STAR: Strategy-Aware Refinement Module in Multitask Learning for Emotional Support Conversations

Anonymous ACL submission

Abstract

Providing effective emotional support requires strategic approaches because it is inherently complex and should account for the diverse situations and needs of each individual. The Emotional Support Conversation framework structures interactions into three phases—exploration, comforting, and action—guiding strategy selection for response generation. Although multitask learning has been used to jointly optimize strategy prediction and response generation, it often suffers from task interference, where conflicting learning objectives hinder optimization. To address this, we propose the **St**rategy-**A**ware **R**efinement Module (**STAR**), which separates and selectively integrates the decoder's hidden states for strategy prediction and response generation through a gating mechanism. This approach preserves task-specific representations while enabling adaptive information exchange, thereby mitigating interference. Experimental results demonstrate that STAR effectively reduces task interference and achieves state-of-the-art performance in both strategy prediction and supportive response generation.

1 Introduction

Emotional support aims to alleviate individual emotional distress while helping individuals understand and resolve their problems (Burleson, 2003; Langford et al., 1997; Heaney and Israel, 2008). However, providing effective emotional support is not inherently intuitive (Burleson, 2003). To address this challenge, the Emotional Support Conversation (ESC) framework is structured into three distinct phases: Exploration, Comforting, and Action (Liu et al., 2021). Within each phase, the strategic selection of appropriate strategy mechanisms and the generation of targeted responses facilitate a more systematic approach to alleviating users' emotional distress.

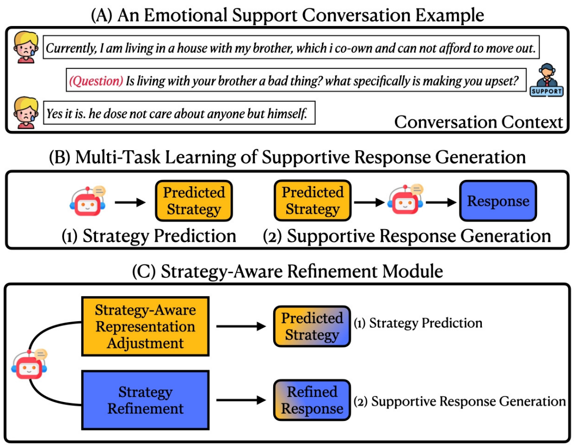


Figure 1: (A) shows an emotional support conversation example, highlighting the dual tasks of strategy prediction and supportive response generation. (B) illustrates the multi-task learning framework, and (C) presents the STAR module that refines hidden representations to mitigate task interference.

Motivated by the findings in medical research, recent AI systems for emotional support have adopted multitask learning (MTL) to simultaneously select appropriate strategies and generate supportive responses (Tu et al., 2022; Zhou et al., 2023; Peng et al., 2022; Cheng et al., 2022; Zhao et al., 2023; Deng et al., 2023; Xu et al., 2024; Li et al., 2024).

However, while the MTL approach is designed to leverage shared information across tasks to enhance learning efficiency, it can sometimes lead to adverse effects (Zhao et al., 2018). This issue arises due to task interference, where the representational requirements of different tasks may be inherently misaligned (Gurulingan et al., 2022a), or when conflicting gradients from multiple tasks disrupt the optimization process during backpropagation (Yu et al., 2020). As a result, instead of facilitating knowledge transfer, MTL can sometimes hinder model performance by introducing conflicts between tasks.

To mitigate task interference, various approaches have been proposed, including independent subnets to isolate task-specific representations (Strezoski et al., 2019), task-specific parameterization to adjust model capacity per task (Kanakis et al., 2020), and task grouping to cluster related tasks and reduce negative transfer (Gurulingan et al., 2022b). However, despite these advancements, effective interference suppression strategies tailored to the Emotional Support Conversation (ESC) domain—particularly for response strategy selection and supportive response generation—remain an open challenge.

To address these limitations, we propose the Strategy-Aware Refinement (STAR) module, which effectively mitigates task interference between strategy prediction and supportive response generation while leveraging contextual and strategic cues. STAR consists of two key components: Strategy-Aware Representation Adjustment (SARA) and Strategy Refinement (SR). Specifically, SR splits the decoder's hidden states into two separate representations—one dedicated to strategy prediction and the other to supportive response generation. To prevent unnecessary entanglement between these two tasks, SARA dynamically integrates the representations only when necessary, ensuring that strategy-related signals remain distinct from linguistic representations. This design prevents the overmixing of strategy cues with linguistic features, allowing each task to fully exploit its unique strengths. As a result, our approach effectively minimizes task conflicts and consistently outperforms existing methods.

Our work makes two key contributions:

- We provide an in-depth analysis revealing that existing multitask learning models for emotional support conversations frequently suffer from task interference, characterized by conflicting gradients and entangled representations.

- We propose the Strategy-Aware Refinement (STAR) module, which effectively mitigates interference between strategy prediction and supportive response generation by dynamically adjusting hidden state representations. Our approach preserves the distinctiveness of strategy-related signals, reducing negative transfer between tasks.

- By minimizing task conflicts, our approach improves both strategy prediction accuracy and the quality of supportive response generation. Experimental results validate these improvements, demonstrating substantial gains over existing methods.

