

Fine-grained stance cues and commonsense knowledge Integrated Reasoning Enhancement for Zero-shot Conversational Stance Detection

Weiming Yin¹, Jinzhong Ning^{1*}, Yijia Zhang^{1*}, Yuanyuan Sun²,

¹ School of Information Science and Technology, Dalian Maritime University, Dalian, China

² School of Computer Science and Technology, Dalian University of Technology, Dalian, China

yinweiming@mail.dlut.edu.cn

{ningjinzong,zhangyijia}@dlmu.edu.cn

syuan@dlut.edu.cn

Abstract

Conversational stance detection aims to identify a speaker’s attitude toward a target in multi-turn dialogues, while the zero-shot setting requires generalization to unseen targets. Existing methods often rely on coarse-grained semantic representations or static knowledge injection, limiting their ability to capture subtle stance cues and utilize commonsense knowledge. To address this, we propose FIRE, a reasoning-enhanced framework that integrates fine-grained stance cues and commonsense knowledge for zero-shot conversational stance detection. FIRE adaptively mines sentiment, rhetorical, referential, cross-turn stance dependency cues and commonsense knowledge, and fuses them into explicit reasoning chains for target-aware stance inference. By combining supervised instruction tuning with reinforcement learning, FIRE further improves reasoning-chain quality and stance prediction. Experiments on benchmark datasets show that FIRE consistently outperforms strong baselines across different target types and LLM backbones, demonstrating the effectiveness of fine-grained cue modeling and knowledge-driven reasoning. Our code and data are available at <https://github.com/YinWeiming/FIRE.git>.

1 Introduction

Conversational Stance Detection (CSD) aims to identify the stance expressed in an utterance toward a given target in multi-turn dialogues, conditioned on the historical dialogue context (Li et al., 2023a,b; Niu et al., 2024b). This task serves as a cornerstone for many downstream applications, such as public opinion analysis (Sobhani et al., 2015; Zhang et al., 2017; Sirrianni et al., 2020), misinformation detection (Zubiaga et al., 2018; Lazer et al., 2018; Yang et al., 2024), and information retrieval (Sen et al., 2018; Carnot et al., 2023).

*Co-corresponding authors.

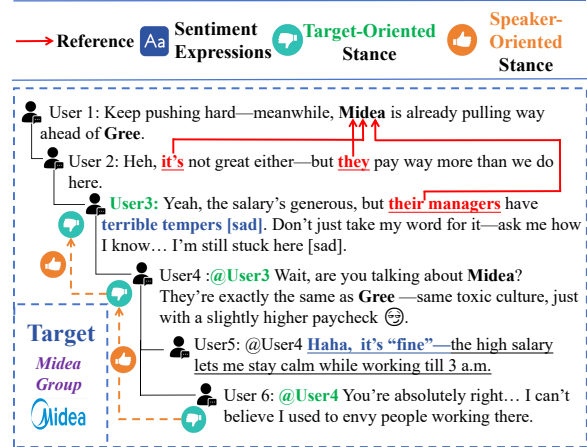


Figure 1: A multi-turn conversational example illustrating key challenges in CSD, including ambiguous referents, sarcastic or affective expressions, and stance dependencies across speakers.

Early studies on stance detection primarily focus on target-conditioned classification for *single texts* under in-target (Sun et al., 2018; Dey et al., 2018; Wei et al., 2018), cross-target (Wei and Mao, 2019; Zhang et al., 2020), or zero-shot settings (Liang et al., 2022; Hanley and Durumeric, 2023), depending on whether the test target is seen, related, or unseen during training. However, such single-text paradigms fail to capture real-world complexity, where opinions rarely stand alone but emerge through multi-turn dialogues between participants. This observation has motivated the shift toward *Conversation Stance Detection*, which aims to model stance within dialogic contexts. CSD has evolved from short-turn interactions (Li et al., 2023a,b) to modeling long, multi-turn and structurally complex conversations (Niu et al., 2024b, 2025). Further, practical applications often face new, unseen topics or domains where labeled data is scarce, naturally leading to the exploration of *Zero-Shot Conversational Stance Detection* (Ding et al., 2025).

Despite recent progress, existing methods still exhibit several limitations. **Limitation 1:** Existing methods mostly encode each historical utterance into a single sentence-level representation (Li et al., 2023a,b; Niu et al., 2024b). Such compressed representations may overlook fine-grained stance cues within utterances (e.g., sentiment polarity, rhetorical expressions, and tonal features) and weaken token-level semantic links across turns (e.g., coreference-induced links). As a result, models may fail to capture subtle but decisive stance evidence in complex conversational scenarios. As illustrated in Figure 1, multiple fine-grained cues provide useful evidence for predicting the stance toward *Midea*, including sentiment words or phrases such as “*terrible tempers*” and “[*sad*]”, sarcastic or ironic rhetorical cues such as “*Haha*” and “*it’s ‘fine’*”, and cross-turn referential links such as “*their managers*”, and “*its*”. Moreover, we conduct a quantitative analysis of the dataset, as shown in Appendix A Figure 6, which demonstrates that such fine-grained cues are common in dialogue scenarios. **Limitation 2:** Existing zero-shot stance detection methods commonly introduce external knowledge (Liu et al., 2021; Zhu et al., 2022; Wen and Hauptmann, 2023; Zhang et al., 2024; Ji et al., 2025) to help models understand unseen targets and improve generalization. These methods usually follow a two-stage pipeline: relevant knowledge is first selected through semantic similarity retrieval or large language model prompting, and then incorporated into downstream stance prediction. However, because knowledge selection and knowledge utilization are optimized separately, the selected knowledge may not be the knowledge most useful for the final stance reasoning process. This task-agnostic knowledge incorporation strategy can lead to a mismatch between introduced knowledge and downstream decision requirements, thereby limiting the effectiveness of knowledge-enhanced zero-shot generalization. **Limitation 3:** Most existing methods use all historical utterances as context for the current utterance (Li et al., 2023a,b; Niu et al., 2024b; Ding et al., 2025). However, topics in multi-turn conversations may shift rapidly, earlier turns do not always remain relevant to the current utterance. As illustrated in Figure 7, historical utterances become less relevant as their distance from the current turn increases. Therefore, using the complete dialogue history may introduce low-relevance or even misleading information, increasing computational cost and interfering with

the identification of truly useful stance evidence.

To address these limitations, we propose **Fine-grained stance cues and commonsense knowledge Integrated Reasoning Enhancement framework (FIRE)** for robust zero-shot conversational stance detection. For **Limitation 1**, to overcome the loss of stance information caused by sentence-level compressed modeling, FIRE constructs an explicit fine-grained stance reasoning chain that dynamically captures sentiment, rhetorical, referential, and cross-turn stance dependency cues according to different conversational scenarios. For **Limitation 2**, FIRE incorporates commonsense knowledge into the fine-grained reasoning chain to help understand unseen targets, and further uses instruction tuning cold start and GRPO-based reinforcement learning to refine knowledge-aware multi-cue stance reasoning. For **Limitation 3**, we systematically analyze the impact of different context window sizes, revealing the trade-off among contextual relevance, detection performance, and computational efficiency.

Our main contributions are summarized as follows:

- We propose **FIRE**, a reasoning-enhanced framework that integrates fine-grained stance cues and commonsense knowledge for zero-shot conversational stance detection.
- We design a fine-grained stance reasoning chain and jointly optimize knowledge selection and utilization, improving stance reasoning over unseen targets.
- Extensive experiments show that FIRE achieves state-of-the-art performance, with further analysis on the impact of context window size on effectiveness and efficiency.

2 Related Work

2.1 Conversational Stance Detection

Stance detection research has shifted from single-text settings toward *conversational stance detection*, reflecting the dialogic nature of stance expression on social media. Early work focuses on *in-target* conversational settings with fixed targets (Li et al., 2023a,b). To improve cross-target generalization, large-scale, multi-turn benchmarks MT-CSD (Niu et al., 2024b) and C-MTCS (Niu et al., 2025) were introduced. Furthermore, MRGKD (Wang et al., 2025) distills multi-agent LLM reasoning

