Emotion-o1: ADAPTIVE LONG REASONING FOR EMOTION UNDERSTANDING IN LLMS

Anonymous authors
Paper under double-blind review

ABSTRACT

Long chain-of-thought (CoT) reasoning has shown great promise in enhancing the emotion understanding performance of large language models (LLMs). However, current fixed-length CoT methods struggle to balance reasoning depth and efficiency. Simple tasks (e.g., sentiment classification) are over-reasoned, while complex tasks (e.g., sarcasm understanding) lack depth. To fill this gap, we present Emotion-o1, an adaptive CoT framework that dynamically adjusts reasoning length based on task complexity. Emotion-o1 is trained by distilling adaptive CoT patterns from a large reasoning model (LRM), followed by supervised fine-tuning and reinforcement learning with a four-part reward targeting accuracy, brevity, structure, and redundancy. Experimental results on four emotion tasks highlight: (1) Emotion-o1 demonstrates significant improvements over its backbone, with F1 score increases of 11%↑(Sentiment), 14%↑(Emotion), 18%↑(Humor), and 27%↑(Sarcasm). (2) In sentiment and emotion tasks, our 8B model demonstrates superior performance against SoTA LLMs, outperforming Grok-3 by 2.1% in sentiment and within 1% of OpenAI-o1 in emotion. (3) The framework maintains accuracy while reducing reasoning length by 83% compared to OpenAI-o1, demonstrating effective precision-efficiency optimization. From a lower-cost perspective, the framework also empowers SLMs to achieve reasoning capabilities comparable to larger ones.

1 INTRODUCTION

CoT reasoning, which elaborates a series of intermediate steps, has significantly improved the ability of LLMs to solve complex problems Yao et al. (2025). This has led to the rise of a new class of models known as large reasoning models (LRMs), such as DeepSeek-R1 Guo et al. (2025), OpenAI-o1 Jaech et al. (2024), and Qwen-QwQ Team (2025). Such LRMs demonstrate that scaling CoT length to hundreds or even thousands of steps can yield continual gains in reasoning accuracy, interpretability, and robustness across a wide range of tasks.

Despite these advances, fixed-length CoT strategies are poorly suited for emotion understanding tasks. For instance, simple tasks such as binary sentiment classification (e.g., "Is this review positive or negative?"), often elicit excessively verbose reasoning, resulting in substantial computational overhead and inefficient overthinking Xia et al. (2025). In contrast, complex tasks such as sarcasm detection suffer from shallow reasoning, failing to capture nuanced pragmatic and contextual cues, as shown in Fig. 1. This disconnect between fixed reasoning lengths and the inherently dynamic nature of emotion understanding leads to both performance bottlenecks and wasted computation.

We posit that effective emotion reasoning demands adaptive flexibility. Simple emotion tasks benefit from short, efficient reasoning paths, while complex emotional phenomena such as irony, ambiguity, and humor require deeper, reflective chains of thought. However, existing CoT-based emotion understanding approaches lack the ability to dynamically adjust the length of the reasoning according to the complexity of the task, limiting their generalization across different emotion domains.

To fill this gap, we introduce **Emotion-o1**, an adaptive reasoning framework that dynamically adjusts CoT length according to the complexity of the emotional task. Specifically, our approach first distills variable-length, structurally diverse reasoning paths, such as backtracking and self-reflection, etc., from SoTA LRMs (e.g., DeepSeek-R1). After supervised fine-tuning the model to acquire comprehensive reasoning capabilities, we further optimize reasoning quality via reinforcement learning,

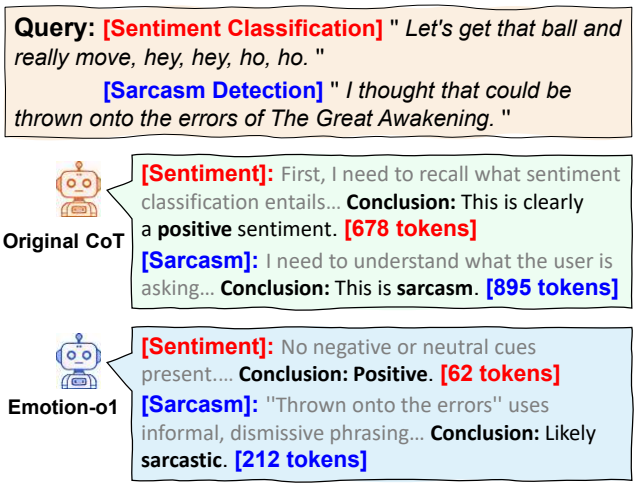

Figure 1: Original long CoT may lead to redundant computations or insufficient reasoning against our Emotion-o1.

guided by a multi-objective reward function across four dimensions: prediction accuracy, depth adaptability, structural diversity, and redundancy suppression. This allows Emotion-o1 to develop emotionally aligned, length-adaptive reasoning strategies tailored to the demands of each task.