2 Related Work

2.1 Emotional Support Conversation

Recent advancements in natural language processing have increasingly focused on enhancing the ability of dialogue systems to understand and respond empathetically, akin to human interlocutors (Ramírez, 2024). Within this domain, Emotional Support Conversation has emerged as a rapidly growing field, finding applications in mental health support, customer service, and motivational interviewing, where empathetic and context-aware dialogue is crucial (Van der Zwaan et al., 2012; Zhou et al., 2020).

A foundational ESC framework structures response generation into three distinct phases: Exploration, Comforting, and Action, where the system leverages predefined strategy tokens to guide response generation (Liu et al., 2021). Building upon this framework, most subsequent studies have adopted MTL to jointly perform strategy prediction and response generation. However, while MTL enables shared learning across tasks, it also introduces task interference, which can hinder overall model performance.

2.2 Multitask Learning for ESC

Many recent ESC studies have leveraged COMET (Bosselut et al., 2019), a commonsense knowledge-based language model, to enhance strategy selection and supportive response generation by incorporating external knowledge (Liu et al., 2021; Tu et al., 2022; Zhou et al., 2023; Peng et al., 2022; Cheng et al., 2022; Zhao et al., 2023; Deng et al., 2023; Xu et al., 2024; Li et al., 2024; Peng et al., 2023; Zhao et al., 2018).

A common strategy in MTL-based ESC systems is to introduce auxiliary subtasks that reinforce the model's core functionalities. For instance, (Li et al., 2024) and (Zhou et al., 2023) incorporated emotional change prediction as a subtask to facilitate more accurate strategy selection and context-aware supportive response generation. Similarly, (Peng et al., 2023) introduced a task for predicting the primary cause of conversation initiation, helping the model capture the underlying psychological intent of user utterances. Meanwhile, (Xu et al., 2024) proposed a backward supportive response generation task to further enhance system performance by refining historical context comprehension.

Despite these advancements, mitigating task interference and optimizing knowledge integration in MTL-based ESC systems remain open research challenges. Addressing these limitations will be crucial for the development of more robust and contextually aware emotional support dialogue systems.

2.3 Task Interference

Task interference is a fundamental challenge in multitask learning, often leading to degraded model performance and suboptimal knowledge transfer. Various strategies have been proposed to mitigate this issue. Some approaches introduce task-specific parameters within the encoder (Liu et al., 2019), while others employ independent subnetworks or update only task-specific parameters based on task loss (Strezoski et al., 2019; Kanakis et al., 2020). While these methods help reduce interference, they do so at the cost of restricting cross-task knowledge transfer, which in turn limits model adaptability and generalization (Gurulingan et al., 2022b).

Alternatively, task grouping strategies have been explored, where tasks are clustered based on similarity to reduce interference (Gurulingan et al., 2022b). However, these methods rely on statically predefined task similarity, making them less adaptable to ESC systems, where task relationships can evolve dynamically based on conversational context.

To overcome these limitations, we propose the strategy-aware refinement approach, which builds upon task-specific parameterization by dynamically regulating task separation and integration. Unlike conventional approaches that completely isolate task parameters or rely on static groupings, STAR adaptively controls the degree of independence and interaction between task-specific representations within shared parameters. Specifically, it preserves distinct strategy-related representations to prevent interference while maintaining sufficient information exchange, enabling both task specialization and effective knowledge sharing. This design effectively mitigates task interference while enhancing overall model performance.

3 Method

3.1 Overview

The task of emotional support conversation generation inherently involves alternating between strategy prediction and response generation. In each decoding cycle, the model first predicts a special strategy token and then generates the corresponding response using that token alongside the dialogue context. This alternating process introduces task interference, as the model must simultaneously learn to predict effective strategies and generate coherent, supportive responses.

In response, our approach incorporates the STAR module. As illustrated in Figure 2, STAR employs a gating mechanism to dynamically refine hidden representations. Initially, each user utterance u_t (with $U = (u_1, u_2, \ldots, u_T)$ representing the dialogue history) is processed together with an initially predicted strategy token τ. The response generation process is defined as generating the optimal response Y given the dialogue history X, situation description s, and τ (i.e., $\max p(Y \mid X, s, \tau)$). The STAR module then refines τ into an optimized strategy τ' through a strategy-aware adjustment process. The model is ultimately trained to maximize

$$\max p(Y \mid X, s, \tau').$$

By mitigating task interference through dynamic gating and refinement, our method enhances response fluency, emotional appropriateness, and strategic coherence, resulting in more effective emotional support conversations.

3.2 Strategy-Aware Refinement Module

The STAR Module dynamically adjusts the decoder's hidden representations, ensuring that supportive response generation is guided by an optimal strategy. The STAR module consists of two key stages: (1) Strategy-Aware Representation Adjustment (SARA) modifies the hidden state to align with the initially predicted strategy and (2) Strategy Refinement (SR) leverages the adjusted representation to refine the strategy and enhance response generation.

Figure 2 illustrates the STAR Module's architecture, which operates within the decoder to regulate the interaction between task-specific and general knowledge while mitigating interference between strategy prediction and response generation.

