

Emotion Trajectory-aware Retrieval for Markov-driven Emotion Anticipation in LLM-based Emotional Support Conversation

Hongyan Wu^{1*}, Zhiliang Tian^{1*}, Zhen Huang^{1†}, Tengyue Hu²,
Linbo Qiao¹, Yifu Gao¹, Feng Liu¹, Dongsheng Li¹

¹College of Computer Science and Technology, National University of Defense Technology,

²College of Computer, Chongqing University

{wuhongyan99, tianzhiliang, huangzhen, qiao.linbo, gaoyifu, richardlf, dsli}@nudt.edu.cn

hutengyue@cqu.edu.cn

Abstract

Emotional support conversation (ESC) aims to alleviate users' psychological stress. Selecting the appropriate strategy is crucial for effective emotional support. Current strategy planner-based methods prioritize immediate responses while neglecting users' future reactions. Some studies retrieve historical examples with similar emotions to the current utterance, then anticipating future emotions based on next-turn emotions of historical examples. However, their retrievals focus on the current emotion (i.e. a single-turn emotion state), while they ignore the evolution of user's emotion before the current state. We argue that retrievals considering the whole emotional trajectories enables models to capture the dynamic emotional needs, thereby enhancing the anticipation of future emotions. To this end, we propose Markov-driven emotion anticipation framework with emotion trajectory-aware retrieval for LLM-based ESC, which anticipates future emotion states to guide strategy planning and achieve sustained emotional support. First, we construct a dynamic emotion memory and perform hierarchical retrieval that combines semantic matching and emotion trajectory alignment. Then, we model emotional transitions as Markov chains, leveraging trajectory-aware retrieval to estimate future emotion. Finally, we use the anticipated emotion to steer LLMs in generating candidate strategies and introduce active online learning to optimize the planner, boosting its robustness on diverse users. Experiments on two datasets with two models shows that our method excels all baselines.

1 Introduction

As large language models (LLMs) demonstrate advanced conversational capabilities, the Emotional Support Conversation (ESC) task has emerged as a promising research domain (Daley et al., 2020).

* contributed equally to this work

† corresponding author

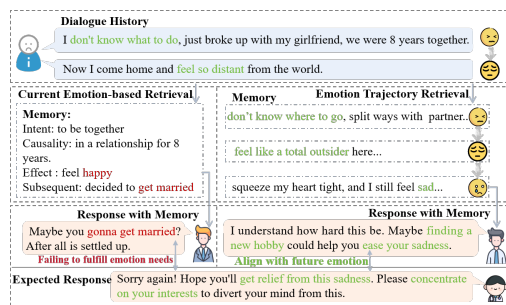


Figure 1: Our emotion trajectory retrieval and current emotion-based retrieval for emotional support. Red indicates a mismatch with current emotion needs and green indicates alignment with emotion needs.

ESC aims to provide users with professional emotional support to alleviate emotional distress (Liu et al., 2021). This technology demonstrates broad application prospects in diverse scenarios, including social companionship, mental health support, and customer service (Wang et al., 2025b). Current mainstream ESC methods primarily focus on planning the next-turn dialogue strategies to guide LLMs in generating empathetic responses.

Existing strategy planning approaches mainly focus on either LLM-based prompt learning or developing an external strategy planner. (1) **LLM-based prompt learning** mainly designs heuristic prompts (Cao et al., 2024; Yu et al., 2023) to stimulate LLMs to perform self-thinking (Deng et al., 2023), self-reflection (Zhang et al., 2025b) and critical feedback (Fu et al., 2023), thereby enhancing the ability of strategy planning. These methods leverage the inherent knowledge of LLMs. However, the pre-training of LLMs seldom explicitly learn strategy planning for ESC (Chen et al., 2023; Farhat, 2024; Kang et al., 2024). Such limited planning capabilities make LLMs struggle to select strategies accurately, thus constraining their performance in ESC. (2) **Developing an external strategy planner** involves training a specialized planner to select appropriate support strategies for the current dia-

logue state via supervised fine-tuning (Deng et al., 2024; He et al., 2025b) or reinforcement learning (Zhang et al., 2024; Rakib et al., 2025). This approach decouples strategy planning from the LLMs, focusing on the improvements in the planner’s strategy capabilities. While these methods select appropriate strategies in ESC, they primarily plan strategies based on the current dialogue state. This prioritizes immediate responses while overlooking potential user reactions in subsequent interactions, potentially leading to a suboptimal strategy.

To address above issues, some studies (Cheng et al., 2022; Zhao et al., 2025) have attempted to anticipate users’ reactions to steer strategy planning. Such approaches (He et al., 2024; Yang et al., 2025) simulate future dialogues to infer potential user feedback, thereby performing lookahead strategy planning. To further enhance the model’s capability to anticipate future reactions, some retrieval-based anticipation methods (Hao and Kong, 2025; Xu et al., 2024) retrieve historical examples with similar emotion states to the current utterance. Then, these methods anticipate the users’ future emotion reaction based on the next-turn emotion within these retrieved historical examples. However, as shown in Fig. 1, these methods focus solely on the emotion state in current turn, overlooking the users’ emotional evolution preceding the current turn.

We argue that it is significant to consider the users’ emotion evolution (i.e. emotion trajectory before the current turn) to capture the dynamic emotion needs of users and cross-turn emotional shifts. Thus, trajectory-based retrieval preserves dependencies between emotion states, which helps anticipate users’ future emotions.

In this paper, we propose emotion trajectory-aware retrieval for Markov-driven emotion anticipation in LLM-based ESC, aiming to anticipate users’ dynamic emotional reactions and thereby enhance strategy planning and empathetic response generation. Specifically, we construct a dynamic emotion memory to retrieve memories with similar emotional evolutions, facilitating LLMs to understand the user’s evolving emotions. Subsequently, we model the user’s emotional trajectory as a Markov process. We estimate emotion change to update the Markov transition matrix based on the retrieved emotion trajectory. Then, we anticipate the future emotion via the updated transition matrix, which guides the prior LLMs to generate candidate strategies tailored to the user’s emotional evolution. Furthermore, we design an active online learning

based on reinforcement learning, which performs weighted sampling of challenging user personas to boost the planner’s robustness on diverse users. Extensive experiments across two datasets on two backbones show that our methods significantly outperform all baselines. Our contributions are three-fold: (1) We propose emotion trajectory-aware retrieval for Markov-driven emotion anticipation in LLM-based ESC, which models emotion trajectories as a Markov processes to proactively guide the emotion change. (2) We design an emotion-driven strategy planning for ESC that integrates future emotion anticipation and introduces active online learning to boost the planner’s robustness for diverse users. (3) Extensive experiments across two datasets demonstrate that our method consistently outperforms other baselines across all metrics.

2 Related Work

2.1 Emotional Support Conversation

ESC aims to provide emotional support to users through multi-turn interactions (Liu et al., 2021). Recent research primarily focuses on strategy planning, which can be divided into three categories: (1) *Traditional deep learning* treats strategy planning as a strategy prediction task, generating the next strategy text span based on the current state (Zhao et al., 2023; Lin and Chen, 2025; Wan et al., 2025). (2) *LLM-based strategy planning* employs LLMs as planners to select strategies through prompt learning (Cao et al., 2024; Yu et al., 2023; Fu et al., 2023) or reinforcement learning (Wang et al., 2025b,a). (3) *External strategy planner* trains a lightweight model to optimize its strategy planning capabilities for specialized strategy planning (Deng et al., 2024; Zhang et al., 2024; Rakib et al., 2025). While these methods are effective, their planning is typically based on the current dialogue state, lacking consideration for the user’s long-term reactions.

Thus, some studies *incorporate long-term consideration* into strategy planning by anticipating future reactions (Cheng et al., 2022; He et al., 2024). He et al. (2025b) construct an intrinsic user world model to simulate personalized feedback. Yang et al. (2025) simulate future dialogues to train a future-oriented reward model. Rakib et al. (2025) employ LLMs to infer users’ potential reactions during strategy planning. However, they have not explicitly considered the evolution of emotions across multi-turn dialogue, thus exhibiting limitations in proactively guiding emotion changes.

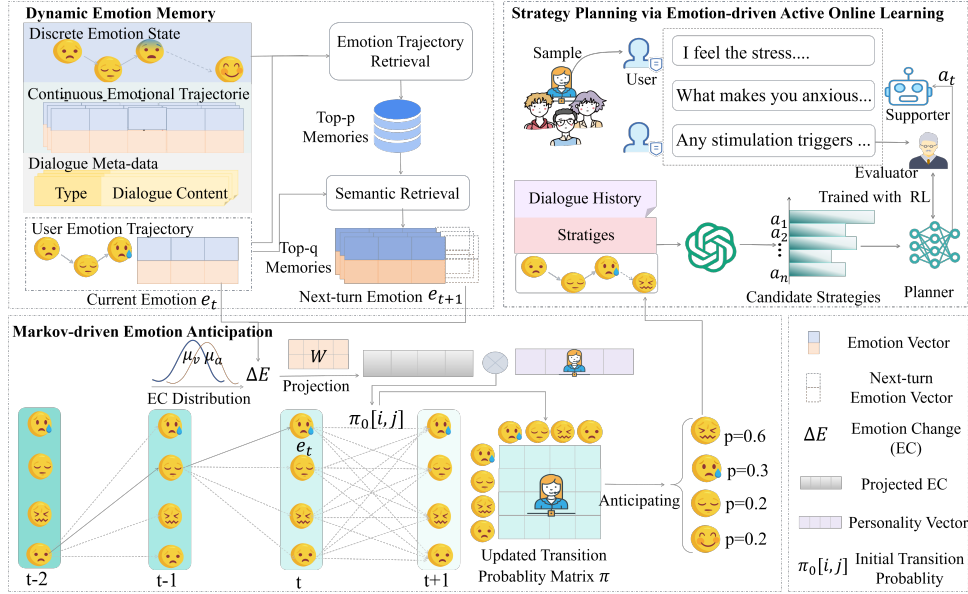


Figure 2: Overview of our framework. The top left is **dynamic emotion memory**, covering dialogue metadata, continuous emotion trajectory and discrete emotion state. The bottom is **Markov-driven emotion anticipation**, which estimates emotion change ΔE based on retrieved emotion trajectory and combines personality vector to update matrix to anticipate future emotion. The top right is **strategy planning via emotion-driven active online learning**, which uses anticipated emotion to generate candidate strategies and optimizes planner via active learning.