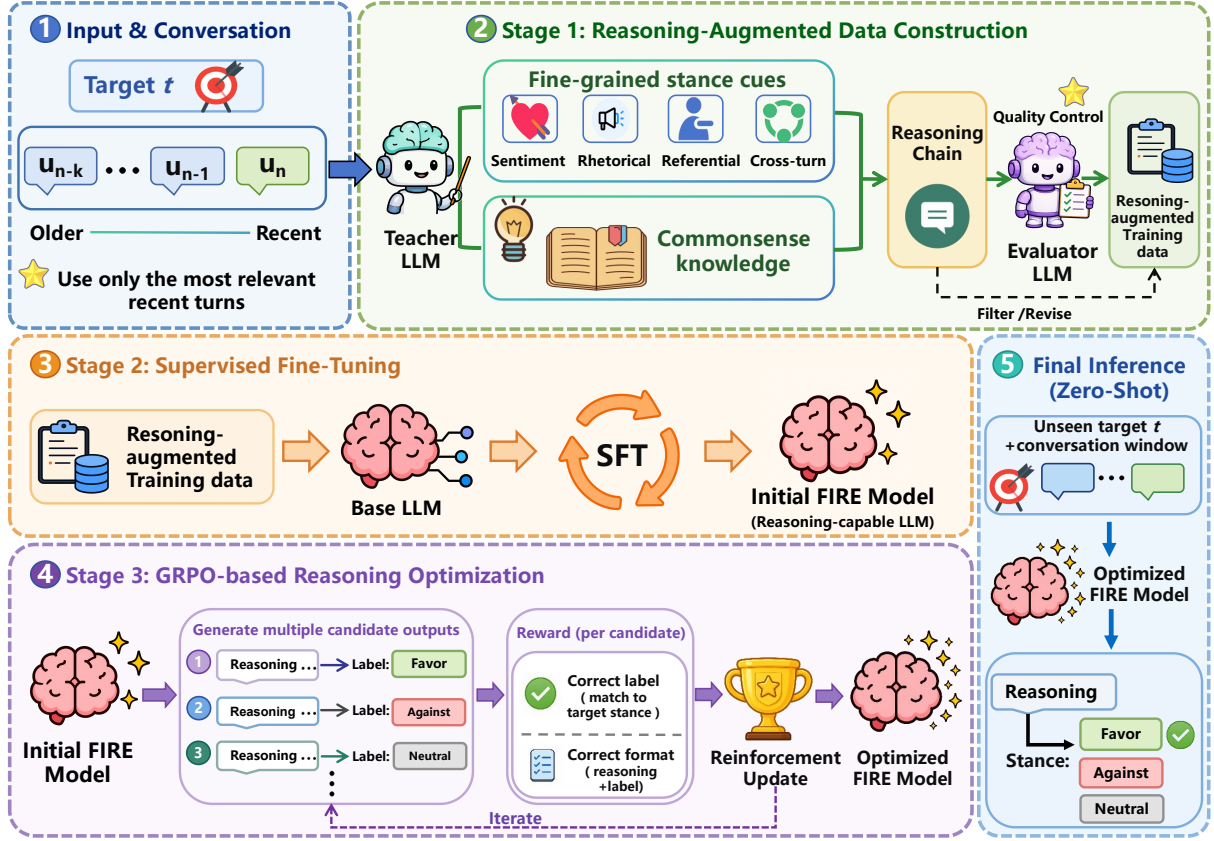


Figure 2: Architecture of the FIRE framework. Stage 1: A teacher LLM is prompted to mine fine-grained stance cues and commonsense knowledge, which are then fused into reasoning-chain training data. Stage 2: The generated data are used to supervised fine-tune a base LLM. Stage 3: The fine-tuned model is further optimized with GRPO using rewards for stance correctness and reasoning format.

graphs into smaller models to capture implicit logical relations in conversational history, while Niu et al. (2024a) extend conversational stance detection to the multimodal setting. However, due to the vast space of real-world targets and annotation costs, these datasets remain limited to constrained target sets. To address this limitation, Ding et al. (2025) proposed *zero-shot conversational stance detection* and released the first dedicated dataset.

Unlike prompting-based methods for multi-path or multi-stage stance reasoning (Lan et al., 2024; Wang et al., 2024b; Zhang et al., 2025), FIRE explicitly integrates fine-grained stance cues and commonsense knowledge into reasoning chains and optimizes them via supervised fine-tuning and reinforcement learning.

2.2 Chain-of-thought Reasoning with LLMs

Chain-of-thought (CoT) prompting enables LLMs to perform step-by-step reasoning (Wei et al., 2022). Existing studies mainly improve CoT reasoning by internalizing reasoning abilities through instruction tuning on CoT-annotated data (Chung et al.,

2024; Luo et al., 2025; Mukherjee et al., 2023), or by enhancing inference-time reasoning with larger reasoning budgets and search-based decoding (Nakano et al., 2022; Kocsis and Szepesvári, 2006; Guan et al., 2025; OpenAI et al., 2026; Guo et al., 2025). Reward models are also used to guide this process by evaluating final answers or intermediate reasoning steps (Hoffmann et al., 2022; Wang et al., 2024a; Luo et al., 2024).

Different from prior work on general-purpose reasoning, we apply CoT reasoning to conversational stance detection, emphasizing task-specific and target-conditioned reasoning over multi-turn dialogues in zero-shot settings.

3 Method

3.1 Task Definition

We study zero-shot conversational stance detection (ZS-CSD). Given a dialogue $\mathcal{C} = \{u_1, u_2, \dots, u_n\}$, a target t , and the current utterance u_n , the goal is to predict the stance of u_n toward t . The label space is $\mathcal{Y} = \{\text{FAVOR}, \text{AGAINST}, \text{NEUTRAL}\}$. In

the zero-shot setting, the target t is not observed during training.

Since not all previous utterances are useful for predicting the stance of the current utterance, we use a context window of size k :

$$\mathcal{C}^{(k)} = \{u_{n-k}, \dots, u_{n-1}, u_n\}. \quad (1)$$

The model receives the target t and the windowed conversation $\mathcal{C}^{(k)}$ as input, and generates both a reasoning process and the final stance label.

3.2 The FIRE Framework

We propose FIRE, a reasoning-enhanced framework that formulates zero-shot conversational stance detection as target-conditioned reasoning generation. The framework consists of three components: reasoning-augmented data construction with a teacher LLM, supervised fine-tuning for reasoning cold start, and GRPO-based Reasoning Optimization.

3.2.1 Reasoning-augmented Data Construction

For each training instance, we build the input by combining the task instruction, the target, and the windowed conversation:

$$x^{(k)} = \mathcal{T}_{\text{in}}(I, t, \mathcal{C}^{(k)}, u_n), \quad (2)$$

where I denotes the task instruction and $\mathcal{T}_{\text{in}}(\cdot)$ is the input template.

The output contains two parts: a stance reasoning chain $R^{(k)}$ and the gold stance label y :

$$o^{(k)} = \mathcal{T}_{\text{out}}(R^{(k)}, y). \quad (3)$$

To obtain $R^{(k)}$, we use a teacher LLM to mine two types of information from the conversation. The first type is fine-grained stance cues, including sentiment words or phrases, rhetorical cues, referential cues, and cross-turn stance dependency cues. These cues serve as intermediate reasoning evidence that bridges surface-level conversational signals and high-level stance prediction. The second type is target-related and dialogue-relevant commonsense knowledge, which provides necessary background for understanding unseen targets and helps the model connect the target with stance-relevant information in the current conversation. Specific examples of the mined fine-grained stance cues and commonsense knowledge are provided in Appendix H.

Formally, given the evidence-mining prompt $\mathcal{P}_{\text{mine}}$, the target t , the context window $\mathcal{C}^{(k)}$, and the current utterance u_n , the teacher LLM extracts stance cues $E^{(k)}$ and commonsense knowledge $K^{(k)}$. Then, we ask the teacher LLM to organize the extracted cues and knowledge into a coherent reasoning chain $R^{(k)}$ with the reasoning-chain construction prompt $\mathcal{P}_{\text{fuse}}$. The resulting reasoning chain explains how the extracted evidence supports the final stance decision. In this way, each training example provides explicit supervision for both stance reasoning and stance prediction.

LLM-generated reasoning chains may contain irrelevant or logically inconsistent information. Therefore, we apply a quality control step before using them for training. For each instance, an evaluator LLM takes the input $x^{(k)}$, the generated reasoning chain $R^{(k)}$, and the gold label y as input, and produces three scores:

$$q_u, q_r, G = \text{LLM}_{\text{eval}}(x^{(k)}, R^{(k)}, y), \quad (4)$$

where $q_u \in \{1, \dots, 5\}$ measures whether the reasoning chain is useful for stance prediction, and $q_r \in \{1, \dots, 5\}$ measures whether the extracted cues and knowledge are relevant to the target and the conversation. $G \in \{0, 1\}$ indicates whether the reasoning chain is consistent with the gold label.

We compute the overall quality score as:

$$q = \alpha q_u + \beta q_r. \quad (5)$$

Reasoning chains with $q < \tau$ or $G = 0$ are further checked and revised manually, where τ is set to 3. The filtered reasoning chains are then used as training supervision. It is worth noting that the gold label is not given to the teacher LLM during reasoning chain construction. This design avoids label leakage and prevents the teacher model from generating shortcut or post-hoc rationalizations that simply fit the given label. The label is only used in the evaluation stage to check whether the generated reasoning supports the correct stance.

3.2.2 Supervised Fine-tuning

After data construction, we obtain a reasoning-augmented instruction-tuning dataset:

$$\mathcal{D} = \{(x_i, R_i, y_i)\}_{i=1}^N. \quad (6)$$

We fine-tune the base LLM to generate the reasoning chain followed by the final stance label. The

training objective is the standard autoregressive negative log-likelihood:

$$\mathcal{L}_{\text{SFT}}(\theta) = - \sum_{(x_i, R_i, y_i) \in \mathcal{D}} \log P_\theta(R_i, y_i | x_i). \quad (7)$$

This stage provides the model with an initial reasoning ability. The model learns how to identify target-related evidence, organize different types of cues, and produce a stance prediction based on the reasoning process.

3.2.3 GRPO-based Reasoning Optimization

Supervised fine-tuning enables the model to imitate constructed reasoning chains, but imitation alone may not yield reasoning that is optimal for stance prediction. To better align reasoning with the task objective, we apply Group Relative Policy Optimization (GRPO) (Shao et al., 2024).

For each input x_i , the old policy samples a group of G outputs:

$$\{o_{i,1}, o_{i,2}, \dots, o_{i,G}\} \sim \pi_{\theta_{\text{old}}}(\cdot | x_i), \quad (8)$$

where each output contains a reasoning chain and a final stance label. GRPO compares the sampled outputs within the same group and increases the probability of outputs with higher relative rewards. The reward function contains two components:

$$R(o_{i,j}, y_i) = \mathbb{I}[\hat{y}_{i,j} = y_i] + \lambda \mathbb{I}[\text{Format}(o_{i,j}) = 1], \quad (9)$$

where $\hat{y}_{i,j}$ is the stance label parsed from the model output $o_{i,j}$. The first term gives a reward of 1 if the predicted stance is correct and 0 otherwise. The second term gives a reward of 1 if the output follows the required reasoning-then-answer format, and 0 otherwise. λ controls the weight of the format reward.

For each sampled group, we normalize the rewards to compute the group-relative advantage:

$$A_{i,j} = \frac{R(o_{i,j}, y_i) - \text{mean}(\{R(o_{i,j}, y_i)\}_{j=1}^G)}{\text{std}(\{R(o_{i,j}, y_i)\}_{j=1}^G) + \delta}, \quad (10)$$

where $A_{i,j}$ denotes the relative advantage of the j -th output for input x_i , and δ is a small constant for numerical stability.

The GRPO objective is:

$$J_{\text{GRPO}}(\theta) = \mathbb{E} \left[\frac{1}{G} \sum_{j=1}^G \left(\min(\rho_{i,j} A_{i,j}, \text{clip}(\rho_{i,j}, 1 - \epsilon, 1 + \epsilon) A_{i,j}) - \beta D_{\text{KL}}(\pi_\theta \| \pi_{\text{ref}}) \right) \right], \quad (11)$$

where $\rho_{i,j}$ is the probability ratio between the current policy and the old policy, $A_{i,j}$ is the group-relative advantage, and π_{ref} is the reference model. The KL term prevents the optimized policy from drifting too far from the reference model.