The STAR module is integrated into a BlenderBot-based decoder, which generates sup-

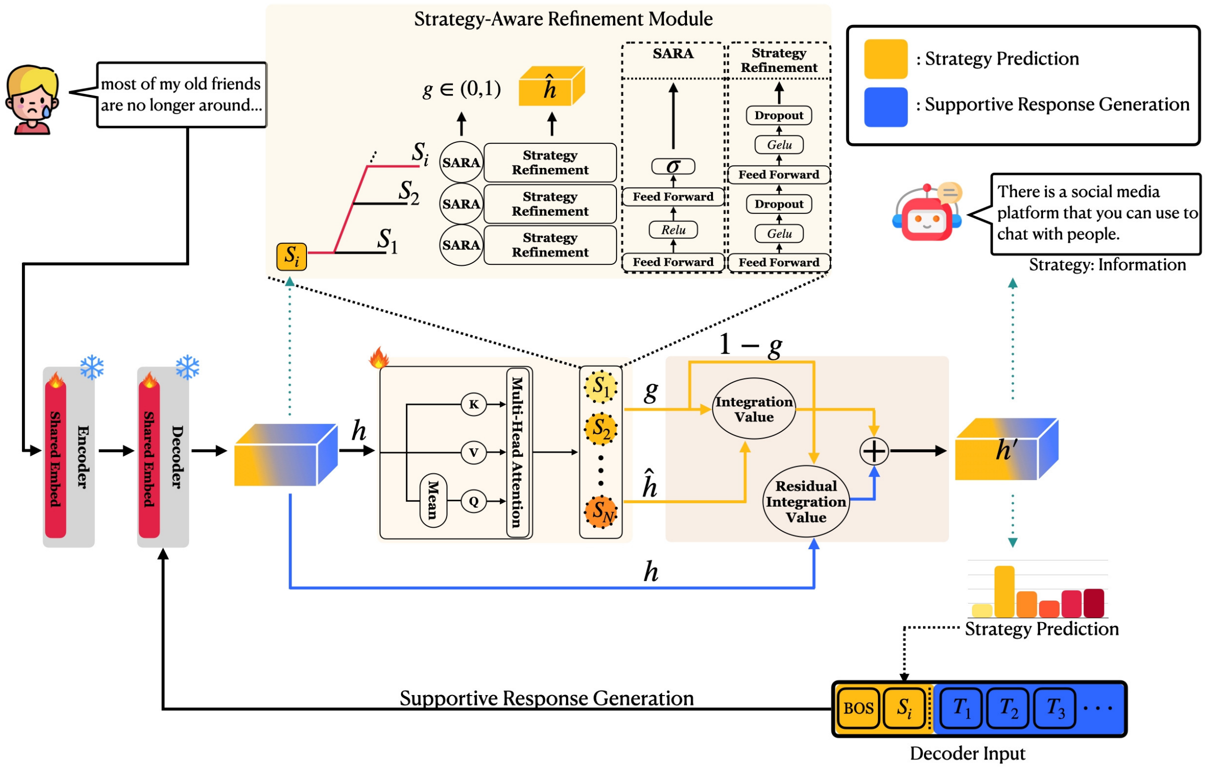


Figure 2: Overall architecture of the STAR module for emotional support conversation. Decoder hidden states from a fine-tuned BlenderBot-Small model are pooled and then fed into two parallel submodules: one computes an integration value, and the other refines the hidden state. The STAR uses the integration value to balance the refined and original hidden states, yielding a strategy-refined state for response generation.

portive responses from the dialogue history. Given a dialogue context, the decoder produces hidden states that encode conversational semantics. Before generating the response, a strategy token corresponding to the initially predicted strategy is appended to guide response formulation. These hidden states and the strategy token are then processed through SARA, the first stage of the STAR module.

The SARA stage modifies the decoder's hidden state to ensure that it aligns with the predicted strategy while preserving general conversational knowledge. Let $h \in R^d$ denote the hidden state of the decoder and let $s \in N$ be the predicted strategy token. SARA first applies a shared attention pooling layer to h, producing a pooled representation z, which captures global contextual information:

$$z = Pooling(h). \tag{1}$$

To regulate the influence of the predicted strategy, the SARA computes an integration value g as:

$$g = \sigma(f(z)), \tag{2}$$

where $f(\cdot)$ is a two-layer network with a hidden layer employing ReLU activation, and σ is the sigmoid function. The integration value $g \in (0, 1)$ determines how much strategy-specific knowledge should contribute to the representation adjustment.

Once the integration value g is computed, the SR stage further processes the decoder's hidden representation to ensure an optimal balance between task-specific knowledge and general dialogue understanding.

The pooled representation z is projected through a two-layer transformation network $P(\cdot)$, generating a transformed representation $\hat{h}$ that maintains the same dimensionality as h:

$$\hat{h} = P(z). \tag{3}$$

The final strategy-refined hidden state h' is then computed as:

$$h' = g \odot \hat{h} + (1 - g) \odot h. \tag{4}$$

This formulation ensures that strategy-relevant information is selectively injected into the hidden state while retaining essential linguistic and contextual features from the original representation.

The strategy-refined hidden representation h' serves as the final hidden state for response generation, incorporating strategy-aware contextual signals to enhance coherence, fluency, and emotional appropriateness. By dynamically adjusting the hidden representation and refining the strategy based on the conversational context, the STAR Module effectively reduces task interference and improves the alignment between strategy selection and supportive response generation.