Given that sentiment classification and emotion recognition mainly involve shallow emotional cues and limited contextual dependencies, we follow prior work in treating them as simple tasks Evans (2002). In contrast, sarcasm detection and humor understanding require complex pragmatic reasoning and deep contextual integration, and are therefore regarded as complex tasks that require deeper reasoning Chauhan et al. (2020; 2022). Using sarcasm detection and sentiment classification as illustrative tasks, we present detailed complexity proofs in Appendix A. We present empirical evaluations of the proposed approach on four emotion understanding tasks, and compare its performance against ten SoTA LLMs (e.g., DeepSeek-R1, GPT-4o, Claude 3.7, etc.). We highlight three key findings: (1) compared to the backbone, Emotion-o1 achieves F1-score improvements of 11%, 14%, 18%, and 27% on the four tasks, demonstrating the effectiveness of incorporating diverse reasoning structures; (2) Emotion-o1 achieves SoTA performance in sentiment and emotion classification. In sarcasm recognition, its F1 score is only 1% lower than that of GPT-4o, demonstrating our cost-efficient 8B model parity with leading large-scale LLMs at substantially lower computational cost. (3) compared to OpenAI-o1 (DeepSeek-R1), Emotion-o1 reduces the average reasoning length by 73% (54%), 52% (27%), 83% (70%), and 70% (58%) across the four tasks, highlighting its efficiency advantage. Our main contributions are as follows:

- We propose **Emotion-o1**, an adaptive CoT reasoning framework that dynamically adjusts reasoning length based on the complexity of emotion understanding tasks.
- We design a multi-objective reward function that jointly optimizes for prediction accuracy, reasoning brevity, structural coherence, and redundancy suppression, enabling the LRM to learn emotionally aligned and task-adaptive reasoning strategies.
- We validate Emotion-o1 on four emotion tasks, achieving SoTA performance with notably reduced reasoning cost.

2 Related Work

Affective Computing Affective computing (AC) enables machines to recognize, interpret, and respond to emotions Zhang et al. (2023). Early methods used feature engineering; with PLMs like BERT Devlin et al. (2019), fine-tuning became dominant for affective understanding (AU) and generation (AG) Verma et al. (2021); Nie & Zhan (2022), but struggled in cross-domain and multi-task reasoning Mao et al. (2022). LLMs Brown et al. (2020); Zhou et al. (2022) offer zero-shot and instruction-based modeling, yet still underperform in fine-grained tasks such as sarcasm or humor

detection Zhang et al. (2024). CoT prompting improves reasoning but often fixes template length. Our Emotion-o1 distills variable-length, structure-rich reasoning traces via multi-stage training and multi-objective rewards for dynamic reasoning depth.

Chain-of-Thought Reasoning CoT reasoning is central to enhancing LLM reasoning. Early short CoT used shallow, linear paths with limited depth and little exploration or error correction Chen et al. (2024), struggling on tasks requiring revisiting steps or exploring alternatives Mirzadeh et al. (2024). Recent work introduced non-linear structures such as Tree-of-Thoughts (ToT) Yao et al. (2023) and Graph-of-Thoughts (GoT) Besta et al. (2024), enabling branching, parallel reasoning, multiple hypotheses, and backtracking—laying the foundation for long CoT. Long-CoT LLMs like OpenAI-O1 Jaech et al. (2024) and DeepSeek-R1 Guo et al. (2025) scale reasoning to thousands of steps with dynamic feedback, achieving SoTA in math, programming, and symbolic inference. While Emotion-o1 bridges short and long CoT by adjusting reasoning depth and structure to task complexity, combining shallow efficiency with deep flexibility.

