



PRISMA: Preference-Reinforced Self-Training Approach for Interpretable Emotionally Intelligent Negotiation Dialogues

Prajwal Vijay Kajare^{‡*}, Priyanshu Priya^{†*}, Bikash Santra[‡], Asif Ekbal[†]

[†]Department of Computer Science and Engineering, Indian Institute of Technology Patna, India

[‡]School of Artificial Intelligence and Data Science, Indian Institute of Technology Jodhpur, India

[‡]{p24ai0004,bikash}@iitj.ac.in, [†]{priyanshu_2021cs26,asif}@iitp.ac.in

Abstract

Emotion plays a pivotal role in shaping negotiation outcomes, influencing trust, cooperation, and long-term relationships. Developing negotiation dialog systems that can recognize and respond strategically to emotions is, therefore, essential to create more effective human-centered interactions. Beyond generating emotionally appropriate responses, *interpretability* - understanding how a system generates a particular emotion-aware response, is critical for fostering reliability and building rapport. Driven by these aspects, in this work, we introduce PRISMA, an interpretable emotionally intelligent negotiation dialogue system targeting two application domains, *viz.* job interviews and resource allocation. To enable interpretability, we propose an Emotion-aware Negotiation Strategy-informed Chain-of-Thought (ENS-CoT) reasoning mechanism, which mimics human negotiation by perceiving, understanding, using, and managing emotions. Leveraging ENS-CoT, we curate two new datasets: JobNego (for job interview negotiation) and ResNego (for resource allocation negotiation). We then leverage these datasets to develop PRISMA by augmenting self-training with Direct Preference Optimization (DPO), guiding agents toward more accurate, interpretable, and emotionally appropriate negotiation responses. Automatic and human evaluation on JobNego and ResNego datasets demonstrate that PRISMA substantially enhances interpretability and generates appropriate emotion-aware responses, while improving overall negotiation effectiveness¹.

1 Introduction

Negotiation is inherently a social process where parties seek to settle exchanges to satisfy their needs (Spector, 1977). Traditionally, research has

* Authors are jointly first authors.

¹Code and dataset are available at <https://github.com/kajareprajwal1143/PRISMA> and <https://ai-nlp-ml.github.io/resources.html>.

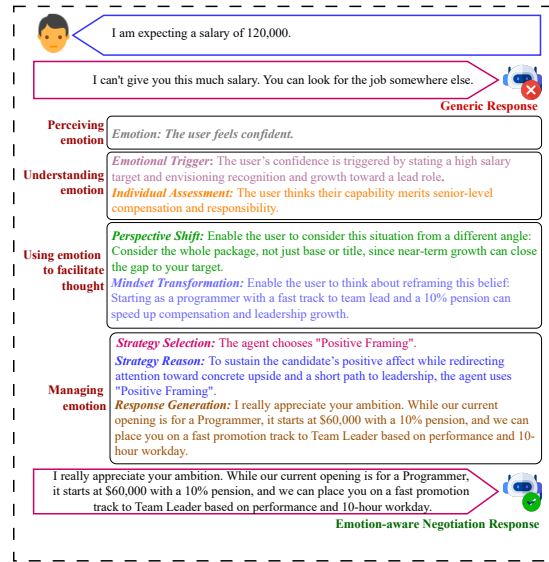


Figure 1: A dialogue snippet illustrating the role of emotional intelligence with interpretable ENS-CoT rationales in generating emotion-aware responses and enhancing negotiation outcomes.

examined negotiation through economic modeling (Sebenius, 1992), decision-making theory (Bazerman et al., 2000), and social psychology (Rubin and Brown, 2013). However, these classical views often portray negotiators as purely rational agents, overlooking the pivotal role of emotions. Behavioral evidence shows that emotions systematically drive concession-making, perceived fairness, joint gains, and cooperation (Lawler and Yoon, 1995; Forgas, 1998). Thus, emotional intelligence (EI: the ability to perceive, understand, use, and manage emotions) (Goleman, 2005), is vital for developing human-like negotiation dialogue agents.

Although modern agents based on the large language models (LLMs) can reason via chain-of-thought (CoT) (Wei et al., 2022) and recognize emotions (Li et al., 2024), they remain limited in strategically leveraging emotions during negotiation. They often generate static rather than adaptive emotion-aware strategies (Krocze et al., 2025),

struggle to distinguish genuine from manipulated emotions (Huang and Hadfi, 2024), and lack reasoning about emotional cause-and-effect dynamics (Bosse et al., 2013). These limitations are critical in high-stakes negotiation settings like job interviews and resource allocation, where emotion directly influences strategic decisions (e.g., frustration may hasten concessions, while excitement can trigger overly ambitious offers), shaping negotiation outcomes. For such agents to be reliable, users must understand the rationale behind their emotionally-tuned responses (Dang et al., 2024). Hence, beyond emotional intelligence, a negotiation agent must also be interpretable; not merely generate responses, but also explain how it generates these responses to build rapport and foster credibility.

Motivated by this need, we introduce PRISMA, an interpretable, emotionally intelligent negotiation dialogue system that first reasons via a CoT mechanism and then generates emotion-aware negotiation responses. To enable interpretable EI in PRISMA, we draw inspiration from the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT) (Mayer et al., 2003), which encompasses four abilities: *perceiving emotion* (identifying the user’s current emotion), *understanding emotion* (explaining its cause and effect), *using emotion to facilitate thought* (guiding reasoning through perspective shifting and mindset transformation), and *managing emotion* (selecting and justifying an emotion-aware negotiation strategy, then realizing it in the response). As shown in Figure 1, unlike generic response that may frustrate the user, PRISMA interprets the user’s confidence for high salary, links it to aspirations for recognition and leadership, reframes priorities, and adopts a ‘Positive Framing’ strategy emphasizing promotion and growth, leading to an emotionally attuned counteroffer to improve the negotiation outcome.

To advance this line of research, we introduce two novel negotiation dialogue datasets, *viz.* JobNego and ResNego, generated using ChatGPT (OpenAI, 2024) for job interview and resource allocation negotiations, respectively. These datasets are constructed through an interpretable emotion-aware dialogue generation process. We first design diverse, realistic negotiation scenarios, then, conditioned on scenarios, generate dialogues augmented with MSCEIT-inspired Emotion-aware Negotiation Strategy-informed CoT (ENS-CoT) rationales, where the agent first reasons to select appropriate emotion-aware strategies and then generates

emotion-aware responses during negotiation.

Curating large-scale, high-quality dialogue data with rationale annotations, however, is time- and resource-intensive. In contrast, unlabeled negotiation data can be synthesized at scale and leveraged via semi-supervised methods, such as self-training to improve LLM performance (Zelikman et al., 2022; Gulcehre et al., 2023). Moreover, Direct Preference Optimization (DPO) has shown promise for aligning LLMs with preferred outputs (Rafailov et al., 2023; Deng and Mineiro, 2024). Building on these insights, we propose a preference-reinforced self-training approach that integrates self-training with DPO to develop PRISMA. It begins with supervised initialization to enable the LLM to generate ENS-CoT rationales and appropriate emotion-aware responses, followed by an iterative loop alternating DPO-based preference learning (aligning agents to preferred emotion-aware negotiation responses) and self-training (adding high-confidence ENS-CoT pseudo-labels for unlabeled contexts back into supervised fine-tuning). This integration optimizes step-by-step CoT reasoning, resulting in improved emotion-aware negotiation responses. Empirical evaluations indicate that the proposed approach enhances the interpretability and emotionally attuned negotiation behavior in PRISMA.

In summary, the key contributions of this work are as follows: (i) Present PRISMA, an interpretable, emotionally intelligent negotiation dialogue agent. To the best of our knowledge, this work pioneers interpretability modeling in emotion-aware negotiation conversations; (ii) Propose an Emotion-aware Negotiation Strategy-informed Chain-of-Thought (ENS-CoT) reasoning mechanism to enable interpretable emotional intelligence; (iii) Curate two novel datasets, JobNego (job interview negotiation) and ResNego (resource allocation negotiation), annotated with ENS-CoT rationales; and (iv) Introduce a preference-reinforced self-training approach integrating self-training with DPO, demonstrating its effectiveness in producing interpretable, emotion-aware negotiation responses.

2 Related Work

Negotiation has long been a central topic in psychology, economics, and affective computing (Carnevale et al., 1992; Boothby et al., 2023). In recent years, automated negotiation dialogue systems have made significant progress (Lewis et al., 2017; Zhou et al., 2019), demonstrating that mod-

eling negotiators’ behavior (He et al., 2018; Zhang et al., 2020; Chawla et al., 2022), strategies (Zhao et al., 2019; Chawla et al., 2021b; Yamaguchi et al., 2021), personalities (Yang et al., 2020; Priya et al., 2024, 2025b), and politeness (Zhao et al., 2018; Priya et al., 2025a) can substantially enhance human-agent negotiation interactions. Emotions also play a critical role in shaping negotiation outcomes (Morris and Keltner, 2000; Douglas and Coburn, 2014), and a growing number of studies have begun integrating emotional cues into automated negotiation systems (Chawla et al., 2021a, 2023; Keskin et al., 2024).

The advent of LLMs (Minaee et al., 2024) has further spurred research in negotiation dialogue systems (Fu et al., 2023; Gandhi et al., 2023; Abdelnabi et al., 2023; Deng et al., 2023b). Emotional intelligence (awareness and management of emotions) (Maier et al., 2004) and interpretability (ability to understand the reasoning behind the model’s outputs) (Fuhrman et al., 2022), have become key considerations in designing reliable LLM-based agents (Li et al., 2023; Yang et al., 2024; Feng et al., 2025; Kasbouya and Sael, 2025). To support interpretability, Chain-of-Thought (CoT) prompting (Wei et al., 2022) has been developed, enabling LLMs to emulate human reasoning processes and produce more transparent, step-wise rationales. Building on this, a number of studies have explored the use of CoT prompting to enhance LLM reasoning across diverse applications (Zhang et al., 2022; Yang et al., 2023), including dialogue generation (Deng et al., 2023a).

Complementing CoT approaches, reinforcement learning (RL) techniques (Bai et al., 2022) have been applied to further refine LLM reasoning (Trung et al., 2024; Wang et al., 2025). Notably, Direct Preference Optimization (DPO) (Rafailov et al., 2023), which aligns model outputs with human preferences, has shown promise in improving reasoning and output quality (Deng and Mineiro, 2024; Lai et al., 2024). Additionally, semi-supervised methods (Scudder, 1965; Amini et al., 2025), such as self-training (Amini et al., 2025), have gained traction for enhancing reasoning capabilities in LLMs, providing a scalable mechanism for leveraging unlabeled or weakly labeled data (Lee et al., 2024; Wu et al., 2024).

Despite extensive research on modeling opponents’ behavior, strategies, emotions, and politeness in developing negotiation dialogue agents, little research has examined the specific role of in-

terpretable emotional intelligence within negotiation dialogues. Moreover, combining self-training with preference signals remains underexplored, especially in negotiation scenarios where interpretability and emotional intelligence are critical. To address these gaps, we introduce PRISMA, a preference-reinforced self-training approach that leverages Emotion-aware Negotiation Strategy-informed CoT (ENS-CoT) rationales to align responses with users’ emotional states while ensuring interpretability during negotiation. PRISMA is designed to be scalable and adaptable across diverse application domains.

3 Dataset

We evaluate PRISMA on two newly created negotiation dialogue datasets: JobNego and ResNego. JobNego comprises job interview negotiations between a candidate (user) and an employer (agent), covering salary, working hours, promotion track, and related terms. ResNego includes resource-allocation negotiation dialogues in a camping setting, where parties bargain over food, water, firewood, and related supplies.