3.3 Model Training

To effectively train our multitask model for both strategy prediction and supportive response generation, we define separate loss functions for each task and combine them into a final loss using a weighted sum. A dynamic weighting factor, denoted as λ, is introduced to balance the contribution of the supportive response generation loss and the strategy prediction loss, adapting throughout training.

Our weighting strategy is designed to address the different learning dynamics of the two tasks. First, supportive response generation benefits from the pretrained BlenderBot architecture, which is already optimized for open-domain conversations, including empathetic responses (Roller, 2020). Thus, it converges faster during training. Strategy prediction, on the other hand, requires high-level decision-making and a stronger contextual understanding, making it harder to optimize early in training (Yerukola et al., 2023; Vaish and Monroy-Hernández, 2017).

To mitigate task interference and promote balanced learning, we initially assign a lower weight to strategy prediction loss. As training progresses, we gradually increase its weight, allowing the model to first refine response generation before optimizing abstract strategic decisions.

Given a context c and a situation description s, our model generates a response

$$\mathbf{r} = \{r_1, r_2, \ldots, r_{|\mathbf{r}|}\},$$

conditioned on a strategy token τ' produced by the STAR module. In our framework, the response generation objective is to produce an optimal response Y given X, s, and τ' (i.e., $\max p(Y \mid X, s, \tau')$).

The language modeling loss is defined as the negative log-likelihood (NLL) of the generated response:

$$\mathcal{L}_{LM} = -\sum_{t=1}^{n_r} \log p(r_t \mid r_{<t}, c, s, x),$$

where n_r is the length of the response.

The strategy prediction loss is computed using cross-entropy:

$$\mathcal{L}_{ST} = -\log p(\tau' \mid c, s, x),$$

with τ' representing the strategy token used to condition response generation.

To balance these losses during training, we introduce a dynamic weighting factor λ defined as:

$$\lambda = \lambda_0 \cdot \frac{\log(E+1)}{\log(E_{\max})},$$

where E is the current epoch, $E_{\max}$ is the total number of epochs, and λ_0 is a scaling hyperparameter.

Initially, λ is set low to prioritize response generation and preserve fluency; as training progresses, λ increases to place greater emphasis on strategy prediction. The final multitask loss function is given by:

$$\mathcal{L} = (1-\lambda)\,\mathcal{L}_{LM} + \lambda\,\mathcal{L}_{ST}.$$

4 Experiment

4.1 Datasets

We evaluated our model using the ESConv benchmark dataset, which contains 1,300 dialogues and a total of 38,365 utterances, each annotated with eight distinct support strategies. This dataset serves as a well-established benchmark for assessing emotional support conversation systems, providing a structured setting for evaluating both strategy prediction and supportive response generation.

4.2 Baselines

To assess the effectiveness of our approach, we compared it against a range of state-of-the-art models previously evaluated on the ESConv benchmark. For models with publicly available code, we reproduced their implementations and evaluated them under identical conditions. Baseline models include BlenderBot-Joint (Roller, 2020), MISC (Tu et al., 2022), SUPPORTER (Zhou et al., 2023), GLHG (Peng et al., 2022), MultiESC (Cheng et al., 2022), TransESC (Zhao et al., 2023), SCBG (Xu et al., 2024), KEMI (Deng et al., 2023), and Emstremo (Li et al., 2024). These baselines cover diverse architectures, from multitask frameworks to knowledge-enhanced models for emotional support generation.

Model	F1 ↑	PPL ↓	B2 ↑	B4 ↑	R-L ↑
SCBG (Xu et al., 2024)	-	-	5.61	2.91	14.83
GLHG (Peng et al., 2022)	-	15.67	7.57	2.13	16.37
TransESC (Zhao et al., 2023)	-	15.85	7.64	2.43	17.51
SUPPORTER (Zhou et al., 2023)	-	**15.37**	7.49	-	-
MultiESC (Cheng et al., 2022)	-	15.41	**9.18**	**3.09**	**20.41**
BlenderBot-Joint (Roller, 2020)	19.23	16.15	5.52	1.29	15.51
MISC (Tu et al., 2022)	19.89	16.08	7.62	2.19	16.40
Emstremo (Li et al., 2024)	21.30	16.12	8.22	2.53	18.04
KEMI (Deng et al., 2023)	**22.70**	16.34	8.08	2.60	17.05
Models with STAR					
KEMI + STAR	23.17	17.42	8.56	2.65	17.42
Emstremo + STAR	22.48	15.96	8.43	2.28	**18.14**
BlenderBot-Joint + STAR	**24.81**	**15.96**	**8.58**	**2.71**	17.20

Table 1: Performance comparison of various models on the emotional support conversation task. The table reports F1 score (↑), perplexity (PPL, ↓), BLEU-2 (B2, ↑), BLEU-4 (B4, ↑), and ROUGE-L (R-L, ↑) metrics. Models with STAR were reproduced using the proposed method and publicly available code. Specifically, after fine-tuning the base model, all parameters were frozen except for the STAR module and the shared embedding layer within the encoder-decoder, which were further trained to integrate the refined strategy into the response generation process.