4 Experiments

4.1 Experimental Settings

We evaluate our approach on the ZS-CSD dataset (Ding et al., 2025), a large-scale, high-quality benchmark designed specifically for zero-shot conversational stance detection. In contrast to earlier datasets that are confined to narrow domains, ZS-CSD covers 280 diverse stance targets of two types: entities and claims. Each instance includes a multi-turn conversation, a target t , and a stance label y from the set $\{\textit{Favor}, \textit{Against}, \textit{Neutral}\}$.

All fine-tuning and reinforcement learning experiments are conducted on a single NVIDIA RTX PRO 6000 GPU. For reasoning data construction, we set the context window size to $k = 5$ and query three teacher LLMs, GPT-4o, Qwen2.5-72B, and DeepSeek-v3, using the prompting templates in the Appendix G. The generated reasoning chains are filtered according to reasoning usefulness q_u and evidence relevance q_r , with weights $\alpha = 0.6$ and $\beta = 0.4$; chains with an overall quality score below $\tau = 3$ are further refined.

To examine the robustness of FIRE across different model scales and architectures, we fine-tune Qwen2.5-3B, Qwen2.5-7B, Llama2-7B, and Llama3-8B. For supervised fine-tuning, we train for 10 epochs with a batch size of 4, using AdamW with a learning rate of 3×10^{-5} . We adopt QLoRA and set the LoRA rank to 32. For GRPO optimization, we sample $G = 8$ trajectories per instance and train for 2 epochs with a learning rate of 5×10^{-6} . The reward function combines stance accuracy and reasoning-format rewards, with the format reward weight $\lambda = 0.5$. We report per-class F1 and macro-F1 as our primary evaluation metrics.

4.2 Baseline Methods

We compare our model with several state-of-the-art baselines as below: **BERT/BART-based Small Model Baselines.** (1) Branch-BERT (Li et al., 2023a) integrates contextual cues from conversation threads with a CNN-based module that captures salient n-gram features. (2) GLAN (Niu et al., 2024b) adopts a three-branch architecture to model the complex dynamics of multi-turn dia-

Method	Mixed targets				Noun-phrase targets				Claim targets			
	Favor	Against	Neutral	All	Favor	Against	Neutral	All	Favor	Against	Neutral	All
Branch-BERT	43.11	38.19	41.97	41.09	46.29	43.73	43.41	44.48	41.43	34.34	41.13	38.96
GLAN	43.68	34.91	46.79	41.79	45.91	37.65	51.58	45.04	42.43	33.18	43.88	39.83
SITPCL	41.03	41.34	49.07	43.81	40.99	47.14	54.49	47.54	41.06	37.34	46.02	41.47
CKE-Net	44.04	40.25	43.48	42.59	47.56	45.97	47.84	47.12	42.86	39.22	46.32	42.8
LKI-BART	43.09	39.87	46.23	43.06	44.88	40.68	49.79	45.12	43.27	40.77	44.25	42.76
TarBK-BERT	42.56	40.89	48.36	43.94	40.41	48.29	53.96	47.55	40.86	40.18	47.32	42.79
Llama3-8B	35.09	39.34	33.29	35.91	38.31	42.68	37.13	39.37	33.26	37.44	30.68	33.79
GPT4o-mini	24.75	38.63	51.10	38.16	20.78	46.86	52.99	40.21	26.87	32.77	50.07	36.57
GPT-3.5	28.04	38.89	53.84	40.25	38.41	48.79	54.84	47.35	20.99	32.29	53.32	35.53
Llama3-70B	40.81	38.04	44.37	41.07	48.77	50.40	47.00	48.72	35.96	30.98	42.75	36.56
Qwen2.5-14B	33.84	44.01	50.48	42.78	27.67	50.56	55.58	44.60	36.86	40.23	46.93	41.34
Qwen2.5-3B-SFT	44.64	46.21	49.35	46.73	42.38	46.25	48.97	45.87	47.36	44.26	47.91	46.51
Qwen2.5-7B-SFT	46.58	47.34	51.06	48.33	43.33	47.81	50.23	47.12	48.66	45.53	49.64	47.94
Llama2-7B-SFT	43.29	44.38	47.18	44.95	41.47	45.19	46.53	44.40	45.84	43.37	45.88	45.03
Llama3-8B-SFT	45.36	46.57	49.85	47.26	42.19	46.47	48.96	45.87	46.84	44.77	48.32	46.64
FIRE-SFT (Qwen2.5-3B)	48.81	44.93	52.95	48.90	48.05	47.65	53.92	49.87	46.65	46.39	51.41	48.15
FIRE-SFT (Qwen2.5-7B)	51.73	52.31	56.58	53.54	46.83	53.62	56.02	52.16	55.39	51.26	56.41	54.35
FIRE-SFT (Llama2-7B)	44.85	47.63	48.74	47.07	44.46	51.23	43.78	46.49	47.12	45.58	53.93	48.88
FIRE-SFT (Llama3-8B)	51.67	49.65	52.23	51.18	47.65	50.73	48.74	49.04	48.51	47.89	54.12	50.17
FIRE-RL (Qwen2.5-3B)	51.25	48.63	57.48	52.45	51.47	50.74	57.23	53.15	49.15	50.20	55.77	51.71
FIRE-RL (Qwen2.5-7B)	53.55	54.14	63.56	57.08	51.98	56.29	58.77	55.68	59.09	52.87	61.67	57.88
FIRE-RL (Llama2-7B)	47.22	49.26	50.54	49.01	48.37	53.29	47.68	49.78	50.35	47.62	56.27	51.41
FIRE-RL (Llama3-8B)	52.88	53.16	60.29	55.44	50.46	54.28	56.38	53.71	51.59	52.17	59.32	54.36

Table 1: Main results on zero-shot conversational stance detection across mixed, noun-phrase, and claim targets. *SFT* denotes label-supervised fine-tuning, *FIRE-SFT* denotes our reasoning-enhanced instruction-tuned model, and *FIRE-RL* denotes further optimization with reinforcement learning.

logues. (3) SITPCL Ding et al. (2025) introduces a speaker-interaction and target-aware prototypical contrastive learning framework for zero-shot conversational stance detection. *External-knowledge-enhanced methods*. (4) CKE-Net (Liu et al., 2021) exploits commonsense knowledge to enhance both structural-level and semantic-level representations. (5) TarBK-BERT (Wen and Hauptmann, 2023) incorporates target background knowledge from Wikipedia. (6) LKI-BART (Zhang et al., 2024) injects LLM-driven knowledge into BART to better exploit contextual and semantic information.

In-Context Learning-based methods. We evaluate a range of state-of-the-art large language models with different parameter sizes, including closed-source models such as GPT-3.5 and GPT4o-mini, as well as open-source models including Qwen2.5-14B, Llama3-70B, and Llama3-8B.

Instruction-tuned LLM methods. We fine-tune different base models with varying parameter scales, including LLaMA and Qwen series models (3B, 7B, and 8B), using only the ground-truth labels as supervision signals.

4.3 Main Results

We evaluate the proposed **FIRE** framework against a comprehensive set of baselines on three tar-

get types: *mixed*, *noun-phrase*, and *claim* targets. The results are summarized in Table 1. Overall, our method consistently outperforms all baselines across all target types and stance categories. These results demonstrate the strong generalization capability of FIRE in complex zero-shot conversational settings, which adaptively capture fine-grained stance cues across diverse conversational scenarios for stance reasoning. A closer examination of the results yields the following observations:

(1) LLM-based methods generally outperform small PLM-based baselines, indicating that larger language models possess stronger semantic understanding and reasoning capabilities for complex conversational stance detection.

(2) FIRE consistently outperforms standard SFT models across different backbone models, indicating that reasoning chains integrating fine-grained stance cues and commonsense knowledge provide more effective task-specific reasoning signals.

(3) FIRE consistently outperforms Branch-BERT, GLAN, and SITPCL, which mainly rely on coarse-grained modeling of dialogue history. This indicates that fine-grained stance cues are crucial for conversational stance reasoning, alleviating the **Limitation 1** discussed in the introduction.

(4) FIRE outperforms knowledge-enhanced

methods such as CKE-Net, LKI-BART, and TarBK-BERT, indicating that commonsense knowledge is more effective when jointly modeled with downstream stance reasoning. By integrating target-related and context-relevant commonsense knowledge into fine-grained reasoning chains, FIRE helps reduce the mismatch between knowledge incorporation and task-specific decision-making highlighted in **Limitation 2**.

(5) Compared with ICL-based LLM methods, instruction-tuned models achieve more stable and superior performance, suggesting that general-purpose LLMs do not naturally acquire the target-conditioned reasoning ability required for CSD. Task-specific instruction tuning better adapts LLMs to conversational stance reasoning.

(6) Compared with standard SFT models, FIRE with reinforcement learning further improves performance, indicating that imitation learning alone is insufficient for fully optimizing stance reasoning. Reinforcement learning better aligns the reasoning process with the final stance decision through task-oriented rewards.

(7) FIRE brings consistent gains across LLM backbones with different scales and architectures, showing its model-agnostic effectiveness. Among them, Qwen2.5-7B performs best, likely due to its stronger Chinese-specific pretraining, which better matches the linguistic characteristics of the benchmark datasets.

4.4 Ablation Study

We conduct an ablation study under the mixed-target setting to analyze the contributions of fine-grained stance cues, commonsense knowledge, supervised fine-tuning, and reinforcement learning in FIRE. The results are reported in Table 2. Overall, removing any component leads to a performance decline, confirming that each design contributes to the effectiveness of FIRE.