2.2 Retrieval-augmented Generation for Dialogues

Retrieval-Augmented Generation (RAG) enhances a dialogue system’s knowledge awareness by integrating external information, which mainly includes two categories: (1) *Integrating structured knowledge* (Tu et al., 2022; Yu et al., 2024; Xu et al., 2024; Hao and Kong, 2025) primarily incorporate knowledge graphs containing commonsense to enhance the model’s understanding of context and user intent. (2) *Incorporating personality information* (Cheng et al., 2023; Chen et al., 2025; Wu et al., 2025) focus on leveraging personality traits such as the Big-Five Personality (Goldberg, 1992) or communication styles (Norton, 1978; Zhao et al., 2024) to deliver personalized responses. However, these methods primarily rely on semantic to retrieve relevant knowledge, overlooking the dynamic evolution of emotions within dialogues, making it struggle to align with users’ dynamic emotional needs. Detailed related work is shown in App. A.

3 Method

3.1 Overview

We propose emotion trajectory-aware retrieval for Markov-driven emotion anticipation to track users’ dynamic emotions in conversations. As illustrated in Fig. 2, our framework contains three modules:

(1) **Dynamic emotion memory** stores users’ emotional trajectories from multi-turn dialogues and corresponding emotion states, and retrieves memories with similar emotional evolution based on emotional trajectories. (2) **Markov-driven emotion anticipation** estimates emotion change from retrieved emotional trajectories to update emotion transition probability, then anticipating the future emotion to guide strategy planning. (3) **Strategy planning via emotion-driven active online learning** steers the LLMs to generate candidate strategies based on the anticipated emotion, and optimizes the planner to select an appropriate strategy via active online learning.

Given user utterance as input, we retrieve historical memories (§ 3.2) to estimate emotion change, then anticipate the user’s future emotion (§ 3.3). Finally, guided by the anticipated emotion, we select a near-optimal strategy with the planner to guide the LLM to generate empathetic responses (§ 3.4).

3.2 Dynamic Emotion Memory

To enable the model to learn from historical dialogues with aligned emotional evolution, we construct a dynamic emotion memory $\mathcal{M} = \{(C^n, L^n, d_n, \mathcal{T}_n)\}$ comprising numerous items, where each item contains dialogue’s metadata (dialogue content d_n and problem type \mathcal{T}_n), the con-

tinuous emotional trajectory C^n , and discrete emotion state L^n . Retrieving similar dialogues enables the model to learn effective empathetic responses. However, such retrieval overlooks the dynamic evolution of emotions, which struggles to fulfill the natural progression logic of human emotions. To this end, we construct a dynamic emotion memory to provide supplementary information with relevant semantic and whole emotion evolution.

3.2.1 Dynamic Emotion Memory Extraction

To provide LLMs with emotion memories to draw from, we extract continuous emotion trajectories and discrete emotion states to construct dynamic emotion memory. The continuous emotional trajectory tracks cross-turn dynamic emotion evolution in dialogue through vector sequences. The discrete emotion state explicitly annotates emotion labels, providing intuitive representations of emotions.

Continuous Emotion Trajectory. Prior studies (Ohman et al., 2024; Teodorescu and Mohammad, 2023) demonstrate that lexicon-based methods achieve high reliability in capturing emotional dynamics in narratives. Thus, we adopt sliding window aggregation based on emotion lexicon to construct an emotional trajectory for each dialogue d_n from the UniConv training data (Jin et al., 2025)¹ with following steps: (1) For the t -th turn user utterance $usr_t^n = \{w_1, w_2, \dots, w_m\}$ in dialogue d_n , we slide a window (size ϵ , step size s) across the utterances to obtain window sequences $\mathcal{W}_{t,i}^n$, where $i = \{1, 2, \dots, k_t^n\}$ and k_t^n denotes the total number of windows. We select valid windows $\mathcal{W}_{t,i}^*$ from these windows based on the threshold θ for the proportion of emotional words. (2) For each valid window $\mathcal{W}_{t,i}^*$, we retrieve the Valence-Arousal-Dominance (VAD) vector of each word from the NRC_VAD emotion lexicon (Mohammad, 2025), which covers VAD vectors for numerous emotional words. Then, we calculate the mean values of all VAD vectors in the window to yield the window-level emotional vector $u_{t,i}^n = (v_{t,i}^n, a_{t,i}^n, d_{t,i}^n)$. (3) We compute the average of emotional vectors $u_{t,i}^n$ across all valid windows for the t -th turn to obtain the turn-level continuous emotional vector $e_t^n = \frac{1}{|k_t^*|} \sum_{i=1}^{|k_t^*|} u_{t,i}^n$. (4) We concatenate turn-level emotional vectors e_t^n in temporal order to construct the emotional trajectory $C^n = \{e_1^n, e_2^n, \dots, e_T^n\}$.

Discrete Emotion State. We discretize the continuous emotional space into a finite set of inter-

vals $\{E_1, E_2, \dots, E_K\}$ based on the valence and arousal² dimensions via the emotion circumplex model (Russell, 1980). Each interval corresponds to an emotional state, represented by a feature vector $E_i = (v_i, a_i)$. We calculate the Euclidean distance between the utterance’s emotional vector $e_t^n = (v_t^n, a_t^n)$ and the feature vector of each discrete emotion state. Then, we assign the utterance the discrete emotion E_i with the closest distance, which provides an explicit emotional anchor for empathetic response. Finally, we obtain the discrete emotion state sequence $L^n = \{E_1^n, E_2^n, \dots, E_T^n\}$.

3.2.2 Dynamic Emotion Memory Retrieval

To retrieve memories with similar emotional trajectories, we employ a hierarchical retrieval strategy to combine semantic retrieval and emotion trend correlation retrieval. This method first performs emotional trend matching to narrow the search scope, then retrieves memories from a refined subset based on the semantic relevance of metadata.

We follow two steps to retrieve historical memories in a hierarchical manner: (1) **Emotion trend correlation (ETC) retrieval** computes two Spearman correlation coefficients r_v and r_a . r_v and r_a are the correlation coefficients between the valence sequence and arousal sequence of the current dialogue and that of each historical dialogue in dynamic emotion memory \mathcal{M} , respectively. Then, we calculate $ETC = \sqrt{\frac{1}{2}(r_v^2 + r_a^2)}$, which quantifies the similarity in emotional trend between the two emotional trajectories, thereby retrieving the top- p memories \mathcal{M}_p based on sorted ETC . (2) **Semantic retrieval** employs an semantic encoder $f(\cdot)$ to map current users’ problem and each problem in \mathcal{M}_p to embeddings. Then, we calculate the cosine similarity $s_{\text{type}} = \cos(f(\widetilde{\mathcal{T}}_n), f(\mathcal{T}_n))$ to return the top- q most similar retrieval items \mathcal{M}_q .

3.3 Markov-driven Emotion Anticipation

To anticipate the future emotions, we model the emotion trajectory as a Markov chain and employ the retrieved memory with similar emotion trajectory to update the emotional transition probabilities. LLMs are typically trained to passively follow user instructions, which leads them to fail to proactively steer users toward positive emotional shifts. To this end, we propose a Markov emotional transition

²Arousal refers to the intensity of our emotional experience, whereas emotional valence refers to whether our emotional experience is positive or negative.

¹Details of the UniConv dataset are provided in App. B.

model to enable the models to shift from passive response to proactive adaptation.

3.3.1 Formulate Emotion Trajectory via Markov Chain

To discern trends in users’ emotions, we employ Markov chains to model emotional trajectories, driven by three advantages of Markov chains. First, Markov chains with a gradual nature capture the progressive evolution of human emotions to avoid an unreasonable shift (Cipresso et al., 2023). Second, the next state in Markov chains depends only on the current state. This aligns with the short-term emotion dependency of humans, where a user’s future emotional state is mainly influenced by their current emotional state. Finally, for an adjustable transition probability matrix in a Markov Chain, we can easily integrate personality traits and dynamic emotion memory to control emotion anticipation.

We treat discrete emotion states as finite state space $S = \{E_1, E_2, \dots, E_K\}$ of the Markov process. Thereafter, as shown in Eq. 1, we count the number of transitions $f_{i,j}$ from state E_i in turn t to E_j in turn $t + 1$ based on the historical emotion state sequence $L^n = \{E_1^n, E_2^n, \dots, E_T^n\}$, where $\mathbb{I}(\cdot)$ denotes the indicator function. We normalize $f_{i,j}$ using the softmax function to obtain the initial Markov emotion transition matrix π_0 .

$$f_{i,j} = \sum_{n=1}^N \sum_{t=1}^{T-1} \mathbb{I}(E_t^n = E_i \wedge E_{t+1}^n = E_j), \quad (1)$$

$$\pi_0[i, j] = \frac{\exp(f_{i,j})}{\sum_{j'} \exp(f_{i,j'})}. \quad (2)$$

3.3.2 Retrieval-Augmented Emotion Estimation

To estimate emotion change, we obtain future turn emotion based on retrieved emotional trajectory and update the transition probability matrix. This involves following steps:

Step 1: Emotional Adjustment Direction Estimation. We retrieve the top- q emotional trajectories as described in § 3.2.2, denoted as \mathcal{M}_q , and align turns between retrieved trajectories and current dialogue. Then we calculate the vector difference ΔE^i as emotion change between the current emotion state and the future emotion state within the trajectories in \mathcal{M}_q . The q independent emotion change ΔE are approximately formalized as a Gaussian distribution \mathcal{N} with mean $\mu = (\mu_v, \mu_a)$, which represents the ideal adjustment direction during emotion transitions. Thus, we use mean μ as

the adjustment. This probabilistic modeling balances multiple valid trajectories, yielding a robust data-driven transition.

