Thoughts to Target: Enhance Planning for Target-driven Conversation

Anonymous ACL submission

Abstract

In conversational AI, large-scale models excel in various tasks but struggle with target-driven conversation planning. Current methods, such as chain-of-thought reasoning and tree-search policy learning techniques, either neglect plan rationality or require extensive human simulation procedures. Addressing this, we propose a novel two-stage framework, named **EnPL**, to improve the LLMs' capability in planning conversations towards designated targets, including (1) *distilling natural language plans* from target-driven conversation corpus and (2) *generating new plans* with demonstration-guided in-context learning. Specifically, we first propose a filter approach to distill a high-quality plan dataset, ConvPlan[1]. With the aid of corresponding conversational data and support from relevant knowledge bases, we validate the quality and rationality of these plans. Then, these plans are leveraged to help guiding LLMs to further plan for new targets. Empirical results demonstrate that our method significantly improves the planning ability of LLMs, especially in target-driven conversations. Furthermore, EnPL is demonstrated to be quite effective in creating large-scale target-driven conversation datasets, paving the way for constructing extensive target-driven conversational models.

1 Introduction

Target-driven conversation is a crucial aspect of conversational AI. The dialogue systems are required to lead the conversation to the target flexibly and coherently. Due to its purpose and flexibility, target-driven dialogue agents have a broad-based demand, e.g., conversational recommendation (Li et al., 2019; Kang et al., 2019a), psychotherapy (Sharma et al., 2020), and education (Clarizia et al., 2018). These conversations, usually characterized by defined user requirements, rely on precise planning capabilities, making it crucial to build autonomous conversational AI.

[1]Resources of this paper can be found at https://anonymous.4open.science/r/ConvPlan-2023

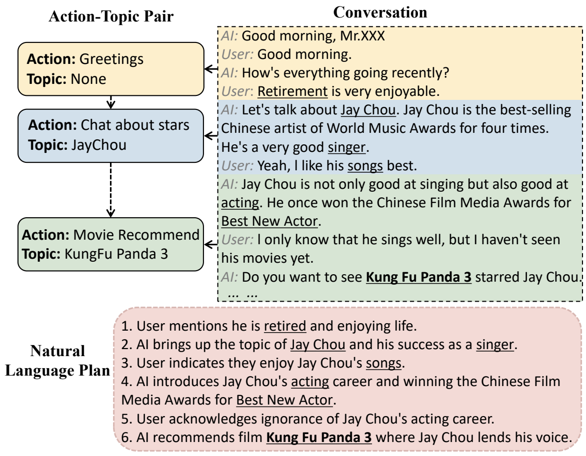


Figure 1: The structured plan (e.g., Action-Topic Pairs) generated by traditional dialogue planning methods hinders both human and LLMs understanding.

In traditional target-driven conversation methods, many studies control dialogue generation through next-turn transition prediction (Tang et al., 2019), subgoal generation (Zhang et al., 2021; Kishinami et al., 2022), and knowledge path reasoning (Gupta et al., 2022). To accomplish this task, effective conversation planning is crucial (Wang et al., 2023a), which requires reasonable actions to smoothly guide the dialogue topics to targets. Different from summarizing a conversation, the process of planning requires not only capturing the key content but also ensuring logically coherent and natural. However, previous studies have employed greedy strategies with single-round topic prediction mechanisms that lack global planning of the conversation process (Yang et al., 2022). These approaches tend to be short-sighted and lead to incoherent topic cues. The generated plan is also too structured (e.g., a sequence of entities or action topic pairs) and not conducive to human understanding. This inherent rigidity prompts a shift in focus toward emergent conversational frameworks, a realm dominated by Large Language Models (LLMs).

Recent advancements have propelled LLMs to the forefront of conversational AI due to their exceptional generation capabilities (Aher et al., 2023). However, LLMs fall short of proactively planning the conversation process (Zheng et al., 2023b; Deng et al., 2023), making it insufficient in handling target-driven conversation. This is because target-driven conversations aim to achieve a global target that often cannot be explicitly defined as a subtask. Conversation agents are required to be able to direct the conversation to the target flexibly and the process must be coherent.

Nevertheless, to enhance the planning and reasoning ability of LLMs, many researchers have investigated Chain-of-thought (CoT) (Kojima et al., 2023; Zhou et al., 2023; Zelikman et al., 2022; Wei et al., 2023; Yao et al., 2023b) and Tree of Thoughts approach (ToT) Yao et al. (2023a), known as reasoning chains or rationales, to eventually lead to the final answer. However these works usually only apply to some well-defined tasks (such as Game of 24), focusing on the evaluation of the final task and neglecting the measurement of the rationality of the plan. In addition, many works use the tree-search approach to improve planning capabilities of LMs (Zhang et al., 2023; Cheng et al., 2022; Yu et al., 2023; Wang et al., 2023b). For example, Yu et al. (2023) treat policy planning as a stochastic game and use prompting for every stage of an open-loop tree search. However, when these methods are faced with the complexity of real-world applications, they require a lot of user simulation.

In this paper, we aim to improve the constrained planning ability of LLMs in the task of target-driven conversation. LLMs have strong comprehension and generation capability but weak planning capability (Yuan et al., 2023; Xie et al., 2023). As illustrated in Figure 1, the structured plan could be difficult to understand by both human and LLMs. To mitigate this issue, we propose a novel two-stage planning construction framework, named Enhance Planning framework (EnPL). EnPL first leverages the existing manually collected conversation dataset to distill the plan describing the conversation process through LLMs. We propose a filter approach, which calculates the entity consistency score between the distilled plans and the conversations, to select high-quality plans for constructing a target-driven conversation plan dataset, named ConvPlan. It consists of 12K high-quality plans with targets, user settings, and plans. Given a new user setting and target, the distilled plans can then serve as demonstrations for generating a new plan as thought to the target with the exceptional in-context learning capability of LLMs. We fully verify the rationality and intelligence of the newly generated plan and reveal that these plans can further guide the generation of target-driven conversation datasets, pointing out feasible directions for constructing large-scale target-driven conversation datasets and model training.

Our contributions are summarized as follows: (1) We propose a novel two-stage framework, named EnPL, to improve the LLMs' capability in planning conversations towards designated targets, including distilling natural language plans from target-guided dialogue corpus and generating new plans with demonstration-guided in-context learning. (2) We propose a filter approach to select high-quality plans distilled by LLMs and introduce a novel evaluation metric, named EntityCov, based on entity-coverage for plan validation. (3) Based on EnPL, we first create a high-quality plan dataset (ConvPlan) for constrained language planning. By leveraging the ConvPlan, we validate that the generated plans play a guiding role in collecting large-scale datasets, suggesting a feasible direction for addressing the issue of scarce dataset availability.