3 METHODOLOGY

As shown in Fig. 2, our framework include three stages: (1) Structured Emotion Reasoning Distillation extracts variable-length reasoning paths from leading LRMs; (2) Adaptive CoT-Augmented SFT initializes the model with structured emotional reasoning ability; (3) Reward-based RL refines reasoning quality via multi-objective optimization.

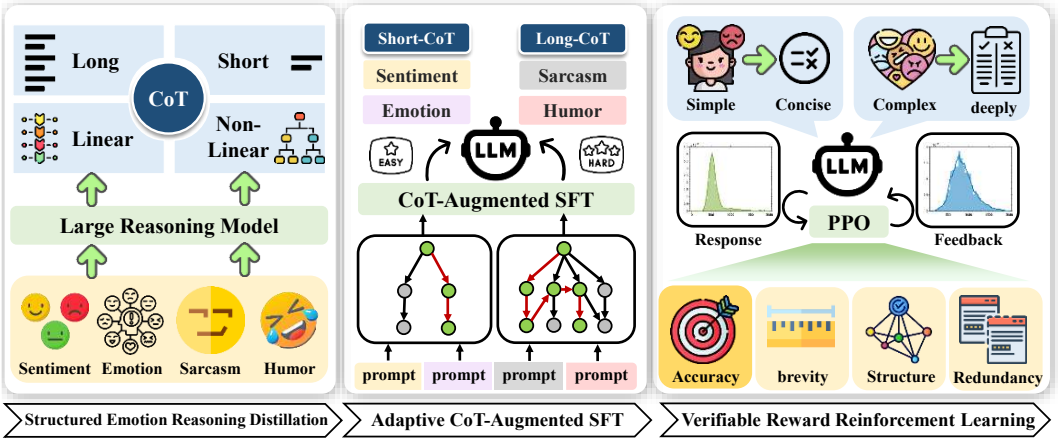

Figure 2: Overview of the proposed framework.

3.1 STRUCTURED EMOTION REASONING DISTILLATION

We construct labeled samples with diverse reasoning paths by distilling a leading LRM. Specifically, we select four key emotion understanding tasks, each paired with a widely used benchmark dataset: MELD Poria et al. (2018) for sentiment classification and emotion recognition, Sarcasm Corpus V2 Oraby et al. (2017) for sarcasm detection, and Reddit Humor Detection Weller & Seppi (2019) for humor recognition. Each instance consists of a text input x_i and its matching label y_i.

For each sample (x_i, y_i), we construct a prompt template $p(x_i, y_i, c)$, where c specifies the reasoning strategy, including *structure type* (linear or non-linear) and *length type* (short or long), more details are provided in Appendix C. We then use the DeepSeek-R1 for conditional sampling and generate N candidate reasoning paths:

$$\{r_{i,j}\}_{j=1}^{N} \sim LLM(p(x_i, y_i, c)) \tag{1}$$

where N is the number of candidate responses generated per prompt. We employ rejection sampling, in which we use $label(\cdot)$ to extract the predicted label from each generated CoT and retain only those whose labels match the ground-truth y_i as correct reasoning processes:

$$\mathcal{R}_i = \{r_{i,j} \mid label(r_{i,j}) = y_i, j \in [1, N]\} \tag{2}$$

where $\mathcal{R}_i$ denotes the set of valid CoT reasoning responses.

Different reasoning strategies exhibit task-specific efficacy across emotion understanding tasks. Our prompt explicitly steers the model to generate responses with distinct reasoning strategies c. Specifically, we considered two primary dimensions of reasoning structure:

- **Length Type**: Short (concise, direct reasoning) and Long (thorough, detailed analysis).
- **Reasoning Type**: Linear (step-by-step reasoning) and Non-linear (multi-path, branched reasoning structures).

Thus, each textual input x_i could yield multiple valid CoT responses across these dimensions, the final dataset $\mathcal{D}$ was constructed by aggregating all valid CoT reasoning instances across input samples and reasoning dimensions:

$$\mathcal{D} = \left\{ (x_i, y_i, r_{i,j}, c_{i,j}, l_{i,j}) \;\middle|\; \left\{ \begin{array}{c} r_{i,j} \in \mathcal{R}_i \\ c_{i,j} \in \{\text{linear}, \text{non-linear}\} \\ l_{i,j} \in \{\text{short}, \text{long}\} \end{array} \right\} \right\} \tag{3}$$

Next, we conducted SFT using the dataset $\mathcal{D}$. More details are provided in the Appendix D.