Step 2: Personality-constrained Emotional Adjustment. We introduce a personality vector $P = \{p_j | j = 1, 2, \dots, K\}$, where K denotes the number of emotion states and P_j represents the current persona’s preference value for each emotion state. This constraint ensures the adjusted emotion remains consistent with the user’s traits, instead of generic transitions. Detailed initialization of the personality vector P is provided in the App. C.

Step 3: Updating Emotional Transition Probability Matrix. According to Eq. 3, we compute the dot product between the emotion change ΔE and the matrix $W \in \mathbb{R}^{2 \times K}$ composed of the feature vectors for K emotional states, thereby obtaining the projection of the emotion change onto each emotional state. Then, we integrate the personality vector and emotion changes into initial transition probability matrix π_0 to update the transition probability with the update rate of η . We normalize the updated matrix with row-wise softmax to maintain a valid probability distribution for each row.

$$\pi[i, j] = \pi_0[i, j] + \eta \cdot \Delta E \cdot W_j \cdot P_j. \quad (3)$$

3.3.3 Temporal Emotion Anticipation via Markov Chain

To anticipate potential future emotions, we leverage updated transition probability matrix to anticipate future emotions based on the Markov property, with the user’s current emotion state E_i as the initial state. We proceed as follows: First, we extract the representation $\pi[i, :]$ of current emotion state. Then, to align the transition with the user’s traits, we calculate a weighted average of the transition probability to obtain emotion vector for the future emotion $e_{t+1}^n = \sum_{j=1}^K P_j \cdot \pi[i, j]$, then mapping the emotion vector e_{t+1}^n to the nearest discrete emotion as user’s future emotion E_{t+1}^n based on the Euclidean distance. Finally, when the anticipated emotion is close to the user’s true emotion, we adaptively update personality vector to align with the user’s traits and improve emotional transition accuracy. The updated details are shown in App. D

3.4 Strategy Planning via Emotion-driven Active Online Learning

To select strategies suitable for the emotional evolution, we use the anticipated emotion to steer the strategy planning and optimize the planner via ac-

tive online learning based on multi-agent dialogue interaction. This involves the user LLM, supporter LLM, evaluator LLM and a language model combined with a deep Q-network (DQN) as the planner.

We optimize the planner with following steps: (1) *User Profile Sampling*. To address the limited behaviors of single-user simulators, we sample diverse user profiles for the user LLM. Detailed user persona design is provided in App. E.1. (2) *Strategy Planning*. To select strategies aligned with the emotional transition trend, we first leverage the anticipated emotion to guide the LLM to compress the strategies into high-quality candidate strategies \mathcal{A} , reducing decision-making complexity. The candidate strategy generation is detailed in App. E.2. Second, the planner estimates the strategy value and selects the strategy with the maximum value $a_t = \operatorname{argmax}_{a \in \mathcal{A}} Q^\theta(s_t, a)$. (3) *Response generation*. The supporter LLM generates an empathetic response s_{t+1} by incorporating the selected strategy. (4) *Goal-oriented rewards*. Consistent with prior studies (Deng et al., 2024; He et al., 2025b), the evaluator LLM assesses the goal achievement degree of current dialogue turn and yields a reward r_t . Detailed reward feedback and agreement analysis of evaluator LLM are shown in App. E.3 and G. (5) *Optimization*. We store the dialogue tuples (s_t, a_t, r_t, s_{t+1}) in the buffer, then sampling a mini-batch to minimize the action-value loss $\mathcal{L}(\theta)$ as Eq. 4 via temporal difference. To enhance the planner’s robustness, we introduce an active online training to optimize the planner. We assign equal initial sampling weights ω_i for all user personas, then dynamically adjust the weights of personas based on interaction outcomes (Eq. 5). This enables the planner to trial more times with challenging users.

$$\mathcal{L}(\theta) = \mathbb{E}_{(s_t, a_t, r_t, s_{t+1}) \sim B} [r_t + \gamma \max_{a' \in \mathcal{A}} Q^\theta(s_{t+1}, a') - Q^\theta(s_t, a)]^2, \quad (4)$$

$$\omega_i = \begin{cases} \omega_i - \mathbb{I}(i = j), & \text{if dialogue is successful,} \\ \omega_i + \mathbb{I}(i = j), & \text{otherwise.} \end{cases} \quad (5)$$

4 Experiments

4.1 Experiment Setting

Datasets. We focus on two datasets: **ESConv** (Liu et al., 2021) and **ExTES** (Zheng et al., 2023).

Baselines. We use following methods as baselines: **LLM-based prompt methods**, including

Standard, Proactive (Deng et al., 2023), ProCOT (Deng et al., 2023) and ICL-AIF (Fu et al., 2023).

External strategy planner-based methods, including PPDPP (Deng et al., 2024), TRIP (Zhang et al., 2024) and DialogXpert (Rakib et al., 2025).

Retrieval-based strategy planning methods, including DialogXpert w/ Retrieval, PRINCIPLES (Kim et al., 2025) and PRINCIPLES w/ Retrieval.

Evaluation Metrics. Consistent with prior studies (Deng et al., 2024; He et al., 2025a), we employ three dialogue-level metrics: **Success Rate (SR)** and **Soft Success Rate (SSR)** and **Average Turn (AT)**. Following Zhang et al. (2025a), we also employ **Fluency (Flu.)**, **Professionalism (Pro.)**, **Empathy (Emp.)** and **Helpfulness (Hel.)** to evaluate response quality.

Details of datasets, baselines, evaluation metrics and implementation are presented in App. F.

4.2 Main Results

The results in Tab. 1 show that our method surpasses all baselines across all metrics on two datasets, achieving improvements of 4.0%-7.5%, 4.0%-14.0%, and 4.7%-16.3% in SR, AT and SSR, respectively. Our method generates more professional and empathetic responses. We also report the results of human evaluation and time latency in App. J and App. K. LLM-based prompt planning methods generally achieve moderate SR while requiring more dialogue turns. Responses exhibit lower helpfulness and insufficient empathy. This demonstrates that pure prompts struggle to select effective support strategies. Overall, external planner-based methods outperform prompt-based approaches, demonstrating that a specialized planner improves strategy rationality. PRINCIPLES achieves further improvement by retrieving historical experiences, which provides the model with successful learning insights. Integrating our dynamic emotion memory into PRINCIPLES and DialogXpert yields significant improvements across most metrics. This indicates that historical scenarios with similar emotion trajectories better align with users’ dynamic emotional needs.

4.3 Ablation Study

Tab. 2 presents ablation studies of our proposed components. Replacing emotional trajectory retrieval (w/o Retrieval) with only semantic retrieval results in poor performance, highlighting the importance of retrieving memories with similar emo-

Table 1: Main results on two datasets. **Bold** denotes best results and underline indicates suboptimal results. Lower values of the AT indicate better performance, while higher values indicate better performance for all other metrics. $\Delta\%$ denotes the percentage difference between the best and the suboptimal results. D.X. denotes DialogXpert and PR. is PRINCIPLES. Our improvements are significant under the t-test with $p < 0.05$ (See details in App. H)

Dataset	Methods	GPT-4o								Qwen2.5-7B-Instruct							
		SR	AT	SSR	Flu.	Pro.	Emp.	Hel.	SR	AT	SSR	Flu.	Pro.	Emp.	Hel.		
<i>LLM-based prompt planning</i>																	
ESConv	Standard	64.00	8.10	61.20	3.75	3.88	3.35	3.88	20.77	9.63	36.00	3.98	3.85	3.52	2.03		
	Proactive	41.00	8.94	30.40	3.73	3.96	3.50	3.40	30.00	9.92	33.46	4.15	4.04	3.50	2.14		
	ProCOT	47.50	8.67	39.20	2.60	2.69	2.58	2.81	27.69	8.34	30.15	4.19	4.01	3.74	3.13		
	ICL-AIF	74.40	7.43	70.90	4.00	3.96	3.67	3.35	28.46	7.43	35.07	4.40	4.47	3.72	3.22		
	<i>External strategy planner</i>																
	PPDPP	70.90	7.65	66.40	3.56	4.01	3.39	3.34	46.92	7.55	59.61	4.63	4.65	4.03	3.32		
	TRIP	80.00	6.84	69.36	3.69	3.86	3.29	3.18	42.31	8.68	48.03	4.23	4.62	3.32	2.71		
	DialogXpert	78.46	<u>5.75</u>	<u>74.80</u>	4.33	3.93	3.47	3.45	76.92	6.70	61.60	4.28	4.65	3.57	2.82		
	<i>Retrieval-based strategy planning</i>																
	D.X. w/ Retrieval	<u>80.76</u>	6.06	73.30	4.56	4.37	3.89	2.96	<u>80.00</u>	<u>5.72</u>	<u>66.90</u>	4.64	<u>4.66</u>	3.74	3.20		
	PRINCIPLES	73.08	6.36	64.92	4.52	4.37	3.41	<u>3.60</u>	46.15	8.38	57.82	4.65	4.46	4.37	3.92		
	PR. w/ Retrieval	74.62	7.15	66.10	<u>4.59</u>	<u>4.61</u>	<u>3.95</u>	3.28	60.00	7.64	62.12	4.59	4.60	4.00	3.26		
Ours	86.23	5.13	79.53	4.70	4.90	4.40	3.68	85.38	5.32	74.75	4.72	4.70	4.48	3.97			
$\Delta\%$	5.47	6.20	4.73	2.20	5.80	9.00	1.60	5.38	4.00	7.85	1.40	0.80	2.20	1.00			
<i>LLM-based prompt planning</i>																	
EXTES	Standard	66.67	7.30	53.63	4.01	4.00	3.90	3.02	58.50	7.13	56.70	3.52	3.95	3.25	2.76		
	Proactive	65.38	7.82	54.40	4.23	3.93	3.54	2.03	39.50	8.45	30.05	4.34	4.21	4.21	2.32		
	ProCOT	66.92	7.58	50.81	3.85	1.82	3.82	3.04	56.00	6.85	47.3	4.68	4.66	<u>4.24</u>	2.71		
	ICL-AIF	71.54	7.37	47.40	4.62	4.13	<u>4.02</u>	2.95	52.00	8.56	48.79	4.58	4.73	3.96	3.43		
	<i>External strategy planner</i>																
	PPDPP	68.46	6.99	51.10	4.34	4.01	3.70	<u>3.30</u>	53.50	7.95	42.57	4.70	4.76	4.18	3.60		
	TRIP	74.50	6.97	65.55	4.56	4.33	3.44	2.68	51.50	7.75	59.69	4.11	4.38	3.30	2.78		
	DialogXpert	76.50	<u>5.79</u>	70.30	4.12	3.85	3.52	2.30	74.50	6.62	65.05	4.36	4.32	3.57	3.54		
	<i>Retrieval-based strategy planning</i>																
	D.X. w/ Retrieval	78.50	5.99	<u>73.55</u>	4.59	<u>4.75</u>	3.75	2.98	<u>80.00</u>	<u>6.31</u>	<u>68.10</u>	4.69	<u>4.78</u>	3.91	3.26		
	PRINCIPLES	83.00	5.95	67.90	4.45	4.65	3.28	3.25	58.50	7.99	59.05	4.69	4.70	3.91	3.65		
	PR. w/ Retrieval	<u>85.00</u>	6.25	68.76	<u>4.62</u>	4.70	3.76	3.20	76.50	6.70	63.49	<u>4.71</u>	<u>4.78</u>	3.80	3.06		
Ours	89.00	4.93	81.00	4.70	4.82	4.36	3.32	87.50	4.91	84.45	4.78	4.83	4.59	3.69			
$\Delta\%$	4.00	8.60	7.45	1.60	1.40	6.80	0.40	7.50	14.00	16.35	1.40	1.00	7.00	0.80			