2 Related Work

2.1 Target-driven Conversation

Target-driven conversation systems focus on how to naturally lead users to accept the designated targets gradually through conversations. Previous research has explored various approaches for using keywords and topics as guided targets (Tang et al., 2019; Qin et al., 2020). The advancement of research in this field was catalyzed by the emergence of several datasets such as DuRecDial (Liu et al., 2021), GoRecDial (Kang et al., 2019b), TG-ReDial (Zhou et al., 2020), and INSPIRED (Hayati et al., 2020). Additionally, external commonsense knowledge graphs were used to facilitate keyword transition (Wu et al., 2019; Ma et al., 2021) and response retrieval using GNNs (Zhong et al., 2020; Liang et al., 2021). These datasets typically feature structured plans comprising sequences of keywords or action-topic pairs. While methodical, these structures lack interpretability and miss crucial conversational details, posing challenges for both human users and LLMs. To address this, there is an in-

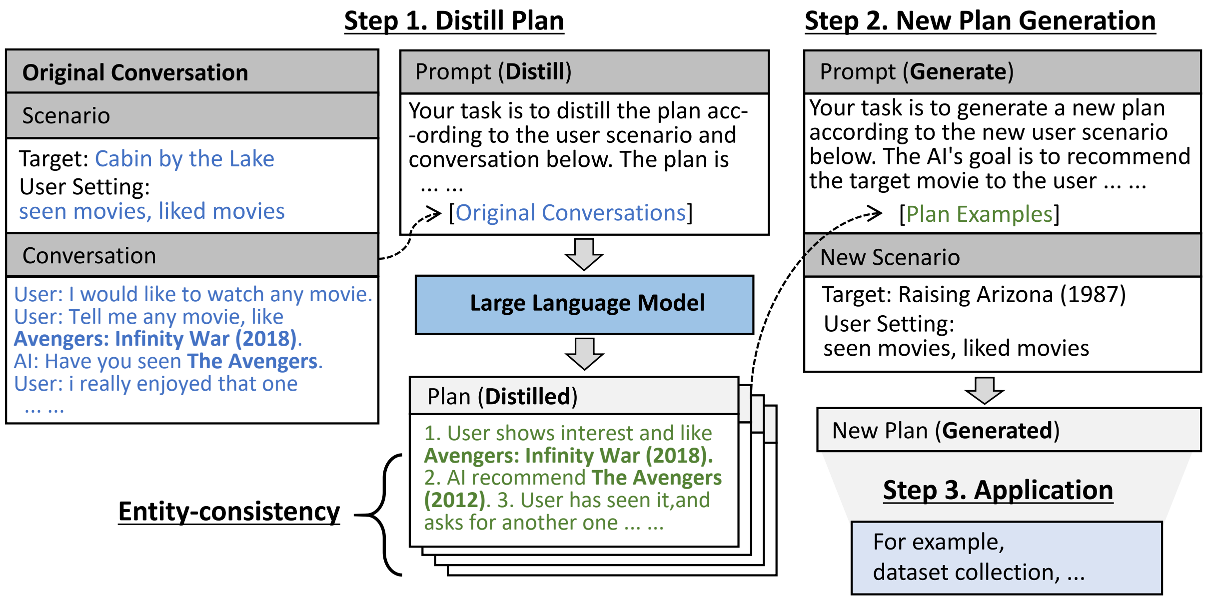


Figure 2: Detailed overview of our proposed two-stage framework (EnPL). Step 1: a large language model is prompted to distill plans (blue) from the existing dataset. Step 2: (green) The distilled plans are used to compose a prompt comprised of other descriptions. The prompt and a new scenario will guide LLM to generate new plans. Step 3: The generated plans can be used for data collection, fine-tuning planning models, or other applications.

creasing emphasis on generating plans in natural language, offering greater clarity and ease of understanding.

2.2 Goal-oriented Planning Script Generation

Prompting in the field of LLM research has seen significant developments towards generating more flexible and efficient outputs. Many researchers have investigated Chain-of-thought (CoT) prompting (Wei et al., 2023; Yao et al., 2023b; Wang et al., 2023d) and Tree of Thoughts approach (ToT) (Yao et al., 2023a). However, these efforts focus on improving the reasoning power of LLMs, while neglecting to measure the rationality of the plan, and are not suitable for planning dialogue process. In order to improve the planning capabilities of LMs, many previous works have investigated how to perform content planning (such as selecting key entities and arranging their sequence) for text generation (Puduppully et al., 2019; Hua and Wang, 2019; Moryossef et al., 2019; Su et al., 2021). Currently, multiple planning frameworks have been proposed for complex generation tasks (Hua et al., 2021; Hu et al., 2022; Li et al., 2022). Our work is more relevant to dialogue generation planning (Kishinami et al., 2022; Yang et al., 2022; Cohen et al., 2022). Wang et al. (2023a) introduced COLOR model to guide goal-oriented dialogue generation using Brownian bridge processes to generate dialogue-level planning. However, this approach is susceptible to error propagation, and when the model fails to plan an appropriate dialogue path, the performance of dialogue generation significantly deteriorates. Our proposed EnPL framework is a novel method to enhance the planning capabilities of large models and can be used to guide target-driven conversation generation.

2.3 LLM for Dialogue Generation

The field of LLMs for dialogue generation has seen remarkable progress. Several recent studies have explored this approach, highlighting its potential across various dialogue applications, such as conversational question-answering (Xu et al., 2023), emotional support dialogues (Zheng et al., 2023b,a), open-domain social dialogues (Chen et al., 2023; Kim et al., 2022), tutoring dialogues (Macina et al., 2023), and more. Despite the remarkable quality of LLM-synthetic dialogue data, this type of data inevitably inherits the limitation of LLMs in handling proactive dialogues, such as inappropriate content, limited understanding of user intent, inability to clarify uncertainty, limited ability to make strategic decisions and plans, etc. In target-driven dialogues, there is a need for the system to proactively plan the conversation process, set targets, and take actions (Wang et al., 2023c), that goes beyond the current capabilities of LLMs. So our approach aims to enhance the planning ability of LLMs.

3 The EnPLAN Framework

As illustrated in Figure 2, the proposed framework can be decomposed into two stages: (1) plan distillation and (2) plan generation. In stage 1, aiming at the existing LLMs with weak planning capability but strong comprehension and generation capability, we use the existing manually collected conversation dataset DuRecDial (Liu et al., 2021)[2] to distill plans describing the conversation process through LLMs. In stage 2, we employ the distilled plans as examples. Then, given a new user setting and target, we can select the plan examples in different ways and generate a new plan as thoughts to target by combining the powerful in-context learning capability of LLMs.

3.1 Distill Plan from Existing Conversation

3.1.1 Problem Formulation

Denote $D = (s_i, c_i)^N$ to be a dataset with N training instances, where s_i is a scenario which is a tuple of user setting and target item (u_i, t_i) and c_i is the corresponding target-driven conversation. Also, we have a handful of human-written instances $E = (s'_i, c'_i, p'_i)^M$, where p'_i is a free-text plan to describe the conversation plan sketch to the target item and $(s'_i, c'_i)^M \in D$ with $M << N$ (we set $M = 30$ in our experiments). Our goal is to fully leverage LLM with E as examples to distill reasonable plans p_i for all (s_i, c_i), where $1 \leq i \leq N$, so that we can utilize these distilled plans from LLM to enhance its planning capability for new scenarios.

3.1.2 Distill Plan with Entity-consistency

Based on the examples E given, we explain to ChatGPT what a plan is and specify the criteria for distilling the plan by referring to the Chain of Thought (CoT) approach (Yao et al., 2023b; Wang et al., 2023d). We then guided ChatGPT to distill plans (prompts are shown in Appendix B).