Model	F1	B2	B4	R-L	D1	D2	BERT-P	BERT-R	BERT-F
BlenderBot-Joint + STAR	**24.81**	**8.58**	**2.71**	**17.20**	**2.71**	**19.38**	**0.8613**	**0.8561**	**0.8585**
GPT4o-mini (0-shot)	23.35	3.54	0.66	12.13	3.59	24.12	0.8380	0.8492	0.8434
GPT4o-mini (5-shot)	23.35	3.97	0.80	12.99	3.59	24.45	0.8423	0.8514	0.8467
GPT4o-mini (10-shot)	23.35	4.09	0.79	13.12	3.59	24.67	0.8427	0.8520	0.8472
using BlenderBot-Joint + STAR as an strategy classifier									
SC+GPT4o-mini (0-shot)	24.81	3.96	0.84	13.08	3.57	23.65	0.8411	0.8526	0.8467
SC+GPT4o-mini (5-shot)	24.81	4.20	0.87	13.77	3.65	25.10	0.8442	0.8532	0.8485
SC+GPT4o-mini (10-shot)	24.81	4.29	0.96	13.76	3.65	25.12	0.8446	0.8538	0.8491

Table 2: Experimental results using the GPT4o-mini model. The table reports for GPT4o-mini in zero-shot, 5-shot, and 10-shot settings, both when used directly and when combined with a strategy classifier (SC)

We also conducted experiments with GPT4o-mini (OpenAI et al., 2024) under zero-, five-, and ten-shot settings. In one configuration, GPT4o-mini performed both strategy prediction and response generation simultaneously. In another, the best-performing model from our experiments was used as a strategy classifier to provide strategy labels for GPT4o-mini's response generation.

4.3 Evaluation Metrics

To comprehensively assess the performance of our model on ESConv, we employ multiple evaluation metrics covering strategy prediction accuracy, response fluency, content preservation, response diversity, and semantic similarity. Strategy prediction accuracy is measured using the Macro F1 score, which evaluates the model's ability to classify support strategies across different categories. To assess response fluency, we use perplexity (PPL), where lower values indicate more fluent and coherent text generation. Content preservation is quantified using BLEU-2, BLEU-4, and ROUGE-L, which measure the degree of lexical overlap between generated responses and reference responses. Response diversity is evaluated through Distinct-1 (D1) and Distinct-2 (D2)(Deng et al., 2023; Liu et al., 2021; Tu et al., 2022), which compute the ratio of unique n-grams to total n-grams, reflecting lexical variety and reducing generic responses. Additionally, we employ BERT-based metrics (BERT-P, BERT-R, and BERT-F)(Zhang* et al., 2020) to measure semantic similarity, capturing how well the generated responses align with reference responses in a contextual embedding space.

4.4 Implementation Details

For our experiments, we fine-tuned the **BlenderBot-Small** model under carefully optimized hyperparameters. The model was trained with a learning rate of 3×10^{-5}, employing a **linear warmup strategy** with 120 warmup steps. To manage input constraints, we set the **maximum input sequence length to 160 tokens** and the **maximum target sequence length to 40 tokens**. During decoding, we applied **Top-p sampling** ($p = 0.3$) and **Top-k sampling** ($k = 30$), with a **temperature setting of 0.7** to control response randomness and a **repetition penalty of 1.03** to mitigate excessive repetition in generated text. The optimization process was carried out using the **AdamW optimizer**, configured with $\beta_1 = 0.9$ and $\beta_2 = 0.999$, while the loss function was modulated with a **gamma value of 0.8**, ensuring effective gradient scaling throughout training. All experiments were conducted on a **single NVIDIA RTX A6000 GPU** with a batch size of 128, and training was performed for a total of 10 epochs.

5 Results

5.1 Quantitative Performance Evaluation

As shown in Table 1, our method establishes a new benchmark in strategy prediction and supportive response generation across multiple ESC models. Notably, it achieves substantial gains in BLEU and ROUGE scores, indicating improved alignment with ground truth responses.

When applied to BlenderBot-Joint, our approach yields an absolute increase of 5.58% in Macro F1, along with improvements of 3.06% in BLEU-2 (B2), 1.42% in BLEU-4 (B4), and 1.69% in ROUGE-L (R-L). For KEMI, it improves F1 by 0.47%, B2 by 0.48%, B4 by 0.05%, and R-L by 0.37%. On Emstremo, it achieves gains of 1.18% in F1, 0.21% in B2, and 0.1% in R-L.

These results demonstrate STAR's ability to consistently enhance performance across architectures by mitigating task interference. By dynamically regulating strategy-conditioned and general response generation, our method achieves more robust and contextually appropriate responses. Additional qualitative analysis can be found in Appendix A.

5.2 Benchmarking GPT-4o-mini

As shown in Table 2, responses generated by GPT-4o-mini yield lower similarity scores compared to those from our proposed method and other state-of-the-art approaches. However, GPT-4o-mini demonstrates superior response diversity, as indicated by higher Distinct-n (D1 and D2) scores. This highlights a trade-off between lexical diversity and reference alignment, suggesting that increased variability may reduce similarity with human-annotated ground truths.