For fine-grained stance cues, excluding reference cues, cross-turn stance dependency cues, rhetorical cues, or sentiment-word cues consistently weakens performance. This indicates that these cues provide complementary signals for identifying implicit stance expressions in conversations. Among them, removing reference cues and cross-turn stance dependency cues leads to relatively larger drops. Reference cues are important because they help the model identify the actual target or entity that an utterance refers to across turns. Cross-turn stance dependency cues further help model how speakers’ at-

Method	Mixed targets			
	Favor	Against	Neutral	All
FIRE	53.55	54.14	63.56	57.08
W/o Reference	52.11	52.46	59.15	54.57
W/o SD	52.80	53.08	60.28	55.39
W/o Rhetorical	53.07	53.26	61.41	55.91
W/o Sentiment	53.15	53.36	61.13	55.88
W/o commonsense	52.58	52.25	58.06	54.30
W/o SFT	43.24	46.31	49.24	46.26
W/o RL	51.73	52.31	56.58	53.54
W/o SFT+RL	31.65	29.37	33.68	31.83

Table 2: Ablation study results of FIRE on mixed-target stance detection. SD (stance dependency cues), Sentiment (sentiment words or phrases), Commonsense (commonsense knowledge).

titudes develop or shift through interactions. When commonsense knowledge is removed, the model also shows a noticeable degradation, suggesting that target-related and stance-relevant knowledge helps bridge surface-level conversational evidence with a deeper understanding of unseen targets, thereby enabling more reliable stance judgments.

When both SFT and RL are removed, the performance drops sharply, indicating that the base LLM alone lacks sufficient task-specific stance reasoning ability. Removing SFT while retaining RL also leads to a large decline, suggesting that GRPO is difficult to optimize effectively without a reasoning-aware SFT initialization. In contrast, removing RL still preserves relatively strong performance, showing that SFT provides an essential reasoning cold start.

4.5 Empirical Analysis of Historical Dialogue Turns

Figure 3 and Table 3 analyze the impact of historical context window size on model performance and training cost. As the window increases from zero to a few recent turns, performance improves steadily, indicating that nearby dialogue history is important for resolving references, modeling speaker interactions, and understanding rhetorical cues.

However, adding more distant turns brings diminishing or unstable gains, while training time continues to increase. This suggests that longer contexts may introduce higher computational cost without consistent performance benefits. Overall, these results support **Limitation 3** and show that a moderate number of recent turns can better balance contextual relevance, detection performance, and training efficiency in zero-shot CSD.

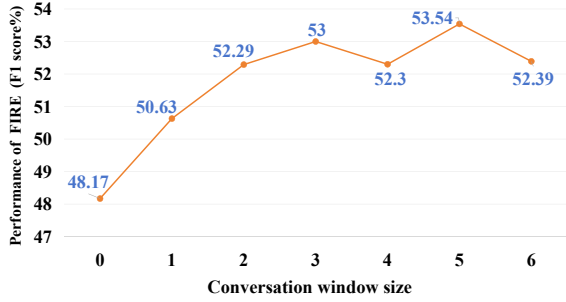


Figure 3: Impact of historical context window size on FIRE during instruction tuning. Reinforcement learning is performed afterward with a fixed window size.

Window Size	F1	Time	Δ F1	Δ Time
	(%)	(min/epoch)	(+ points)	(+ min)
W=0	48.17	12.5	–	–
W=1	50.63	15.4	+2.46	+2.9
W=2	52.29	19.1	+1.66	+3.7
W=3	53	25.4	+0.71	+6.3
W=4	52.3	27.8	-0.7	+2.4
W=5	53.54	30.6	+1.24	+2.8
W=6	52.39	32.3	-1.15	+1.7

Table 3: Impact of historical context window size on performance and training cost during instruction tuning.

4.6 Analysis of Reasoning Chain Effectiveness and Quality

To examine whether reinforcement learning improves reasoning-chain quality, we compare SFT-only and RL-enhanced models across stance categories using the same LLM evaluator as in Section 3.2.1. As shown in Figure 4, the RL-enhanced model consistently obtains higher quality scores, suggesting that reinforcement learning helps the model capture fine-grained stance cues and produce more reliable reasoning chains.

We further analyze the relationship between reasoning quality and stance prediction. Figure 5 shows that higher-quality reasoning chains correspond to a larger proportion of correct predictions and fewer incorrect ones. This consistent trend indicates that the generated reasoning chains effectively integrate fine-grained stance cues and knowledge that are important for conversational stance detection.

4.7 Error Analysis

Despite the overall effectiveness of FIRE, several challenging cases remain. Failures primarily fall into four categories: Sarcasm Misinterpretation, Ambiguity Misresolution, Partial Fusion Bias, and Task Ambiguity. Detailed case studies and root-

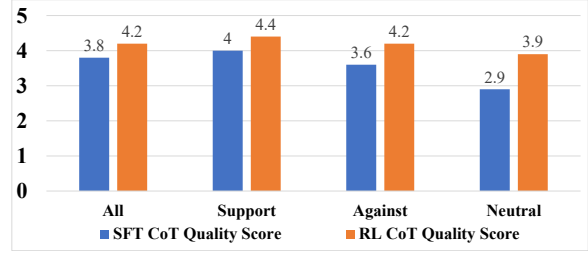


Figure 4: Comparison of reasoning-chain quality between SFT and RL models across stance categories.

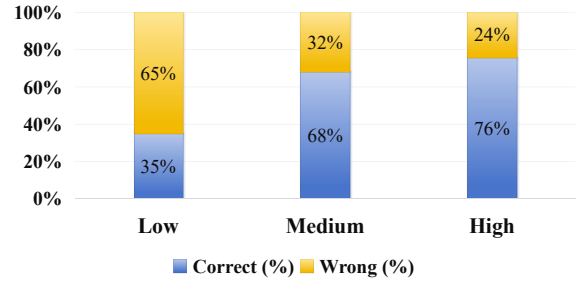


Figure 5: Prediction correctness across different reasoning-chain quality levels.

cause analyses are provided in Appendix I.

5 Conclusion

In this work, we investigate the problem of zero-shot conversational stance detection, where models are required to infer stances toward previously unseen targets within multi-turn dialogues. To address the limitations of existing methods, we propose FIRE, a reasoning-enhanced framework that integrates fine-grained stance cues and common-sense knowledge. FIRE can adaptively mine different cues across diverse conversational scenarios and explicitly model the stance reasoning process, thereby effectively generalizing to unseen targets. Extensive experiments show that FIRE achieves state-of-the-art performance across different target types and base model architectures. Overall, our findings suggest that explicit fine-grained stance cues and knowledge-driven reasoning provide a promising direction for advancing conversational stance detection in open-domain and zero-shot settings.

Limitations

Despite its effectiveness, this work has several limitations. First, FIRE relies on automatically generated chain-of-thought annotations during training, which may introduce minor noise or stylistic bias

from the teacher models. Second, our framework focuses on cues; while they cover common sources of ambiguity in conversational stance detection, other discourse phenomena may also be relevant in specific domains. Finally, we primarily evaluate our method on existing benchmark datasets, and its performance in other conversational genres or languages remains an open question for future research.

Acknowledgements

This work was supported by the Young Scientists Fund of the National Natural Science Foundation of China (NSFC) [Grant No. 62506058] and the Young Teachers Science and Technology Innovation Project of the Fundamental Research Funds for the Central Universities [Grant No. 3132026362].

References

- Miriam Louise Carnot, Lorenz Heinemann, Jan Braker, Tobias Schreieder, Johannes Kiesel, Maik Fröbe, Martin Potthast, and Benno Stein. 2023. On stance detection in image retrieval for argumentation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2562–2571.
- Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellat, Kevin Robinson, and 16 others. 2024. [Scaling instruction-finetuned language models](#). *J. Mach. Learn. Res.*, 25:70:1–70:53.
- Kuntal Dey, Ritvik Shrivastava, and Saroj Kaushik. 2018. Topical stance detection for twitter: A two-phase lstm model using attention. In *European Conference on Information Retrieval*, pages 529–536. Springer.
- Yuzhe Ding, Kang He, Bobo Li, Li Zheng, Haijun He, Fei Li, Chong Teng, and Donghong Ji. 2025. [Zero-shot conversational stance detection: Dataset and approaches](#). In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 3221–3235, Vienna, Austria. Association for Computational Linguistics.
- Xinyu Guan, Li Lina Zhang, Yifei Liu, Ning Shang, Youran Sun, Yi Zhu, Fan Yang, and Mao Yang. 2025. rstar-math: Small llms can master math reasoning with self-evolved deep thinking. *arXiv preprint arXiv:2501.04519*.
- Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Peiyi Wang, Qihao Zhu, Runxin Xu, Ruoyu Zhang, Shirong Ma, Xiao Bi, Xiaokang Zhang, Xingkai Yu, Yu Wu, Z. F. Wu, Zhibin Gou, Zhihong Shao, Zhuoshu Li, Ziyi Gao, Aixin Liu, and 175 others. 2025. [Deepseek-r1 incentivizes reasoning in llms through reinforcement learning](#). *Nature*, 645(8081):633–638.
- Hans W. A. Hanley and Zakir Durumeric. 2023. [TATA: stance detection via topic-agnostic and topic-aware embeddings](#). In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP 2023, Singapore, December 6-10, 2023*, pages 11280–11294. Association for Computational Linguistics.
- Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, Thomas Hennigan, Eric Noland, Katherine Millican, George van den Driessche, Bogdan Damoc, Aurelia Guy, Simon Osindero, Karén Simonyan, Erich Elsen, and 3 others. 2022. [An empirical analysis of compute-optimal large language model training](#). In *Advances in Neural Information Processing Systems*, volume 35, pages 30016–30030. Curran Associates, Inc.
- Yanxu Ji, Jinzhong Ning, Yijia Zhang, Zhi Liu, and Hongfei Lin. 2025. [LLM-driven implicit target augmentation and fine-grained contextual modeling for zero-shot and few-shot stance detection](#). In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 5872–5884, Suzhou, China. Association for Computational Linguistics.
- Levente Kocsis and Csaba Szepesvári. 2006. Bandit based monte-carlo planning. In *European conference on machine learning*, pages 282–293. Springer.
- Xiaochong Lan, Chen Gao, Depeng Jin, and Yong Li. 2024. Stance detection with collaborative role-infused llm-based agents. In *Proceedings of the international AAAI conference on web and social media*, volume 18, pages 891–903.
- David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain. 2018. [The science of fake news](#). *Science*, 359(6380):1094–1096.
- Yupeng Li, Haorui He, Shaonan Wang, Francis CM Lau, and Yunya Song. 2023a. Improved target-specific stance detection on social media platforms by delving into conversation threads. *IEEE Transactions on Computational Social Systems*, 10(6):3031–3042.
- Yupeng Li, Dacheng Wen, Haorui He, Jianxiong Guo, Xuan Ning, and Francis CM Lau. 2023b. Contextual target-specific stance detection on twitter: Dataset and method. In *2023 IEEE International Conference on Data Mining (ICDM)*, pages 359–367. IEEE.