We further utilize entity-consistency to improve the quality of the distilled plans. The main idea is to select high-quality ones from multiple distilled plans. We first extract the set of key entities from the distilled plan K_{plan} and the original conversation K_{conv} using TextRank algorithm (Mihalcea and Tarau, 2004). Then, we calculate the consistency score between the plan and the original conversation using the Levenshtein distance algorithm[3]. Unlike the original Levenshtein distance algorithm, we treat key entities as the smallest units instead of individual characters. The Levenshtein distance between K_{plan} and K_{conv} (of length i and j respectively) is given by $Leven_{p,c} = L(i, j)$:

$$L(i,j) = \begin{cases} \max(i,j), & \text{if} \min(i,j) = 0 \\ s, & \text{otherwise} \end{cases} \quad (1)$$

where K_{plan} and K_{conv} are noted as p and c, respectively, for simplicity. Then s is computed by

$$s = \min\{\mathrm{L}(i-1,j)+1, \mathrm{L}(i,j-1)+1, \mathrm{L}(i-1,j-1)+1_{(p_i \neq c_j)}\} \quad (2)$$

We calculate the consistency score via:

$$consistency = 1 - \frac{\mathrm{L}(i,j)}{\max(i,j)} \quad (3)$$

An example is shown in Appendix C. The Levenshtein distance directly reflects the degree of difference between the distilled plan and the original conversation, considering the order of entity occurrences. We filter out the top 2 plans with the highest consistency scores from the 10 distilled plans in each round to form the plan repository (ConvPlan).

3.2 Demonstrated Planning for New Scenario

We construct new scenarios each includes a user setting and a target item $s_j = (u_j, t_j)$, and then select (s_i, p_i) as an example from the distilled plans. Our goal is to give new s_j under the guidance of example (s_i, p_i) to generate new plan p_j.

3.2.1 Demonstration Selection Strategies

For better guiding LLM to generate new plans, it is important to select examples for new user scenarios. We explore three different strategies for selecting examples.

Random-based. Randomly select scenarios and plans as example (s_i, p_i) in ConvPlan. This setup does not consider the similarity and diversity between the new user scenario s_j and the user scenarios s_i in existing plans.

Similarity-based. Based on the similarity, we select the similar user scenarios and plans as example s_i, p_i. Specifically, we select the plan with the largest overlap ($max\left(|s_j \bigcap s_i|\right)$) between the movie in the current user scenarios s_j and the movie contained in s_i.

[2]Note that our framework also can be applied to other target-driven conversation datasets.

[3]https://en.wikipedia.org/wiki/Levenshtein_distance

Diversity-based. We use K-means++ clustering (Chang et al., 2021) to select the most representative and diverse plan samples, which will maximize the possibility of maximizing the large models to generate diverse plans. We first map each data point into a vector, then cluster the vectors with the K-means algorithm. The objective is the sum of the squared errors (SSE), called cluster inertia:

$$SSE = \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{K} w_{i,j} ||x^i - \mu^j||_2^2, \quad (4)$$

where μ^j is the centroid of the j-th cluster, x^i is the embedding vector of U_i, and $w_{i,j} = 1$ if x^i belongs to the cluster j and 0 otherwise. We optimize the objective function with the EM algorithm (Dempster et al., 1977) which iteratively assigns each data point into its closest cluster centroid. The initial centroid points are chosen based on the K-means++. The first cluster center is chosen uniformly at random from the data points, after which each subsequent cluster center is chosen from the remaining data points with probability proportional to its squared distance from the point's closest existing cluster center. By this means, we maximize the chance of spreading out the K initial cluster centers. We use 50 random seeds for selecting initial centers and the clustering with the minimum SSE is chosen.

3.2.2 Usage of Generated Plan

The generated plans can map out a complete and logical dialogue path, outlining how to achieve the target step by step. Under the guidance of planning, LLMs can be guided to better connect domain knowledge, dialogue context, and goals, know when to discuss what content, and thus guide the topic to target. Our method has the potential to contribute significantly to the field of conversational AI. By incorporating the generated plans into response generation models or exploring various domains, we can create intelligent and coherent conversational agents. And our work can pave the way for more extensive and meaningful dataset collection in target-driven conversations. Additionally, utilizing our method can improve the planning ability of LLMs, while distilling a dataset from LLMs to train specialized models.

4 Evaluating Stage 1: Distill Plan

4.1 Baselines

We explore prompting for three different ways of distilling plans (Appendix B).

GPT4-abs. GPT4-abs (Liu et al., 2023b) is a method that utilizes GPT4 for text summarization and quality assessment.

Direct Prompt. Directly gives the LLM instructions to generate a plan describing the conversation process, including zero-shot and one-shot settings. The one-shot demonstration is randomly selected from 30 manually constructed plan examples.

CoT+Prompt. Based on the manual examples given, explain to LLM what a plan is and specify the criteria for generating the plan by referring to the Chain of Thought (CoT) method (Yao et al., 2023b; Wang et al., 2023d), also including zero-shot and one-shot settings.

4.2 Proposed Evaluation Metrics

Entity-centered Protocol . The quality and rationality of the plan can be measured and verified through the correspondence of the conversation data and the support of the related knowledge base. Referring to (Mihalcea and Tarau, 2004), we designed the entity-coverage evaluation metric **EntityCov**. First, the text is divided into nodes $V_1, V_2, \ldots, V_n$, and the edges $E(i, j)$ between nodes are constructed to represent the association strength between nodes. Initially, the weight of each node is $W(i) = 1$. Then, TextRank uses an iterative method to calculate the weight of the node. Taking into account the correlation between nodes, the formula is as follows:

$$W(i) = (1-d) + d \cdot \sum_j \left(\frac{W(j) \cdot W(i,j)}{\sum_k W(k)} \right), \quad (5)$$

where j is the neighbor node of node i, and d is the damping coefficient (usually 0.85). Iteratively calculating weight values until convergence, this process enables the identification of the most important words or phrases in the conversation as keywords. Then extract the first 20 keywords K_{conv} based on the final weight value of the node. On this basis, we take the union of the keywords K_{user} and K_{conv} in user information and get $K_{conv+user} = K_{user} \bigcup K_{conv}$. We then use the above principle to get the keyword list K_{plan} in the plan, and calculate the entity-coverage score:

$$EntityCov = \frac{|K_{plan} \bigcap K_{conv+user}|}{|K_{conv+user}|}. \quad (6)$$

Human-centered Protocol . In general, the best method for evaluating such texts is still human evaluation, where human annotators assess the generated plans' quality. This evaluation can be done

Methods	EntityCov	BERTScore	BARTScore	Coherence
GPT4-abs	0.4385	0.5676	-3.610	0.3485
Direct Prompt	0.3961	0.6143	-3.586	0.3986
w/ *example*	0.4657	0.5874	-3.395	0.4252
CoT+Prompt	0.4551	0.6197	-3.384	0.4167
w/ *example*	0.5142	0.6251	**-3.282**	0.4348
EnPL	**0.5509**	**0.6630**	-3.3559	**0.4597**

Table 1: Results of automatic evaluation in plan distillation. The results in bold indicate significant superiority over all the competitors.