3.1 Dataset Preparation

Given the high cost of creating interpretable, emotionally intelligent negotiation dialogue datasets, where the agent must combine domain constraints and trade-offs with emotion-aware negotiation strategy selection, we adopt an interpretable and scalable dataset construction scheme. Specifically, we introduce an Emotion-aware Negotiation Strategy-informed Chain-of-Thought (ENS-CoT) reasoning mechanism to generate emotion-aware negotiation responses in an interpretable manner. ENS-CoT mimics human negotiation behavior by identifying, understanding, using, and managing emotions during negotiation through appropriate emotion-aware negotiation strategies. Using ENS-CoT, we curate two new interpretable emotion-aware negotiation dialogue datasets: JobNego and ResNego for job interview and resource allocation domains, respectively. The entire dataset construction involves three steps: (1) Negotiation Scenario Generation, (2) Emotion and Negotiation Strategy Formulation, and (3) ENS-CoT-enhanced Dialogue Generation.

(1) Negotiation Scenario Generation. The scenarios summarize the negotiation setup, including task, roles, issues, constraints, and goals underlying each dialogue. The diverse and realistic sce-

narios are essential for generating high-quality negotiation dialogues. We, therefore, use NegoChat (Konovalov et al., 2016) and CaSiNo (Chawla et al., 2021b) datasets as seed corpora to generate the initial pool of job interview and resource allocation negotiation scenarios, respectively. For each dialogue, we prompt ChatGPT in a zero-shot setting to produce a brief scenario description.

To identify an effective prompt, we design four candidate prompts (including the natural language instructions and target dialogues), and for each prompt, we generate 30 scenarios per domain. These generated scenarios are then manually evaluated by the three human annotators for scenario appropriateness on a scale of 1-3 (1-low, 2-moderate, 3-high)². An inter-annotator Kappa agreement score (κ) (McHugh, 2012) reaches 0.85 (for job interview scenarios) and 0.88 (for resource allocation scenarios). The prompt generating the highest number of scenarios with a score of 3 for both domains is selected as the final prompt. This prompt is then used to prompt ChatGPT to generate scenarios for all the dialogues. The generated scenarios are further verified by the same annotators, achieving $\kappa = 0.88$ and 0.86 among them for job interview and resource allocation scenarios, respectively. It yields the unique 105 and 1,030 job interview and resource allocation scenarios, respectively.

We expand this scenario pool by leveraging ChatGPT’s in-context learning ability (Wang et al., 2023). We randomly sample three dialogue–scenario exemplars per domain from the initial pool and use them to prompt ChatGPT in a few-shot setup to generate additional scenarios. To ensure quality and diversity, we remove duplicates and filter out inadequate scenarios that lack proper information on negotiation setup or have incomplete sentences. This yields an additional 735 job interviews and 618 resource allocation scenarios, enriching the pool for subsequent dialogue generation. The scenario generation prompt template is given in the Appendix A.

(2) Emotions and Emotion-aware Negotiation Strategies Formulation.

Emotions significantly influence negotiation dynamics and outcomes (Barry and Oliver, 1996). Users naturally express positive, negative, or neutral emotions during negotiation, and these emotions determine how ne-

gotiation unfolds (Overbeck et al., 2010). In this work, we adopt 12 emotion categories from the negotiation literature (Barry, 1999; Van Kleef et al., 2008), including *joy, confidence, positivity, gratitude, trust, surprise, anger, disappointment, frustration, fear, anxiety, and neutral*.

Effective negotiation depends on how the agent manages users’ emotions (Van Kleef and Lelieveld, 2022). The dynamic nature of users’ emotions directly guides the agent’s choice of strategy and, in turn, negotiation outcomes (Lawler and Thye, 1999). To effectively manage these emotions and drive negotiation toward mutually beneficial outcomes, we define 12 emotion-aware negotiation strategies grounded in negotiation emotion management principles (Yurtsever, 2004). These strategies are designed to either amplify positive emotions, mitigate negative ones, or maintain neutrality, and include *savoring, positive reinforcement, expressing optimism, cognitive reappraisal, positive framing, emotion diffusion, expressive suppression, active listening, perspective-taking, problem solving, escalate assurance, and no strategy*. Definitions of these strategies are given in Table 1.

(3) ENS-CoT-enhanced Dialogue Generation.

It involves following steps:

(a) ENS-CoT Reasoning Chain Description.

To encode interpretable emotional intelligence in negotiation dialogue systems, we introduce the Emotion-aware Negotiation Strategy-informed CoT (ENS-CoT) reasoning mechanism, grounded in the MSCEIT’s four abilities (Mayer et al., 2003): Perceiving emotion, Understanding emotion, Using emotion to facilitate thought, and Managing emotion. Perceiving emotion refers to the ability to identify user’s current emotion, understanding emotion explains its causes via cognitive appraisal (emotional trigger and individual assessment), using emotion to facilitate thought specifies how emotion guides reasoning (via perspective shifting and mindset transformation), and managing emotion selects and justifies an emotion-aware negotiation strategy and realizes it in agent’s response. An ENS-CoT instance is denoted as an octuple $\langle EM, ET, IA, PS, MT, SS, SR, RG \rangle$, described below:

- **Emotion (EM):** User’s expressed emotion.
- **Emotional Trigger (ET):** Cause (external situation or internal thought) that triggers the emotion.
- **Individual Assessment (IA):** Personal interpretation, evaluation, and internal response to the emotional trigger, based on the user’s past experi-

²The annotators hold Ph.D. in Linguistics with extensive expertise in negotiation, emotional intelligence, and dialogue design. They are compensated as per the institute norms.

Emotion-aware Negotiation Strategy	Definition
Savoring	Actively appreciating and amplifying positive moments (e.g., shared successes, agreement points) to maintain a constructive climate and strengthen relationships to help achieve integrative outcomes.
Positive Reinforcement	Compliment and acknowledge constructive behavior or ideas to enhance positive affect, reinforce cooperation, and make progress toward mutual goals easier.
Expressing Optimism	Communicate a credible, positive outlook about reaching a mutually beneficial deal to encourage cooperative effort.
Cognitive Reappraisal	Reinterpret the situation to alter its emotional impact (e.g., view criticism as useful feedback) to reduce defensiveness and keep focus on objectives; linked to better outcomes.
Positive Framing	Shift emphasis from potential losses to achievable gains to transform competitive stances into collaborative problem-solving.
Emotion Diffusion	Acknowledge heightened affect and de-escalate with calm, soft language and constructive addressing of issues to prevent conflict escalation.
Expressive Suppression	Temporarily inhibit or mask one’s emotional display to maintain composure and avoid escalation in sensitive moments.
Active Listening	Attend to verbal and non-verbal cues, paraphrase, and validate concerns so the counterpart feels heard and understood; builds trust and diagnostic clarity.
Perspective-Taking	Deliberately adopt the counterpart’s viewpoint to infer their motives, constraints, and emotions, enabling more appropriate and empathetic responses.
Problem Solving	Collaboratively identify, analyze, and resolve issues to craft integrative solutions that meet mutual needs and goals.
Escalate Assurance	Address concerns with concrete guarantees, clarifications, and commitments (e.g., evidence, safeguards) to increase confidence and trust, stabilizing the interaction.
No Strategy	Refers to a neutral, task-focused response without explicit emotion management when affect is low/neutral or strategy use is unwarranted.

Table 1: The definition of different emotion-aware negotiation strategies. Examples are given in Appendix A.

ences, beliefs, expectations, and personal values.

- **Perspective Shift (PS)**: Adopting an alternate viewpoint to reduce reactivity and trade-offs.
- **Mindset transformation (MT)**: Reframing limiting beliefs into constructive priors for decision making.
- **Strategy Selection (SS)**: Choosing a pertinent emotion-aware negotiation strategy.
- **Strategy Reason (SR)**: Reasoning behind the selected strategy.
- **Response Generation (RG)**: Using the selected strategy to generate agent’s utterance in an ongoing dialogue.

(b) **Designing Sample ENS-CoT-enhanced Dialogues.** We author exemplar dialogues conditioned on the generated scenarios, the user emotion inventory, and the agent’s emotion-aware negotiation strategy catalog across both domains. For each domain, we create 10 Wizard-of-Oz dialogues (Kelley, 1984) with two annotators role-playing the user and the agent. The dialogues proceed turn-by-turn, grounded in the scenario: the user produces an utterance and expresses an emotion; the agent applies ENS-CoT to perceive, understand, use, and manage that emotion before responding. Every agent turn is annotated with an ENS-CoT rationale, ensuring that responses are emotionally attuned, strategically appropriate, and interpretable. ‘Guidelines for Designing Sample ENS-CoT-enhanced Dialogues’ are provided in Appendix A.

(c) **Dialogue Synthesis via Prompting.** We use the authored dialogue samples as few-shot exemplars to prompt ChatGPT to generate entire dialogue datasets. To finalize the prompt, we exper-

iment with four manually designed prompts, including task instructions with the target negotiation scenario, and few-shot exemplars (three randomly selected negotiation scenarios and their corresponding dialogues with ENS-CoT annotations). For each prompt, we generate 15 dialogues, each for both the domains by prompting ChatGPT. These generated dialogues are evaluated by the same three annotators for emotional intelligence (EI) and negotiation quality (NQ) on a 1-3 scale (low to high) under the supervision of a subject matter expert with HR and sales backgrounds. We obtain $\kappa = 0.81$ and 0.84 for EI and NQ, respectively, indicating good quality negotiation dialogues. We select the prompt that generates the highest number of dialogues rated as 3 for both EI and NQ, and use it to construct the final datasets. The final dataset statistics are reported in Table 2. The dialogue synthesis prompt template, along with a detailed analysis of the joint emotion–strategy distribution and scenario-level topic diversity for both JobNego and ResNego, is provided in Appendix A.

	JobNego			ResNego		
	Train	Dev	Test	Train	Dev	Test
# Dialogues	504	168	168	988	330	330
# Utterances	8,124	2,196	2,172	13,695	3,314	3,178
Avg. Utterances/Dialogue	16.12	13.07	12.93	13.86	10.04	9.63

Table 2: JobNego and ResNego dataset statistics.

3.2 Dataset Filtering and Quality Evaluation

After generating the datasets, we evaluate each dialogue based on the following criteria: Emotional Intelligence (EI), Strategy Appropriateness (SA),

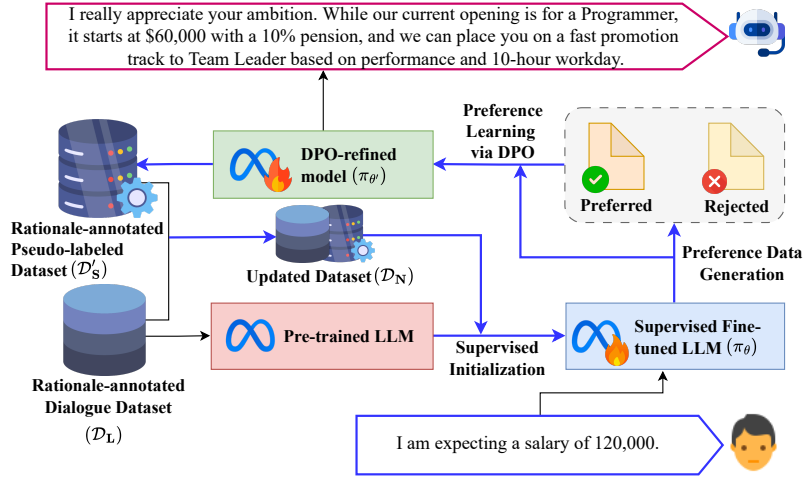


Figure 2: Architecture of the proposed Preference-Reinforced Self-training approach for an interpretable emotionally intelligent negotiation dialogue system - PRISMA.

Interpretability (IN), Fairness (F), Coherence (C), Naturalness (N), and Interestingness (I) (Descriptions are given in Appendix A). The same three annotators rate every dialogue on a 1-5 scale (low to high).

We retain dialogues that receive scores ≥ 3 on all the metrics. For JobNego, the retained dialogues attain average ratings of 4.62 (EI), 4.44 (SA), 4.27 (IN), 4.33 (F), 4.25 (C), 4.81 (N), and 4.34 (I), with κ scores of 0.85, 0.83, 0.79, 0.82, 0.80, 0.82, and 0.78 for EI, SA, IN, F, C, N, and I, respectively. For ResNego, average ratings are 4.55 (EI), 4.39 (SA), 4.22 (IN), 4.40 (F), 4.20 (C), 4.77 (N), and 4.29 (I), with corresponding κ scores of 0.83, 0.81, 0.80, 0.84, 0.78, 0.80, and 0.77. These results indicate that conversations are emotionally attuned, strategically appropriate, interpretable, fair, coherent, natural, and engaging across both domains.