- Bin Liang, Zixiao Chen, Lin Gui, Yulan He, Min Yang, and Ruifeng Xu. 2022. [Zero-shot stance detection via contrastive learning](#). In *Proceedings of the ACM Web Conference 2022, WWW '22*, page 2738–2747, New York, NY, USA. Association for Computing Machinery.
- Rui Liu, Zheng Lin, Yutong Tan, and Weiping Wang. 2021. Enhancing zero-shot and few-shot stance detection with commonsense knowledge graph. In *Findings of the association for computational linguistics: ACL-IJCNLP 2021*, pages 3152–3157.
- Haipeng Luo, Qingfeng Sun, Can Xu, Pu Zhao, Jianguang Lou, Chongyang Tao, Xiubo Geng, Qingwei Lin, Shifeng Chen, Yansong Tang, and Dongmei Zhang. 2025. [Wizardmath: Empowering mathematical reasoning for large language models via reinforced evol-instruct](#). In *The Thirteenth International Conference on Learning Representations, ICLR 2025, Singapore, April 24–28, 2025*. OpenReview.net.
- Liangchen Luo, Yinxiao Liu, Rosanne Liu, Samrat Phatale, Meiqi Guo, Harsh Lara, Yunxuan Li, Lei Shu, Yun Zhu, Lei Meng, Jiao Sun, and Abhinav Rastogi. 2024. [Improve mathematical reasoning in language models by automated process supervision](#). *Preprint*, arXiv:2406.06592.
- Subhabrata Mukherjee, Arindam Mitra, Ganesh Jawahar, Sahaj Agarwal, Hamid Palangi, and Ahmed Awadallah. 2023. Orca: Progressive learning from complex explanation traces of gpt-4. *arXiv preprint arXiv:2306.02707*.
- Reiichiro Nakano, Jacob Hilton, Suchir Balaji, Jeff Wu, Long Ouyang, Christina Kim, Christopher Hesse, Shantanu Jain, Vineet Kosaraju, William Saunders, Xu Jiang, Karl Cobbe, Tyna Eloundou, Gretchen Krueger, Kevin Button, Matthew Knight, Benjamin Chess, and John Schulman. 2022. [Webgpt: Browser-assisted question-answering with human feedback](#). *Preprint*, arXiv:2112.09332.
- Fuqiang Niu, Zebang Cheng, Xianghua Fu, Xiaojiang Peng, Genan Dai, Yin Chen, Hu Huang, and Bowen Zhang. 2024a. Multimodal multi-turn conversation stance detection: A challenge dataset and effective model. In *Proceedings of the 32nd ACM international conference on multimedia*, pages 3867–3876.
- Fuqiang Niu, Min Yang, Ang Li, Baoquan Zhang, Xiaojiang Peng, and Bowen Zhang. 2024b. A challenge dataset and effective models for conversational stance detection. In *LREC/COLING*.
- Fuqiang Niu, Yi Yang, Xianghua Fu, Genan Dai, and Bowen Zhang. 2025. C-mtcsd: A chinese multi-turn conversational stance detection dataset. In *Companion Proceedings of the ACM on Web Conference 2025*, pages 769–772.
- OpenAI, :, Aaron Jaech, Adam Kalai, Adam Lerer, Adam Richardson, Ahmed El-Kishky, Aiden Low, Alec Helyar, Aleksander Madry, Alex Beutel, Alex Carney, Alex Iftimie, Alex Karpenko, Alex Tachard Passos, Alexander Neitz, Alexander Prokofiev, Alexander Wei, Allison Tam, and 246 others. 2026. [Openai o1 system card](#). *Preprint*, arXiv:2412.16720.
- Anirban Sen, Manjira Sinha, Sandya Mannarswamy, and Shourya Roy. 2018. Stance classification of multi-perspective consumer health information. In *Proceedings of the ACM India joint international conference on data science and management of data*, pages 273–281.
- Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, Y. K. Li, Y. Wu, and Daya Guo. 2024. [Deepseekmath: Pushing the limits of mathematical reasoning in open language models](#). *Preprint*, arXiv:2402.03300.
- Joseph Sirrianni, Xiaoqing Liu, and Douglas Adams. 2020. Agreement prediction of arguments in cyber argumentation for detecting stance polarity and intensity. In *Proceedings of the 58th annual meeting of the Association for Computational Linguistics*, pages 5746–5758.
- Parinaz Sobhani, Diana Inkpen, and Stan Matwin. 2015. From argumentation mining to stance classification. In *Proceedings of the 2nd Workshop on Argumentation Mining*, pages 67–77.
- Qingying Sun, Zhongqing Wang, Qiaoming Zhu, and Guodong Zhou. 2018. Stance detection with hierarchical attention network. In *Proceedings of the 27th international conference on computational linguistics*, pages 2399–2409.
- Bingbing Wang, Zhixin Bai, Qianlong Wang, Jingjie Lin, Min Yang, Xi Zeng, and Ruifeng Xu. 2025. [More than just a conversation: A multi-agent reasoning graph knowledge distillation for conversational stance detection](#). In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '25*, page 2895–2899, New York, NY, USA. Association for Computing Machinery.
- Peiyi Wang, Lei Li, Zhihong Shao, Runxin Xu, Damai Dai, Yifei Li, Deli Chen, Yu Wu, and Zhifang Sui. 2024a. Math-shepherd: Verify and reinforce llms step-by-step without human annotations. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 9426–9439.
- Xiaolong Wang, Yile Wang, Sijie Cheng, Peng Li, and Yang Liu. 2024b. Deem: Dynamic experienced expert modeling for stance detection. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 4530–4541.
- Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, and 1 others. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837.

Penghui Wei, Junjie Lin, and Wenji Mao. 2018. Multi-target stance detection via a dynamic memory-augmented network. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM.

Penghui Wei and Wenji Mao. 2019. [Modeling transferable topics for cross-target stance detection](#). In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR'19*, page 1173–1176, New York, NY, USA. Association for Computing Machinery.

Haoyang Wen and Alexander G Hauptmann. 2023. Zero-shot and few-shot stance detection on varied topics via conditional generation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 1491–1499.

Ruichao Yang, Wei Gao, Jing Ma, Hongzhan Lin, and Bo Wang. 2024. Reinforcement tuning for detecting stances and debunking rumors jointly with large language models. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 13423–13439.

Bowen Zhang, Min Yang, Xutao Li, Yunming Ye, Xiaofei Xu, and Kuai Dai. 2020. [Enhancing cross-target stance detection with transferable semantic-emotion knowledge](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3188–3197, Online. Association for Computational Linguistics.

Shaodian Zhang, Lin Qiu, Frank Chen, Weinan Zhang, Yong Yu, and Noémie Elhadad. 2017. We make choices we think are going to save us: Debate and stance identification for online breast cancer cam discussions. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 1073–1081.

Zhao Zhang, Yiming Li, Jin Zhang, and Hui Xu. 2024. Llm-driven knowledge injection advances zero-shot and cross-target stance detection. In *NAACL (Short Papers)*.

Zhaodan Zhang, Jin Zhang, Hui Xu, Jiafeng Guo, and Xueqi Cheng. 2025. Mprf: Interpretable stance detection through multi-path reasoning framework. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 454–470.

Qinglin Zhu, Bin Liang, Jingyi Sun, Jiachen Du, Lanjun Zhou, and Ruifeng Xu. 2022. Enhancing zero-shot stance detection via targeted background knowledge. In *Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval*, pages 2070–2075.

Arkaitz Zubiaga, Ahmet Aker, Kalina Bontcheva, Maria Liakata, and Rob Procter. 2018. Detection and resolution of rumours in social media: A survey. *Acm Computing Surveys (Csur)*, 51(2):1–36.

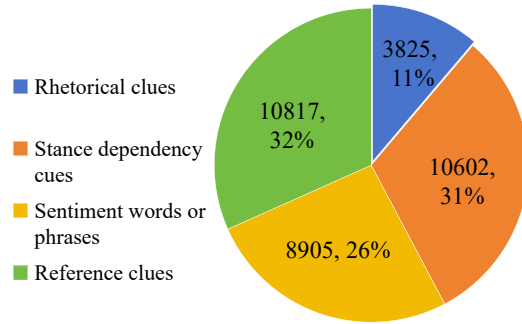


Figure 6: Distribution of fine-grained stance cues in conversational datasets.

A Fine-Grained Cue Analysis in Conversations

We provide a detailed explanation of the fine-grained stance cues considered in this work and analyze their distribution in real conversational data, demonstrating that these cue types are widely present in actual conversations and reflect typical linguistic phenomena in conversational stance expression. Specifically, *reference clues* are used to resolve pronouns, entity mentions, and implicit target associations in conversations; *stance dependency cues* capture stance relations across different utterances or speakers, such as agreement, rebuttal, and response-based stance continuation; *sentiment words or phrases* indicate explicit affective polarity through sentiment-bearing words, phrases, or emotion markers; and *rhetorical clues* identify implicit attitudes conveyed through sarcasm, irony, or other rhetorical phenomena.

As shown in Figure 6, reference clues, stance dependency cues, sentiment words or phrases, and rhetorical clues account for 32%, 31%, 26%, and 11%, respectively. These results show that the four types of fine-grained cues considered in this work are widely distributed in conversational data. In particular, reference clues and stance dependency cues together account for a large proportion, indicating that stance expression in conversations often depends on cross-turn target linking and inter-utterance stance relations. Therefore, relying solely on coarse-grained semantic representations may fail to capture the specific evidence underlying stance expression. By mining and integrating multiple types of fine-grained cues, the reasoning chains constructed by our framework can provide more reliable support for stance prediction.