Methods	Coh.	Rel.	Int.	Con.	Ove.
GPT4-abs	2.02	2.45	2.31	1.97	2.07
Direct Prompt	1.95	2.46	2.23	2.39	2.22
w/ *example*	2.24	2.40	2.42	2.51	2.41
CoT+Prompt	2.13	2.47	2.35	2.40	2.38
w/ *example*	2.15	2.54	2.51	2.42	2.43
EnPL	**2.30**	**2.63**	**2.74**	**2.55**	**2.58**
κ	0.45	0.35	0.33	0.47	0.42

Table 2: Human evaluation results in plan distillation. The scores (from 0 to 3) are averaged over all the samples rated by three annotators. κ denotes Fleiss' Kappa (Fleiss, 1971), indicating fair or moderate inter-annotator agreement ($0.2 < \kappa < 0.6$).

from different perspectives, and we propose a few common varieties: (1) **Coherence (Coh.)**: Is the overall logic of the plan coherent and clear? (2) **Relevance (Rel.)**: Can the plan capture the key information and discussion process of the original conversation? (3) **Intelligence (Int.)**: whether the plan to guide the conversation process to target is smart. (4) **Concise (Con.)** Is the language of the plan concise? (5) **Overall (Ove.)**: Which version do you prefer overall?

Other Metrics To evaluate the performance of plans distilled, we adopt **BERTScore** (Zhang et al., 2019) and **BARTScore** (Yuan et al., 2021) to measure the semantic similarity between the plan and the original conversation. Following (Yang et al., 2022), we also use **Coherence** as another global evaluation metric. BERTScore calculates the cosine similarity between two sentences based on BERT model. BARTScore computes a similarity score for each token in the candidate sentence with each token in the reference sentence. Coherence is a global evaluation metric, that measures the average contextual semantic similarity between the last utterance in the context and generated utterance.

4.3 Quality Analysis for Distilled Plans

To demonstrate the effectiveness of distilled plans within our EnPL framework, we carried out both automatic evaluation compared to other methods and human evaluation involving five master's students. We randomly selected 50 distilled plans from ConvPlan for comparative analysis. For human evaluation, participants were prompted with the questions in Section 4.2. The comparison outcomes presented in Table 1 and Table 2 reveal the following findings: (a) Our method demonstrates a capacity to include more key entities and clearer logical structures compared to directly summarizing dialogues. (b) We find that the Direct Prompt lacks comprehensive examples and guidance, leading LLM to struggle in understanding the task of plan distillation, resulting in unsatisfactory responses and formatting inconsistencies. (c) Compared to CoT+Prompt, under similar examples and guidance, the plans we distilled closely resemble the original conversations due to our utilization of entity-consistency, filtering the distilled plans to ensure their quality. Overall, our approach effectively guides LLMs in distilling dialogue plans and efficiently filters them, affirming the high quality and practicality of our ConvPlan (distilled plans).

5 Evaluating Stage 2: Generate New Plan

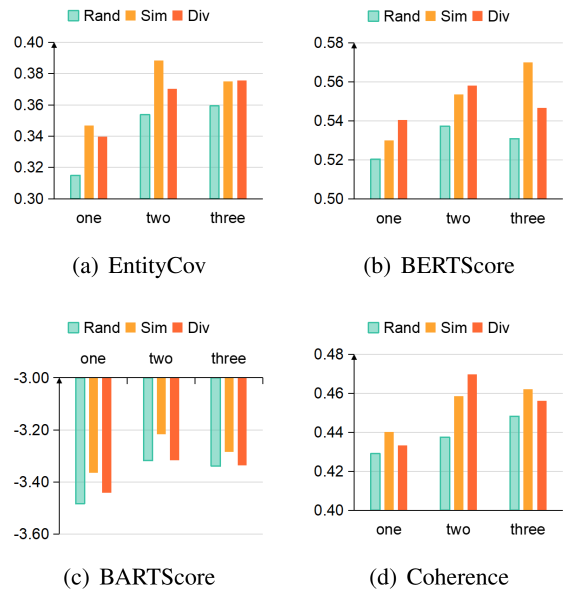
In this section, we fully verify the rationality and intelligence of the newly generated plan and reveal that these generated plans can further guide the generation of target-driven conversation datasets.

5.1 New Scenarios Setting

To create a scenario similar to the real case, we use the 2k scenarios in the DuRecDial testset (Liu et al., 2021) as new scenarios to guide LLM to generate new plans. These scenarios include target movie, user profile, and knowledge graph. The user profile contains personal information (e.g. name, gender, age, residence city, occupation, etc.) and his/her preference And the knowledge graphs include star, movie, music, news, food, and so on. LLMs could generate more realistic and content-rich plans with the assistance of this information.

5.2 Baselines

For plan generation, our baselines include:
Direct prompting (Brown et al., 2020) is a standard method of prompting that makes a request directly to the LLM, including ChatGPT (175B) and LLaMA2 (70B) (Touvron et al., 2023).
CoT prompting (Liu et al., 2023a) use a new



(a) EntityCov (b) BERTScore

(c) BARTScore (d) Coherence

Figure 3: The impact of the number of examples (one, three, and five) and selection strategy on our framework. We select the best version EnPL w/ similarity giving 3 examples for subsequent experiments.

Baselines	EntityCov	BERTScore	BARTScore	Coherence
LLaMA2	0.2556	0.3743	-3.675	0.3137
Direct prompting	0.2125	0.4823	-3.652	0.3169
CoT prompting	0.3273	0.5017	-3.5062	0.3809
COLOR	0.2976	0.5145	-3.5447	0.2731
EnPL	**0.3882**	**0.5535**	**-3.215**	**0.4584**

Table 3: Results of automatic evaluation in plan generation. The results in bold indicate significant superiority over all the competitors.

CoT prompting paradigm of text summarization that considers LLMs as the reference on commonly used summarization datasets such as the CNN/DailyMail dataset (Liu et al., 2023a).
COLOR (Wang et al., 2023b) uses the Brownian bridge stochastic process to plan dialogue process, which models global coherence and incorporates user feedback in goal-directed dialogue planning.
Our variations. We analyze the following variants of our method: (1) w/ Random, which randomly selects context examples in ConvPlan; (2) w/ Similarity, which selects plans with similar scenarios; (3) w/ Diversity, which uses K-means++ clustering to select diverse and representative examples.

5.3 Effect of Demonstration Selection

Regarding the impact of selection strategies and the number of examples used in guiding LLMs to generate new plans, shown in Figure 3, we observe that the strategy based on scenario similarity outperforms the diversity-based strategy overall. This is because the similarity-based strategy selects plans in similar scenarios from our ConvPlan for the current new scenario, resulting in higher scores on BERTScore and BARTScore metrics. The diversity-based strategy, by clustering, selects diverse examples, which are beneficial in enriching the content of LLM-generated plans. Both strategies significantly outperform random selection, confirming the effectiveness of our selection strategies. Furthermore, compared to providing three examples, there is a gap in effectiveness when only one example is provided. This indicates that more examples aid LLMs in more comprehensive learning, thereby generating more refined plans. However, as the number of examples increases to five, the improvement in generation is not substantial and even shows a decline. Considering input length limitations and cost factors, further increasing the number of examples is unnecessary. Therefore, we opt for the version EnPL w/ sim giving 3 examples for subsequent experiments.