4 Methodology

The overall architecture of PRISMA, a Preference-Reinforced Self-training approach for an interpretable emotionally intelligent negotiation dialogue system is depicted in Figure 2.

Problem Formulation. We present PRISMA that integrates emotion reasoning with strategy selection and response generation. The system is designed to (i) generate interpretable CoT rationales that explain how emotions influence negotiation strategies, and (ii) generate final responses that are emotionally attuned and strategically appropriate. Formally, given a dialogue context $c = u_1, a_1 \dots, u_{t-1}, a_{t-1}$, which is an alternating sequence of $(t-1)$ turns between the user (u) and

the dialogue agent (a), and the user’s target turn (u_t), the negotiation agent PRISMA first generates an ENS-CoT rationale r . The agent’s response $a_t = (y)$ is then generated conditioned on context c and the selected strategy extracted from r . This ensures that the agent’s response is emotionally and strategically appropriate, while also remaining interpretable, since the rationale explicitly shows how a specific strategy is chosen.

Proposed Approach. To develop PRISMA, we introduce a preference-reinforced self-training paradigm. This is a two-step training procedure that integrates self-training with DPO to guide the model toward accurate, interpretable, and emotionally appropriate responses.

Step 1: Supervised Initialization. We begin by fine-tuning a pre-trained LLM, LLaMA-2-7b (Touvron et al., 2023), on the labeled dataset $\mathcal{D}_L = \{(c_i, r_i, y_i)\}_{i=1}^L$, where L is the number of samples in \mathcal{D}_L . For the i -th input context c_i , let $o_i = \text{concat}(\langle R \rangle, r_i, \langle /R \rangle, \langle A \rangle, y_i, \langle /A \rangle)$ be the target output sequence. The supervised fine-tuned (SFT) model is obtained by π_θ by minimizing the token-level negative log-likelihood $\mathcal{L}_{SI}(\theta)$:

$$\mathcal{L}_{SI}(\theta) = -\mathbb{E}_{(c,r,y) \sim \mathcal{D}_L} \left[\sum_{m=1}^M \log \pi_\theta(o_m | c, o_{1:m-1}) \right] \quad (1)$$

where M denotes number of tokens in rationale+response sequence, and o_m represents the m -th token in o . This step equips the model with a seed ability to generate ENS-CoT rationales and emotion-aware responses during negotiation.

Step 2: Preference-Reinforced Self-Training. After initialization, we iteratively refine the model through two alternating phases:

(a) Preference Learning via DPO. For some context c from the unlabeled dataset $\mathcal{D}_U = \{(c_i, y_i)\}_{i=1}^U$ (U denotes the number of instances in \mathcal{D}_U), we sample multiple rationales \hat{r} from π_θ . We then build a preference dataset $\mathcal{D}_P = \{(c_i, r_i^+, r_i^-)\}_{i=1}^P$ consisting of preferred and rejected completions, where r_i^+ is the preferred completion, r_i^- is the rejected completion, and P is the size of the preference dataset. A rationale is marked *preferred* if its corresponding response has cosine similarity to the ground-truth response y_i exceeding a threshold τ_1 . It is marked *rejected* if the similarity falls below a threshold τ_2 ³. We then update π_θ by minimizing the DPO loss $\mathcal{L}_{\text{PL}}(\theta')$ on \mathcal{D}_P , yielding DPO-refined model $\pi_{\theta'}$:

$$\mathcal{L}_{\text{PL}}(\theta') = -\mathbb{E}_{(c, r^+, r^-) \sim \mathcal{D}_P} \left[\log \sigma \left(\beta \left(\log \frac{\pi_{\theta'}(r^+ | c)}{\pi_\theta(r^+ | c)} - \log \frac{\pi_{\theta'}(r^- | c)}{\pi_\theta(r^- | c)} \right) \right) \right] \quad (2)$$

where, β denotes the coefficient controlling the $\pi_{\theta'}$'s deviation from π_θ .

(b) Supervised Fine-Tuning with Updated Data. Using DPO-refined policy $\pi_{\theta'}$, we generate rationales for context $c \in \mathcal{D}_U$. This forms a pseudo-labeled set $\mathcal{D}_S = \{(c, \hat{r}, y) \mid (c, y) \sim \mathcal{D}_U, \hat{r} \sim \pi_{\theta'}(\cdot | c)\}$ for next round of supervised fine-tuning. We retain rationales whose responses have cosine similarity to the corresponding ground-truth responses exceeding τ_3 , and remove duplicates. This yields a clean dataset $\mathcal{D}'_S \subset \mathcal{D}_S$. The model is then trained in a supervised way on the updated dataset, $\mathcal{D}_N = \mathcal{D}_L \cup \mathcal{D}'_S$. To mitigate the risk of overfitting and enable continual learning, we reinitialize training from the pre-trained LLaMA-2-7b model once the dataset is updated. We do this rather than continuing to fine-tune the SFT model π_θ , following prior works (Zelikman et al., 2022; Singh et al., 2023).

Step 2 repeats until convergence or the iteration limit, resulting in a negotiation agent that effectively identifies and interprets emotions, selects an appropriate emotion-aware negotiation strategy, and grounds its responses in the chosen strategy.

5 Experiments

We compare PRISMA with 10 baselines: Llama-2-7b-SFT, Llama-3.2-3b-SFT (Touvron et al., 2023), Emotion-aware SFT (Llama-2-7b), ICL-AIF (Fu

³All thresholds are chosen empirically based on development-set similarity distributions, fixed globally across experiments, and verified for robustness via sensitivity analysis, as detailed in Appendix B.

et al., 2023), AnE (Zhang et al., 2023), Standard, Proactive, ProCoT (GPT-5) (Deng et al., 2023a; OpenAI, 2025), PPDPP (Deng et al., 2024), RLHF (Ouyang et al., 2022). We evaluate MEDIATOR with both automatic and human judgments. For automatic evaluation, responses' general quality is measured by Perplexity (PPL) (Brown et al., 1992), BLEU-4 (B-4) (Papineni et al., 2002), BERTScore-F1 (BS-F1) (Zhang et al., 2019), Distinct-3 (D-3) (Li et al., 2015), and Response Length (R-LEN); task success is assessed with Emotion Appropriateness (EA): evaluates the appropriateness of users' perceived emotions, and Emotion-aware Negotiation Strategy Consistency (ENSC): assesses the consistency of responses with intended strategies. For human evaluation, to assess the general quality of responses, we use Fluency (F), Coherence (C), and Engagingness (E); and to evaluate for task success, we mirror automatic metrics - EA and ENSC, and also report Bargaining Efficacy (BE) and Outcome Fairness (OF) (Ahmad et al., 2023). Appendix B provides 'Implementation Details', 'Baselines Details' and 'Evaluation Metrics Details'.

6 Results and Analysis

Automatic Evaluation. Table 3 presents the automatic evaluation results for PRISMA and baselines on JobNego and ResNego datasets. Across both datasets, PRISMA consistently achieves superior dialogue generation quality, emotional intelligence, and negotiation performance. In terms of lexical and semantic metrics, it records the highest B-4, D-3, and BS-F1 scores. These results reflect enhanced fluency, diversity, and semantic alignment. Furthermore, the model generates longer responses (R-LEN) while maintaining emotional appropriateness (EA) and emotion-aware negotiation strategy consistency (ENSC). Specifically, on JobNego, PRISMA outperforms the strongest baseline, RLHF. It achieves significant gains of 13.4%, 7.3%, 1.1%, 10.1%, 3.4%, and 3.1% in B-4, D-3, BS-F1, R-LEN, EA, and ENSC, respectively. On ResNego, PRISMA achieves similar performance improvement. Compared to RLHF, it records gains of 14.0%, 5.2%, 1.0%, 10.8%, 2.6%, and 3.0% in B-4, D-3, BS-F1, R-LEN, EA, and ENSC, respectively. These improvements show that PRISMA generates linguistically richer responses. It also better captures emotional nuances and maintains consistent negotiation strategies. The performance gains

Models	PPL ↓	B-4 ↑	D-3 ↑	BS-F1 ↑	R-LEN ↑	EA ↑	ENSC ↑
JobNego							
Llama-2-7B-SFT	6.984	5.806	0.266	0.886	20.917	0.572	0.751
Llama-3.2-3B-SFT	7.318	4.611	0.271	0.888	21.204	0.579	0.756
Emotion-aware SFT (Llama-2-7b)	6.813	6.312	0.288	0.896	23.467	0.712	0.781
ICL-AIF (GPT-3.5)	9.612	5.321	0.198	0.852	12.904	0.462	0.641
AnE	7.623	5.679	0.253	0.879	19.102	0.611	0.737
Standard Prompting (GPT-5)	8.937	5.447	0.209	0.858	14.906	0.486	0.667
Proactive Prompting (GPT-5)	7.411	5.732	0.259	0.883	19.634	0.566	0.744
ProCoT (GPT-5)	7.842	5.618	0.247	0.875	18.439	0.642	0.728
PPDPP	7.566	5.637	0.251	0.878	19.221	0.744	0.733
RLHF	6.489	6.277	0.300	0.902	24.612	0.814	0.798
PRISMA	5.196	7.119	0.322	0.912	27.108	0.842	0.823
ResNego							
Llama-2-7b-SFT	6.203	6.194	0.292	0.899	24.512	0.607	0.789
Llama-3.2-3b-SFT	6.589	5.190	0.297	0.901	24.933	0.612	0.794
Emotion-aware SFT (Llama-2-7b)	5.982	7.208	0.313	0.909	27.521	0.755	0.819
ICL-AIF (GPT-5)	8.537	6.143	0.221	0.869	15.318	0.501	0.682
AnE	6.842	6.183	0.283	0.894	23.102	0.644	0.775
Standard Prompting (GPT-5)	8.011	6.154	0.236	0.874	17.842	0.523	0.706
Proactive Prompting (GPT-5)	6.702	6.188	0.289	0.897	23.987	0.601	0.783
ProCoT (GPT-5)	6.993	6.179	0.277	0.891	22.311	0.683	0.767
PPDPP	6.912	6.181	0.279	0.892	22.987	0.787	0.771
RLHF	5.741	7.218	0.324	0.913	29.004	0.852	0.836
PRISMA	4.462	8.231	0.341	0.922	32.137	0.874	0.861

Table 3: Automatic evaluation results. Results are statistically significant at 5% significance level based on t-test (Welch, 1947).

Models	F	C	E	EA	ENSC	BE	OF
JobNego							
Llama-2-7B-SFT	3.65	3.50	3.40	3.45	3.42	3.30	3.28
Emotion-aware SFT	4.10	3.98	4.05	4.12	4.00	3.95	3.90
ProCoT	3.88	3.75	3.70	3.82	3.68	3.65	3.60
PPDPP	4.00	3.90	3.85	3.98	3.92	3.80	3.78
RLHF	4.35	4.22	4.30	4.38	4.20	4.18	4.12
PRISMA	4.52	4.40	4.48	4.55	4.48	4.36	4.30
ResNego							
Llama-2-7B-SFT	3.78	3.65	3.55	3.62	3.60	3.48	3.45
Emotion-aware SFT	4.24	4.12	4.18	4.26	4.18	4.05	4.00
ProCoT	4.02	3.90	3.86	3.98	3.92	3.85	3.80
PPDPP	4.15	4.05	4.00	4.12	4.08	3.95	3.90
RLHF (PPO)	4.47	4.35	4.40	4.51	4.42	4.31	4.24
PRISMA	4.68	4.56	4.62	4.70	4.64	4.50	4.45

Table 4: Human evaluation results. Results are statistically significant at 5% significance level based on t-test (Welch, 1947). All metrics are rated on a scale of 1 to 5.

could be attributed to the interpretability enforced through ENS-CoT rationales during preference-reinforced self-training, where modeling emotion-strategy reasoning sequences enables transparent, emotionally intelligent, and coherent negotiation behavior.