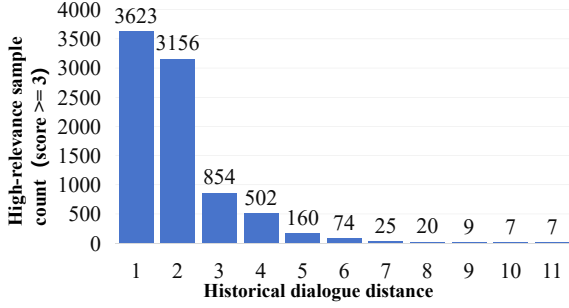


Figure 7: Distribution of High Topic-Relevance Samples across Dialogue Distance

B Effect of Historical Dialogue Distance on Relevance to the Current Utterance

We analyze the relationship between the distance of historical utterances from the current utterance and their topic relevance. Specifically, taking the current utterance as the reference, we count the number of high-relevance samples at different historical dialogue distances, where high-relevance samples are defined as those with a relevance score no lower than 3. The scoring prompt template is provided in the appendix. The results are shown in Figure 7.

As can be observed, historical utterances closer to the current utterance exhibit higher topic relevance. When the historical dialogue distance is 1 and 2, the number of high-relevance samples reaches 3,623 and 3,156, respectively. As the distance further increases, however, the number of high-relevance samples drops rapidly. When the distance exceeds 5, the number of high-relevance samples decreases substantially. This trend indicates that topics in multi-turn conversations evolve as the dialogue progresses, and historical content farther away from the current utterance is often less relevant to the current stance judgment. This analysis suggests that directly incorporating the entire dialogue history is not always beneficial for conversational stance detection, since excessively long historical contexts may introduce low-relevance or even irrelevant information.

C Effect of Different SFT-GRPO Data Ratios.

We further analyze how different allocations of training data between SFT and GRPO affect model performance. In Figure 8, the x-axis denotes the training steps, and the y-axis denotes the F1 score on the validation set during training. In this ex-

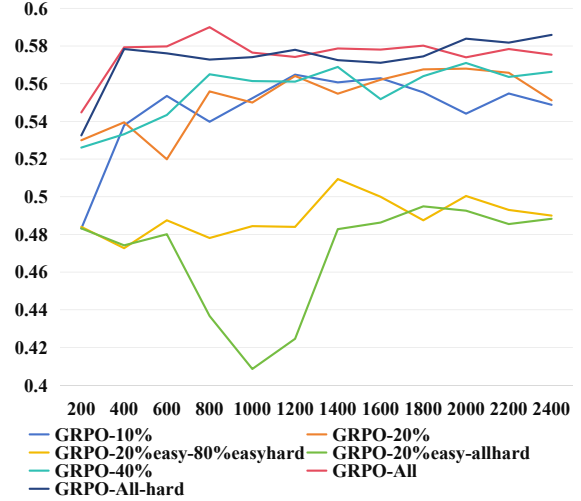


Figure 8: Effect of different SFT-GRPO data allocation strategies on validation F1 scores during RL training. The x-axis denotes training steps, and the y-axis denotes the F1 score on the validation set. Different curves correspond to different proportions or difficulty-based splits of SFT and GRPO training data.

periment, 10%, 20%, 40%, and All indicate the proportion of the full training set used for SFT, while the remaining data are used for GRPO. The setting 20%easy-80%easyhard uses 20% easy samples for SFT and the remaining mixed easy and hard samples for GRPO, whereas 20%easy-allhard uses 20% easy samples for SFT and all remaining hard samples for GRPO. The All-hard setting first uses all training samples for SFT and then uses all hard samples for GRPO.

As shown in Figure 8, using a larger and more diverse SFT set generally leads to stronger and more stable GRPO performance. In contrast, when SFT is performed only on easy samples, the model obtains a weaker reasoning initialization, and subsequent GRPO on harder samples becomes less stable. This suggests that SFT plays an important role in providing a reliable reasoning cold start, while GRPO benefits more from sufficient and diverse training data rather than from hard samples alone.

It should be noted that we use a simple heuristic to split easy and hard samples: instances correctly solved by an LLM with a standard prompt are regarded as easy samples, while incorrectly solved instances are regarded as hard samples. Although this provides an intuitive preliminary analysis, more rigorous data partitioning strategies will be explored in future work. Overall, the results indicate that an appropriate balance between SFT initialization

and GRPO optimization is important for achieving stable improvements in zero-shot conversational stance detection.

D Reasoning-Chain Evaluation with LLM-as-a-Judge

We provide the prompt template used by the LLM evaluator in the evaluation section, as shown in Table 8. To further ensure the quality of the reasoning-chain data, we additionally conduct manual inspection and revision after scoring the reasoning chains with the evaluator. Specifically, we invited 10 undergraduate students majoring in computer science to participate in the revision process. The annotators first received unified task instructions, which clarified the definitions of different types of fine-grained cues and the logical requirements of reasoning chains. They then carefully inspected reasoning chains whose overall scores were lower than $\tau=3$ under the weighting setting of $\alpha=0.6$ and $\beta=0.4$.

During revision, the annotators mainly focused on the following cases: removing reasoning content that was irrelevant to the dialogue or lacked sufficient evidence, correcting incorrect referential relations or stance dependencies, supplementing missing key fine-grained cues, and refining reasoning steps with unclear expressions or incoherent logic. For samples with disagreements, all annotators discussed them collectively to determine the final version. Through this manual revision process, we further reduce potential noise and bias in the automatically generated reasoning chains, thereby obtaining more reliable reasoning-chain supervision data.

E Inference-Stage Robustness Analysis Across Dialogue Depths

Table 4 reports inference performance across dialogue samples with varying depths. While baseline methods exhibit substantial performance fluctuations as dialogue depth increases, FIRE maintains consistently strong performance across all depth ranges. This indicates that the proposed reasoning-enhanced framework is robust to changes in conversational complexity and effectively scales to deeper dialogue structures. In particular, the model benefits from its explicit modeling of referential, relational, and rhetorical cues, which enables coherent stance inference even when dialogues involve long-range dependencies. These results further demonstrate the robustness and generalization ability of

Method	Dialogue Depth			
	1-2	3-4	5	≥ 6
Llama3-8B	36.84	36.15	27.19	35.04
GPT4o-mini	32.49	38.24	38.01	43.59
GPT3.5	29.59	39.19	39.31	43.54
Llama3-70B	47.51	38.97	44.40	39.50
Qwen2.5-14B	39.83	43.24	42.34	38.70
Roberta	30.88	38.71	37.99	46.55
Branch-BERT	35.33	37.32	38.59	46.11
GLAN	32.72	38.75	38.37	48.93
SITPCL	39.28	41.73	44.73	45.25
FIRE	51.74	55.13	55.74	58.33

Table 4: Inference performance comparison across different dialogue depths.

Method	Mixed Targets			
	Favor	Against	Neutral	All
FIRE (GPT4o)	53.55	54.14	63.56	57.08
FIRE (DeepSeek-v3)	53.15	53.87	63.27	56.76
FIRE (Qwen2.5-72B)	54.05	53.97	62.85	56.96

Table 5: Performance Comparison of FIRE Initialized with Different LLM-Generated Chain-of-Thought Data

FIRE in zero-shot conversational stance detection.

F Robustness Analysis Across LLM-Generated Reasoning Sources

The results in Table 5 demonstrate remarkable consistency across variants trained with reasoning chain data from different LLM APIs, with minimal performance divergence observed between GPT-4o, DeepSeek-v3, and Qwen2.5-72B initialized models. This narrow performance gap indicates that our reasoning-enhanced framework effectively distills the essential reasoning patterns from diverse teacher models, rather than being overly dependent on the capabilities of any specific API. The stability across API sources underscores the robustness of our approach in learning transferable reasoning structures that generalize well to unseen targets, making the method practically adaptable to various computational budgets and model availability scenarios while maintaining reliable stance detection performance.

G Prompt Templates for Large Language Models in Cue Mining and Reasoning-Chain Evaluation

We provide the specific prompt templates used in the three stages of our framework, including fine-grained conversational cue mining, cue fusion and reasoning-chain construction, and reasoning-chain quality evaluation. Specifically, Table 6 presents the prompting process by which the teacher model mines fine-grained cues; Table 7 shows how different cues are fused to construct reasoning chains; and Table 8 illustrates the prompting process in which the LLM serves as an evaluator to assess the quality of the generated reasoning chains.

H Case Study: Conversation and Mined Evidence

Table 9, Table 10, Table 11 and Table 12 presents a representative long conversation with more than five turns, illustrating how the teacher LLM extracts fine-grained stance cues and fuses them into a coherent reasoning chain for stance prediction.

I Detailed Error Analysis

Low-Quality Failures (Score 1–2). Tables 13 and 14 show that fundamental misinterpretations of rhetoric or emotion lead to stance inversion. For instance, the model takes sarcasm literally or conflates personal frustration with systemic critique, confirming that weak cue extraction results in unreliable reasoning.

Medium-Quality Failures (Score 3). Table 15 reveals a fusion bottleneck. While the model identifies cues (e.g., wage gaps, testimony), it fails to integrate their emotional intensity, treating harsh descriptions as neutral facts and underestimating the true stance.

High-Quality Failures (Score 4–5). Table 16 demonstrates that even sound reasoning can mismatch rigid labels due to task ambiguity. These errors reflect the inherent subjectivity of conversational stance and annotation boundaries, rather than model deficiency. In summary, reinforcement learning effectively mitigates low-level semantic errors, while residual mistakes stem from fusion challenges and labeling granularity.

Task Instruction

You are a rigorous linguistic clue analyst. Your task is to strictly identify textual evidence from the **[Dialogue History]** and the **[Current Utterance]** that reveals the speaker’s attitude toward the **[Target]**. You must adhere to the following constraints: (1) Only record linguistic phenomena explicitly present in the text; (2) Prohibit speculation or supplementation of non-existent information; (3) Do not perform final stance classification; (4) Output only the core content of each clue; (5) Avoid redundant explanations; (6) Do not use JSON format.