Table 2: Ablation studies on the ESConv. w/o Retrieval replaces emotional trajectory retrieval with only semantic retrieval. w/o Anticipating removes the Markov emotion transition and directly prompts LLMs to anticipate emotions. w/o Active Learning samples user profiles with equal probability. w/o RL optimizes planner without RL. w/o PV discards the personality vector to update the transition probability matrix.

Methods	GPT-4o								Qwen2.5-7B-Instruct							
	SR	AT	SSR	Flu.	Pro.	Emp.	Hel.	SR	AT	SSR	Flu.	Pro.	Emp.	Hel.		
w/o Retrieval	76.92	5.23	76.03	4.57	4.50	3.78	2.82	70.77	6.58	65.61	4.46	4.45	3.67	3.01		
w/o Anticipating	76.15	5.24	67.55	4.67	4.79	4.20	3.62	70.46	5.56	60.81	4.70	4.56	3.77	3.68		
w/o Active Learning	85.15	5.16	75.34	4.64	4.87	4.21	3.60	76.15	5.84	70.99	4.65	4.00	3.88	3.78		
w/o RL	79.07	5.21	73.38	4.68	4.80	4.31	3.39	73.07	5.62	69.92	4.58	4.62	4.05	3.89		
w/o PV	81.53	5.25	76.73	4.30	4.81	4.21	3.56	70.23	5.55	68.00	4.16	4.66	3.98	3.70		
Ours	86.23	5.13	79.53	4.70	4.90	4.40	3.68	85.38	5.32	74.75	4.72	4.70	4.48	3.97		

tional evolution. w/o Anticipating discards the Markov emotion anticipation and directly prompts the LLM to infer future emotions. The results reveal remarkable performance drop, which indi-

cates that Markov chains effectively anticipate emotion. w/o Active Learning sampling user profiles with equal probability degrades the model’s performance, demonstrating that active learning fa-

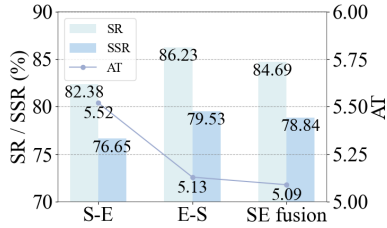


Figure 3: The results of different retrieval strategies. S-E denotes semantic first strategy and E-S indicates emotion first strategy. SE fusion denotes the combining semantic scores and ETC scores to rank retrieval results.

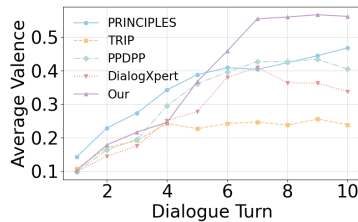


Figure 4: Average valence of user utterance across multi-turn dialogues under different methods.

cilitates the planner to adaptively learn strategy selection for diverse users. w/o RL without reinforcement learning to optimize planner yields varying degrees of performance degradation, confirming that using reinforcement learning enables more promising strategy selection. w/o PV discarding the personality vector leads to a notable performance decline, indicating that personality vectors effectively constrain the emotion transition to align with the user’s traits. Ablation study on EXTES dataset is shown in App. I.

4.4 Analysis of Planning Strategies

Table 3: Distances of strategies on same personas versus different personas.

Models	Intra-Persona↓	Inter-Persona↑
ICL-AIF	32.13	23.45
PPDPP	17.74	20.39
TRIP	19.97	22.69
PRINCIPLES	17.29	18.02
DialogXpert	25.02	23.11
Ours	13.10	29.88

In Tab. 3, we analyze the effectiveness of our strategy planning for diverse users. Following Zhang et al. (2024), we collect turn-level strategies to form sequential strategy sequences and encode them as an embedding vectors with BERT model

(Devlin et al., 2019), then reducing the dimensionality of vectors via the t-SNE. Finally, we calculate the average distance between strategy sequences of the same persona (Intra-Persona) and the average distance between strategy sequences of different personas (Inter-Persona) via Euclidean distance. The results show that our method achieves the lowest Intra-Persona and the highest Inter-Persona, enabling user-aware strategy planning.

4.5 Analysis of Emotion Memory Retrieval

Fig. 3 illustrates the results of three retrieval methods: **S-E** first retrieves the memories with the most similar semantic, then re-ranking according to ETC scores in § 3.2.2. **E-S** first retrieves memories with the most similar emotional evolution based on ETC scores, then re-ranking according to their semantic similarity. **SE fusion** combines semantic scores and ETC scores to rank and yield top retrieval results. The results reveal that E-S achieves the overall optimal performance, demonstrating that retrieval methods prioritizing emotional similarity are more effective in emotional support.

4.6 Visualization of Emotional Trajectories

The Fig. 4 illustrates the visualization of average valence of user utterances. We extract the emotional trajectory for each dialogue and calculate the average valence for each dialogue turn in the emotional trajectory. Valence indicates whether the current emotion is positive or negative, with higher values representing greater positivity (Cipresso et al., 2023). Results show that our method elevates users’ emotion to the highest positive level at the end of the conversation. While some methods show emotion decline in later turns, our approach effectively maintains positive emotion, indicating its advantage in sustained emotional support.

5 Conclusion

We propose an emotion trajectory retrieval for Markov emotion anticipation in ESC. We construct dynamic emotion memory to capture temporal emotion dynamics and retrieve memories via emotion trajectory retrieval. We model emotional trajectories as Markov chains and estimate emotion change via trajectory-aware retrieval for lookahead strategy planning. We further introduce active online learning to enhance the planner’s robustness for diverse users. Experimental results on two datasets show that our method excels all baselines.

Limitations

Despite the remarkable performance of our approach in the emotional support conversation, there are a few limitations that should be acknowledged:

First, our method focuses on text-based interactions, limiting its applicability in real-time, multi-modal scenarios, where audio, visual, and contextual cues play a crucial role in emotion support and response generation. Future work will try to integrate multi-modal emotion interaction, improving adaptability to diverse contexts.

Second, although our method achieves a relatively high success rate in delivering emotional support, it cannot guarantee consistent efficacy in alleviating users' negative emotions or providing actionable, effective guidance in all scenarios. We advise users not to rely entirely on our method for emotional support, while also considering the potential ethical considerations.

Ethics Statement

We acknowledge several critical ethical considerations to ensure its responsible and ethical deployment in practical scenarios. In this study, we employ publicly available datasets and pre-trained models and provide corresponding citations. All datasets are without incorporating any identifiable personal information. For the human evaluation involved in our work, we perform proactive ethical safeguards by explicitly informing all annotators in advance that the content to be labeled may contain expressions of negative emotions. As for the AI assistant, we utilize ChatGPT to identify textual errors and polish papers. Moreover, our method aims to provide convenient emotional support in daily conversational settings. This is not intended to replace professional human psychological counseling or clinical treatment services.

Acknowledgements

This work is supported by the following foundations: the National Natural Science Foundation of China under Grant No.62306330 and No.62376284.

References

M. S. Bartlett. 1937. [Properties of sufficiency and statistical tests](#). *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 160(901):268–282.

Yaru Cao, Zhuang Chen, Guanqun Bi, Yulin Feng, Min Chen, Fucheng Wan, Minlie Huang, and Hongzhi Yu. 2024. Enhancing emotional support conversation with cognitive chain-of-thought reasoning. In *CCF International Conference on Natural Language Processing and Chinese Computing*, pages 175–187. Springer.

Yirong Chen, Xiaofen Xing, Jingkai Lin, Huimin Zheng, Zhenyu Wang, Qi Liu, and Xiangmin Xu. 2023. Soulchat: Improving llms' empathy, listening, and comfort abilities through fine-tuning with multi-turn empathy conversations. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 1170–1183.

Zhuang Chen, Yaru Cao, Guanqun Bi, Jincenzi Wu, Jinfeng Zhou, Xiyao Xiao, Si Chen, Hongning Wang, and Minlie Huang. 2025. Socialsim: Towards socialized simulation of emotional support conversation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 1274–1282.