5.4 Automatic Evaluation

To demonstrate the effectiveness of our two-stage EnPL framework, we compare generated new plans by EnPL with other models (shown in Table 3). We find that across most metrics, our EnPL consistently outperforms other baselines. Specifically, Direct prompting slightly outperforms LLaMA2, indicating an advantage of ChatGPT (175B) over LLaMA2 (70B), potentially due to ChatGPT's larger generation space and better comprehension compared to LLaMA2. Furthermore, when evaluating similarity-based metrics (BERTScore and BARTScore), EnPL, when guided with examples, could generate longer and more comprehensive content than other methods, encompassing a broader range of key entities. We noticed that adhering closely to our defined scenario prompts the LLM to generate more diverse content, leveraging its inherent knowledge. Additionally, COLOR's reliance on traditional methods for plan generation results in overly structured plans, leading to a lower Coherence score. Conversely, our newly generated plans exhibit a significant improvement in Coherence, a trend also reflected in Table 4. Overall, the two-stage framework we propose offers dual benefits: it distills a high-quality plan dataset (ConvPlan) and selects comprehensive examples within it to guide LLMs in generating high-quality new

Baselines	Coh.	Rel.	Int.	Con.	Ove.
LLaMA2	2.03	2.41	2.23	2.32	2.30
Direct prompting	2.18	2.59	2.51	2.74	2.46
CoT prompting	2.37	2.80	2.56	2.67	2.64
COLOR	1.72	2.07	1.72	2.35	2.11
EnPL	**2.46**	**2.81**	**2.56**	**2.78**	**2.71**
κ	0.42	0.37	0.35	0.40	0.41

Table 4: Human evaluation results in plan generation. The Fleiss' Kappa is a fair or moderate inter-annotator agreement ($0.2 < \kappa < 0.6$).

	DuRecDial	Ours	κ
Appr.	2.54	**2.65**	0.48
Info.	**2.64**	2.62	0.43
Proact.	**2.61**	2.58	0.39
Coh.	2.77	**2.85**	0.52
Succ.	2.83	**2.95**	0.37

Table 5: Human evaluation of conversation quality. The scores (from 0 to 3) are averaged over all the samples rated by three annotators. κ denotes Fleiss' Kappa (Fleiss, 1971), indicating fair to moderate inter-annotator agreement ($0.2 < \kappa < 0.6$).

plans. EnPL stands as a promising method, laying a strong foundation for further leveraging plans to generate conversations or train plan models.

5.5 Human Evaluation

We further conduct a human evaluation on the generated plans with five annotators. From each baseline, we randomly selected 50 generated plans for comparison. Annotators were tasked with rating the performance of different baselines, and the outcomes of this comparison (as shown in Table 4) reveal the following findings: (1) LLaMA2 slightly underperforms compared to our EnPL, which is understandable considering our method builds upon ChatGPT, offering a larger generation space and better comprehension. (2) The COLOR's performance in plan generation is unsatisfactory. We observed that COLOR, relying on an external knowledge graph, lacks the capability for comprehensive planning, resulting in lower scores. (3) Detailing to explain the plan proves crucial; otherwise, the LLM lacks an understanding of the task's goal. Direct prompting may provide ambiguous guidance, leading to struggles in generating plans, thereby affecting scores in Clarity and Intelligent metrics. Overall, the results of the human evaluation align with those of the automatic evaluation. It reveals that our EnPL achieves better performances than other baselines on most metrics. Our method adeptly guides LLMs in generating reasonable new plans.

5.6 Evaluation on Generated Conversations

We further validate the quality of plans via conducting comprehensive human evaluation and comparing the generated conversations with DuRecDial dataset. The details on the metrics and evaluation procedure are described in Appendix D.2.

Table 5 shows human evaluation results. We observed that our method shows advantages over manually constructed DuRecDial. Although DuRecDial scores slightly better than us in terms of informativeness, the difference is not significant. Our method can generate more appropriate discourses based on the dialogue context. Moreover, our method achieves a higher coherence score and target success rate, which may be due to the fact that manual chats often use abbreviated or omissions discourse, which leads to a decrease in coherence. In general, our plans can guide conversations to reach their targets as well as effectively keep them coherent. Our approach allows us to plan a dialogue path through reasonable actions and appropriate topics, outlining how to reach the target step by step. Under the guidance of planning, the model can be guided to know better what to talk about when, so as to proactively advance the conversation and ultimately succeed in achieving the target.

6 Conclusion

This paper introduces a novel two-stage enhance planning framework to overcome challenges in target-driven conversation planning via LLMs. Our method involved harnessing the generative capabilities of LLM in distilling plans from existing human-curated datasets. We filter the over-generate plans and introduce comprehensive method for plan validation. We further guides LLM to generate plans according to new user scenarios and targets via in-context learning. Our approach not only advances the capabilities of LLMs in planning target-driven conversations but also provides a scalable strategy for generating large-scale datasets. Consequently, this is a significant step towards building sophisticated target-driven conversational models. Future research will focus on refining the plan generation and validation process for even greater processing efficiency and accuracy.

Limitations

Our framework significantly advances LLM-based conversation planning but faces limitations inherent to LLMs, such as biases in training data and tendencies to produce incorrect information. While we enhance LLMs' planning capabilities, our focus isn't on modifying the model architecture itself, and our reliance on automatic evaluation metrics might lead to overestimations or underestimations, despite attempts to balance these with human evaluations. Currently, our ConvPlan dataset is limited to English, restricting multilingual applicability. A notable area we will explore in the near future is the dynamic generation of conversation plans mid-dialogue, which would address our framework's current limitation of only generating plans at the conversation's outset and significantly enhance adaptability in real-time interactions.

Ethical Considerations

We protect the privacy rights of crowd-sourced workers and pay them above the local minimum wage (pay at a rate of $7 per hour). We acknowledge that constructing datasets from large language models may suffer from toxic language and cause severe risks for social society (Weidinger et al., 2021; Baldini et al., 2022). Factuality, Toxicity and Biases We recognize that the factuality of generated content is crucial, especially in high-stakes scenarios. Therefore, we ask the annotators to discard the offensive and harmful data when reviewing the ConvPlan. They also assess and revise the content to minimize hallucinations, factual errors, and any inappropriate or misleading information. However, there may still be prejudicial data in our final dataset that goes unnoticed. We highlight that our ConvPlan dataset is not intended for safety-critical applications or as a substitute for expert advice in such domains. Significant further progress needs to be made in areas like debiasing, grounding in actuality, and efficient serving before we can safely deploy this type of system in a production setting.

References

Gati V Aher, Rosa I. Arriaga, and Adam Tauman Kalai. 2023. Using large language models to simulate multiple humans and replicate human subject studies. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202, pages 337–371. PMLR.

I Elaine Allen and Christopher A Seaman. 2007. Likert scales and data analyses. *Quality progress*, 40(7):64–65.

Ioana Baldini, Dennis Wei, Karthikeyan Natesan Ramamurthy, Moninder Singh, and Mikhail Yurochkin. 2022. Your fairness may vary: Pretrained language model fairness in toxic text classification. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 2245–2262, Dublin, Ireland. Association for Computational Linguistics.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners.

Ernie Chang, Xiaoyu Shen, Hui-Syuan Yeh, and Vera Demberg. 2021. On training instance selection for few-shot neural text generation.