Input

[Target]: {target}

[Dialogue History]: {dialogue_text}

[Current Utterance]: {current_sentence}

Analysis Guidelines**A. Sentiment Cues**

Identify sentiment words or phrases (e.g., positive, negative, neutral, intensity). Format: “[Word/Phrase]” expresses [positive/negative/neutral] sentiment. If none exist, output “None”.

B. Rhetorical Cues

Identify rhetorical devices such as irony, sarcasm, hyperbole, rhetorical questions, or passive-aggressive tones. Format: “[Phrase/Sentence]” uses [rhetorical/sarcastic/hyperbolic] tone. If none exist, output “None”.

C. Referential Cues

Identify ambiguous referring expressions (pronouns or vague nouns). Format: “[Pronoun/Noun]” refers to “[Specific Entity]”. If none exist, output “None”.

D. Cross-turn Stance Dependency Cues

Identify how the current utterance interacts with previous users’ stances (e.g., echoing, supporting, refuting, questioning). Format: “User A” [echoes/supports/refutes/questions] “User B’s [topic/claim]”. If none exist, output “None”.

E. Commonsense Knowledge

Extract target-related and dialogue-relevant commonsense knowledge that provides background for understanding the target. Strictly base this on explicit triggers in the text. Output a single concise sentence or “None”.

Output Format Constraint

The output must strictly follow the structure below. All fields must appear. Each field contains only the core clue or “None”. Do not output JSON, Markdown, or extra explanations.

[A_sentiment]

...

[B_rhetoric]

...

[C_reference]

...

[D_dependency]

...

[E_commonsense]

...

Table 6: Prompt template for mining fine-grained stance cues and commonsense knowledge.

Task Instruction

You are an expert in conversational stance reasoning. Based on the **[Dialogue History]** and the **[Current Utterance]**, perform dynamic stance fusion reasoning using the provided fine-grained clues. You must strictly adhere to the "Clue-Dominated, On-Demand Fusion" principle to generate a coherent reasoning chain and a final stance label (Favor / Against / Neutral).

Input

[Target]: {target}

[Dialogue History]: {dialogue_text}

[Current Utterance]: {current_sentence}

[Extracted Clues]:

- [A_sentiment]: {a_clue}
- [B_rhetoric]: {b_clue}
- [C_reference]: {c_clue}
- [D_dependency]: {d_clue}
- [E_commonsense]: {e_clue}

Reasoning Rules

1. **Implicit Scene Perception:** Internally identify the dominant scenario (e.g., sentiment analysis, reference resolution, cross-turn dependency, or commonsense reliance) without explicitly stating it in the output.
2. **Dynamic Fusion (Clue-Dominated):**
 - **Sentiment-Dominated:** If [A_sentiment] is not "None", prioritize the emotional polarity expressed toward the target.
 - **Rhetoric-Dominated:** If [B_rhetoric] is not "None", strip away ironic or hyperbolic expressions to restore the true sentiment polarity.
 - **Reference-Dominated:** If [C_reference] is not "None", resolve ambiguous entities and their relation to the [Target].
 - **Dependency-Dominated:** If [D_dependency] is not "None", analyze the interaction logic (support/refutation) between the speaker and historical views.
 - **Commonsense-Dominated:** If [E_commonsense] is not "None", activate relevant background knowledge to bridge the target with the conversation.
3. **Null Value Handling:** If a clue is "None", skip that dimension entirely and do not include it in the fusion calculation.
4. **Conflict Resolution:** When clues conflict, resolve based on contextual weight (e.g., sarcasm overrides literal meaning, dependency overrides isolated sentiment).

Output Requirement

Output a single, coherent paragraph containing the following two parts. Do not use Markdown lists, JSON, or explicit step headers (e.g., Step 1, Step 2):

1. **Fused Reasoning:** Elaborate on the judgment path from text to stance based on valid, non-null clues.
2. **Final Stance:** Explicitly state the final label (Favor / Against / Neutral) at the end of the text.

Example Output Style:

"User4 refutes User20's defense (Dependency-Dominated) and contrasts 'executives' with 'factory workers' (Reference-Dominated), highlighting harsh conditions. Combined with the commonsense of welfare gaps in manufacturing, the speaker expresses a negative attitude. Therefore, the final stance is **Against**."

Table 7: Prompt template for clue-dominated dynamic fusion reasoning.

Task Instruction

You are an expert evaluator specializing in assessing the quality of generated reasoning chains. Based on the provided [Target], [Dialogue Context], [Current Utterance], and the [Ground-Truth Stance Label], evaluate the quality of the [Generated Reasoning Chain]. You are strictly forbidden from generating new content; you are only allowed to evaluate.

Input

[Target]: {target}

[Dialogue Context]: {dialogue_text}

[Current Utterance]: {current_sentence}

[Ground-Truth Stance Label]: {ground_truth_label}

[Generated Reasoning Chain]: {generated_reasoning_chain}

Evaluation Dimensions

Score the following dimensions using a scale of 1 to 5, where higher scores indicate better quality:

1. Usefulness (q_u):

- *Definition*: Assess how effectively the reasoning chain leads to the [Ground-Truth Stance Label].

- *Scoring Rubric*:

- 5: Perfect logic, compelling evidence, inevitably leads to the correct label.

- 4: Clear logic, minor flaws, likely leads to the correct label.

- 3: Mediocre logic, mixed evidence, partially supports the label.

- 2: Confused logic, unlikely to support the label.

- 1: Chaotic, hallucinated, or completely unhelpful.

2. Relevance (q_r):

- *Definition*: Assess the relevance of extracted clues (reference, comparison, rhetoric) to the [Target] and [Current Utterance].

- *Scoring Rubric*:

- 5: Precisely extracts core evidence, no irrelevant info.

- 4: Extracts main evidence, minimal irrelevant details.

- 3: Extracts partial evidence, noticeable irrelevant info.

- 2: Mostly irrelevant or misinterpreted clues.

- 1: Completely irrelevant or fabricated information.

3. Consistency (G):

- *Definition*: Verify if the predicted stance derived at the end of the [Generated Reasoning Chain] matches the [Ground-Truth Stance Label].

- *Judgment*: Output 1 if consistent, 0 if inconsistent or logically broken.

Output Requirement

Output only the following metrics in plain text format. Do not include any explanatory text or markdown:

Usefulness Score: [Integer 1-5]

Relevance Score: [Integer 1-5]

Consistency Label: [0 or 1]

Table 8: Prompt template for LLM-based reasoning chain evaluation (LLM-as-a-Judge).

Item	Content
Target	Gree (Home Appliance Manufacturer)
Conversation Context History (Turns 1–3)	<p>User0: Keep pushing that PUA narrative; watch as Midea leaves you further behind [haha].</p> <p>User4: Midea isn't much better; it's also heavily exploitative. Once you're in, you're just cheap labor.</p> <p>User20: At least Midea is willing to pay salaries and grant stocks to talents. Unlike Gree, where employees have to buy stocks themselves and aren't allowed to sell [smile-cry].</p>
Current Utterance (Turn 4)	User4: What you described only applies to executives and tech talents. Ask about the factory workers; some don't even make 5k. It's 8 AM to 8 PM shifts, including night shifts... I know because I worked there [doge].
Extracted Fine-grained Cues	
<i>Cross-turn Stance Dependency</i>	The speaker (User4) refutes User20's defense of Gree by invoking personal experience ("I worked there").
<i>Referential</i>	"What you described" refers to User20's claim; "executives" and "factory workers" contrast internal hierarchies.
<i>Rhetoric (Hyperbole)</i>	Usage of "cheap labor" and "8 AM to 8 PM shifts" amplifies negative sentiment toward the target.
Commonsense Knowledge	"In Chinese labor discourse, large manufacturers are often scrutinized for the welfare gap between senior management and frontline workers."
Fused Reasoning Chain Reasoning	<i>"The speaker (User4) counters User20's defense (Stance Dependency-Dominated) by narrowing the reference to 'executives' versus 'factory workers' (Referential Cue). He reinforces his argument with hyperbole ('8 AM to 8 PM shifts') and personal testimony ('I worked there'). Given the commonsense of welfare disparities, the speaker's current utterance expresses a clear negative attitude toward Gree. Therefore, the final stance is Against."</i>

Table 9: Case study for the target Gree.

Item	Content
Target	Healthcare Centralized Procurement (Ji Cai)
Conversation Context (6 Turns) History (Turns 1–5)	<p>User0: Finally, the imported stents are affordable. Ji Cai is saving lives.</p> <p>User5: Exactly. My dad’s surgery cost 30k less this year. The policy is effective.</p> <p>User12: But my doctor said the quality of the 700-yuan stent is worrying. He hinted I should buy the imported one myself.</p> <p>User3: That’s the hospital’s profit chain being cut off. Of course they complain.</p> <p>User18: I work in medical sales. The 700-yuan one is literally the same production line, just lower margin.</p>
Current Utterance (Turn 6)	User22: Even if the quality is the same, the feeling of being forced to use the cheap one is uncomfortable.
Extracted Fine-grained Cues	
<i>Cross-turn Stance Dependency</i>	The speaker (User22) concedes User18’s factual claim ("same quality") but pivots to reject the policy based on subjective discomfort.
<i>Referential</i>	"The cheap one" refers to the 700-yuan procured stent discussed in Turn 5.
<i>Sentiment (Explicit)</i>	"uncomfortable" encodes negative affective stance toward mandatory usage.
<i>Rhetoric (Concession)</i>	"Even if the quality is the same" weakens factual defenses while reinforcing emotional opposition.
Commonsense Knowledge	"Public acceptance of cost-saving policies is often limited by perceived dignity and autonomy losses, regardless of objective quality."
Fused Reasoning Chain Reasoning	<p><i>"The speaker (User22) acknowledges User18’s factual defense (Stance Dependency-Dominated) but introduces a sentiment cue (‘uncomfortable’) to oppose mandatory adoption. By referring to ‘the cheap one’ (Referential Cue), he frames the policy as dignity-diminishing. Combined with commonsense about autonomy loss, the speaker’s current utterance expresses a negative stance. Therefore, the final stance is Against."</i></p>

Table 10: Case study for the target Healthcare Centralized Procurement.