Jiale Cheng, Sahand Sabour, Hao Sun, Zhuang Chen, and Minlie Huang. 2023. Pal: Persona-augmented emotional support conversation generation. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 535–554.

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. *arXiv preprint arXiv:2210.04242*.

Pietro Cipresso, Francesca Borghesi, and Alice Chirico. 2023. Affects affect affects: A markov chain. *Frontiers in Psychology*, 14:1162655.

Kate Daley, Ines Hungerbuehler, Kate Cavanagh, Heloisa Garcia Claro, Paul Alan Swinton, and Michael Kapps. 2020. [Preliminary evaluation of the engagement and effectiveness of a mental health chatbot](#). *Frontiers Digit. Health*, 2:576361.

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*, pages 10602–10621, Singapore. Association for Computational Linguistics.

Yang Deng, Wenxuan ZHANG, Wai LAM, See-Kiong NG, and Tat-Seng CHUA. 2024. Plug-and-play policy planner for large language model powered dialogue agents.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. 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 4171–4186.

- Faiza Farhat. 2024. Chatgpt as a complementary mental health resource: a boon or a bane. *Annals of Biomedical Engineering*, 52(5):1111–1114.
- Yao Fu, Hao Peng, Tushar Khot, and Mirella Lapata. 2023. Improving language model negotiation with self-play and in-context learning from ai feedback. *arXiv preprint arXiv:2305.10142*.
- Lewis R Goldberg. 1992. The development of markers for the big-five factor structure. *Psychological assessment*, 4(1):26.
- Jiawang Hao and Fang Kong. 2025. Enhancing emotional support conversations: A framework for dynamic knowledge filtering and persona extraction. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 3193–3202.
- Tao He, Lizi Liao, Yixin Cao, Yuanxing Liu, Ming Liu, Zerui Chen, and Bing Qin. 2024. [Planning like human: A dual-process framework for dialogue planning](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4768–4791, Bangkok, Thailand. Association for Computational Linguistics.
- Tao He, Lizi Liao, Yixin Cao, Yuanxing Liu, Yiheng Sun, Zerui Chen, Ming Liu, and Bing Qin. 2025a. Simulation-free hierarchical latent policy planning for proactive dialogues. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 24032–24040.
- Tao He, Lizi Liao, Ming Liu, and Bing Qin. 2025b. Simulating before planning: Constructing intrinsic user world model for user-tailored dialogue policy planning. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 645–655.
- Le Huang, Hengzhi Lan, Zijun Sun, Chuan Shi, and Ting Bai. 2024. Emotional rag: Enhancing role-playing agents through emotional retrieval. In *2024 IEEE International Conference on Knowledge Graph (ICKG)*, pages 120–127. IEEE.
- Mengzhao Jia, Qianglong Chen, Liqiang Jing, Dawei Fu, and Renyu Li. 2024. Knowledge-enhanced memory model for emotional support conversation. In *MLACMH@ AAAI*.
- Zhengda Jin, Bingbing Wang, Geng Tu, and Ruifeng Xu. 2025. Overview of the nlpcc 2025 shared task 8: Personalized emotional support conversation. In *CCF International Conference on Natural Language Processing and Chinese Computing*, pages 512–520. Springer.
- Dongjin Kang, Sunghwan Mac Kim, Taeyoon Kwon, Seungjun Moon, Hyunsouk Cho, Youngjae Yu, Dongha Lee, and Jinyoung Yeo. 2024. Can large language models be good emotional supporter? mitigating preference bias on emotional support conversation. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15232–15261.
- Namyoun Kim, Kai Tzu-iunn Ong, Yeonjun Hwang, Minseok Kang, Iiseo Jihn, Gayoung Kim, Minju Kim, and Jinyoung Yeo. 2025. Principles: Synthetic strategy memory for proactive dialogue agents. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 21329–21368.
- Yi-Feng Lin and Arbee LP Chen. 2025. Predicting the following support strategy during the emotional support dialogue by integrating the dialogue history and supporter features. *Multimedia Tools and Applications*, 84(24):28353–28373.
- Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang, and Minlie Huang. 2021. [Towards emotional support dialog systems](#). 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 3469–3483, Online. Association for Computational Linguistics.
- Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.
- Saif M Mohammad. 2025. Nrc vad lexicon v2: Norms for valence, arousal, and dominance for over 55k english terms. *arXiv preprint arXiv:2503.23547*.
- Robert W Norton. 1978. Foundation of a communicator style construct. *Human communication research*, 4(2):99–112.
- Emily Ohman, Yuri Bizzoni, Pascale Feldkamp Moreira, and Kristoffer Nielbo. 2024. [EmotionArcs: Emotion arcs for 9,000 literary texts](#). In *Proceedings of the 8th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (LaTeCH-CLfL 2024)*, pages 51–66, St. Julians, Malta. Association for Computational Linguistics.
- Tazeek Bin Abdur Rakib, Ambuj Mehrish, Lay-Ki Soon, Wern Han Lim, and Soujanya Poria. 2025. Dialogxpert: Driving intelligent and emotion-aware conversations through online value-based reinforcement learning with llm priors. *arXiv preprint arXiv:2505.17795*.
- James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161.
- Yixuan Tang, Yi Yang, and Ahmed Abbasi. 2025. Personafuse: A personality activation-driven framework for enhancing human-llm interactions. *arXiv preprint arXiv:2509.07370*.
- Daniela Teodorescu and Saif Mohammad. 2023. [Evaluating emotion arcs across languages: Bridging the global divide in sentiment analysis](#). In *Findings of the Association for Computational Linguistics: EMNLP*

- 2023, pages 4124–4137, Singapore. Association for Computational Linguistics.
- Zhiliang Tian, Wei Bi, Dongkyu Lee, Lanqing Xue, Yiping Song, Xiaojiang Liu, and Nevin L. Zhang. 2020. [Response-anticipated memory for on-demand knowledge integration in response generation](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 650–659, Online. Association for Computational Linguistics.
- Quan Tu, Yanran Li, Jianwei Cui, Bin Wang, Ji-Rong Wen, and Rui Yan. 2022. Misc: A mixed strategy-aware model integrating comet for emotional support conversation. *arXiv preprint arXiv:2203.13560*.
- Chenwei Wan, Matthieu Labeau, and Chloé Clavel. 2025. Emodynamix: Emotional support dialogue strategy prediction by modelling mixed emotions and discourse dynamics. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 1678–1695.
- Xiaoyu Wang, Yue Zhao, Qingqing Gu, Zhonglin Jiang, Yong Chen, and Luo Ji. 2025a. Convert language model into a value-based strategic planner. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 6: Industry Track)*, pages 1444–1456.
- Yunxiao Wang, Meng Liu, Wenqi Liu, Kaiyu Jiang, Bin Wen, Fan Yang, Tingting Gao, Guorui Zhou, and Liqiang Nie. 2025b. Compeer: Controllable empathetic reinforcement reasoning for emotional support conversation. *arXiv preprint arXiv:2508.09521*.
- Shenghan Wu, Yimo Zhu, Wynne Hsu, Mong-Li Lee, and Yang Deng. 2025. From personas to talks: Revisiting the impact of personas on llm-synthesized emotional support conversations. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 5439–5453.
- Zhe Xu, Daoyuan Chen, Jiayi Kuang, Zihao Yi, Yaliang Li, and Ying Shen. 2024. Dynamic demonstration retrieval and cognitive understanding for emotional support conversation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 774–784.
- Ting Yang, Li Chen, and Huimin Wang. 2025. Towards open-ended emotional support conversations in llms via reinforcement learning with future-oriented rewards. *arXiv preprint arXiv:2508.12935*.
- Jiong Yu, Sixing Wu, Jiahao Chen, and Wei Zhou. 2024. Llms as collaborator: Demands-guided collaborative retrieval-augmented generation for common-sense knowledge-grounded open-domain dialogue systems. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 13586–13612.
- Xiao Yu, Maximillian Chen, and Zhou Yu. 2023. Prompt-based monte-carlo tree search for goal-oriented dialogue policy planning. *arXiv preprint arXiv:2305.13660*.
- Chao Zhang, Xin Shi, Xueqiao Zhang, Yifan Zhu, Yi Yang, and Yawei Luo. 2025a. Decoupledsc: Enhancing emotional support generation via strategy-response decoupled preference optimization. *arXiv preprint arXiv:2505.16995*.
- Tong Zhang, Chen Huang, Yang Deng, Hongru Liang, Jia Liu, Zujie Wen, Wenqiang Lei, and Tat-Seng Chua. 2024. Strength lies in differences! improving strategy planning for non-collaborative dialogues via diversified user simulation. *arXiv preprint arXiv:2403.06769*.
- Xinjie Zhang, Wenxuan Wang, and Qin Jin. 2025b. Intentionesc: An intention-centered framework for enhancing emotional support in dialogue systems. *arXiv preprint arXiv:2506.05947*.
- Suifeng Zhao, Tong Zhou, Zhuoran Jin, Hongbang Yuan, Yubo Chen, Kang Liu, and Sujian Li. 2024. [Awecita: Generating answer with appropriate and well-grained citations using llms](#). *DATA INTELLIGENCE*, 6(4):1134–1157.
- Weixiang Zhao, Xingyu Sui, Xinyang Han, Yang Deng, Yulin Hu, Jiahe Guo, Libo Qin, Qianyun Du, Shijin Wang, Yanyan Zhao, et al. 2025. Chain of strategy optimization makes large language models better emotional supporter. *arXiv preprint arXiv:2503.05362*.
- Weixiang Zhao, Yanyan Zhao, Shilong Wang, and Bing Qin. 2023. Transesc: smoothing emotional support conversation via turn-level state transition. *arXiv preprint arXiv:2305.03296*.
- Zhonghua Zheng, Lizi Liao, Yang Deng, and Liqiang Nie. 2023. Building emotional support chatbots in the era of llms. *arXiv preprint arXiv:2308.11584*.
- Nana Zhu, Caihai Zhu, Qingfu Zhu, and Yuanxing Liu. 2025. [Style rewriter: A three-stage approach for stylistic generation of open-domain conversational responses](#). *DATA INTELLIGENCE*, 7(2):397–415.