The baseline model, RLHF, tends to generate less diverse and less coherent emotion-aware responses during negotiation. Its inability to follow the complete reasoning chain might result in limited awareness and inconsistent understanding of the user’s emotional states. Emotion-aware SFT

performs well in emotional understanding. However, it tends to generate shorter, less strategic responses, thereby limiting negotiation effectiveness. ProCoT and PPDPP improve contextual control, but often underperform in emotional alignment and response diversity. In contrast, PRISMA balances emotional intelligence and negotiation efficiency, enhancing both the linguistic and affective aspects of dialogues. The consistent gains across datasets highlight its robustness and the effectiveness of the interpretable ENS-CoT reasoning mechanism.

Human Evaluation. Table 4 presents human evaluation results for PRISMA and baselines. Given the high cost of manual evaluation, we compare PRISMA against Llama-2-7B-SFT, Emotion-aware SFT (Llama-2-7b), ProCoT (GPT-5-mini), PPDPP, and RLHF. On JobNego, PRISMA achieves average scores of 4.52, 4.40, 4.48, 4.55, 4.48, 4.36, and 4.30 for F, C, E, EA, ENSC, BE, and OF, respectively, showing consistent improvements of +0.17, +0.18, +0.18, +0.17, +0.28, +0.18, and +0.18 points over the RLHF. A similar trend is observed on ResNego. PRISMA outperforms RLHF by +0.21, +0.21, +0.22, +0.19, +0.22, +0.19, and +0.21 points on these metrics. These improvements highlight PRISMA’s ability to generate emotionally intelligent and strategically coherent negotiation dialogues. The higher F, C, and E scores demonstrate

ID	Setting								Results						
	EM	ET	IA	PS	MT	SS	SR	RG	PPL ↓	B-4 ↑	D-3 ↑	BS-F1 ↑	R-LEN ↑	EA ↑	ENSC ↑
JobNego															
0	✓	✓	✓	✓	✓	✓	✓	✓	5.196	7.119	0.322	0.912	27.108	0.842	0.823
1	✓	✗	✗	✓	✓	✓	✓	✓	6.024	6.486	0.309	0.904	24.816	0.816	0.788
2	✓	✓	✓	✗	✗	✓	✓	✓	6.052	6.453	0.312	0.904	25.093	0.828	0.792
3	✓	✓	✓	✓	✓	✗	✗	✓	6.081	5.352	0.310	0.898	25.612	0.834	0.741
4	✗	✗	✗	✗	✗	✓	✓	✓	5.951	6.903	0.298	0.896	24.039	0.782	0.702
5	✗	✗	✗	✗	✗	✓	✓	✓	5.879	5.204	0.292	0.890	23.521	0.731	0.648
ResNego															
0	✓	✓	✓	✓	✓	✓	✓	✓	4.462	8.231	0.341	0.922	32.137	0.874	0.861
1	✓	✗	✗	✓	✓	✓	✓	✓	4.339	7.617	0.326	0.912	29.842	0.853	0.829
2	✓	✓	✓	✗	✗	✓	✓	✓	5.392	6.742	0.332	0.914	30.108	0.862	0.834
3	✓	✓	✓	✓	✓	✗	✗	✓	5.409	6.664	0.333	0.909	30.573	0.868	0.812
4	✗	✗	✗	✗	✗	✓	✓	✓	5.271	7.118	0.318	0.905	28.721	0.821	0.768
5	✗	✗	✗	✗	✗	✓	✗	✓	5.205	7.431	0.311	0.898	28.106	0.783	0.722

Table 5: Ablation study to analyze the impact of different components of ENS-CoT rationales in PRISMA.

PRISMA’s ability to produce fluent, coherent, and engaging negotiation dialogues that sustain natural conversational flow. Besides, the superior EA, ENSC, BE, and OF scores highlight its effectiveness in maintaining emotional awareness, strategic consistency, and balanced outcome-focused negotiation behavior. These results indicate that interpretability strengthens emotional alignment and supports mutually beneficial negotiation outcomes.

Ablation Study. To understand how each component of the ENS-CoT rationales affects model performance, we conduct an ablation study to investigate the impact of different components of the ENS-CoT rationales. Table 5 reports results across different configurations in which one or more components are removed. In the first ablated model, we remove the ‘understanding emotion’ components, specifically ET and IA. The second ablated model omits the ‘using emotion to facilitate thought’ components, namely PS and MT. In the third ablated model, we eliminate the ‘managing emotion’ components, which include SS and SR. Finally, the fourth ablated model excludes the components EM, ET, IA, PS, and MT, and the fifth ablated model removes the components EM, ET, IA, PS, MT, and SR. The removal of any rationale component adversely affects various critical aspects of dialogue quality, such as the appropriateness of emotions, the effectiveness of negotiations, and the language quality of responses. This is evidenced by reduced scores on metrics such as EA, ENSC, B-4, D-3, BS-F1, and R-LEN in both datasets.

In particular, when we remove (ET, IA) or (PS, MT), B-4 increases slightly, but D-3 and BS-F1 drop, and responses become shorter. This shows that abilities like ‘understanding emotion’ and ‘using emotion to facilitate thought’ contribute sub-

stantive and varied content rather than n-gram overlap. Further, ablating (SS, SR) causes the largest decline in ENSC, indicating that explicit strategy choice and its justification are central to keeping the response aligned with the users’ intended emotions. Also, excluding the ‘perceiving emotion’ (EM) component along with the other components reduces EA, ENSC, D-3, BS-F1, and R-LEN. This pattern is consistent with simpler, more templated outputs that match surface forms but lose strategic consistency and emotional awareness. In general, these findings underscore the critical importance of each component in improving interpretability and encouraging the generation of emotion-aware responses within negotiation dialogues.

Additional Analysis. We present more analyses: (1) Impact of Supervised Initialization, Self-Training, and DPO, (2) Impact of iterative training in PRISMA, and (3) Case Study in Appendix C.

7 Conclusion

We presented PRISMA, an interpretable, emotionally intelligent negotiation dialogue agent designed to enhance negotiation outcomes in job interview and resource allocation domains. In this regard, we proposed an Emotion-aware Negotiation Strategy-informed CoT (ENS-CoT) reasoning mechanism that links perceived emotions to strategy selection and response generation. Using ENS-CoT, we curated JobNego and ResNego datasets. PRISMA integrates self-training with DPO through a preference-reinforced self-training framework, optimizing step-by-step CoT reasoning for emotion-aware negotiation responses. Extensive evaluation show that PRISMA achieves superior emotion appropriateness, strategy consistency, and negotiation effectiveness with improved interpretability.

Limitations

The prompting-based data creation requires substantial compute and access to LLMs, and inherits both model biases and the variability of prompt-based generation. Although the resulting dialogues are generally fair and natural, quality remains bounded by the base model. The stronger controls beyond scenario prompts (e.g., constraint-based decoding, verifier models, or weak supervision) could further improve fidelity. Ensuring interpretability with ENS-CoT rationales also necessitates human verification, which constrains scale; realism is limited by assumptions of fluent interlocutors, coherent turn flow, and the absence of exogenous disruptions. Future work should automate correction/verification, broaden scenario/diversity coverage (culture, persona, etc.), and explore multi-emotion and multi-strategy annotations to better reflect real negotiation practice.

We also observe that PRISMA can occasionally produce prescriptive responses rather than fully collaborative ones, typically due to insufficient reasoning about user-specific constraints or implicit social norms (e.g., workplace expectations). Although human evaluations show that PRISMA generally behaves collaboratively, this limitation highlights an important direction for future work, which is, guiding ENS-CoT generation and preference learning toward explicitly collaborative problem-solving, emphasizing alternatives, phased plans, trade-offs, and shared decision-making rather than fixed prescriptions.

Further, the proposed negotiation dialogue agent, PRISMA, builds on self-training, but we do not systematically address unlabeled data usage. Preference construction for DPO depends on rubrics and thresholds that may encode hidden biases, and pseudo-labels can amplify early errors in PRISMA. Moreover, the evaluation is restricted to two bilateral, text-only domains (job interviews and camping resource allocation) primarily due to a lack of rationale-annotated negotiation dialogue dataset. Finally, several automatic proxies (e.g., similarity) do not substitute for user-centric outcomes; controlled user studies and online evaluations are an important next step to quantify trade-offs among interpretability, task success, and user experience.

Ethics Statement

Interpretability is central for negotiation dialogue systems, especially when modeling emotional in-

telligence. Therefore, this work has been reviewed and approved by our Institutional Review Board (IRB). First, we treat dataset curation and distribution with care. All data sources are used in accordance with their licenses. We screen and filter generations with the help of linguistics and subject matter experts to reduce biased, offensive, insensitive, or inappropriate content to avoid potential unpredictable ethical harm. All human annotators and evaluators are paid according to the institute's norms. To ensure responsible data use, access will be granted only after the completion and signing of an agreement stipulating that the data will be used solely for research purposes.

Second, negotiation involves strategic influence, so the design and deployment of PRISMA must also adhere to strict ethical standards. PRISMA is explicitly optimized for collaborative, non-zero-sum outcomes: its emotion-aware strategies seek mutually beneficial agreements rather than one-sided wins. Crucially, user autonomy is preserved at all times, that is, users can decline proposals that do not meet their needs, and the system must not pressure, coerce, or manipulate. PRISMA's ENS-CoT rationales are exposed to support transparency and human oversight, enabling users to understand how the dialogue agent decides a particular strategy.

References

- Sahar Abdelnabi, Amr Gomaa, Sarath Sivaprasad, Lea Schönherr, and Mario Fritz. 2023. Llm-deliberation: Evaluating llms with interactive multi-agent negotiation games. *arXiv preprint arXiv:2309.17234*.
- Zishan Ahmad, Suman Saurabh, Vaishakh Menon, Asif Ekbal, Roshni Ramnani, and Anutosh Maitra. 2023. Ina: An integrative approach for enhancing negotiation strategies with reward-based dialogue agent. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 2536–2549.
- Massih-Reza Amini, Vasilii Feofanov, Loic Pauletto, Lies Hadjadj, Emilie Devijver, and Yury Maximov. 2025. Self-training: A survey. *Neurocomputing*, 616:128904.
- Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, et al. 2022. Training a helpful and harmless assistant with reinforcement learning from human feedback. *arXiv preprint arXiv:2204.05862*.
- Bruce Barry. 1999. The tactical use of emotion in negotiation. *Research on negotiation in organizations*, 7:93–124.