Maximillian Chen, Alexandros Papangelis, Chenyang Tao, Seokhwan Kim, Andy Rosenbaum, Yang Liu, Zhou Yu, and Dilek Hakkani-Tur. 2023. PLACES: Prompting language models for social conversation synthesis. In *Findings of the Association for Computational Linguistics: EACL 2023*, pages 844–868, Dubrovnik, Croatia. Association for Computational Linguistics.

Yi Cheng, Wenge Liu, Wenjie Li, Jiashuo Wang, Ruihui Zhao, Bang Liu, Xiaodan Liang, and Yefeng Zheng. 2022. Improving multi-turn emotional support dialogue generation with lookahead strategy planning.

Fabio Clarizia, Francesco Colace, Marco Lombardi, Francesco Pascale, and Domenico Santaniello. 2018. Chatbot: An education support system for student. In *International Conference on Cryptography and Security Systems*.

Deborah Cohen, Moonkyung Ryu, Yinlam Chow, Orgad Keller, Ido Greenberg, Avinatan Hassidim, Michael Fink, Yossi Matias, Idan Szpektor, Craig Boutilier, et al. 2022. Dynamic planning in open-ended dialogue using reinforcement learning. *arXiv preprint arXiv:2208.02294*.

Arthur P Dempster, Nan M Laird, and Donald B Rubin. 1977. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22.

Yang Deng, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, and Tat-Seng Chua. 2023. Prompting and evaluating large language models for proactive dialogues: Clarification, target-guided, and non-collaboration. In *Findings of the Association for*

Computational Linguistics: EMNLP 2023, Singapore, December 6-10, 2023, pages 10602–10621. Association for Computational Linguistics.

Joseph L. Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76:378–382.

Prakhar Gupta, Harsh Jhamtani, and Jeffrey Bigham. 2022. Target-guided dialogue response generation using commonsense and data augmentation. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 1301–1317, Seattle, United States. Association for Computational Linguistics.

Shirley Anugrah Hayati, Dongyeop Kang, Qingxiaoyang Zhu, Weiyan Shi, and Zhou Yu. 2020. INSPIRED: Toward sociable recommendation dialog systems. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8142–8152, Online. Association for Computational Linguistics.

Zhe Hu, Hou Pong Chan, Jiachen Liu, Xinyan Xiao, Hua Wu, and Lifu Huang. 2022. PLANET: Dynamic content planning in autoregressive transformers for long-form text generation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2288–2305, Dublin, Ireland. Association for Computational Linguistics.

Xinyu Hua, Ashwin Sreevatsa, and Lu Wang. 2021. DYPLOC: Dynamic planning of content using mixed language models for text generation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 6408–6423, Online. Association for Computational Linguistics.

Xinyu Hua and Lu Wang. 2019. Sentence-level content planning and style specification for neural text generation. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 591–602, Hong Kong, China. Association for Computational Linguistics.

Dongyeop Kang, Anusha Balakrishnan, Pararth Shah, Paul Crook, Y-Lan Boureau, and Jason Weston. 2019a. Recommendation as a communication game: Self-supervised bot-play for goal-oriented dialogue.

Dongyeop Kang, Anusha Balakrishnan, Pararth Shah, Paul Crook, Y-Lan Boureau, and Jason Weston. 2019b. Recommendation as a communication game: Self-supervised bot-play for goal-oriented dialogue. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 1951–1961, Hong Kong, China. Association for Computational Linguistics.

Hyunwoo Kim, Jack Hessel, Liwei Jiang, Ximing Lu, Youngjae Yu, Pei Zhou, Ronan Bras, Malihe Alikhani, Gunhee Kim, Maarten Sap, and Choi Yejin. 2022. Soda: Million-scale dialogue distillation with social commonsense contextualization.

Yosuke Kishinami, Reina Akama, Shiki Sato, Ryoko Tokuhisa, Jun Suzuki, and Kentaro Inui. 2022. Target-guided open-domain conversation planning. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 660–668, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2023. Large language models are zero-shot reasoners.

Qintong Li, Piji Li, Wei Bi, Zhaochun Ren, Yuxuan Lai, and Lingpeng Kong. 2022. Event transition planning for open-ended text generation. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 3412–3426, Dublin, Ireland. Association for Computational Linguistics.

Raymond Li, Samira Kahou, Hannes Schulz, Vincent Michalski, Laurent Charlin, and Chris Pal. 2019. Towards deep conversational recommendations.

Zujie Liang, Huang Hu, Can Xu, Jian Miao, Yingying He, Yining Chen, Xiubo Geng, Fan Liang, and Daxin Jiang. 2021. Learning neural templates for recommender dialogue system. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 7821–7833, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Yixin Liu, Alexander R. Fabbri, Pengfei Liu, Dragomir Radev, and Arman Cohan. 2023a. On learning to summarize with large language models as references.

Yixin Liu, Kejian Shi, Katherine S He, Longtian Ye, Alexander R. Fabbri, Pengfei Liu, Dragomir Radev, and Arman Cohan. 2023b. On learning to summarize with large language models as references.

Zeming Liu, Haifeng Wang, Zheng-Yu Niu, Hua Wu, and Wanxiang Che. 2021. DuRecDial 2.0: A bilingual parallel corpus for conversational recommendation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 4335–4347, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Zeming Liu, Haifeng Wang, Zheng-Yu Niu, Hua Wu, Wanxiang Che, and Ting Liu. 2020. Towards conversational recommendation over multi-type dialogs.

Wenchang Ma, Ryuichi Takanobu, and Minlie Huang. 2021. CR-walker: Tree-structured graph reasoning and dialog acts for conversational recommendation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 1839–1851, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Jakub Macina, Nico Daheim, Sankalan Pal Chowdhury, Tanmay Sinha, Manu Kapur, Iryna Gurevych, and Mrinmaya Sachan. 2023. Mathdial: A dialogue tutoring dataset with rich pedagogical properties grounded in math reasoning problems. *ArXiv*, abs/2305.14536.

Rada Mihalcea and Paul Tarau. 2004. TextRank: Bringing order into text. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pages 404–411, Barcelona, Spain. Association for Computational Linguistics.

Amit Moryossef, Yoav Goldberg, and Ido Dagan. 2019. Step-by-step: Separating planning from realization in neural data-to-text generation. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 2267–2277, Minneapolis, Minnesota. Association for Computational Linguistics.

Ratish Puduppully, Li Dong, and Mirella Lapata. 2019. Data-to-text generation with content selection and planning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 6908–6915.

Jinghui Qin, Zheng Ye, Jianheng Tang, and Xiaodan Liang. 2020. Dynamic knowledge routing network for target-guided open-domain conversation. In *AAAI Conference on Artificial Intelligence*.

Ashish Sharma, Adam Miner, David Atkins, and Tim Althoff. 2020. A computational approach to understanding empathy expressed in text-based mental health support. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 5263–5276, Online. Association for Computational Linguistics.