A Detailed Related Work

A.1 Emotional Support Conversation

ESC aims to provide emotional support to users through multi-turn interactions to alleviate their psychological stress. (Liu et al., 2021). Recent ESC research primarily focuses on strategy planning methods, which can be divided into three categories: (1) *Traditional deep learning* treats strategy planning as a strategy prediction task, generating the next strategy text span based on the current state (Zhao et al., 2023; Lin and Chen, 2025; Wan et al., 2025). Lin and Chen (2025) employ pre-trained models to encode history dialogue and integrate three supporter-related features from the dialogue to predict support strategies. Wan et al. (2025) use a heterogeneous graph to capture the discourse dynamics between user emotions and system strategies to predict strategies. (2) *LLM-based strategy planning* employs LLMs as planners to select strategies through prompt learning (Cao et al., 2024; Yu et al., 2023; Fu et al., 2023) or reinforcement learning (Wang et al., 2025b,a). Zhang et al. (2025b) prompt LLMs to infer the emotion intentions of users and map them to appropriate strategies. Wang et al. (2025a) serve LLMs as a value-based strategy planner and adopt Q-learning to optimize the ability to select the optimal strategy. (3) *External strategy planner* trains a lightweight model to optimize its strategy planning capabilities for specialized strategy planning (Deng et al., 2024; Zhang et al., 2024). Deng et al. (2024) fine-tune a strategy planner with supervised learning and reinforcement learning to predict the strategy for the next turn, and then they merge the strategy planner and LLMs to generate dialogue. While these methods are effective, their planning is typically based on the current dialogue state, lacking consideration for the user’s long-term reactions.

Thus, some studies *incorporate long-term consideration* into strategy planning by anticipating future reactions (Cheng et al., 2022; He et al., 2024; Tian et al., 2020). He et al. (2025b) construct an intrinsic user world model to simulate personalized feedback. Yang et al. (2025) utilize a multi-agent framework to simulate future dialogues to train a future-oriented reward model. Rakib et al. (2025) employ LLMs to infer users’ potential reactions during strategy planning. Although these methods perform strategy planning with a prospective perspective, they have not explicitly considered the evolution of user emotional states across multi-turn

dialogue, thus exhibiting limitations in proactively and precisely guiding emotion changes.

A.2 Retrieval-augmented Generation for Dialogues

Retrieval-Augmented Generation (RAG) enhances a dialogue system’s knowledge awareness by integrating external information, which mainly includes two categories: (1) *Integrating structured knowledge* (Tu et al., 2022; Yu et al., 2024; Xu et al., 2024; Zhao et al., 2024) primarily incorporate knowledge graphs containing commonsense to enhance the model’s understanding of context and user intent. (Hao and Kong, 2025) select commonsense knowledge aligned with the semantic of user discourse to enhance the empathetic response generation. Jia et al. (2024) collect knowledge-enriched dialogue context and select context-related concepts from ConceptNet to generate responses. (2) *Incorporating personality information* (Cheng et al., 2023; Chen et al., 2025; Zhu et al., 2025) focus on leveraging personality traits such as the Big-Five Personality (Goldberg, 1992) or communication styles (Norton, 1978) to deliver personalized responses. Cheng et al. (2023) extract personality information and then train the model to infer the user’s personality based on historical conversations, thereby generating responses consistent with that personality. Wu et al. (2025) incorporate persona traits into LLMs based on established psychological frameworks, then generating emotional support responses. However, these methods primarily rely on semantic matching to retrieve relevant knowledge. Semantic retrieval overlooks the dynamic evolution of emotions within dialogues, making it struggle to align with users’ dynamic emotional needs.

B Details of UniConv Dataset

We collect the UniConv dataset³ (Jin et al., 2025) for personalized emotional support conversation and construct dynamic emotion memory based on dialogues from training data. This dataset covers a wide range of real-world personal scenarios and contains rich content. This dataset consists of multi-turn dialogues generated by a multi-agent role-playing framework. After quality filtering and manual verification, it comprises 1,500 dialogues: 1,200 for the training set, 150 for the validation set, and 150 for the test set. Thus, our dynamic emotion

³<https://github.com/Jin-zd/NLPCC-2025-Shared-Task8>

memory covers 1,200 dialogues from training set. Specifically, each dialogue consists of 10 turns, covering user profiles, emotion types, problem types and situation.

C Personality Vector Initialize

Motivated by (Tang et al., 2025; Huang et al., 2024), we initialize the personality vector P for each user persona by prompting LLMs. Specifically, we provide detailed descriptions of the user persona’s background, preferences, and behaviors, and prompt the LLM to output structured weights that represent the degree of match between the persona personality and different emotion states. Subsequently, we normalize the values of the personality vector to constrain the dynamic emotion transition according to user traits, thereby maintaining emotional coherence. The initialization results of the personality vector for each user persona are illustrated in Fig. 5.

D Adaptive Update of Personality Vector

To align the personality vector’s matching feature for different emotions with the user’s traits, we adaptively adjust the personality vector P when the anticipated emotion is close to the user’s true emotion. During training, we take the user’s real-time utterance as input and employ a classifier to identify the true emotion. As shown in Eq. 6, we strengthen the dimension j_{true} in the personality vector P that matches the user’s true emotions and weaken the weights of the mismatched dimensions. This adaptively updates the personality vector P to enhance alignment with the user’s personality traits. To ensure optimization stability, we set a minimum weight threshold $P_j \geq \tau$ ($\tau = 0.05$) to prevent users from failing to express a particular emotion state. Additionally, we maintain the sum of weights at 1. This data-driven optimization approach enhances emotional transition accuracy while preserving consistent character personality.

$$P_j = \begin{cases} P_j + \alpha(1 - P_j), & \text{if } j = j_{true}, \\ P_j - \alpha P_j, & \text{otherwise.} \end{cases} \quad (6)$$

E Active Online Learning

E.1 User Persona Designing

To simulate the diverse users in real-world scenarios, we design user personas based on three key dimensions of the Big Five personality model

(Goldberg, 1992): extroversion, openness, and neuroticism. Each dimension comprises two distinct attributes: [Extravert, Introvert], [Adventurous, Conservative], and [Rational, Neurotic]. By combining attributes across these dimensions, we design eight distinct user types. Similar to (Zhang et al., 2024), we employ LLMs to generate eight detailed user profile descriptions for each persona type with instructions, encompassing occupation, background, age, and other relevant information. The detailed instructions for generating profile description are shown in the Tab. 4. Ultimately, we obtain a total of 64 user profile descriptions, where 40 profile descriptions are used for the user simulator during the training process, 8 profile descriptions are allocated for validation, and 8 profile descriptions are reserved for testing.

E.2 Candidate Strategy Generation

To select strategies aligned with the user’s current emotional evolution, we utilize anticipated emotions to guide the LLM to generate a high-quality candidate strategy set. This compresses the action space into a small number of high-quality candidate strategies, thereby reducing decision-making complexity. Given a predefined action space \mathcal{A}' , we employ a prior LLM to infer potential strategies based on the current conversational state s_t and anticipated future emotional state E_{k+1}^n , generating strategies as free-form text. Then, we use rule-based functions \mathcal{P} to parse open text strategies into valid actions. As shown in Eq. 7, we construct an action probability distribution $p(a|s_t, E_{k+1}^n)$ using the logit probabilities of the generated text, selecting the top- k strategies that are semantically coherent with the current state and align with the user’s dynamic emotional evolution as the candidate strategy set \mathcal{A} . This narrows the action space into a small number of high-quality candidate strategies.

$$p(a|s_t, E_{k+1}^n) = \sum_{o: \mathcal{P}(o)=a} LLM(o | s_t, E_{k+1}^n, \mathcal{A}') \quad (7)$$

E.3 Evaluation Workflow

Consistent with (Deng et al., 2024; Kim et al., 2025), for each dialogue turn, we employ GPT-4o-mini with temperature=1.0 as an evaluator to determine the completion status of the emotional support goal. The evaluator classifies the each dialogue turn state based on the dialogue history and user emotional trajectory into one of four categories:

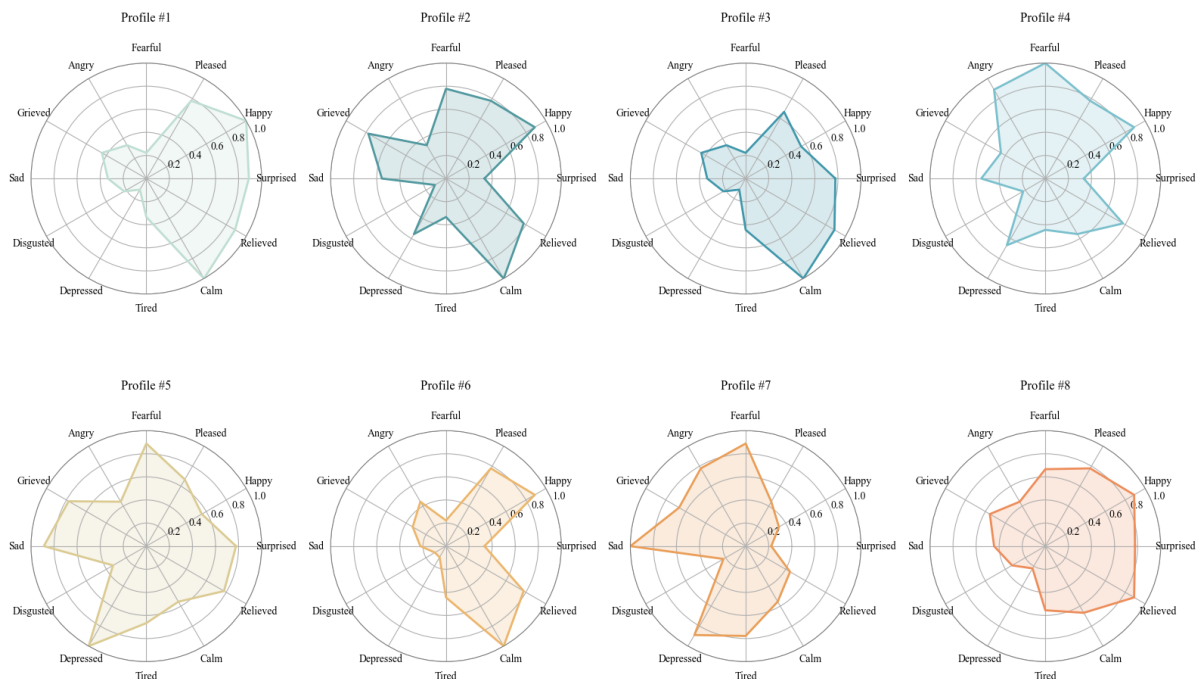


Figure 5: Visualization of initialization weights for personality vectors.