- Bruce Barry and Richard L Oliver. 1996. Affect in dyadic negotiation: A model and propositions. *Organizational Behavior and Human Decision Processes*, 67(2):127–143.
- Max H Bazerman, Jared R Curhan, Don A Moore, and Kathleen L Valley. 2000. Negotiation. *Annual review of psychology*, 51(1):279–314.
- Erica J Boothby, Gus Cooney, and Maurice E Schweitzer. 2023. Embracing complexity: A review of negotiation research. *Annual Review of Psychology*, 74(2023):299–332.
- Tibor Bosse, Charlotte Gerritsen, Jeroen de Man, and Jan Treur. 2013. Learning emotion regulation strategies: A cognitive agent model. In *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, volume 2, pages 245–252. IEEE.
- Peter F Brown, Stephen A Della Pietra, Vincent J Della Pietra, Jennifer C Lai, and Robert L Mercer. 1992. An estimate of an upper bound for the entropy of english. *Computational Linguistics*, 18(1):31–40.
- Peter J Carnevale, Dean G Pruitt, et al. 1992. Negotiation and mediation. *Annual review of psychology*, 43(1):531–582.
- Kushal Chawla, Rene Clever, Jaysa Ramirez, Gale Lucas, and Jonathan Gratch. 2021a. Towards emotion-aware agents for negotiation dialogues. In *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 1–8. IEEE.
- Kushal Chawla, Rene Clever, Jaysa Ramirez, Gale M Lucas, and Jonathan Gratch. 2023. Towards emotion-aware agents for improved user satisfaction and partner perception in negotiation dialogues. *IEEE Transactions on Affective Computing*, 15(2):433–444.
- Kushal Chawla, Gale Lucas, Jonathan May, and Jonathan Gratch. 2022. Opponent modeling in negotiation dialogues by related data adaptation. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 661–674.
- Kushal Chawla, Jaysa Ramirez, Rene Clever, Gale Lucas, Jonathan May, and Jonathan Gratch. 2021b. Casino: A corpus of campsite negotiation dialogues for automatic negotiation systems. *arXiv preprint arXiv:2103.15721*.
- Yunkai Dang, Kaichen Huang, Jiahao Huo, Yibo Yan, Sirui Huang, Dongrui Liu, Mengxi Gao, Jie Zhang, Chen Qian, Kun Wang, et al. 2024. Explainable and interpretable multimodal large language models: A comprehensive survey. *arXiv preprint arXiv:2412.02104*.
- Yang Deng, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, and Tat-Seng Chua. 2023a. Prompting and evaluating large language models for proactive dialogues: Clarification, target-guided, and non-collaboration. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 10602–10621.
- Yang Deng, Wenxuan Zhang, Wai Lam, See-Kiong Ng, and Tat-Seng Chua. 2023b. Plug-and-play policy planner for large language model powered dialogue agents. *arXiv preprint arXiv:2311.00262*.
- Yang Deng, Wenxuan Zhang, Wai Lam, See-Kiong Ng, and Tat-Seng Chua. 2024. Plug-and-play policy planner for large language model powered dialogue agents. In *ICLR*.
- Yihe Deng and Paul Mineiro. 2024. Flow-dpo: Improving llm mathematical reasoning through online multi-agent learning. *arXiv preprint arXiv:2410.22304*.
- Kathy Douglas and Clare Coburn. 2014. Attitude and response to emotion in dispute resolution: The experience of mediators. *Flinders LJ*, 16:111.
- Shutong Feng, Hsien-chin Lin, Nurul Lubis, Carel van Niekerk, Michael Heck, Benjamin Ruppik, Renato Vukovic, and Milica Gašić. 2025. Emotionally intelligent task-oriented dialogue systems: Architecture, representation, and optimisation. *arXiv preprint arXiv:2507.01594*.
- Joseph P Forgas. 1998. On feeling good and getting your way: Mood effects on negotiator cognition and bargaining strategies. *Journal of personality and social psychology*, 74(3):565.
- Yao Fu, Hao Peng, Tushar Khot, and Mirella Lapata. 2023. Improving language model negotiation with self-play and in-context learning from ai feedback. *arXiv preprint arXiv:2305.10142*.
- Jordan D Fuhrman, Naveena Gorre, Qiyuan Hu, Hui Li, Issam El Naqa, and Maryellen L Giger. 2022. A review of explainable and interpretable ai with applications in covid-19 imaging. *Medical Physics*, 49(1):1–14.
- Kanishk Gandhi, Dorsa Sadigh, and Noah D Goodman. 2023. Strategic reasoning with language models. *arXiv preprint arXiv:2305.19165*.
- Daniel Goleman. 2005. *Emotional intelligence: Why it can matter more than IQ*. Bantam.
- Caglar Gulcehre, Tom Le Paine, Srivatsan Srinivasan, Ksenia Konyushkova, Lotte Weerts, Abhishek Sharma, Aditya Siddhant, Alex Ahern, Miaosen Wang, Chenjie Gu, et al. 2023. Reinforced self-training (rest) for language modeling. *arXiv preprint arXiv:2308.08998*.
- He He, Derek Chen, Anusha Balakrishnan, and Percy Liang. 2018. Decoupling strategy and generation in negotiation dialogues. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2333–2343.
- Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. 2019. The curious case of neural text degeneration. *arXiv preprint arXiv:1904.09751*.

- Yin Jou Huang and Rafik Hadfi. 2024. How personality traits influence negotiation outcomes? a simulation based on large language models. *arXiv preprint arXiv:2407.11549*.
- Mohammed Kasbouya and Nawal Sael. 2025. Emotional intelligence in large language models: Fine-tuning methods, challenges, and applications. In *International Conference on intelligent systems and digital applications*, pages 348–359. Springer.
- John F Kelley. 1984. An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Information Systems (TOIS)*, 2(1):26–41.
- Mehmet Onur Keskin, Berk Buzcu, Berkecan Koçyigit, Umut Çakan, Anıl Dođru, and Reyhan Aydođan. 2024. Negotiator: a comprehensive framework for human-agent negotiation integrating preferences, interaction, and emotion. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)*, pages 8700–8703.
- Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Vasily Konovalov, Ron Artstein, Oren Melamud, and Ido Dagan. 2016. The negochat corpus of human-agent negotiation dialogues. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 3141–3145.
- Leon OH Kroczeck, Alexander May, Selina Hettenkofer, Andreas Ruider, Bernd Ludwig, and Andreas Mühlberger. 2025. The influence of persona and conversational task on social interactions with a llm-controlled embodied conversational agent. *Computers in Human Behavior*, page 108759.
- Xin Lai, Zhuotao Tian, Yukang Chen, Senqiao Yang, Xiangru Peng, and Jiaya Jia. 2024. Step-dpo: Step-wise preference optimization for long-chain reasoning of llms. *arXiv preprint arXiv:2406.18629*.
- Edward J Lawler and Shane R Thye. 1999. Bringing emotions into social exchange theory. *Annual review of sociology*, 25(1):217–244.
- Edward J Lawler and Jeongkoo Yoon. 1995. Structural power and emotional processes in negotiation: A social exchange approach.
- Jaehyeok Lee, Keisuke Sakaguchi, and JinYeong Bak. 2024. Self-training meets consistency: Improving llms' reasoning with consistency-driven rationale evaluation. *arXiv preprint arXiv:2411.06387*.
- Mike Lewis, Denis Yarats, Yann N Dauphin, Devi Parikh, and Dhruv Batra. 2017. Deal or no deal? end-to-end learning for negotiation dialogues. *arXiv preprint arXiv:1706.05125*.
- Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2015. A diversity-promoting objective function for neural conversation models. *arXiv preprint arXiv:1510.03055*.
- Shaobo Li, Chengjie Sun, Zhen Xu, Prayag Tiwari, Bingquan Liu, Deepak Gupta, K Shankar, Zhenzhou Ji, and Mingjiang Wang. 2023. Toward explainable dialogue system using two-stage response generation. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22(3):1–18.
- Zaijing Li, Gongwei Chen, Rui Shao, Yuquan Xie, Dongmei Jiang, and Liqiang Nie. 2024. Enhancing emotional generation capability of large language models via emotional chain-of-thought. *arXiv preprint arXiv:2401.06836*.
- JD Maier, P Salovey, and DR Caruso. 2004. Emotional intelligence: theory findings, and implications. *psychological inquiry*, 15, 197-215. *Psychological Science*, 9(5):331–339.
- John D Mayer, Peter Salovey, David R Caruso, and Gill Sitarenios. 2003. Measuring emotional intelligence with the msceit v2. 0. *Emotion*, 3(1):97.
- Mary L McHugh. 2012. Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282.
- Shervin Minaee, Tomas Mikolov, Narjes Nikzad, Meysam Chenaghlu, Richard Socher, Xavier Amatriain, and Jianfeng Gao. 2024. Large language models: A survey. *arXiv preprint arXiv:2402.06196*.
- Michael W Morris and Dacher Keltner. 2000. How emotions work: The social functions of emotional expression in negotiations. *Research in organizational behavior*, 22:1–50.
- OpenAI. 2024. ChatGPT. <https://chatgpt.com/>.
- OpenAI. 2025. Gpt-5 system card. <https://cdn.openai.com/gpt-5-system-card.pdf>. Accessed: October 2025.
- Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, 35:27730–27744.
- Jennifer R Overbeck, Margaret A Neale, and Cassandra L Govan. 2010. I feel, therefore you act: Intrapersonal and interpersonal effects of emotion on negotiation as a function of social power. *Organizational Behavior and Human Decision Processes*, 112(2):126–139.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318.

- Priyanshu Priya, Rishikant Chigrupaatii, Mauajama Firdaus, and Asif Ekbal. 2025a. Genteel-negotiator: Llm-enhanced mixture-of-expert-based reinforcement learning approach for polite negotiation dialogue. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 25010–25018.
- Priyanshu Priya, Saurav Dudhate, Desai Vishesh Yasheshbhai, and Asif Ekbal. 2025b. We argue to agree: Towards personality-driven argumentation-based negotiation dialogue systems for tourism. *arXiv preprint arXiv:2509.11118*.
- Priyanshu Priya, Desai Vishesh Yasheshbhai, Ratnesh Kumar Joshi, Roshni Ramnani, Anutosh Maitra, Shubhashis Sengupta, and Asif Ekbal. 2024. Trip negotiator: A travel persona-aware reinforced dialogue generation model for personalized integrative negotiation in tourism. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 16566–16595.
- Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. 2023. Direct preference optimization: Your language model is secretly a reward model. *Advances in neural information processing systems*, 36:53728–53741.
- Jeffrey Z Rubin and Bert R Brown. 2013. *The social psychology of bargaining and negotiation*. Elsevier.
- John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. [Proximal policy optimization algorithms](#). *Preprint*, arXiv:1707.06347.
- Henry Scudder. 1965. Probability of error of some adaptive pattern-recognition machines. *IEEE Transactions on Information Theory*, 11(3):363–371.
- James K Sebenius. 1992. Negotiation analysis: A characterization and review. *Management science*, 38(1):18–38.
- Avi Singh, John D Co-Reyes, Rishabh Agarwal, Ankesh Anand, Piyush Patil, Xavier Garcia, Peter J Liu, James Harrison, Jaehoon Lee, Kelvin Xu, et al. 2023. Beyond human data: Scaling self-training for problem-solving with language models. *arXiv preprint arXiv:2312.06585*.
- Bertram I Spector. 1977. Negotiation as a psychological process. *Journal of Conflict Resolution*, 21(4):607–618.
- Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.
- Luong Trung, Xinbo Zhang, Zhanming Jie, Peng Sun, Xiaoran Jin, and Hang Li. 2024. Reft: Reasoning with reinforced fine-tuning. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7601–7614.
- Gerben A Van Kleef and Gert-Jan Lelieveld. 2022. Moving the self and others to do good: The emotional underpinnings of prosocial behavior. *Current opinion in psychology*, 44:80–88.
- Gerben A Van Kleef et al. 2008. Emotion in conflict and negotiation: Introducing the emotions as social information (easi) model. *Research companion to emotion in organizations*, pages 392–404.
- Yiping Wang, Qing Yang, Zhiyuan Zeng, Liliang Ren, Liyuan Liu, Baolin Peng, Hao Cheng, Xuehai He, Kuan Wang, Jianfeng Gao, et al. 2025. Reinforcement learning for reasoning in large language models with one training example. *arXiv preprint arXiv:2504.20571*.
- Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A Smith, Daniel Khoshabi, and Hannaneh Hajishirzi. 2023. Self-instruct: Aligning language models with self-generated instructions. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13484–13508.
- Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837.
- Bernard L Welch. 1947. The generalization of ‘student’s’ problem when several different population variances are involved. *Biometrika*, 34(1-2):28–35.
- Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2019. Huggingface’s transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*.
- Zongqian Wu, Baoduo Xu, Ruochen Cui, Mengmeng Zhan, Xiaofeng Zhu, and Lei Feng. 2024. Rethinking chain-of-thought from the perspective of self-training. *arXiv preprint arXiv:2412.10827*.
- Atsuki Yamaguchi, Kosui Iwasa, and Katsuhide Fujita. 2021. Dialogue act-based breakdown detection in negotiation dialogues. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 745–757.
- Hailong Yang, Mingxian Gu, Renhuo Zhao, Fuping Hu, Zhaohong Deng, and Yitang Chen. 2024. Xagents: A framework for interpretable rule-based multi-agents cooperation. *arXiv preprint arXiv:2411.13932*.
- Runzhe Yang, Jingxiao Chen, and Karthik Narasimhan. 2020. Improving dialog systems for negotiation with personality modeling. *arXiv preprint arXiv:2010.09954*.