Yixuan Su, David Vandyke, Sihui Wang, Yimai Fang, and Nigel Collier. 2021. Plan-then-generate: Controlled data-to-text generation via planning. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 895–909, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Jianheng Tang, Tiancheng Zhao, Chenyan Xiong, Xiaodan Liang, Eric Xing, and Zhiting Hu. 2019. Target-guided open-domain conversation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5624–5634, Florence, Italy. Association for Computational Linguistics.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami, Naman Goyal, Anthony Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madian Khabsa, Isabel Kloumann, Artem Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, Ranjan Subramanian, Xiaoqing Ellen Tan, Binh Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zheng Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurelien Rodriguez, Robert Stojnic, Sergey Edunov, and Thomas Scialom. 2023. Llama 2: Open foundation and fine-tuned chat models.

Jian Wang, Dongding Lin, and Wenjie Li. 2023a. Dialogue planning via brownian bridge stochastic process for goal-directed proactive dialogue. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 370–387, Toronto, Canada. Association for Computational Linguistics.

Jian Wang, Dongding Lin, and Wenjie Li. 2023b. Dialogue planning via brownian bridge stochastic process for goal-directed proactive dialogue. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 370–387, Toronto, Canada. Association for Computational Linguistics.

Jian Wang, Dongding Lin, and Wenjie Li. 2023c. A target-driven planning approach for goal-directed dialog systems. *IEEE Transactions on Neural Networks and Learning Systems*, pages 1–13.

Lei Wang, Wanyu Xu, Yihuai Lan, Zhiqiang Hu, Yunshi Lan, Roy Ka-Wei Lee, and Ee-Peng Lim. 2023d. Plan-and-solve prompting: Improving zero-shot chain-of-thought reasoning by large language models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2609–2634, Toronto, Canada. Association for Computational Linguistics.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. 2023. Chain-of-thought prompting elicits reasoning in large language models.

Laura Weidinger, John Mellor, Maribeth Rauh, Conor Griffin, Jonathan Uesato, Po-Sen Huang, Myra Cheng, Mia Glaese, Borja Balle, Atoosa Kasirzadeh, Zac Kenton, Sasha Brown, Will Hawkins, Tom Stepleton, Courtney Biles, Abeba Birhane, Julia Haas, Laura Rimell, Lisa Anne Hendricks, William Isaac, Sean Legassick, Geoffrey Irving, and Iason Gabriel. 2021. Ethical and social risks of harm from language models.

Wenquan Wu, Zhen Guo, Xiangyang Zhou, Hua Wu, Xiyuan Zhang, Rongzhong Lian, and Haifeng Wang. 2019. Proactive human-machine conversation with explicit conversation goal. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3794–3804, Florence, Italy. Association for Computational Linguistics.

Yaqi Xie, Chen Yu, Tongyao Zhu, Jinbin Bai, Ze Gong, and Harold Soh. 2023. Translating natural language to planning goals with large-language models.

Canwen Xu, Daya Guo, Nan Duan, and Julian McAuley. 2023. Baize: An open-source chat model with parameter-efficient tuning on self-chat data.

Zhitong Yang, Bo Wang, Jinfeng Zhou, Yue Tan, Dongming Zhao, Kun Huang, Ruifang He, and Yuexian Hou. 2022. TopKG: Target-oriented dialog via global planning on knowledge graph. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 745–755, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L. Griffiths, Yuan Cao, and Karthik Narasimhan. 2023a. Tree of thoughts: Deliberate problem solving with large language models.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. 2023b. React: Synergizing reasoning and acting in language models.

Xiao Yu, Maximillian Chen, and Zhou Yu. 2023. Prompt-based monte-carlo tree search for goal-oriented dialogue policy planning.

Siyu Yuan, Jiangjie Chen, Ziquan Fu, Xuyang Ge, Soham Shah, Charles Robert Jankowski, Yanghua Xiao, and Deqing Yang. 2023. Distilling script knowledge from large language models for constrained language planning.

Weizhe Yuan, Graham Neubig, and Pengfei Liu. 2021. Bartscore: Evaluating generated text as text generation. *ArXiv*, abs/2106.11520.

Eric Zelikman, Yuhuai Wu, Jesse Mu, and Noah D. Goodman. 2022. Star: Bootstrapping reasoning with reasoning.

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 *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 1092–1101, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Shun Zhang, Zhenfang Chen, Yikang Shen, Mingyu Ding, Joshua B. Tenenbaum, and Chuang Gan. 2023. Planning with large language models for code generation.

Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. 2019. Bertscore: Evaluating text generation with bert. *ArXiv*, abs/1904.09675.

Chujie Zheng, Sahand Sabour, Jiaxin Wen, Zheng Zhang, and Minlie Huang. 2023a. AugESC: Dialogue augmentation with large language models for emotional support conversation. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 1552–1568, Toronto, Canada. Association for Computational Linguistics.

Zhonghua Zheng, Lizi Liao, Yang Deng, and Liqiang Nie. 2023b. Building emotional support chatbots in the era of llms.

Peixiang Zhong, Yong Liu, Hongya Wang, and Chunyan Miao. 2020. Keyword-guided neural conversational model. In *AAAI Conference on Artificial Intelligence*.

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*, pages 4128–4139, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Yongchao Zhou, Andrei Ioan Muresanu, Ziwen Han, Keiran Paster, Silviu Pitis, Harris Chan, and Jimmy Ba. 2023. Large language models are human-level prompt engineers.

A Case Study

Table 6 shows some cases generated from LLaMA2, COLOR, and our EnPL.

B Prompt Details

The prompts used in our experiments are as follows:

B.1 Plan Distillation (Stage 1)

B.1.1 GPT4-abstract

Conversation: ${**Conversation**}
Please summarize the conversation.
Summary:

B.1.2 Direct Prompt

Here is an example:
Conversation: ${**Conversation**}
Distilled Plan:${**Plan**}

Please distill the plan according to the target-driven conversation below. The plan shows the process of the conversation AI recommending the target movie to the user.
Conversation: ${**Conversation**}
Plan:

B.1.3 EnPL Prompt (CoT Prompt)

Here is an example:
Conversation: ${**Conversation**}
Distilled Plan: ${**Plan**}

Your task is to distill the plan according to the target-driven conversation below. The AI's goal is to recommend the target movie to the user. The plan shows the process of the conversation AI recommending the target movie to the user. The conversation between recommendation AI and the user is target-driven, gradually shifting the topic to the target movie. And the plan should be as short as possible to reflect the focus of the conversation. Attention to entities mentioned in the reservations dialogue. Only return the plan.
The following is the conversation you need to use in distilling plan:
Conversation: ${**Conversation**}
Plan:

B.2 Generate New Plans (Stage 2)

B.2.1 Direct Prompting

Please generate a conversation plan according to the "Target" and "User Setting" below. The AI's goal is to recommend the target movie to the user. The plan shows the process of the conversation AI recommending the target movie to the user.
Target: ${**Target**}
User Setting: ${**User Setting**}
Plan:

B.2.2 EnPL Prompt (CoT Prompting)

Examples:
Target: ${**Target**}
User Setting: ${**User Setting**}
Plan: ${**Plan**}

Your task is to generate a conversation plan according to the "Target" and "User Setting" below. The AI's goal is to recommend the target movie to the user. The plan shows the process of the conversation AI recommending the target movie to the user. The conversation process between conversation AI and the user is target-driven, gradually shifting the topic to the target movie. You can expand on the information you know to make the conversation process richer. You can refer to the Example above. Only return the plan. The following are the "Target" and "User Setting" you need to use in generating a new plan:
Target: ${**New Target**}
User Setting: ${**New User Setting**}
Plan:

C An Example of Entity-consistency

Figure 4 shows the workflow of entity-consistency to filter distilled plans. The K_{plan} and K_{conv} are the lists of key entities extracted from the distilled plan and the original conversation using TextRank (Mihalcea and Tarau, 2004). Then, we calculate the consistency score between the plan and conversation using the Levenshtein distance algorithm.