Table 4: The instruction used to generate the detailed user profile description.

You need to incorporate the following persona attributes and generate a cohesive persona description.
 Note: Each persona description must include two core elements:
 1. Identity background: A plausible fictional identity (e.g., ‘35-year-old male entrepreneur’, ‘29-year-old female teacher’);
 2. Personality trait interpretation: Translating attributes into specific behaviours (e.g., ‘Neurotic personality → Prone to anxiety under pressure; requires reassurance when seeking emotion support’);
 Personality: *[User Persona]*
 An Example:
 You are a 28-year-old female software developer. You are extrovert, which means you can express yourself clearly. Your personality characterized by adventurous to experience, which means you are willing to try new things. And you are rational, in your occupation, you excel at analyzing problems and finding logical solutions.
 Persona description: ...

[solved, better, same, worse]. Each category corresponds to a scalar reward value [1.0, 0.5, -0.5, -1.0]. To ensure the robustness of evaluation, we perform 10 independent judgment for each dialogue turn and calculate mean scores as the reward value r_t . We consider the dialogue goal to be fulfilled when the reward value r_t exceeds a threshold of 0.5.

E.4 Prompts for Multi-agent Interaction

In multi-agent interactions, we prompt the LLMs to engage in dialogues as a user, supporter, and evaluator, as illustrated in the Fig. 6, Fig. 7 and Fig. 8.

F Experiment Setting Details

F.1 Dataset Details

We conduct experiments on two emotional support conversation datasets:

- **ESConv dataset (Liu et al., 2021)**: This consists of 1,300 crowd-sourced dialogues and annotations of user problem types, emotion types, and situation descriptions. This dataset contains 8 distinct emotional support strategies. For the ESConv dataset, we adopt its official data split.
- **ExtES dataset (Zheng et al., 2023)**: This comprises 11,177 dialogues generated by ChatGPT and validated by human annotators,

covering 16 emotional support strategies. We partition this dataset into train/dev/test using a random 8:1:1 split ratio.

F.2 Baseline Details

We compare our approach with four **LLM-based prompt methods**:

- **Standard**: This directly invokes LLMs for zero-shot response generation without strategies.
- **Proactive** (Deng et al., 2023): This instructs LLMs to select the most appropriate strategy to guide response generation.
- **ProCOT** (Deng et al., 2023): This introduces a chain-of-thought reasoning into the strategy planning.
- **ICL-AIF** (Fu et al., 2023): This provides real-time feedback to the LLMs to determine dialogue strategy before response generation.

We also compare our approach with three **external strategy planner-based methods**:

- **PPDPP** (Deng et al., 2024): This designs a specialized strategy planner to select dialogue strategies.
- **TRIP** (Zhang et al., 2024): This trains a strategy planner by introducing diverse user simulators.
- **DialogXpert** (Rakib et al., 2025): This integrates prior LLMs with a trainable Q-adaptor to learn strategy planning.

We also compare our approach with three **retrieval-based strategy planning methods**:

- **DialogXpert w/ Retrieval**: This integrates our dynamic emotion memory and generates empathetic responses by retrieving similar examples based on emotion trajectories.
- **PRINCIPLES** (Kim et al., 2025): This constructs a synthetic strategy memory bank and retrieves similar principles to guide strategy planning.
- **PRINCIPLES w/ Retrieval**: This replaces the static strategy memory bank with our dynamic emotion memory, and retrieves examples with similar emotional evolution to generate response.

F.3 Details of Evaluation Metrics

Consistent with prior studies (Deng et al., 2024; He et al., 2025a), we employ three dialogue-level metrics: **Success Rate (SR)**, **Average Turn (AT)**, and **Soft Success Rate (SSR)**. SR evaluates the performance of achieving a goal and is defined as the ratio of the number of successful cases to the total number of test cases. In practice, we employ LLM-based evaluator to assess dialogue completion status and return rewards. Rewards exceeding the threshold of 0.5 are deemed successful. AT measures the efficiency of goal completion by calculating the average dialogue turns of all test cases. SSR complements the SR to more accurately evaluate effectiveness by averaging final turn rewards for all test cases. Following Zhang et al. (2025a), we also measure the quality of generated response following four metrics: **Fluency (Flu.)**, **Professionalism (Pro.)**, **Empathy (Emp.)** and **Helpfulness (Hel.)**. We employ GPT-4o with temperature=0 to evaluate the response on these four metrics. Specifically, we prompt LLMs three times using a 5-point Likert scale instruction and calculate the average scores. The detailed instructions are presented in the Table 8.

F.4 Implementation Details

We conduct all experiments on an A800 GPU with 80GB memory. In § 3.2.1, we adopt the NRC VAD Lexicon v2 (Mohammad, 2025), covering over 55,000 English words. We set the sliding window size to 10 and the stride to 1 to construct dynamic emotion memory. We filter out invalid windows with an emotional word ratio below 0.2. In the Markov emotion anticipation, we set 12 distinct emotion states. Each emotional state coordinates in the valence-arousal space is shown in the App. F.5. We update the transition matrix with a learning rate of 0.5 and update the personality vector with a learning rate of 0.1. For the emotional trajectory retrieval, the numbers p and q of retrieved examples are 10 and 5, respectively. Consistent with prior research (Kim et al., 2025), for the multi-agent interaction, we employ GPT-4o and Qwen2.5-7B-Instruct as backbones to simulate supporters, and consistently adopt GPT-4o to simulate users. We also prompt GPT-4o-mini to serve as an evaluator to determine the completion status of the emotional support goal. The agreement analysis of evaluator LLM is shown in App. G. More detailed prompts for user LLM, supporter LLM and evalu-

Table 5: The p values of t-test on our method with baselines. The p values are all smaller than 0.05, indicating the significance of our improvements.

Datasets	t-test						
	SR	AT	SSR	Flu.	Pro.	Emp.	Hel.
ESConv	4.80e-7	1.00e-5	1.00e-8	1.76e-2	9.18e-6	1.45e-6	1.65e-5
EXTES	4.01e-12	1.32e-6	1.50e-7	6.32e-4	1.96e-5	3.79e-6	9.02e-3

Table 6: Human evaluation results comparing our methods with the baselines across five metrics.

Ours vs.	Fluency		Empathy		Suggestion		Identification		Overall	
	Win	Lose	Win	Lose	Win	Lose	Win	Lose	Win	Lose
Standard	80%	20%	78%	22%	60%	40%	74%	26%	90%	10%
Proactive	100%	0%	74%	26%	72%	28%	98%	2%	94%	6%
ProCOT	92%	8%	70%	30%	60%	40%	64%	36%	86%	14%
ICL-AIF	78%	22%	80%	20%	72%	28%	82%	18%	84%	16%
PPDPP	94%	6%	94%	6%	80%	20%	78%	22%	88%	12%
TRIP	70%	30%	78%	22%	72%	28%	72%	28%	82%	18%
PRINCIPLES	72%	28%	94%	6%	62%	38%	60%	40%	72%	28%
PRINCIPLES w/ retrieval	60%	40%	60%	40%	54%	46%	68%	32%	64%	36%
DialogXpert	76%	24%	82%	18%	70%	30%	56%	44%	74%	26%
DialogXpert w/ retrieval	70%	30%	84%	16%	68%	32%	60%	40%	70%	30%

ator LLM are presented in the Fig. 6, Fig. 7 and Fig. 8. We employ a pretrained RoBERTa⁴ (Liu et al., 2019) followed by a lightweight adaptor network (4-layer MLP) as the plug-and-play strategy planner and train the Q-network for 5 epochs with a learning rate of 1e-6, a discount factor of 0.999, and a maximum conversation turn of 10.

F.5 Discrete Emotional Coordinates

Following Russell (1980), as shown in Tab. 7, we establish the coordinates of 12 emotional states in the emotional space based on the twelve equal divisions of the unit circle, ensuring each state aligns with the definition of the emotion circumplex model. This aims to construct an ideal emotional space for emotion states, facilitating a Markov chain modeling and guaranteeing the smoothness and controllability of Markov transitions.

G Agreement Analysis

To validate the reliability of the LLM-based evaluator, we perform an alignment analysis against human preferences. Specifically, we randomly sample 100 dialogues from the ESConv dataset. Then, we require three human annotators with a psychology background, as well as the LLM-based evaluator, to independently evaluate whether the patient’s emotional issues had been resolved, adhering to

⁴<https://huggingface.co/FacebookAI/roberta-base/tree/main>

Table 7: valence-arousal coordinates for each emotion state in the valence-arousal space.

Emotion State	Coordinate (Valence, Arousal)
Surprised	(0.383, 0.924)
Happy	(0.707, 0.707)
Pleased	(0.924, 0.383)
Fearful	(-0.383, 0.924)
Angry	(-0.707, 0.707)
Grieved	(-0.924, 0.383)
Sad	(-0.924, -0.383)
Disgusted	(-0.707, -0.707)
Depressed	(-0.383, -0.924)
Tired	(0.383, -0.924)
Calm	(0.707, -0.707)
Relieved	(0.924, -0.383)

unified assessment criteria. Our analysis demonstrates strong inter-rater agreement among human annotators (Krippendorff’s $\alpha = 0.792$; Spearman’s $\rho = 0.784$). Additionally, the LLM-based evaluator exhibits a high degree of alignment with the human majority vote (Krippendorff’s $\alpha = 0.806$; Spearman’s $\rho = 0.789$, $p < 0.0001$). These findings indicate that the feedback generated by the LLM evaluator is highly consistent with the judgments of human evaluators.