Furthermore, when our method is applied to the BlenderBot-Joint model as a strategy classifier, it yields an average improvement of 0.42% in BLEU-2, 0.18% in BLEU-4, 0.06% in D1, 0.65% in D2, 0.002% in BERT-P, 0.004% in BERT-R, and 0.002% in BERT-F. These results indicate that our approach not only preserves response diversity but also enhances similarity and consistency through strategy-aware calibration.

5.3 Impact of Task Interference in MTL-Based ESC

ESC systems inherently rely on MTL for strategy selection via special tokens. However, a major challenge in this setting is task interference, which arises when representation learning for multiple tasks conflicts, or when gradients exhibit opposing directions, leading to suboptimal optimization. To address this, we analyze both representation clustering and gradient conflicts to evaluate the impact of STAR.

5.3.1 Gradient Conflicts

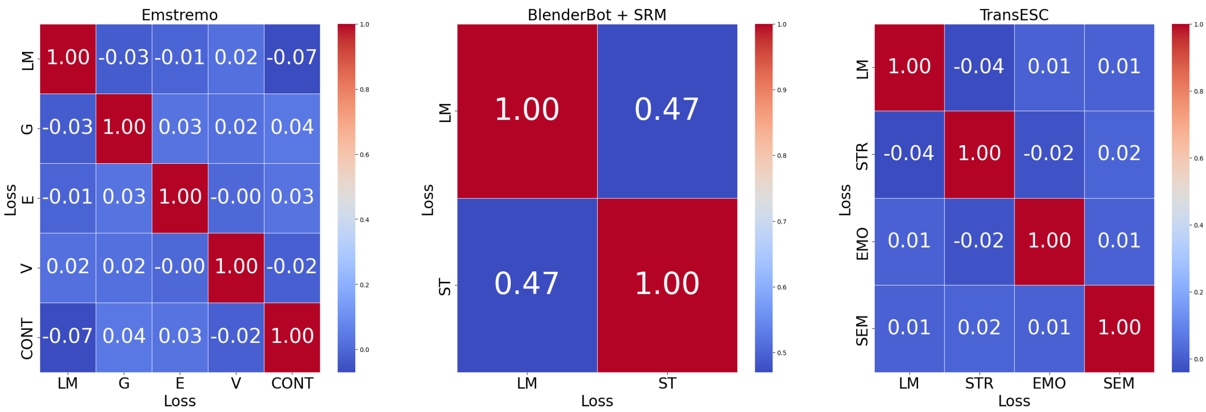
As shown in Figure 3b, applying STAR to the BlenderBot-Joint model (see Section 3.3) results in a cosine similarity of 0.47 between the response generation loss (L_{LM}) and strategy classification loss (L_{ST}), indicating a significant reduction in gradient conflict. In contrast, Figure 3c shows that TransESC exhibits predominantly negative or near-zero cosine similarities across tasks (e.g., STR, EMO, SEM), suggesting substantial gradient conflicts. Similarly, Figure 3a illustrates that Emstremo's task similarities (LM, G, V, CONT, etc.) remain close to zero or negative, reinforcing the presence of gradient conflicts.

5.3.2 Representation Conflicts

Moreover, Figure 4a indicates that the original BlenderBot-Joint model struggles with clear task separation, as evidenced by overlapping cluster boundaries. A similar issue is observed in the KEMI model (see Figure 4c). In contrast, Figure 4b shows that applying STAR leads to distinctly separated clusters with tighter intra-cluster cohesion, demonstrating its effectiveness in enforcing task separation. These results confirm that STAR effectively reduces task interference by preserving independent and well-structured task representations.

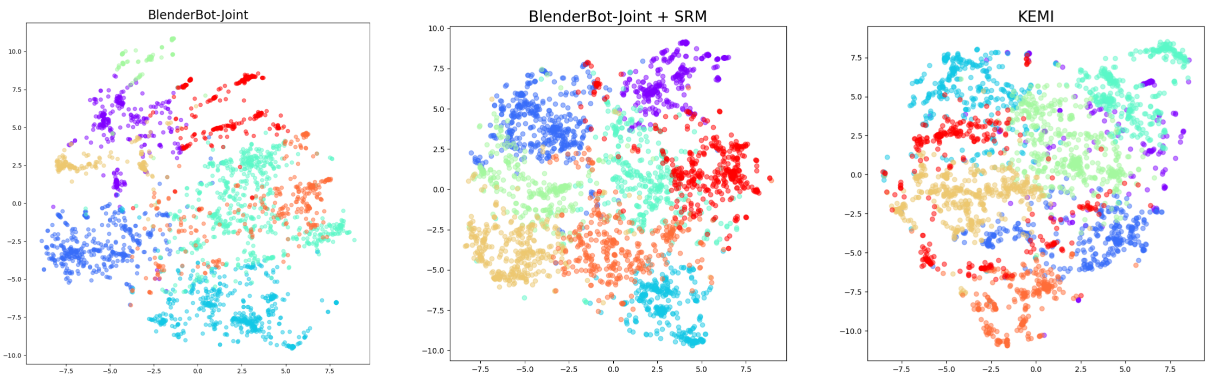
6 Conclusion

This paper analyzes and addresses the task interference problem in MTL-based ESC systems, a fundamental challenge that arises from conflicts between strategy-conditioned token generation and general text generation. To mitigate this issue, we



(a) Emstremo (b) BlenderBot + STAR (c) TransESC

Figure 3: Loss correlation matrices for three different models. Each matrix is obtained by storing the gradient vectors corresponding to each loss component, then computing the cosine similarity among these gradient vectors. A higher correlation value indicates reduced gradient conflict between losses, suggesting more harmonious multi-task optimization.