3.2 Adaptive CoT-Augmented SFT

We propose an Adaptive CoT-Augmented SFT method to enhance the reasoning capabilities of LLMs across different emotion tasks. Given an input text sequence $x = \{x_1, x_2, \ldots, x_T\}$ and a corresponding CoT rationale r, we construct task-adaptive instruction prompts according to the specific task type τ, where $\tau \in \{\textit{sentiment}, \textit{emotion}, \textit{sarcasm}, \textit{humor}\}$. The adaptive prompt construction function is formally defined as:

$$\mathcal{P}(x, r, y, \mathcal{D}, \tau) = \Phi(\tau) \oplus \Gamma(x) \oplus \Psi(r) \oplus \Omega(y, \mathcal{D}, \tau) \tag{4}$$

here, $\Phi(\tau)$ denotes the task-specific identifier that contextualizes the objective (e.g., Emotion Classification Task for $\tau = \textit{emotion}$); $\Gamma(x)$ formats the input text; $\Psi(r)$ incorporates the reasoning steps; and $\Omega(y, \mathcal{D}, \tau)$ encodes the label y along with class definitions customized according to the task type τ drawn from the dataset $\mathcal{D}$. The operator $\oplus$ represents string concatenation.

The training objective maximizes the conditional likelihood of the complete reasoning path and label prediction, where the model implicitly adapts the reasoning depth and structure according to τ:

$$\mathcal{L}_{\text{SFT-CoT}} = -\sum_{t=1}^{L} \log P(w_t \mid \mathcal{P}(x, r, y, \mathcal{D}, \tau), w_{<t}, \tau; \theta) \tag{5}$$

Here, w_t denotes the t-th token in the response, L is the total length of the reasoning and label sequence, and θ represents the model parameters. Incorporating τ as a conditioning variable in the likelihood term allows the model to progressively adapt its reasoning strategy to diverse tasks.

3.3 Verifiable Reward RL

We propose a verifiable reward RL approach to optimize the reasoning quality of the model further. Initialized with a fine-tuned SFT model, our method employs the proximal policy optimization (PPO) algorithm for training. By sampling multiple candidate responses during each update, the model learns to adjust its reasoning length according to task complexity adaptively.

The reward function is constructed as a weighted sum of prediction accuracy, depth adaptivity, structural diversity, and redundancy reduction. The components are given as follows:

3.3.1 Accuracy Reward

The first and most important reward is the accuracy reward, ensuring that the model prioritizes generating correct predictions aligned with the ground-truth labels. It is written as:

$$r_{\text{acc}} = \begin{cases} +1.0, & \text{if} \quad \hat{y} = y \\ -1.0, & \text{if} \quad \hat{y} \neq y \\ -\epsilon_{acc}, & \text{if prediction is missing} \end{cases} \tag{6}$$

where $\hat{y}$ is the predicted label, y is the ground truth label, and ϵ_{acc} is a small constant introduced to impose an appropriate penalty when the prediction is missing, thereby encouraging the model to generate correct labels.

3.3.2 TASK-AWARE VARIABLE-LENGTH REWARD

To enable the model to flexibly choose between **long** and **short** inference modes for each emotion task, we first define an expected length L_{base}, a maximum length L_{max}, and a minimum length L_{min} for each inference mode within each task. Specifically, we statistically analyze the length distribution of responses under different inference modes within our constructed SFT dataset $\mathcal{D}$. And utilize quantile-based statistics derived from the response length distribution L_l to establish the inference length boundaries. Thus, the minimum and maximum lengths are defined as follows:

$$L_{min} = \lfloor P_5(L_l) \rfloor, L_{max} = \lceil P_{95}(L_l) \rceil \tag{7}$$

where P_k denotes the k-th percentile of the distribution. Given the varying complexity of different emotion tasks, we empirically set an expected length L_{exp} based on domain expertise. Meanwhile, we adopt the statistical median of the length distribution as $L_{sts} = median(L_l)$. Thus, the final expected length L_{base} is formulated as a weighted combination of empirical and statistical values:

$$L_{base} = \alpha \cdot L_{exp} + (1 - \alpha) \cdot L_{sts} \tag{8}$$

where α is a tunable parameter that trades off empirical knowledge and statistical observations.