H Statistic Significance Test

We conduct a t-test (Bartlett, 1937) to examine whether the improvement in our method is statistically significant. The p -values in Tab. 5 are all less than 0.05, indicating that our improvement is statistically significant.

Table 8: Response quality evaluation criteria across four dimensions.

Dimension	Description	Criterion	Score
Fluency	Fluency evaluates whether language expression is natural, coherent, and comprehensible.	1.1 Unclear expression; patient may struggle to grasp the meaning.	1
		1.2 Some parts are confusing, though the main point can be inferred.	2
		1.3 Mostly clear and coherent with minor ambiguities.	3
		1.4 Fluent and well-structured; logically organized and easy to follow.	4
		1.5 Concise and impactful language; precise and elegant communication that conveys ideas efficiently.	5
Professionalism	Professionalism evaluates whether the model demonstrates psychological knowledge, follows ethical principles, and avoids misleading or inappropriate advice.	2.1 Shows serious misunderstanding or misuse of psychological concepts, or provides inappropriate advice. .	1
		2.2 Minor factual inaccuracies or advice that lacks evidence, but does not pose direct harm.	2
		2.3 No major errors; advice is acceptable and somewhat aligned with psychological principles.	3
		2.4 Demonstrates solid understanding of psychological concepts and appropriate intervention techniques.	4
		2.5 Highly professional, reflects strong psychological insight, maintains boundaries, and communicates in a grounded, ethical manner.	5
Empathy	Empathy evaluates whether the model genuinely understands the user’s emotions, expresses care, and provides emotional support.	3.1 Fails to provide emotional comfort or assist the patient in analyzing their problems.	1
		3.2 Either lacks emotional comfort or fails to support problem analysis.	2
		3.3 No significant issues, but empathy and analysis remain surface-level.	3
		3.4 Demonstrates a warm, human-like tone—like a friend—offering both emotional relief and analytical support.	4
		3.5 Deep emotional insight with sincere and stable empathy, conveyed through attentive and flexible language.	5
Helpfulness	Helpfulness evaluates the effectiveness of an AI assistant’s suggestions by considering both the number of recommendations provided per interaction and the relevance or usefulness of each suggestion in addressing the user’s question.	4.1 Irrelevant, misleading, or potentially harmful suggestions.	1
		4.2 Ineffective or generic advice that does not respond to the patient’s needs.	2
		4.3 Weakly relevant suggestions with limited practical value.	3
		4.4 Somewhat helpful; suggestions are relevant and usable, but some of them are ineffective.	4
		4.5 There are many clear and practical suggestions that align well with the user’s issue, and all of them are effective.	5

I Detailed Ablation Study

As shown in Tab. 9, we investigate the contribution of our proposed components on the EXTES dataset. Replacing emotional trajectory retrieval (w/o Retrieval) with semantic retrieval results in poor performance, specifically lower empathy and helpfulness for responses, highlighting the importance of retrieving memories with similar emotional evolution. w/o Anticipating discards the Markov emotion anticipation and directly prompts the LLM to infer future emotions based on the user’s current state. The results reveal noticeable performance drop across all metrics on both models, which indicates that Markov chains effectively capture temporal dependencies in dynamic emotional trajectories, providing clear evolutionary logic for emotion anticipation. w/o Active Learning sampling

user profiles with equal probability distribution during training degrades the emotional support performance. This demonstrates that active online training facilitates the planner to adaptively learn strategy selection for challenging users, enhancing its robustness. w/o RL without reinforcement learning to optimize planner yields varying degrees of performance degradation, confirming that using reinforcement learning to optimize an external strategy planner enables more promising strategy selection.

J Human Evaluation

J.1 Instruction for Human Evaluation

We recruit six annotators with backgrounds in psychology and compare the responses generated by our method and the baseline following these in-

User Prompt

Now enter the role-playing mode. In the following conversation, you will play as a patient in a counseling conversation with a therapist. You are seeking help from the therapist, because you are dealing with emotional issues related to *[emotion_type]* regarding *[situation]*.

Your Persona:...

You must follow the instructions below during chat.

1. Your utterances and seeking behavior need to strictly follow your persona. Varying your wording and avoid repeating yourself verbatim!
2. You can decide to your emotion state flexibly based on your persona and the conversation. If you feel support and the issue has been resolved, please indicate that the emotion issue has been settled.

Full conversation:...

The therapist just said: ...

Express how you feel in a natural, emotional way. Please reply with only one short and succinct sentence.

Figure 6: User prompt in the multi-turn emotional support conversation dialogues.

Supporter Prompt

You are role playing as a therapist in a counseling conversation with a patient. Your goal is to help the patient resolve their emotional issues and assist them in understanding and working through their challenges.

This is an example dialogue: *****[example]*****

Current conversation: ...

The patient just said: ...

Based on the Patient’s input, follow these steps:

1. Read the dialogue between the therapist and patient in the example to identify how the therapist facilitates problem-solving and provides support for the patient.
2. The anticipated next emotion about patient is *[next_emotion]*. You should provide a response that helps mitigate this potential negative emotion state (or move towards or maintain positive emotion state) and offers support.
3. Respond the patient with the strategy: *[strategy]*. You can refer to the example when replying to the patient, including emotions and tone.

Your response must start with 'Response:' and be a short, concise sentence, no extra explanation.

(Example: Response: ...)

Figure 7: System prompt in the multi-turn emotional support conversation dialogues.

Table 9: Ablation studies on the EXTES. w/o Retrieval replaces emotional trajectory retrieval with only semantic retrieval. w/o Anticipating removes the Markov emotion transition and directly prompts LLMs to anticipate emotions. w/o Active Learning samples user profiles with equal probability. w/o RL optimizes planner without RL.

Methods	GPT-4o								Qwen2.5-7B-Instruct							
	SR	AT	SSR	Flu.	Pro.	Emp.	Hel.	SR	AT	SSR	Flu.	Pro.	Emp.	Hel.		
w/o Retrieval	76.88	5.36	75.50	4.05	4.47	4.01	3.21	74.50	5.35	76.00	4.10	4.49	3.99	3.47		
w/o Anticipating	78.39	5.40	76.10	4.23	4.69	4.12	3.05	76.38	5.34	75.41	4.21	4.76	4.03	3.39		
w/o Active Learning	79.90	5.17	79.44	4.26	4.74	4.11	3.15	79.39	5.38	80.70	4.12	4.70	4.06	3.44		
w/o RL	80.21	5.09	76.65	4.17	4.61	4.14	3.19	82.00	5.12	74.60	4.40	4.56	4.01	3.52		
Ours	89.00	4.93	81.00	4.70	4.82	4.36	3.32	87.50	4.91	84.45	4.78	4.83	4.59	3.69		

structions:

smooth sentence flow?

- Fluency: Which model produces the response that is more natural, well-structured, and logically coherent, with correct grammar and
- Empathy: Which model responds with greater emotional understanding, showing warmth, compassion, or concern that aligns with the

Evaluator Prompt
<p>The following is a conversation:...</p> <p>User’s emotion history:...</p> <p>Question: What’s the Patient’s current emotion state?</p> <p>Requirement: You can ONLY reply with ONE of the following 4 descriptive levels WITHOUT EXPLANATION:</p> <p>Solved: The patient’s feelings show clear positive changes, indicating that the issue has been resolved.</p> <p>Same: The patient’s feelings remain unchanged.</p> <p>Better: The patient feels slightly better, but a small improvement.</p> <p>Worse: The patient feels an increase in tension and is more worried, with noticeable decline.</p>

Figure 8: Evaluator prompt in the multi-turn emotional support conversation dialogues.

user’s feelings?

- Identification: Which model explores the user’s situation more effectively to identify the problem?
- Suggestion: Which model offers more relevant, practical, and emotionally sensitive suggestions that could help the user cope or take action?
- Overall: Overall, which model provides better emotional support, considering empathy, understanding, helpfulness, and communication quality?

J.2 Results Analysis of Human Evaluation

Consistent with prior studies (Deng et al., 2024; Yang et al., 2025), we conduct a human evaluation on 100 randomly sampled dialogues from the ES-Conv dataset. Three annotators with psychology backgrounds are recruited to compare our method’s generated responses against baselines across five dimensions. As shown in Tab. 6, our approach outperforms all baselines on all human evaluation metrics. This demonstrates our method generates more effective emotional support responses, particularly excelling in empathy, fluency, problem identification, and providing suggestions to help users overcome emotional challenges.

K Comparison of Time Consumption

Tab. 10 presents the time consumption of various baselines and our proposed method when deployed on the GPT-4o model. To ensure a fair comparison, mitigating effects from variations in generated text length, we report two metrics: (1) the average

inference latency (in seconds) required to generate per token, (2) the relative latency ratio of each method normalized to the Standard baseline. Our results demonstrate that our method incurs only marginal overhead in inference time. Specifically, its per-token latency is merely 0.004 seconds higher than the average latency of the Standard baseline, which is well within acceptable limits for practical deployment.

Table 10: Time consumption of baselines and our method on GPT-4o.

Methods	Seconds per Token	Relative Ratio
Standard	0.023	1×
Proactive	0.029	1.26×
ProCOT	0.026	1.13×
ICL-AIF	0.022	0.96×
PPDPP	0.025	1.09×
TRIP	0.026	1.13×
DialogXpert	0.024	1.04×
D.X. w/ Retrieval	0.026	1.13×
PRINCIPLES	0.068	2.95×
PR. w/ Retrieval	0.030	1.30×
Ours	0.027	1.17×