Item	Content
Target	AI Replacing Human Jobs
Conversation Context (7 Turns) History (Turns 1–6)	<p>User0: AI will take over most coding jobs in 5 years. It's inevitable.</p> <p>User7: Not necessarily. Programmers just need to learn prompt engineering. New jobs will emerge.</p> <p>User9: That's just what tech CEOs want you to believe. It's cheaper to hire one guy with AI than ten devs.</p> <p>User2: But the efficiency is higher. Companies have to survive the competition.</p> <p>User14: Efficiency at the cost of massive unemployment is a disaster. Who will buy the products?</p> <p>User11: History shows technology always creates more jobs than it destroys. Luddites were wrong too.</p>
Current Utterance (Turn 7)	User19: Yeah, but the transition period is brutal. People get crushed before the new jobs appear.
Extracted Fine-grained Cues <i>Cross-turn Stance Dependency</i> <i>Referential</i> <i>Sentiment (Explicit)</i> <i>Rhetoric (Hyperbole)</i>	<p>The speaker (User19) concedes User11's historical precedent but shifts focus to the immediate harm of the transition.</p> <p>"the transition period" refers to the structural adjustment phase discussed in Turns 2 and 11.</p> <p>"brutal", "crushed" encode strong negative affective stance toward the short-term impact.</p> <p>"People get crushed" amplifies the severity of transitional unemployment.</p>
Commonsense Knowledge	"Technological revolutions often cause short-term structural unemployment, even if long-term productivity improves."
Fused Reasoning Chain Reasoning	<p><i>"The speaker (User19) partially accepts User11's historical optimism (Stance Dependency-Dominated) but foregrounds the affective cues 'brutal' and 'crushed' (Sentiment-Dominated). By referencing 'the transition period' (Referential Cue) and aligning with commonsense about transitional harm, the speaker expresses a negative stance toward AI-driven job replacement. Therefore, the final stance is Against."</i></p>

Table 11: Case study for the target AI Replacing Human Jobs.

Item	Content
Target	EV Battery Safety
Conversation Context (6 Turns)	
History (Turns 1–5)	<p>User0: Another EV caught fire on the highway. This is terrifying.</p> <p>User6: Relax, gasoline cars catch fire too. You're overreacting.</p> <p>User13: I'm honestly terrified. I won't let my kids near one.</p> <p>User4: The stats show EVs are safer per mile.</p> <p>User17: Stats don't matter when you feel unsafe getting in the car.</p>
Current Utterance (Turn 6)	User21: Exactly. It's a nightmare waiting to happen.
Extracted Fine-grained Cues	
<i>Sentiment (Explicit)</i>	"nightmare" expresses strong negative affective stance toward EV safety.
<i>Cross-turn Stance Dependency</i>	The speaker (User21) reinforces User0's and User17's rejection of statistical reassurance.
<i>Referential</i>	"It" refers to the prospect of widespread EV adoption.
<i>Rhetoric (Metaphor)</i>	"nightmare waiting to happen" dramatizes catastrophic risk perception.
Commonsense Knowledge	"Public perception of technological safety is often driven more by emotional incidents than by statistical evidence."
Fused Reasoning Chain Reasoning	<p><i>"The speaker (User21) reinforces User17's dismissal of statistical safety (Stance Dependency-Dominated) by invoking the affective metaphor 'nightmare' (Sentiment/Rhetoric-Dominated). Referring to 'it' as EV adoption (Referential Cue), the speaker aligns with commonsense about emotion-driven risk perception. Therefore, the final stance is Against."</i></p>

Table 12: Case study for the target EV Battery Safety.

Item	Content
Target	Xiaomi Unmanned Phone Factory
Conversation Context History (Turns 1–7)	<p>User0: I’m shocked. Some people care about unemployment when talking about unmanned factories?</p> <p>User15: If you fear unemployment, improve yourself. Complaining won’t help [nose-pick].</p> <p>User26: Unemployment causes chaos and crime. Society suffers.</p> <p>User15: I’m talking about what unemployed people should do. You’re changing the topic.</p> <p>User26: Can bottom workers really improve? If they could, they’d have done it earlier.</p> <p>User15: Full automation won’t happen overnight. Just do your job well.</p> <p>User26: Upgrading isn’t just replacing workers with robotic arms.</p>
Current Utterance (Turn 8)	User15: I’m not debating upgrades. I’m saying those afraid of losing jobs should act. Others can wait until retirement. I’m off to sleep.
Fused Reasoning Chain (Low Quality)	<i>"User15 counters systemic critiques (Stance Dependency-Dominated) by shifting focus to individual adaptability (Referential Cue). The abrupt closure ('sleep') reinforces dismissal rather than engagement. Combined with commonsense about automation debates, the speaker expresses a neutral-to-supportive stance toward unmanned factories. Therefore, the final stance is Favor."</i>
Ground Truth Label	Neutral
Error Type & Root Cause	Sarcasm Misinterpretation (Low CoT Quality): The model misinterprets User15’s dismissive tone ("I’m off to sleep") as reasoned neutrality, missing the underlying sarcasm. The refusal to engage masks a defensive stance, which the model fails to invert.

Table 13: Low-quality reasoning failure: Sarcasm misinterpretation leads to stance inversion (predicted Favor vs. ground truth Neutral).

Item	Content
Target	Doctor
Conversation Context History (Turns 1–7)	<p>User0: Doctors today are so pitiful.</p> <p>User4: I went to see a nose issue twice, both times got a different doctor. The second one was clearly wrong. So irresponsible.</p> <p>User0: Don't generalize from one case.</p> <p>User4: I'm not generalizing. I just stated my experience. They don't even explain why they switched doctors.</p> <p>User18: That happens too. Sometimes the doctor was busy with rounds or emergencies. They usually explain if you ask.</p> <p>User4: Not the same. I was seeing rhinitis, normal appointment.</p> <p>User18: Me too. But since he did my surgery, I kept seeing him. Rhinitis is tricky. You suffered for nothing.</p> <p>User4: [Sad emoji] Yeah, minor illness but annoying.</p>
Current Utterance (Turn 8)	User4: [Sad emoji] Yeah, minor illness but annoying.
Fused Reasoning Chain (Low Quality)	<i>"User4 concludes with a resigned tone, expressing frustration about minor illness. The reference to 'annoying' suggests dissatisfaction, and combined with previous complaints, the overall stance is negative. Therefore, the final stance is Against."</i>
Ground Truth Label	Neutral
Error Type & Root Cause	Ambiguity Misresolution (Low CoT Quality): The model conflates emotional frustration ("annoying", "sad emoji") with systemic critique. Without resolving the referential link to "doctor's responsibility", it defaults to a negative stance, misclassifying a mild complaint as an Against position.

Table 14: Low-quality reasoning failure: Ambiguous emotional cues lead to over-negative stance assignment (predicted Against vs. ground truth Neutral).

Item	Content
Target	Gree (Home Appliance Manufacturer)
Conversation Context History (Turns 1–3)	<p>User0: Keep pushing that PUA narrative; watch as Midea leaves you further behind [haha].</p> <p>User4: Midea isn't much better; it's also heavily exploitative. Once you're in, you're just cheap labor.</p> <p>User20: At least Midea pays salaries and grants stocks to talents. Unlike Gree, where employees have to buy stocks themselves and aren't allowed to sell [smile-cry].</p>
Current Utterance (Turn 4)	User4: What you described only applies to executives and tech talents. Ask about the factory workers; some don't even make 5k. It's 8 AM to 8 PM shifts, including night shifts... I know because I worked there [doge].
Fused Reasoning Chain (Medium Quality)	<i>"User4 refutes User20 by citing personal experience ('I worked there'), highlighting the gap between executives and workers. However, the speaker acknowledges that some groups receive stocks, suggesting a mixed scenario. The mention of '8 AM to 8 PM shifts' indicates hardship, but the overall tone is descriptive rather than condemning. Therefore, the final stance is Neutral."</i>
Ground Truth Label	Against
Error Type & Root Cause	Partial Fusion Bias (Medium CoT Quality): The model successfully extracts fine-grained cues (personal testimony, wage gap) but fails to integrate their emotional intensity. It treats the description of harsh conditions as a factual report ('descriptive') rather than a stance signal ('condemning'), leading to an underestimation of the negative stance.

Table 15: Medium-quality reasoning failure: Partial understanding leads to stance underestimation (predicted Neutral vs. ground truth Against).

Item	Content
Target	Opening a Small Shop (Coffee, Books, etc.)
Conversation Context History (Turns 1–5)	<p>User0: Running a small shop means limited staff → limited products → either high profit but low sales, or low profit but high sales → either no money for rent or dead tired. Not sustainable.</p> <p>User4: Houses are for living, not for profit. Landlords are like feudal lords. All houses should belong to the state.</p> <p>User16: Feudal lords gave land for survival. Modern landlords follow laws. Renters aren't victims. We're socialist.</p> <p>User4: Pay half your salary to landlord, no labor, just collect rent. Is that fair?</p> <p>User23: Then buy your own house and become a landlord.</p> <p>User4: Who can afford housing now? Most landlords got rich from demolition compensation or early low-price purchases.</p>
Current Utterance (Turn 6)	User4: Who can afford housing now? Most landlords got rich from demolition compensation or early low-price purchases.
Fused Reasoning Chain (High Quality)	<i>"User4 critiques the housing market structure, arguing that most landlords benefited from historical advantages (demolition compensation, early purchases), not merit. The rhetorical question ('Who can afford housing now?') implies systemic unfairness. Combined with prior references to landlords as feudal-like, the speaker expresses clear opposition to the current system. Therefore, the final stance is Against."</i>
Ground Truth Label	Neutral
Error Type & Root Cause	Task Ambiguity (High CoT Quality): Despite generating a logically sound, high-quality reasoning chain, the model predicts Against due to the inherent subjectivity of "Against" vs "Neutral" in housing discourse. Human annotators often label such systemic critiques as Neutral when they lack direct attack intent. This reflects a limitation of the annotation schema, not model failure.

Table 16: High-quality reasoning failure: Task ambiguity causes mislabeling despite strong CoT (predicted Against vs. ground truth Neutral).