D Generate Target-driven Conversations

D.1 Prompt of Generating Conversations

The following is the prompt template we use the generated plan to guide ChatGPT to generate target-driven conversations. Figure 7 shows an example of this process.

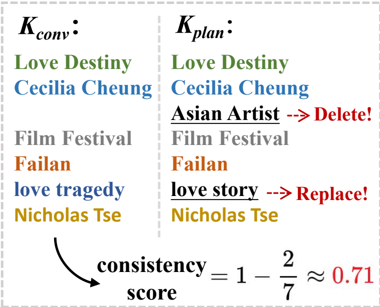


Figure 4: The workflow of entity-consistency to filter distilled plans.

> Here is an example:
> Target: ${**Target**}
> Plan: ${**Plan**}
> Generated conversation: ${**Conversation**}
>
> Your task is to create a movie recommendation conversation between a user and an AI recommender according to the Plan below. The AI's goal is to recommend the target movie to the user. Generate a conversation with as many topic changes as possible to generate more rounds of dialogue. Switch the topic to the target during the chat with the user. Make the conversation more like a real-life chat and be specific. In the example above, where User/AI represents whether the speaker is a User or an AI. Below is the Target and Plan you need to refer to generate conversation.
> Target: ${**Target**}
> Plan: ${**Plan**}
> Generated conversation:

D.2 Evaluation Details on Generated Conversations

We recruited 5 master students to serve as annotators for this project. We randomly selected 50 dialogue examples from DuRecDial and 50 examples from conversations guided by our plans. At least two different annotators rated each dialogue example. For a fair comparison, the examples were randomly renamed as "example-1", "example-2", and so forth. Referring to (Liu et al., 2020), we adopted the following metrics to evaluate the quality of dialogues: The turn-level evaluation measures appropriateness (**Appr.**) and informativeness (**Info.**). The dialogue-level evaluation measures proactivity (**Proact.**), coherence (**Coh.**), and goal success (**Succ.**). The annotators were then asked to mark scores for the compared examples from (1) appropriateness (Appr.), which measures whether the utterance responds to the dialogue context appropriately, and (2) informativeness (Info.), which measures whether the utterance is informative by making full use of the grounded knowledge. (3) proactivity (Proact.), which measures whether a model proactively leads the dialogue; (4) coherence (Coh.), which manually examines whether the whole dialogue is fluent, coherent, and smoothly transited; (5) target success (Succ.), which estimates whether a model effectively reaches the target. All the metrics are rated with the four-level Likert scale (Allen and Seaman, 2007) ranging from 0 to 3, with higher scores indicating better quality. The consensus among annotators was assessed using Fleiss's kappa (Fleiss, 1971). Our final human evaluation results were based on the average scores provided by multiple annotators for each example. Furthermore, we were open and clear with all annotators about the purpose of our research, paid reasonable wages, and provided sufficient time for them to complete the evaluation.

User Scenario	
Target Movie: Who Am I **User Setting:** liked movies: The Foreigner, Jackie Chan: My Story, Miracle seen movies: Chop Socky: Cinema Hong Kong liked music: Piano in the Attic rejected movies: Bruce Lee	
Methods	**Generated Plan**
LLaMA2	1. Ask the user about their favorite movie. 2. User discusses different genres of movies with AI. 3. AI mentions Jackie Chan as a popular action star. 4. Users express interest in Jackie Chan but lack understanding. 5. AI mentions that Jackie Chan received an accolade in a film festival. 6. User agrees and praises the actor. 7. AI recommends "Who Am I" as a movie that showcases Jackie Chan's martial arts skills. 8. User shows interest in the recommended movie. 9. Provide more details about "Who Am I" to persuade the user to give it a try. 10. User expresses gratitude.
COLOR	1. `[Action:]` Respond Q&A `[Topic:]` Jackie Chan: My Story `[Action:]` Chat about stars `[Topic:]` Jackie Chan `[Action:]` Movie recommendation `[Topic:]` Who Am I 2. `[Action:]` Chat about stars `[Topic:]` Jackie Chan `[Action:]` Movie recommendation `[Topic:]` Who Am I 3. (same as 2.) 4. `[Action:]` Movie recommendation `[Topic:]` Who Am I 5. (same as 4.) (too structured, hard to understand)
EnPL	1. User inquires about the star of the movie "Jackie Chan: My Story". 2. AI confirms that Jackie Chan is the star. 3. User expresses admiration for Jackie Chan. 4. AI mentions that Jackie Chan had won the Golden Horse Award for Best Actor. 5. User shows interest in Jackie Chan's achievements. 6. AI recommends another Jackie Chan movie, "Who Am I". 7. User shows interest in the recommended movie. 8. AI describes the recommended kung fu movie. 9. User finds the description interesting and decides to watch the movie. (more related entities and coherence)

Table 6: Generated plan examples in a new scenario. The plans are generated by different methods, including (a) LLaMA2, (b) COLOR, and (c) Ours (EnPL). The topics and topic-related entities that also appear in the domain knowledge are marked with underlines.

Plan
1. User asks about movie "Fly Me to Polaris". 2. AI replies that Cecilia Cheung stars the movie. 3. User asks for more information about Cecilia Cheung. 4. AI praises Cecilia Cheung's talent and mentions her awards in the Hong Kong Film Awards and Golden Bauhinia Awards. 5. AI recommends a movie called "Failan" starred by Cecilia Cheung, which is a touching story of a woman who marries a man she never met before. 6. User expresses interest but mentions that he has already seen it. 7. AI recommends another movie "The Bullet Vanishes" starring Nicholas Tse, who is also one of the user's favorite celebrities, describing it as a thrilling detective story set in 1930s Shanghai. 8. User agrees to watch. 9. AI encourages the user to watch it and expresses confidence that they will enjoy it
Generated Conversation
User: Who was the lead in the movie "Fly Me to Polaris"? AI: That would be Cecilia Cheung. User: Oh, Cecilia Cheung! Tell me more about her. AI: She's an incredibly talented and versatile actress, known not just for her skills but also for her beauty. She's been recognized with awards at the Hong Kong Film Awards and the Golden Bauhinia Awards. User: Impressive! Any particular movie you'd recommend? AI: Definitely! There's a touching one called "Failan" where she plays a woman who marries a stranger. It's a really heartfelt story. User: That sounds interesting, but I have watched it. AI: Ah, got it. How about "The Bullet Vanishes"? It stars Nicholas Tse, someone you're a fan of, right? It's a thrilling detective story set in 1930s Shanghai. User: Nicholas Tse! I'm definitely intrigued. I'll consider watching that. AI: Great choice! I'm confident you'll enjoy it. User: I'm going to grab some food now. AI: Sure thing. Enjoy your meal! Feel free to come back anytime for more recommendations.

Table 7: Example of using the generated plan to guide LLM to generate a target-driven conversation. The target movie is marked with underlines.