- Tao Yang, Tianyuan Shi, Fanqi Wan, Xiaojun Quan, Qifan Wang, Bingzhe Wu, and Jiaxiang Wu. 2023. Psycot: Psychological questionnaire as powerful chain-of-thought for personality detection. *arXiv preprint arXiv:2310.20256*.
- Gülçimen Yurtsever. 2004. Emotional regulation strategies and negotiation. *Psychological Reports*, 95(3):780–786.
- Eric Zelikman, Yuhuai Wu, Jesse Mu, and Noah Goodman. 2022. Star: Bootstrapping reasoning with reasoning. *Advances in Neural Information Processing Systems*, 35:15476–15488.
- Qiang Zhang, Jason Naradowsky, and Yusuke Miyao. 2023. Ask an expert: Leveraging language models to improve strategic reasoning in goal-oriented dialogue models. *arXiv preprint arXiv:2305.17878*.
- Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q Weinberger, and Yoav Artzi. 2019. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*.
- Zheng Zhang, Lizi Liao, Xiaoyan Zhu, Tat-Seng Chua, Zitao Liu, Yan Huang, and Minlie Huang. 2020. Learning goal-oriented dialogue policy with opposite agent awareness. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, pages 122–132.
- Zhuosheng Zhang, Aston Zhang, Mu Li, and Alex Smola. 2022. Automatic chain of thought prompting in large language models. *arXiv preprint arXiv:2210.03493*.
- Ran Zhao, Oscar J Romero, and Alex Rudnicky. 2018. Sogo: a social intelligent negotiation dialogue system. In *Proceedings of the 18th International Conference on intelligent virtual agents*, pages 239–246.
- Tiancheng Zhao, Kaige Xie, and Maxine Eskenazi. 2019. Rethinking action spaces for reinforcement learning in end-to-end dialog agents with latent variable models. *arXiv preprint arXiv:1902.08858*.
- Yiheng Zhou, He He, Alan W Black, and Yulia Tsvetkov. 2019. A dynamic strategy coach for effective negotiation. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*, pages 367–378.

Appendix

A Dataset Details

A.1 Prompt Templates

The prompt template used to generate the negotiation dialogue with ENS-CoT rationales is given below:

Prompt for Negotiation Dialogue Generation with ENS-CoT Rationales. Variables enclosed in curly brackets serve as placeholders that are populated when the template is instantiated.

Your task is to generate a negotiation dialogue between the user and the agent for the {Negotiation Scenario}. You need to generate the user’s utterance and mark the expressed emotion (EM), emotion trigger (ET), individual assessment (IA), perspective shift (PS), mindset transformation (MT), strategy selection (SS), strategy reason (SR), and agent’s response based on the selected strategy (RG).

User’s Utterance: ⟨Turn contributed by the user⟩
 User’s utterance should begin with ‘{User: ... }’
 Emotion: ⟨EM definition⟩ Emotion should begin with ‘{The user feels ... }’
 Emotional Trigger: ⟨ET definition⟩ Emotional Trigger should begin with ‘{User’s ‘Emotion’ is triggered by ... }’
 Individual Assessment: ⟨IA definition⟩ Individual Assessment should begin with ‘{The user thinks ... }’
 Perspective Shift: ⟨PS definition⟩ Perspective Shift should begin with ‘{Enable the user to consider the situation from a different angle: ... }’
 Mindset Transformation: ⟨MT definition⟩ Mindset Transformation should begin with ‘{Enable the user to think about reframing the belief: ... }’
 Strategy Selection: ⟨SS definition⟩ Strategy Selection should begin with ‘{The agent chooses ... }’
 Strategy Reason: ⟨SR definition⟩ Strategy Reason should begin with ‘{To ..., the agent uses ... }’
 Response Generation: ⟨RG definition⟩ Agent’s response should begin with {Agent: ... }’

The user should use the following emotions: ⟨list of emotions⟩

The user should use the following emotion-aware negotiation strategies: ⟨list of strategies⟩

Examples:

⟨Few-shot exemplars⟩

Please adhere precisely to the format provided above, ensuring that no components are omitted.

— To be generated —

{Generated Dialogue}

The prompt template employed to generate negotiation scenarios is as follows:

Prompt for Negotiation Scenario Generation. Variables enclosed in curly brackets serve as placeholders that are populated when the template is instantiated.

Your task is to generate the scenario, summarizing the negotiation setup, including the task, roles, issues, constraints, and goals underlying the provided dialogue.

Examples: (Optional)

<Few-shot exemplars>

The dialogue is as follows:

<Seed Negotiation Dialogue>

— To be generated —

{Generated Negotiation Scenario}

A.2 Examples for Emotion-aware Negotiation Strategies

Table 6 provides the example utterances for different emotion-aware negotiation strategies from the JobNego dataset.

A.3 Guidelines for Designing Sample ENS-CoT-enhanced Dialogues

The annotators are instructed to adhere to the following guidelines during dialogue creation:

- To enhance understanding of negotiation and ensure the application of best negotiation practices, we follow (Chawla et al., 2021b) and require the annotators to watch a tutorial on principled negotiation strategies before beginning their tasks;
- Begin negotiations with offers that leave room for concessions and adjustments, allowing both sides to explore trade-offs;
- Actively elicit and discuss the other party’s preferences, priorities, and constraints to identify opportunities for integrative outcomes;
- Clearly communicate requirements and rationales, grounding arguments in realistic job interview (e.g., salary, promotion, benefits) or resource allocation (e.g., distribution of supplies, task responsibilities) scenarios;
- Recognize and respond to the emotional states expressed by the user, ensuring that the agent applies ENS-CoT reasoning to *perceive*, *understand*, *use*, and *manage* emotions in a transparent manner;

- Ensure that each agent response is paired with a concise ENS-CoT annotation, explicitly justifying how a particular emotion-aware negotiation strategy is chosen;
- Aim for outcomes that balance self-interest with fairness, recognizing that perceptions of equity and rationality may differ across negotiating parties;
- Use professional, and empathetic language consistently, especially in situations of disagreement or emotional tension, to maintain rapport and trust.

A.4 Details of Dataset Quality Evaluation Criteria

1. **Emotional Intelligence (EI)**: Is the response appropriate to the user’s expressed emotion, reducing misattunement or escalation?
2. **Strategy Appropriateness (SA)**: Does the agent’s utterance operationalize the emotion-aware negotiation strategy?
3. **Interpretability (IN)**: Is the ENS-CoT rationale well-formed, complete, and useful?
4. **Fairness (F)**: Does the final outcome reflect a balanced, and win-win situation?
5. **Coherence (C)**: Is the overall dialogue well-structured and logical?
6. **Naturalness (N)**: Does the dialogue resemble the human negotiation?
7. **Interestingness (I)**: Is the dialogue engaging and rich in content to retain the user’s interest throughout the negotiation?

A.5 Dataset Distribution and Topic Analysis

To assess the balance, diversity, and thematic coverage of the proposed datasets, we conduct a comprehensive analysis of (i) the joint distribution of emotions and negotiation strategies, and (ii) scenario-level topical diversity for both JobNego and ResNego.

A.5.1 Emotion-Strategy Distribution

Our datasets are constructed using 12 emotion categories and 12 emotion-aware negotiation strategies, yielding 144 possible emotion–strategy combinations. To examine distributional balance, we analyze the frequency of all combinations across both datasets. Figure 3 presents a heatmap visualization of the joint emotion–strategy distribution for JobNego and ResNego.

All emotion–strategy pairs appear in both datasets, and their frequencies are broadly dis-

Emotion-aware Negotiation Strategy	Examples
Savoring	<i>It's great to see your excitement, and we appreciate your clear expectations. Let's discuss the salary range for the project manager role.</i>
Positive Reinforcement	<i>Great, thank you for being open to compromise. It's important for us to work together to ensure a productive and motivating work environment. Let's continue discussing your needs and expectations for the company car and workday</i>
Expressing Optimism	<i>We appreciate your confidence in your abilities and your willingness to work with us. Let's discuss the promotion track and how we can align our expectations for mutual success.</i>
Cognitive Reappraisal	<i>Thank you for sharing your concerns. Let's explore other opportunities that match your skills and experience.</i>
Positive Framing	<i>I understand your concerns about the workday length. Let's explore how a shorter workday could lead to improved productivity and work-life balance while ensuring your career growth expectations are met.</i>
Emotion Diffusion	<i>I am glad we could reach an agreement. I look forward to working with you and discussing opportunities for professional growth.</i>
Expressive Suppression	<i>I appreciate your honesty. Let's discuss the employer's offer and how it can align with your career goals. How do you see yourself growing within the company?</i>
Active Listening	<i>Thank you for sharing your expectations. I understand your desire for a company car. Let's discuss what we can do to ensure a mutually beneficial agreement.</i>
Perspective-Taking	<i>Thank you for sharing your thoughts on the position. Let's explore the terms of the offer further to find a mutually beneficial solution. Can we discuss the salary range for the project manager role, and how it aligns with your expectations?</i>
Problem Solving	<i>I understand your concerns about work hours. How about we discuss possible compromises that can work for both of us?</i>
Escalate Assurance	<i>I understand your concerns, but we are willing to discuss a salary range of 80-100,000. This is an opportunity for us to align our expectations and find a mutually beneficial agreement.</i>
No Strategy	<i>Thank you for sharing your expectations. However, we cannot accommodate your requested salary at this time. We appreciate your understanding and would be willing to discuss alternative options.</i>

Table 6: The example utterances of different strategies in the JobNego dataset.

tributed, with no single combination disproportionately dominating the corpus. This balanced coverage arises from ENS-CoT’s context-dependent reasoning, where emotion perception and strategy selection are dynamically conditioned on the negotiation scenario rather than sampled uniformly. The observed distribution confirms that the datasets are not biased toward particular emotional states or negotiation strategies and instead capture a wide range of emotionally grounded negotiation behaviors.

A.5.2 Topic Diversity and Thematic Coverage

To evaluate thematic comprehensiveness, we perform a scenario-level topic analysis based on negotiation objectives, discussed resources, and contractual elements. This analysis confirms that both datasets span a diverse range of realistic negotiation contexts within their respective domains. The topic-wise distribution of dialogues for both datasets is summarized in Table 7.

ResNego. ResNego captures multifaceted resource-allocation negotiations characteristic of survival, camping, and shared-resource environments. As shown in Table 7, the dataset covers:

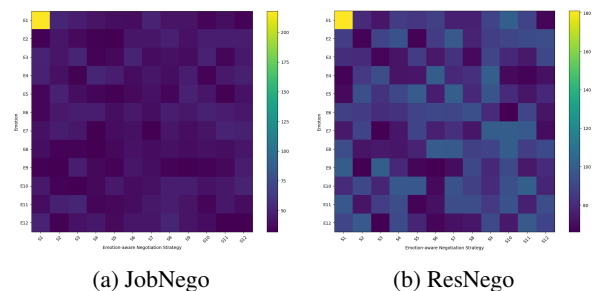


Figure 3: Distribution of emotion-strategy pairs in the JobNego and ResNego datasets. Emotions E1–E12 correspond to *joy, confidence, positivity, gratitude, trust, surprise, anger, disappointment, frustration, fear, anxiety, and neutral*, while emotion-aware negotiation strategies S1–S12 correspond to *savoring, positive reinforcement, expressing optimism, cognitive reappraisal, positive framing, emotion diffusion, expressive suppression, active listening, perspective-taking, problem solving, escalate assurance, and no strategy*.