Furthermore, we introduce a length-based reward r_{length} to penalize responses deviating from the expected length range and reward responses close to the desired length. Specifically, the length reward is computed as follows:

$$r_{\text{length}} = \begin{cases} \left(s_{\min} \frac{L}{L_{\min}}\right)^2, & L < L_{\min} \\ \exp\left(-s_{\max} \frac{L - L_{\max}}{L_{\max}}\right), & L > L_{\max} \\ \exp\left(-\frac{1}{2}\left(\frac{L - L_{\text{base}}}{s_{\text{base}}(L_{\max} - L_{\min})}\right)^2\right), & L_{\min} \le L \le L_{\max} \end{cases} \tag{9}$$

here, L denotes the length of the generated response, while s_{min}, s_{max}, and s_{base} are scaling factors that control the intensity of rewards and penalties within the reward function. Specifically, for responses shorter than the minimum length, we apply a quadratic penalty to encourage the model to produce more comprehensive inferences. For responses exceeding the maximum length, we employ an exponential penalty to discourage the generation of redundant information. For responses within the acceptable length boundaries, a Gaussian-shaped reward is used to incentivize the model to generate outputs close to the expected length.

3.3.3 REASONING STRUCTURE REWARD

During reasoning, whether to use structured reasoning approaches significantly determines the length and clarity of the CoT generated by the model. Thus, we propose a novel structure-oriented reward r_{struct}. This reward explicitly evaluates key reasoning behaviors (e.g., "decomposition", "reflection", and "verification") and the appropriate usage of logical connectives (e.g., "therefore", "however", and "thereby"). Formally, the proposed reward function is defined as follows:

$$r_{struct} = \lambda \cdot \min\left(\frac{|A|}{N_A}, 1\right) + (1 - \lambda) \cdot \min\left(\frac{|C|}{N_C}, 1\right) \tag{10}$$

The set A denotes valid reasoning actions appearing in the response, while C represents the logical connectives used. N_A and N_C are predefined target numbers of reasoning actions and logical connectives, respectively, and λ is a hyperparameter balancing their relative importance. Given the variability of optimal reasoning patterns across tasks, our reward design deliberately refrains from constraining the sequence or positioning of reasoning actions and connectives. A response receives the reward if it contains the targeted number of reasoning actions and connectives. Through this carefully designed reward mechanism, we expect the model to autonomously adjust its inference length based on the varying difficulty levels of different emotion tasks.

3.3.4 REPETITION PENALTY

Additionally, we introduce a repetition penalty r_{repeat}, which discourages redundant content within generated responses. The similarity between sentence pairs within the response is determined by combining lexical overlap and semantic similarity from BERT embeddings by parameter β:

$$Sim(s_i, s_j) = \beta \cdot S_{lex}(s_i, s_j) + (1 - \beta) \cdot S_{sem}(s_i, s_j) \tag{11}$$

where lexical similarity $S_{lex}(s_i, s_j) = \frac{|W(s_i) \cap W(s_j)|}{|W(s_i) \cup W(s_j)|}$ is computed using the Jaccard similarity between two sentences, with $W(s)$ representing the set of words in sentence s. Semantic similarity $S_{sem}(s_i, s_j) = \frac{\text{BERT}(s_i) \cdot \text{BERT}(s_j)}{|\text{BERT}(s_i)||\text{BERT}(s_j)|}$ is calculated as the cosine similarity between BERT embeddings of the two sentences. The similarity threshold τ is set empirically. And the repetition penalty r_{repeat} is defined as the proportion of sentence pairs whose similarity exceeds τ: $r_{repeat} = \min\left(\frac{C_\tau}{T}, 1.0\right)$, where C_τ is the number of such pairs and T is the total number of pairs.

This ensures that responses with higher redundancy receive a greater negative penalty, effectively promoting more concise and diverse generated content.

