system has two research lines. One is from the recommender system and another is from the dialogue system. The research works of the first one (Sun and Zhang, 2018; Konstantina et al., 2018; Zhang et al., 2018; Yu et al., 2019; Lei et al., 2020a,b) aim to infer user's interest by historical interactions and the system generally consider two actions: ask questions or recommend items. The second line from the dialogue system aims to enforce natural semantic transitions in multi-turn human-machine natural language interactions. For example, (Xu et al., 2020a) proposed to leverage the information of global graph structure to enhance goal embedding learning, and (Liu et al., 2020; Zhou et al., 2020) incorporated topic threads to enforce natural semantic transitions towards recommendation.

Hierarchical structure modeling. Hierarchical structure modeling has attracted lots of research attention in many fields including recommender system (Xu et al., 2020b; Qi et al., 2021) and natural language processing (Su et al., 2021; Hu et al., 2021; Wu et al., 2021; Chen et al., 2021; Wang et al., 2021). To name a few, (Xu et al., 2020b; Qi et al., 2021) model the user's interest hierarchy from a higher level to a lower level. To train the matching model in an "easy-to-difficult" scheme, (Su et al., 2021) proposed a hierarchical curriculum learning framework that consists of the corpus-level curriculum and the instance-level curriculum. (Hu et al., 2021) introduced a recursive transformer to model multiple levels of granularity (e.g., words, phrases, and sentences) and (Wu et al., 2021) proposed a Hi-Transformer which models documents in a hierarchical way.

In this paper, we exploit the hierarchical structure of goal sequences from dual spaces, including representation space and optimization space.

6 Conclusion

In this paper, we propose DHL to enhance the proactive goal planning in CRS by exploiting the hierarchical structure of multi-level goal sequences. Specifically, we propose the hierarchical representation learning in the representation space and the hierarchical weight learning in the optimization space to model hierarchical goal sequences. In addition, we develop a novel soft labeling strategy, which can gradually guide the conversation to the final recommendation goal. We conduct extensive experiments on the DuRecDial dataset and the TG-ReDial dataset and the results demonstrate the effectiveness of DHL.

References

Haibin Chen, Qianli Ma, Zhenxi Lin, and Jiangyue Yan. 2021. Hierarchy-aware label semantics matching network for hierarchical text classification. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics.

Szegedy Christian, Vanhoucke Vincent, Ioffe Sergey, Shlens Jon, and Wojna Zbigniew. 2016. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition.

Xiang Hu, Haitao Mi, Zujie Wen, Yafang Wang, Yi Su, Jing Zheng, and Gerard de Melo. 2021. R2D2: Recursive transformer based on differentiable tree for interpretable hierarchical language modeling. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics.

Diederik P Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In 3rd International Conference on Learning Representations.

Christakopoulou Konstantina, Beutel Alex, Li Rui, Jain Sagar, and Chi Ed H. 2018. Q&r: A two-stage approach toward interactive recommendation. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.

Wenqiang Lei, Xiangnan He, Yisong Miao, Qingyun Wu, Richang Hong, Min-Yen Kan, and Tat-Seng Chua. 2020a. Estimation-action-reflection: Towards deep interaction between conversational and recommender systems. In Proceedings of the 13th International Conference on Web Search and Data Mining.

Wenqiang Lei, Gangyi Zhang, Xiangnan He, Yisong Miao, Xiang Wang, Liang Chen, and Tat-Seng Chua. 2020b. Interactive path reasoning on graph for conversational recommendation. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.

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 the 58th Annual Meeting of the Association for Computational Linguistics.

Franceschi Luca, Frasconi Paolo, Salzo Saverio, Grazzi Riccardo, and Pontil Massimiliano. 2018. Bilevel programming for hyperparameter optimization and meta-learning. In International Conference on Machine Learning.

Tao Qi, Fangzhao Wu, Chuhan Wu, Peiru Yang, Yang Yu, Xing Xie, and Yongfeng Huang. 2021. HieRec: Hierarchical user interest modeling for personalized news recommendation. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics.

Hochreiter Sepp and Schmidhuber Jürgen. 1997. Long short-term memory. Neural computation.

Jun Shu, Qi Xie, Lixuan Yi, Qian Zhao, Sanping Zhou, Zongben Xu, and Deyu Meng. 2019. Meta-weight-net: Learning an explicit mapping for sample weighting. In NeurIPS.

Yixuan Su, Deng Cai, Qingyu Zhou, Zibo Lin, Simon Baker, Yunbo Cao, Shuming Shi, Nigel Collier, and Yan Wang. 2021. Dialogue response selection with hierarchical curriculum learning. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics.

Yueming Sun and Yi Zhang. 2018. Conversational recommender system. In The 41st international acm sigir conference on research & development in information retrieval.