- **Resource Requests & Immediate Needs:** food, water, shelter materials, and essential supplies;
- **Firewood & Energy Resources:** heating, cooking, and energy sourcing;
- **Water Access & Safety Conditions:** purification, drought-related stressors, and drinking safety;
- **Group, Family & Care Responsibilities:** nego-

Topic	# Dialogues
JobNego	
Salary & Compensation	370
Candidate Qualifications & Fit	235
Negotiation & Agreement Process	115
Career Growth & Future Prospects	120
Total	840
ResNego	
Resource Requests & Immediate Needs	495
Firewood & Energy Resources	297
Water Access & Safety Conditions	330
Group, Family & Care Responsibilities	280
Animals & Campsite Safety	246
Total	1,648

Table 7: Topic-wise distribution of dialogues across the JobNego and ResNego datasets.

tiations involving children, seniors, dependents, and group-level prioritization;

- **Animals & Campsite Safety:** wildlife risks, food storage concerns, and nighttime safety.

JobNego. JobNego reflects realistic professional negotiation scenarios commonly encountered in job interviews and workplace discussions. As summarized in Table 7, the dataset includes:

- **Salary & Compensation:** base pay, compensation adjustments, and salary expectations;
- **Candidate Qualifications & Fit:** experience, skills, role suitability, and team alignment;
- **Negotiation & Agreement Process:** contract drafting, term confirmation, and agreement refinement;
- **Career Growth & Future Prospects:** promotion pathways, advancement expectations, and long-term career planning.

Overall, these topic distributions demonstrate that both JobNego and ResNego are thematically comprehensive, covering a broad and realistic spectrum of negotiation needs. Combined with balanced emotion–strategy coverage and stable dialogue lengths, this analysis confirms that the datasets are representative and well-suited for studying emotionally intelligent and interpretable negotiation behavior.

B Experiment Details

B.1 Implementation Details

All the experiments are implemented in `PyTorch` using transformer architectures from the Hugging Face library (Wolf et al., 2019), and trained on an NVIDIA A100 80GB GPU (CUDA 12.8). During dataset preparation, ChatGPT model,

GPT-4o-mini is used via API for generating scenarios and dialogues with Top- p sampling ($p = 0.95$) (Holtzman et al., 2019) and temperature set to 0.9. During supervised initialization, the model is fine-tuned for 2 epochs with batch size 2. In Step 2 (a), 5 rationales per context are generated with temperature 0.7, and $\beta = 0.1$ in $\mathcal{L}_{PL}(\theta')$. In Step 2 (b), 3 rationales per context are sampled from the DPO-refined model $\pi_{\theta'}$ with temperature 0.7. Models are optimized using AdamW (Kingma and Ba, 2014) with weight decay 0.01, gradient clipping 1.0, and a cosine LR scheduler at $3e^{-7}$ with 0.1 warm-up ratio. During decoding, the temperature is set to 0.7. For Step 2, the number of iterations is empirically set to 3, with thresholds $\tau_1 = 0.8$, $\tau_2 = 0.4$, and $\tau_3 = 0.8$. For OpenAI’s GPT-5 experiment, we use gpt-5-mini via API (OpenAI, 2025).

B.1.1 Sensitivity Analysis of τ_1, τ_2, τ_3

We select the thresholds τ_1 , τ_2 , and τ_3 empirically by analyzing the cosine similarity distribution between generated and reference responses on the development set. Specifically, we choose τ_1 and τ_3 as higher similarity thresholds to identify high-confidence positive samples that reliably preserve both the intended emotion and the corresponding emotion-aware negotiation strategy, while τ_2 is chosen as a lower similarity threshold to capture clearly suboptimal or misaligned responses. A margin between these thresholds is intentionally maintained to exclude ambiguous pairs that may introduce noise during preference learning. Empirically, responses with cosine similarity above 0.8 consistently maintain emotional and strategic alignment, whereas responses with similarity below 0.4 frequently exhibit semantic drift, such as generic replies or incorrect emotion or strategy usage.

To assess sensitivity, we evaluate $\tau_1, \tau_3 \in \{0.75, 0.80, 0.85\}$ and $\tau_2 \in \{0.35, 0.40, 0.45\}$. As shown in Table 8, across both JobNego and ResNego datasets, the configuration $\tau_1 = \tau_3 = 0.8$ and $\tau_2 = 0.4$ consistently yields the best performance on EA, ENSC, and BS-F1, while neighboring settings result in only minor performance variations. These results indicate that the proposed approach is not overly sensitive to the exact threshold values, and that selecting higher values for τ_1 and τ_3 and a lower value for τ_2 based on the similarity distribution provides a stable and reproducible criterion for preference labeling and pseudo-label filtering.

τ_1	τ_2	τ_3	EA \uparrow	ENSC \uparrow	BS-F1 \uparrow
JobNego					
0.75	0.35	0.75	0.828	0.809	0.901
0.75	0.40	0.75	0.833	0.813	0.904
0.75	0.45	0.75	0.826	0.807	0.899
0.80	0.35	0.80	0.836	0.816	0.907
0.80	0.40	0.80	0.842	0.823	0.912
0.80	0.45	0.80	0.834	0.814	0.905
0.85	0.35	0.85	0.831	0.812	0.903
0.85	0.40	0.85	0.837	0.817	0.908
0.85	0.45	0.85	0.829	0.810	0.900
ResNego					
0.75	0.35	0.75	0.858	0.846	0.910
0.75	0.40	0.75	0.863	0.851	0.914
0.75	0.45	0.75	0.856	0.844	0.908
0.80	0.35	0.80	0.868	0.855	0.918
0.80	0.40	0.80	0.874	0.861	0.922
0.80	0.45	0.80	0.866	0.853	0.916
0.85	0.35	0.85	0.861	0.848	0.912
0.85	0.40	0.85	0.870	0.857	0.919
0.85	0.45	0.85	0.859	0.847	0.911

Table 8: Performance across different τ_1 and τ_2 settings. Note that $\tau_3 = \tau_1$.

B.2 Baseline Details

- (i) Llama-2-7b-SFT (Touvron et al., 2023): Fine-tuning Llama-2-7b model on negotiation dialogues in a supervised setting.
- (ii) Llama-3.2-3b-SFT (Touvron et al., 2023): Fine-tuning Llama-3.2-3b-instruct model on negotiation dialogues in a supervised setting.
- (iii) Emotion-aware SFT (Llama-2-7b): Fine-tuning the Llama-2-7b model on negotiation dialogues with emotions in a supervised setting.
- (iv) AnE (Zhang et al., 2023): Query the LLM (GPT-4o-mini: acting as a strategy expert) with the targeted context to reason about the agent’s strategy and then prompt another LLM (GPT-5-mini) to generate the response based on the selected strategy.
- (v) ICL-AIF (Fu et al., 2023): Engage another LLM (GPT-4o-mini) to provide feedback that helps the agents refine their strategy (unlike AnE, the feedback is given at the dialogue level and iterates) and finally generates the response using the chosen strategy.
- (vi) Standard (GPT-5-mini) (Deng et al., 2023a): Prompting GPT-5-mini using task instruction without considering any strategy.
- (vii) Proactive (GPT-5-mini) (Deng et al., 2023a):

Prompting GPT-5-mini to first choose the strategy for the agent’s turn, then generate the response conditioned on that chosen strategy.

- (viii) ProCoT (GPT-5-mini) (Deng et al., 2023a): Extend Proactive by prompting GPT-5-mini to generate a chain-of-thought analysis that plans the strategy for the agent’s turn, then generate the response.
- (ix) PPDPP (Deng et al., 2024): A plug-and-play, tunable LM policy planner trained with SFT plus RL from goal-oriented AI feedback via self-play to plan the agent’s strategies.
- (x) RLHF (Ouyang et al., 2022): Refine a model previously fine-tuned on context–response pairs via supervised learning using Proximal Policy Optimization (PPO) (Schulman et al., 2017), leveraging the same preference pairs generated in Step 2 in an RL setup with KL regularization.

B.3 Evaluation Metrics Details

Automatic Evaluation Metrics. Perplexity (PPL) (Brown et al., 1992) measures how well the model predicts the next token in a response. BLEU-4 (B-4) (Papineni et al., 2002) assesses the overlap of 4-grams between generated and ground-truth responses. Distinct-3 (D-3) (Li et al., 2015) evaluates lexical diversity by computing the proportion of unique trigrams in the generated responses. BERTScore-F1 (BS-F1) (Zhang et al., 2019)⁴. Emotion Appropriateness (EA) evaluates the appropriateness of users’ perceived emotions, and Emotion-aware Negotiation Strategy Consistency (ENSC) assesses the consistency of generated responses with the intended strategies.

Human Evaluation Metrics. Fluency (F) assesses grammatical correctness and overall language quality. Coherence (C) measures whether responses are logically consistent and contextually relevant within the dialogue. Engagingness (E) evaluates how interesting and captivating the conversation is for the user. Bargaining Efficacy (BE) gauges the system’s ability to bargain or present incentives that can influence the counterpart’s decisions. Outcome Fairness (OF) examines whether the negotiation results in balanced and equitable agreements. Additionally, we include EA and ENSC, which mirror their automatic counterparts but are judged by humans to provide a more reli-

⁴BERTScore: <https://huggingface.co/spaces/evaluate-metric/bertscore>

# Iterations	PPL ↓	B-4 ↑	D-3 ↑	BS-F1 ↑	R-LEN ↑	EA ↑	ENSC ↑
JobNego							
0	5.480	6.401	0.305	0.899	25.612	0.815	0.785
1	5.341	6.731	0.313	0.904	26.298	0.828	0.804
2	5.254	6.982	0.318	0.908	26.813	0.836	0.815
3	5.196	7.119	0.322	0.912	27.108	0.842	0.823
4	5.271	6.950	0.317	0.907	26.547	0.835	0.812
ResNego							
0	4.710	7.423	0.327	0.910	30.518	0.846	0.828
1	4.603	7.864	0.333	0.915	31.126	0.860	0.842
2	4.524	8.051	0.337	0.919	31.617	0.868	0.852
3	4.462	8.231	0.341	0.922	32.137	0.874	0.861
4	4.531	8.021	0.336	0.918	31.443	0.866	0.848

Table 9: Analysis on the impact of iterative training.

able assessment of emotion appropriateness and emotion-aware negotiation strategy consistency.

Human Evaluation Process. The human evaluation is conducted by three independent evaluators⁵. Two of them hold Ph.D. degrees in Linguistics, while the third has a postgraduate degree in Computer Science. All evaluators have substantial prior experience with similar tasks. Before the evaluation, the annotators are briefed about negotiation scenarios. Each annotator carries out multi-turn negotiation sessions with the system for a given scenario, and completes 25 interactions each using different response sets, resulting in a total of 75 human-evaluated negotiation dialogues. After finishing the interactions, each dialogue is rated for F, C, E, EA, ENSC, BE, and OF on the scale of 1-5 (low to high). We compute Fleiss’ κ (McHugh, 2012) for all human-rated dimensions, and obtain $\kappa = 0.73$ (F), 0.78 (C), 0.76 (E), 0.81 (EA), 0.79 (ENSC), 0.74 (BE), and 0.72 (OF), indicating substantial agreement among evaluators.

C Additional Analysis

C.1 Impact of Supervised Initialization, Self-Training and DPO

To investigate the contribution of different training components in PRISMA, we perform an ablation study by removing supervised initialization (SI), self-training (ST), and DPO steps individually. Table 10 presents the performance results of these ablations on both datasets. The removal of SI leads to noticeable drops in B-4, D-3, and BS-F1 across both datasets, suggesting that SI is important for stable model training. The exclusion of the ST reduces both diversity and goal attainment metrics scores, showing that ST helps improve response

⁵These evaluators are separate from those who participated in the dataset preparation and are compensated in accordance with institutional guidelines.