3.3.5 FINAL REWARD FUNCTION

The total reward r_{total} for each response is defined as the weighted sum of four components:

$$r_{\text{total}} = w_{\text{acc}} \cdot r_{\text{acc}} + w_{\text{length}} \cdot r_{\text{length}} + w_{\text{struct}} \cdot r_{\text{struct}} - w_{\text{repeat}} \cdot r_{\text{repeat}} \tag{12}$$

We assign the highest weight to w_{acc} to ensure prediction accuracy remains the top priority. For Long-CoT reasoning, w_{struct} receives the second-highest weight to encourage structured reasoning; for Short-CoT, we set $w_{\text{struct}} = 0$ to disable this term. To encourage appropriate response lengths, w_{length} is given the third-highest weight. Finally, w_{repeat} receives the lowest weight, applying mild penalties to promote output diversity without sacrificing accuracy.

These weights are empirically tuned through extensive experimentation to effectively balance accuracy, structure, length control, and content diversity.

4 EXPERIMENTS

4.1 BASELINE AND EXPERIMENTAL SETUP

Ten SoTA LLMs are included for comparison, including models without intermediate reasoning steps (No-CoT): LLaMA-3.1-8B, Qwen-2.5-7B; models employing a few concise reasoning steps (Short-CoT): DeepSeek-V3, GLM-4, Qwen-3-14B, Grok-3; and models supporting more detailed and structured reasoning (Long-CoT): DeepSeek-R1, OpenAI-o1, GPT-4o, Claude-3.7.

Our experiments are conducted on 4×A100 40G GPUs, running on Ubuntu 22.04, with Python 3.12, PyTorch 2.4.0, and CUDA 12.1. We adopt the **LLaMA-3.1-8B** as our base model and use evaluation metrics widely used for classification tasks: Accuracy, Macro-F1, and Weighted-F1. Additional details of the experimental parameter settings can be found in the Appendix F.

4.2 MAIN RESULT

Table 1 summarizes the performance comparison between Emotion-o1 and 10 baseline models across four emotion tasks, using the average results from 3 experiments. Results show that Emotion-o1 consistently outperforms baselines to varying degrees. Further details about error analysis and limitations are provided in Appendix B and Appendix G.

Specifically, compared with the backbone model LLaMA-3.1-8B, emotion-o1 achieves accuracy improvements of **7%**, **7%**, **14%**, and **21%** on the sentiment, emotion, humor, and sarcasm tasks, respectively. Additionally, significant enhancements are observed in terms of Macro-F1 and Weighted-F1 scores. For simple tasks (sentiment and emotion), the Macro-F1 score increases by up to **14%**, while for more complex tasks (humor and sarcasm), the Weighted-F1 score improves by up to **27%**. These findings confirm the effectiveness of explicitly modeling and adapting varying reasoning lengths and strategies to diverse emotion understanding tasks.

Table 1: Model performance comparison across four emotion recognition tasks, **Bold** and underline indicate the best and second-best results for each task.

Paradigm	Model	Sentiment			Emotion			Humor			Sarcasm		
		Acc	Ma-f1	We-f1	Acc	Ma-f1	We-f1	Acc	Ma-f1	We-f1	Acc	Ma-f1	We-f1
No-CoT	Qwen-2.5-7B	0.6437	0.6086	0.6381	0.5586	0.4158	0.5700	0.6475	0.6092	0.6097	0.5030	0.3460	0.3437
	LLaMA-3.1-8B	0.6073	0.5391	0.5835	0.5613	0.3206	0.5047	0.6270	0.5876	0.5871	0.5360	0.4714	0.4701
Short-CoT	Grok-3	0.6559	0.6295	0.6518	0.6261	0.5058	0.6265	0.8869	0.8858	0.8859	0.7257	0.7126	0.7121
	GLM-4	0.6326	0.6106	0.6326	0.6215	0.4819	0.6146	0.8989	0.8989	0.8989	0.7247	0.7174	0.7171
	Qwen-3-14B	0.6494	0.6007	0.6331	0.5966	0.4261	0.5811	0.7373	0.7233	0.7236	0.7084	0.6895	0.6890
	DeepSeek-V3	0.6314	0.6148	0.6316	0.5801	0.4783	0.5878	0.8506	0.8481	0.8482	0.6802	0.6518	0.6511
Long-CoT	GPT-4o	0.6326	0.5924	0.6196	0.6015	0.4499	0.5812	0.9043	0.9041	0.9041	0.7577	0.7570	0.7569
	Claude-3.7	0.6483	0.6184	0.6408	0.6310	0.4814	0.6098	**0.9488**	**0.9488**	**0.9488**	0.7328	0.7267	0.7270
	OpenAI-o1	0.6379	0.6060	0.6306	**0.6398**	**0.4984**	**0.6341**	0.9231	0.9231	0.9231	**0.7724**	**0.7717**	**0.7718**
	DeepSeek-R1	0.6402	0.6134	0.6381	0.5900	0.4697	0.5967	0.8342	0.8306	0.8307	0.7062	0.6947	0.6943
Adaptive-CoT	**Emotion-o1(ours)**	**0.6770**	**0.6548**	**0.6766**	0.6352	0.4649	0.6141	0.7694	0.7684	0.7684	0.7469	0.7468	0.7469