(a) BlenderBot-Joint (b) BlenderBot-Joint + STAR (c) KEMI

Figure 4: t-SNE visualizations of the final hidden states extracted from three different models. We apply K-means clustering with $k = 8$, reflecting the eight strategy types in the ESConv dataset. As shown, the model variant employing STAR (*middle*) achieves more distinct cluster separation, indicating clearer differentiation among strategies compared to both the baseline (*left*) and KEMI (*right*).

propose the Strategy-Aware Refinement module, which dynamically regulates task separation and integration by disentangling strategy prediction from response generation.

Experimental results demonstrate that STAR achieves significant improvements in both strategy prediction and response generation by effectively mitigating task interference, outperforming existing methods. These findings highlight the importance of task-aware refinement in MTL-based ESC systems, paving the way for future research on optimizing multi-objective dialogue modeling.

Limitations

We acknowledge the following limitations in our study:

- To the best of our knowledge, this is the first study to systematically analyze task interference in ESC. As such, the proposed evaluation metrics may require further refinement for more robust future assessments.
- Our study does not focus on leveraging large language models or exploring various prompt-based in-context learning techniques. However, as indicated in Table 2, incorporating effective prompt-based methods could significantly enhance performance.
- The proposed method relies on a gating mechanism to dynamically regulate task-specific information flow. However, if the gate network fails to optimally balance integration under varying conditions, performance may degrade.

While this issue was not observed on the ESConv dataset (Table 1), further validation on diverse datasets is necessary. Constructing new datasets tailored for ESC systems would be valuable for assessing generalization.

References

Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. 2019. COMET: Commonsense transformers for automatic knowledge graph construction. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4762–4779, Florence, Italy. Association for Computational Linguistics.

Brant R Burleson. 2003. Emotional support skills. In *Handbook of communication and social interaction skills*, pages 569–612. Routledge.

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. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 3014–3026, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Yang Deng, Wenxuan Zhang, Yifei Yuan, and Wai Lam. 2023. Knowledge-enhanced mixed-initiative dialogue system for emotional support conversations. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4079–4095.

Naresh Kumar Gurulingan, Elahe Arani, and Bahram Zonooz. 2022a. Curbing task interference using representation similarity-guided multi-task feature sharing. *Preprint*, arXiv:2208.09427.

Naresh Kumar Gurulingan, Elahe Arani, and Bahram Zonooz. 2022b. Curbing task interference using representation similarity-guided multi-task feature sharing. In *Conference on Lifelong Learning Agents*, pages 937–951. PMLR.

Catherine A Heaney and Barbara A Israel. 2008. Social networks and social support. *Health behavior and health education: Theory, research, and practice*, 4(1):189–210.

Menelaos Kanakis, David Bruggemann, Suman Saha, Stamatios Georgoulis, Anton Obukhov, and Luc Van Gool. 2020. Reparameterizing convolutions for incremental multi-task learning without task interference. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XX 16*, pages 689–707. Springer.

Catherine Penny Hinson Langford, Juanita Bowsher, Joseph P Maloney, and Patricia P Lillis. 1997. Social support: a conceptual analysis. *Journal of advanced nursing*, 25(1):95–100.

Junlin Li, Bo Peng, and Yu-Yin Hsu. 2024. Emstremo: Adapting emotional support response with enhanced emotion-strategy integrated selection. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 5794–5805, Torino, Italia. ELRA and ICCL.

Shikun Liu, Edward Johns, and Andrew J Davison. 2019. End-to-end multi-task learning with attention. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1871–1880.

Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang, and Minlie Huang. 2021. Towards emotional support dialog systems. *arXiv preprint arXiv:2106.01144*.

OpenAI et al. 2024. Gpt-4o system card. *Preprint*, arXiv:2410.21276.

Wei Peng, Yue Hu, Luxi Xing, Yuqiang Xie, Yajing Sun, and Yunpeng Li. 2022. Control globally, understand locally: A global-to-local hierarchical graph network for emotional support conversation. *arXiv preprint arXiv:2204.12749*.

Wei Peng, Ziyuan Qin, Yue Hu, Yuqiang Xie, and Yunpeng Li. 2023. Fado: Feedback-aware double controlling network for emotional support conversation. *Preprint*, arXiv:2211.00250.

José Gabriel Carrasco Ramírez. 2024. Natural language processing advancements: Breaking barriers in human-computer interaction. *Journal of Artificial Intelligence General Science (JAIGS) ISSN: 3006-4023*, 3(1):31–39.

S Roller. 2020. Recipes for building an open-domain chatbot. *arXiv preprint arXiv:2004.13637*.

Gjorgji Strezoski, Nanne van Noord, and Marcel Worring. 2019. Many task learning with task routing. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1375–1384.