Xuepeng Wang, Li Zhao, Bing Liu, Tao Chen, Feng Zhang, and Di Wang. 2021. Concept-based label embedding via dynamic routing for hierarchical text classification. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics.

Chuhan Wu, Fangzhao Wu, Tao Qi, and Yongfeng Huang. 2021. Hi-transformer: Hierarchical interactive transformer for efficient and effective long document modeling. arXiv preprint arXiv:2106.01040.

Jun Xu, Haifeng Wang, Zheng-Yu Niu, Hua Wu, Wanxiang Che, and Ting Liu. 2020a. Conversational graph grounded policy learning for open-domain conversation generation. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics.

Weinan Xu, Hengxu He, Minshi Tan, Yunming Li, Jun Lang, and Dongbai Guo. 2020b. Deep interest with hierarchical attention network for click-through rate prediction. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval.

Tong Yu, Yilin Shen, and Hongxia Jin. 2019. A visual dialog augmented interactive recommender system. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.

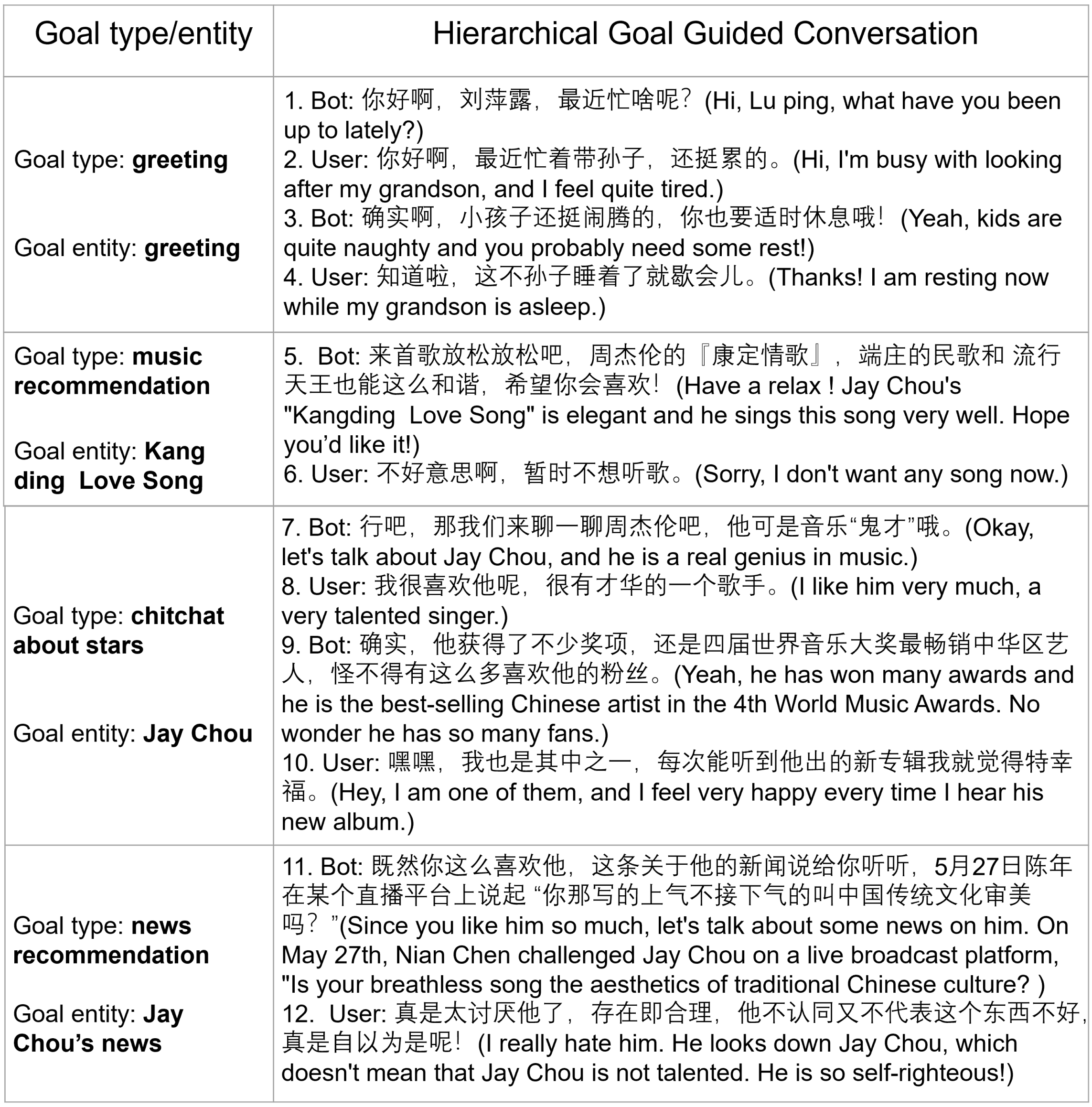
Yongfeng Zhang, Xu Chen, Qingyao Ai, Liu Yang, and W Bruce Croft. 2018. Towards conversational search and recommendation: System ask, user respond. In Proceedings of the 27th acm international conference on information and knowledge management.