We further conducted a comparative evaluation of emotion-o1 against several widely recognized LLMs. Experimental results demonstrate that emotion-o1 achieves SoTA performance on the sentiment task, surpassing established baselines in terms of accuracy, Macro-F1, and Weighted-F1 metrics. In the emotion task, accuracy is on par with OpenAI-o1, demonstrating its generalization ability in the multi-classification task. Regarding the more complex sarcasm task, emotion-o1 obtains competitive yet slightly suboptimal performance, achieving results marginally just below GPT-4o 1%, showing the effectiveness of Long CoT in handling complex tasks. For humor detection, Emotion-o1 achieves parity with larger-scale models like Qwen3-14B, outperforming it by 3%, yet demonstrates a substantial gap against top competitors such as Claude-3.7 in complex humor tasks.

Overall, these results show that our 8B model delivers competitive or superior performance compared to SoTA LLMs across all tasks, and also exhibits strong generalization across model scales.

4.3 CoT Length Analysis

We analyze the impact of each stage on the average length of CoT reasoning, as shown in Table 2. The base model LLaMA-3.1-8B produces shallow rationales with limited variation across tasks. In contrast, SFT introduces task-specific adaptation: it shortens reasoning for simple tasks like Sentiment (↓18%), while substantially increasing CoT depth for complex tasks such as Sarcasm (↑281%) and Humor (↑246%). The RL stage further refines this adaptation. It further reduces CoT length for Sentiment and Humor by **7%** and **13%**, respectively. Appendix E furnishes several typical cases.

Table 2: The average CoT length of the model in each stage. ↓ indicates shorter, ↑ indicates longer.

	Base	SFT	RL
Sentiment	114	93 (↓**18%**)	85 (↓**25%**)
Emotion	123	186 (↑**51%**)	183 (↑**49%**)
Sarcasm	52	198 (↑**281%**)	199 (↑**283%**)
Humor	70	242 (↑**246%**)	233 (↑**233%**)

Based on the results shown in Table 1, it can be further observed that a shorter CoT length achieves the best results in the sentiment task, demonstrating the rationality of reducing redundant reasoning; in the emotion task, merely increasing CoT length by 49% can achieve comparable performance to the SoTA model OpenAI-o1 of Long CoT. In more complex tasks such as sarcasm, Emotion-o1 with an extended structured reasoning (↑283%) outperforms all models using Short CoT, and is only inferior to two larger-scale Long CoT models.

These results confirm that our two-stage pipeline enables Emotion-o1 to learn task-aware reasoning strategies, by using shorter CoTs for straightforward tasks and longer ones for cognitively demanding tasks, and improves both performance and inference efficiency.

4.4 Ablation Study

We conduct ablation experiments to examine the effects of CoT length and structure, as illustrated in Fig. 3. The key findings are summarized as follows:

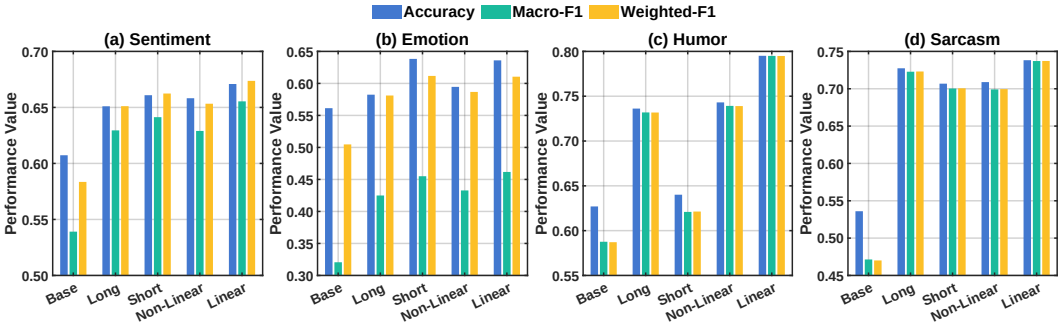

Figure 3: Ablation analysis of CoT variants: Long, Short, Linear and Non-Linear Reasoning