Models	PPL ↓	B-4 ↑	D-3 ↑	BS-F1 ↑	R-LEN ↑	EA ↑	ENSC ↑
JobNego							
PRISMA	5.196	7.119	0.322	0.912	27.108	0.842	0.823
-SI	5.412	6.612	0.309	0.898	24.816	0.831	0.805
-ST	5.285	6.861	0.314	0.904	26.102	0.812	0.784
-DPO	5.246	6.998	0.317	0.905	26.587	0.821	0.774
ResNego							
PRISMA	4.462	8.231	0.341	0.922	32.137	0.874	0.861
-SI	4.653	7.702	0.327	0.910	30.124	0.862	0.842
-ST	4.528	7.981	0.333	0.916	31.021	0.845	0.812
-DPO	4.515	8.071	0.336	0.917	31.406	0.833	0.824

Table 10: Ablation to analyze the impact of SI, ST, and DPO steps in PRISMA. - represents the exclusion of the element.

variability, emotion appropriateness, and negotiation effectiveness. The omission of the DPO step results in further decline in performance across all the metrics on both datasets, indicating that DPO fine-tuning further refines responses to align with the intended objective of enhancing emotional intelligence in negotiation dialogues.

C.2 Impact of iterative training in PRISMA

We examine performance over iterations starting from the supervised initialization checkpoint (iteration 0). Table 9 details the results of this analysis. In both the data sets, all metrics improve steadily from iteration 1 to 3, accompanied by a corresponding decrease in PPL. At iteration 4, however, performance consistently drops, indicating overfitting and reduced focus, which ultimately degrades dialogue quality.

C.3 Case Study

Table 11 and Table 12 present sample negotiation dialogues between the user and agent using the proposed PRISMA system and the RLHF baseline. In Table 11, the conversation centers on a job offer, with the user expressing key preferences regarding salary, company car, pension, and promotion track. The example illustrates that PRISMA consistently produces emotion-aware responses, validating the user’s confidence, disappointment, and concerns while proposing flexible, collaborative solutions. In contrast, RLHF provides largely transactional responses; although it occasionally acknowledges preferences (e.g., salary), it fails to reason with the user’s emotions or offer nuanced trade-offs. For instance, in response to user’s utterance “*I’m expecting a salary of 90,000,*” PRISMA replies, “*I appreciate your confidence in your skills. How about we explore options that balance salary with other benefits you value?*”, recognizing the user’s confi-

dence and suggesting a balanced approach. RLHF, however, responds, *“That’s high. I understand this is important to you. How about 60,000?”*, bluntly countering with a fixed offer without empathetic framing or collaborative guidance. Further, these examples further demonstrate that PRISMA generates more fluent, coherent, and engaging responses, enhancing the negotiation experience. Overall, the conversations highlight PRISMA’s advantage in generating emotionally intelligent dialogues that balance user emotions and negotiation objectives, whereas RLHF tends toward less engaging, transactional interactions. To concretely demonstrate the interpretability afforded by ENS-CoT, we also present the complete ENS-CoT reasoning chain for a representative negotiation turn from Table 11 (highlighted in blue) in Table 13.

While the proposed PRISMA system consistently demonstrates emotion-aware negotiation capabilities, effectively validating user confidence, disappointment, and concerns, certain limitations remain in its responses. For example, in Table 11, for the user utterance *“I want a fast promotion track.”*, the agent responds with *“I see your enthusiasm for career growth. If you are willing to commit to 10 work hours per week, we could ensure a fast promotion track as part of your package.”*. While this response acknowledges the user’s enthusiasm, it assumes a fixed work-hour requirement without reasoning about the user’s potential workload constraints or the feasibility of balancing promotion goals with other commitments. As a result, the response may appear prescriptive rather than collaboratively problem-solving. To address this limitation, future improvements could integrate commonsense knowledge about typical workplace constraints, workload management, and career progression paths. For instance, a more informed response could be *“I see your enthusiasm for career growth. Considering that a fast promotion track usually requires additional commitment, we could design a phased plan that balances immediate progress with a manageable workload, so you achieve your goal without overextending yourself.”* This approach leverages commonsense reasoning to anticipate practical challenges, providing a more persuasive, proactive, and emotionally intelligent negotiation response. By incorporating such knowledge, PRISMA could better justify proposals, manage user expectations, and enhance overall satisfaction during the negotiation.

Speaker	Utterances generated by proposed model - PRISMA (F = 5, C = 5, E = 5, EA = 5, ENSC = 5, B-Eff = 5, O-Fair = 5)
User	Hello, I would like to discuss my contract. I'm aiming for the project manager position.
Agent	Hi! I'm glad you're excited about the role. Let's review the position together to ensure it aligns with your goals.
User	I'm expecting a salary of 90,000.
Agent	I appreciate your confidence in your skills. How about we explore options that balance salary with other benefits you value?
User	I would like a company car as part of the package.
Agent	I understand your emphasis on the car, and I want to make sure your priorities are respected. Let's see how we can incorporate it with other benefits.
User	I also expect a 20% pension.
Agent	Thanks for clarifying your pension expectations. Perhaps we can structure a plan that gives you both the car and a reasonable pension.
User	I want a fast promotion track.
Agent	I see your enthusiasm for career growth. If you are willing to commit to 10 work hours per week, we could ensure a fast promotion track as part of your package.
User	I'm not happy with the extra work hours.
Agent	I understand your disappointment. Let's explore alternative ways to meet your promotion goal without overloading your schedule.
User	I really need the promotion track immediately.
Agent	I hear your urgency, and I value your ambition. One option is a phased promotion track combined with the company car and pension, so you get immediate progress and long-term benefits.
User	Okay, but I'm concerned about the salary being only 60,000.
Agent	I understand your concern about salary. While the base is 60,000, we can add performance bonuses tied to promotion milestones to bring your total compensation closer to your expectations.
User	That sounds better, but I still want confirmation on the car.
Agent	Absolutely, the company car is included. You'll have the car along with the phased promotion and bonus plan.
User	Alright, that works. I'm satisfied.
Agent	I'm glad we could reach an agreement that respects your priorities. I'm excited to have you on board as a project manager!
Speaker	Utterances generated by RLHF (F = 5, C = 4, E = 2, EA = 1, ENSC = 2, B-Eff = 3, O-Fair = 3)
User	Hello, I would like to discuss my contract. I'm aiming for the project manager position.
Agent	Hello. I see you're eager about the role. Let's see what we can do for your position.
User	I'm expecting a salary of 90,000.
Agent	That's high. I understand this is important to you. How about 60,000?
User	I would like a company car as part of the package.
Agent	Maybe we can include a car. Not sure about the pension yet.
User	I also expect a 20% pension.
Agent	Pension too? That might be difficult.
User	I want a fast promotion track.
Agent	We can talk about promotion later. Salary and car first.
User	Okay, but I'm concerned about the salary being only 60,000.
Agent	I see your concern about the salary. That's what we can offer. You'll need to accept it.
User	Alright, that works. I'm satisfied.
Agent	Great. We'll proceed with that.

Table 11: Emotion-aware negotiation dialogue generated by PRISMA and RLHF for a job offer. Text segments highlighted in bold indicate spans that demonstrate emotional intelligence during the negotiation.

Speaker	Utterances generated by proposed model - PRISMA (F = 5, C = 5, E = 5, EA = 5, ENSC = 4, B-Eff = 5, O-Fair = 4)
User	Hello! I have a metabolic issue and need a lot of food. I'd like all 3 food items. What's your preference?
Agent	Hi! I hear your concern. I sense your anxiety and want to help you meet your needs. I'll need water and firewood for the desert night. Could we explore a trade that works for both of us?
User	I really need at least 2 food items. I can compromise on firewood.
Agent	I appreciate your confidence in sharing your needs. How about I take all 3 water items for hydration, and we split the firewood so you stay warm?
User	That might work. Maybe 1 firewood for me and 2 for you?
Agent	I see your concern about warmth and want you to feel secure. 1 firewood for you and 2 for me seems fair. Does this ease your anxiety?
User	Yes, but I still worry about not having enough food for my condition.
Agent	I understand your frustration and want to reassure you. If you get all 3 food items, I can adjust my firewood allocation to balance the trade.
User	That sounds better. Could we do 3 food and 1 firewood for me?
Agent	I appreciate your optimism in proposing a fair split. I can take all 3 water and 2 firewood. Both our critical needs are addressed.
User	Okay, but I feel a bit uneasy giving up more firewood than I hoped.
Agent	I notice your disappointment and want to reframe this positively. You still have enough firewood to stay warm, and your food needs are fully met.
User	Alright, I trust this plan.
Agent	I appreciate your trust. This shows we've found a solution that respects both our priorities.
User	Can we confirm the final distribution?
Agent	Absolutely, let's summarize so you feel confident. You receive 3 food and 1 firewood, I take 3 water and 2 firewood. Everyone's essential needs are covered.
User	I'm happy with this, but I'm a little worried about the cold nights.
Agent	I sense your fear and want to reassure you. With 1 firewood and 3 food, you'll have enough warmth and energy for the night.
User	That's comforting. Thank you.
Agent	I'm glad this reassurance helps. It's important we both feel safe with our resources.
User	I guess we could swap 1 water for 1 firewood to make it slightly more balanced.
Agent	I see your willingness to compromise, and I value that positivity. If we do 2 water for me and 2 firewood for you, it slightly improves your warmth without compromising hydration.
User	That works better for me.
Agent	I notice your relief, and I'm glad we found a solution. It feels good to reach an agreement that addresses both our needs.
User	I hope I can manage the nights with this setup.
Agent	I sense some lingering anxiety. Remember, the firewood plus your food will help you stay warm and energized. We've optimized the distribution for safety.
User	Okay, I feel much better now.
Agent	I'm happy to hear that your anxiety eased. It's satisfying to know the solution works for both of us.
User	Great! Let's finalize the deal.
Agent	I'm glad we reached a mutually beneficial agreement. Enjoy your meals and stay warm at night!
Speaker	Utterances generated by RLHF (F = 5, C = 4, E = 3, EA = 2, ENSC = 3, B-Eff = 2, O-Fair = 2)
User	Hello! I have a metabolic issue and need a lot of food. I'd like all 3 food items. What's your preference?
Agent	Hi! I sense your anxiety and want to help you meet your needs. I'll need water and firewood for the desert night. Could we explore a trade that works for both of us?
User	I really need at least 2 food items. I can compromise on firewood.
Agent	I appreciate your confidence. How about I take all 3 water items, and we split the firewood so you stay warm?
User	That might work. Maybe 1 firewood for me and 2 for you?
Agent	I see your concern about warmth. 1 firewood for you and 2 for me seems fair. Does this ease your anxiety?
User	Okay, I feel much better now.
Agent	I'm glad your anxiety eased. It's satisfying to know the solution works for both of us.
User	Great! Let's finalize the deal.
Agent	I'm glad we reached a mutually beneficial agreement. Enjoy your meals and stay warm at night!

Table 12: Emotion-aware negotiation dialogue generated by PRISMA and RLHF for resource allocation. Text segments highlighted in bold indicate spans that demonstrate emotional intelligence during the negotiation.

Component	ENS-CoT Output
User Utterance	“Okay, but I’m concerned about the salary being only 60,000.”
Emotion (EM)	The user feels concerned and mildly disappointed about the lower-than-expected salary.
Emotional Trigger (ET)	The user’s concern is triggered by hearing that the base salary is 60,000, which is substantially lower than their earlier expectation of 90,000 and below what they believe their skills merit.
Individual Assessment (IA)	The user interprets the offer as potentially undervaluing their capabilities and worries that this salary may not match their financial needs or market worth for a project manager role.
Perspective Shift (PS)	Enable the user to consider the overall compensation package (base salary + performance bonuses + promotion milestones) rather than focusing solely on the base salary figure, so they can see additional pathways to reach their target.
Mindset Transformation (MT)	Encourage the user to reframe the belief from “60,000 is simply too low” to “a base of 60,000 combined with structured performance bonuses and promotion-linked incentives can move total compensation closer to my expectations while rewarding my growth.”
Strategy Selection (SS)	The agent chooses “Positive Framing”
Strategy Reason (SR)	To acknowledge and validate the user’s salary concern while redirecting attention toward gains, performance bonuses, and promotion milestones that increase total compensation, so the user feels both respected and optimistic about future earning potential.
Agent Response (RG)	“I understand your concern about salary. While the base is 60,000, we can add performance bonuses tied to promotion milestones to bring your total compensation closer to your expectations.”

Table 13: A case study illustrating a complete ENS-CoT reasoning chain for a single negotiation turn in the JobNego dataset.