Kun Zhou, Yuanhang Zhou, Wayne Xin Zhao, Xiaoke Wang, and Ji-Rong Wen. 2020. Towards topic-guided conversational recommender system. In Proceedings of the 28th International Conference on Computational Linguistics.

7 Appendix

7.1 Dataset

DuRecDial contains 10.2k dialogues, 15.5k utterances and 1362 seekers and TG-ReDial contains 10.0k dialogues, 129.4k utterances and 1482 seekers. Following the original paper, we randomly



Goal type/entity	Hierarchical Goal Guided Conversation
Goal type: **greeting** Goal entity: **greeting**	1. Bot: 你好啊，刘萍露，最近忙啥呢？(Hi, Lu ping, what have you been up to lately?) 2. User: 你好啊，最近忙着带孙子，还挺累的。(Hi, I'm busy with looking after my grandson, and I feel quite tired.) 3. Bot: 确实啊，小孩子还挺闹腾的，你也要适时休息哦！(Yeah, kids are quite naughty and you probably need some rest!) 4. User: 知道啦，这不孙子睡着了就歇会儿。(Thanks! I am resting now while my grandson is asleep.)
Goal type: **music recommendation** Goal entity: **Kang ding Love Song**	5. Bot: 来首歌放松放松吧，周杰伦的『康定情歌』，端庄的民歌和 流行天王也能这么和谐，希望你会喜欢！(Have a relax ! Jay Chou's "Kangding Love Song" is elegant and he sings this song very well. Hope you'd like it!) 6. User: 不好意思啊，暂时不想听歌。(Sorry, I don't want any song now.)
Goal type: **chitchat about stars** Goal entity: **Jay Chou**	7. Bot: 行吧，那我们来聊一聊周杰伦吧，他可是音乐"鬼才"哦。(Okay, let's talk about Jay Chou, and he is a real genius in music.) 8. User: 我很喜欢他呢，很有才华的一个歌手。(I like him very much, a very talented singer.) 9. Bot: 确实，他获得了不少奖项，还是四届世界音乐大奖最畅销中华区艺人，怪不得有这么多喜欢他的粉丝。(Yeah, he has won many awards and he is the best-selling Chinese artist in the 4th World Music Awards. No wonder he has so many fans.) 10. User: 嘿嘿，我也是其中之一，每次能听到他出的新专辑我就觉得特幸福。(Hey, I am one of them, and I feel very happy every time I hear his new album.)
Goal type: **news recommendation** Goal entity: **Jay Chou's news**	11. Bot: 既然你这么喜欢他，这条关于他的新闻说给你听听，5月27日陈年在某个直播平台上说起 "你那写的上气不接下气的叫中国传统文化审美吗？"(Since you like him so much, let's talk about some news on him. On May 27th, Nian Chen challenged Jay Chou on a live broadcast platform, "Is your breathless song the aesthetics of traditional Chinese culture?) 12. User: 真是太讨厌他了，存在即合理，他不认同又不代表这个东西不好，真是自以为是呢！(I really hate him. He looks down Jay Chou, which doesn't mean that Jay Chou is not talented. He is so self-righteous!)

Figure 6: The sampled dialog in the DuRecDial dataset for case study.

sampling 65%/10%/25% data in DuRecDial at the level of seekers to form the train/dev/test datasets. As for TG-ReDial, we extract the topic as the goal entity, and reduce the number of goal entity by clustering and manual processing. We treat the action in TG-ReDial as the goal type and follow the train/dev/test splitting in the original paper.

7.2 Training Details

The hidden size of LSTM is set to 256. We adopt the cosine learning rate decay schedule for a total of 30 epochs for all comparison methods. We use the dev set to tune learning rate from [1e−4, 1e−3.5, 1e−3, 1e−2.5, 1e−2, 1e−1.5, 1e−1]. The Adam optimizer (Kingma and Ba, 2015) is used in training the model parameters. For the hierarchical weight learning, we set the embedding size of the MLP as 100 and optimize the MLP via the Adam optimizer with a 1e−5 learning rate. We tune the soft parameter from [0, 0.01, 0.02, 0.04, 0.08, 0.16, 0.32] via the performance of the dev set. All experiments are performed on a single Tesla P40.

7.3 Computational Efficiency

As for the computational efficiency, on a single Tesla P40, MHFL takes 28.87s to finish the inference on the whole DuRecDial test set (LSTM: 27.85s; CNN: 24.86s; TG: 24.70s; MGCG: 29.06s). Note that the inference speed of all methods are similar and MHFL is a little slower due to the introduction of the cross attention module.

7.4 Dialog Details

The dialog is shown in Figure 6.