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. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 308–319, Dublin, Ireland. Association for Computational Linguistics.

Rajan Vaish and Andrés Monroy-Hernández. 2017. Crowdtone: Crowd-powered tone feedback and improvement system for emails. *arXiv preprint arXiv:1701.01793*.

Janneke M Van der Zwaan, Virginia Dignum, and Catholijn M Jonker. 2012. A bdi dialogue agent for social support: Specification and evaluation method. In *Proceedings of the 3rd Workshop on Emotional and Empathic Agents@ AAMAS*, volume 2012, pages 1–8.

Yangyang Xu, Zhuoer Zhao, and Xiao Sun. 2024. Scbg: Semantic-constrained bidirectional generation for emotional support conversation. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, 23(7).

Akhila Yerukola, Xuhui Zhou, Elizabeth Clark, and Maarten Sap. 2023. On the need for contextual models and evaluations for stylistic rewriting. In *The 2023 Conference on Empirical Methods in Natural Language Processing*.

Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. 2020. Gradient surgery for multi-task learning. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA. Curran Associates Inc.

Tianyi Zhang*, Varsha Kishore*, Felix Wu*, Kilian Q. Weinberger, and Yoav Artzi. 2020. Bertscore: Evaluating text generation with bert. In *International Conference on Learning Representations*.

Weixiang Zhao, Yanyan Zhao, Shilong Wang, and Bing Qin. 2023. TransESC: Smoothing emotional support conversation via turn-level state transition. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 6725–6739, Toronto, Canada. Association for Computational Linguistics.

Xiangyun Zhao, Haoxiang Li, Xiaohui Shen, Xiaodan Liang, and Ying Wu. 2018. A modulation module for multi-task learning with applications in image retrieval. In *Proceedings of the European Conference on Computer Vision (ECCV)*.

Jinfeng Zhou, Zhuang Chen, Bo Wang, and Minlie Huang. 2023. Facilitating multi-turn emotional support conversation with positive emotion elicitation: A reinforcement learning approach. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1714–1729, Toronto, Canada. Association for Computational Linguistics.

Li Zhou, Jianfeng Gao, Di Li, and Heung-Yeung Shum. 2020. The design and implementation of xiaoice, an empathetic social chatbot. *Computational Linguistics*, 46(1):53–93.

Appendix

A Case Study

Table 3 presents two case studies comparing responses generated by three baseline models and their counterparts after applying our proposed method. Overall, responses generated with STAR exhibit stronger alignment with designated support strategies, ensuring more contextually appropriate and strategically coherent interactions.

In the first case, responses incorporating our method effectively implement the "Affirmation and Reassurance" strategy. These responses not only provide encouragement and support but also incorporate concrete recommendations—such as exercise and meditation—resulting in a more thoughtful and contextually appropriate interaction. In contrast, baseline models fail to achieve this level of strategic refinement. For instance, the BlenderBot-Joint model merely expresses agreement without additional guidance, the KEMI model acknowledges the benefits of exercise but lacks elaboration, and the Emstremo model, despite offering encouragement, introduces contextually misaligned content that may reduce response effectiveness.

A similar pattern emerges in the second case, where our method effectively applies the "Information" strategy by offering relevant details and actionable guidance to help the user form new social connections. In contrast, baseline models fail to fully implement the designated strategy, leading to responses that lack practical guidance and fail to maximize user engagement.

These case studies demonstrate that STAR not only preserves response diversity but also enhances strategic calibration, enabling more effective, contextually coherent, and user-centered interactions. This underscores the importance of strategy-aware refinement in ESC, highlighting its potential to significantly improve both conversational quality and strategic fidelity.

Case 1
Seeker: “okay i will try this and see if this works. shall i also try walking or exercsing” **Supporter:** “yeah, anything to take your mind of it :)” **Strategy:** Affirmation and reassurance
BlenderBot-Joint: “yes, that’s a great idea!” **BlenderBot-Joint + STAR:** “yes, that is a great idea. you can also try to meditate in the morning, it is a good way to relax and get your mind off of things.” **KEMI:** “yes, i think that is a great idea. exercise is a good way to relax” **KEMI + STAR:** “yes, that is a great idea. exercise is great for releasing endorphins and releasing endpressants” **Emstremo:** “Yeah, anything to take your mind of it:)” **Emstremo + STAR:** “i know that it can be a challenge, but it’s a of who knows what you could possibly do with such talents!”
Case 2
Seeker: “most of my old friends are no longer around...” **Supporter:** “social networks online, you can connect with family or old coworkers. and post pictures and see what other people are doing with their lives.” **Strategy:** Information
BlenderBot-Joint: “you can join new groups and meet new people. there are many online forums out there.” **BlenderBot-Joint + STAR:** “There is a social media platform that you can use to chat with people.” **KEMI:** “i understand. i’m sure you’ll find someone.” **KEMI + STAR:** “i understand. i think you are going through a tough time. i’m sure you will find the way in online.” **Emstremo:** “i’m glad to hear that!” **Emstremo + STAR:** “It’s a social networking site that lets you connect and chat with other people.”

Table 3: Two example cases illustrating how different models respond to user queries under distinct situations. The first case focuses on exercise as a coping strategy, while the second highlights social networking for maintaining connections.