- **Sentiment**: Short CoT and linear CoT yield greater improvements. They achieve increases of approximately **6%** and **7%** in accuracy, respectively, suggesting concise reasoning is sufficient for such tasks, offering both effectiveness and computational efficiency.
- **Emotion**: Short CoT achieves **5%** higher accuracy than long CoT, and linear reasoning outperforms nonlinear reasoning by **4%**. This suggests that excessive reasoning may introduce unnecessary complexity or noise, and that direct, focused reasoning is more beneficial.
- **Humor**: Long CoT yields an **11%** improvement in F1 over short CoT, while linear reasoning provides a **5%** gain in accuracy. This confirms the need for deeper and more structured reasoning in humor understanding, where implicit intent and abstract cues are involved.
- **Sarcasm**: Similar to humor, the combination of long and linear CoT delivers the best results, highlighting the importance of detailed, explicit reasoning in handling pragmatic and context-dependent phenomena.

Overall, all CoT-based strategies clearly improve performance over the base model. Moreover, the results validate that sentiment and emotion tasks benefit from concise and direct reasoning, while sarcasm and humor tasks demand longer and deeper reasoning to capture subtle linguistic cues.

4.5 Visualization of Token Consumption

To intuitively compare the computational cost, we perform a comparative analysis of Emotion-o1's token consumption against the performance of mainstream reasoning models, which is linearly correlated with FLOPs, as illustrated in Fig. 4. While achieving superior results, Emotion-o1 also shows higher efficiency across all tasks compared to DeepSeek-R1 and OpenAI-o1. Specifically:

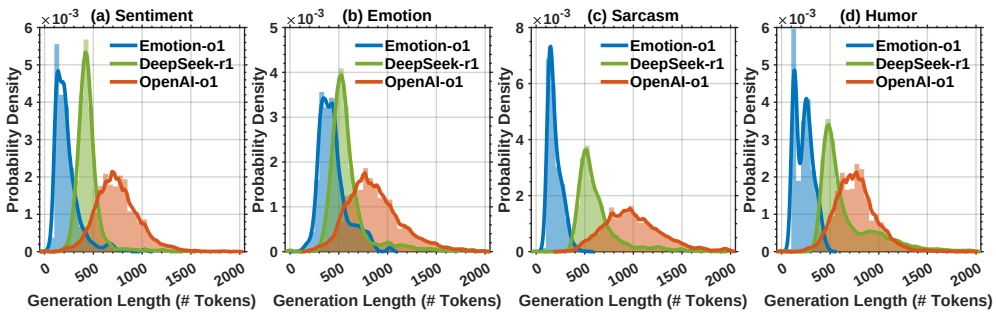

Figure 4: Token consumption distribution across reasoning models

The median tokens of Emotion-o1 consistently rank as the lowest among the models. This trend is particularly pronounced in the Sarcasm task, where the tokens reduces by **83%** compared to OpenAI-o1. Even in the Emotion task, which exhibits the smallest improvement, tokens decrease by approximately **52%** relative to OpenAI-o1.

Furthermore, Emotion-o1 demonstrates a lower maximum tokens compared to its competitors. In the Sarcasm task, the 90th percentile tokens reduces by nearly **4.8** times relative to OpenAI-o1, underscoring the model's ability to effectively suppress extreme values.

The standard deviation of tokens generated by Emotion-o1 across tasks is consistently lower. For example, in the Sarcasm task, the standard deviation is only **1/3** that of OpenAI-o1, indicating Emotion-o1 produces more stable outputs with greater concentration in response values.

In summary, Emotion-o1 consistently exhibits superior output efficiency across all tasks while maintaining competitiveness against mainstream reasoning models.

4.6 Scale Analysis

In this section, we investigate how variations in model size affect performance across different tasks, using LLaMA models of three scales (1B, 3B, and 8B). As illustrated in Fig. 5, all three evaluation metrics show noticeable improvements as the model size increases.

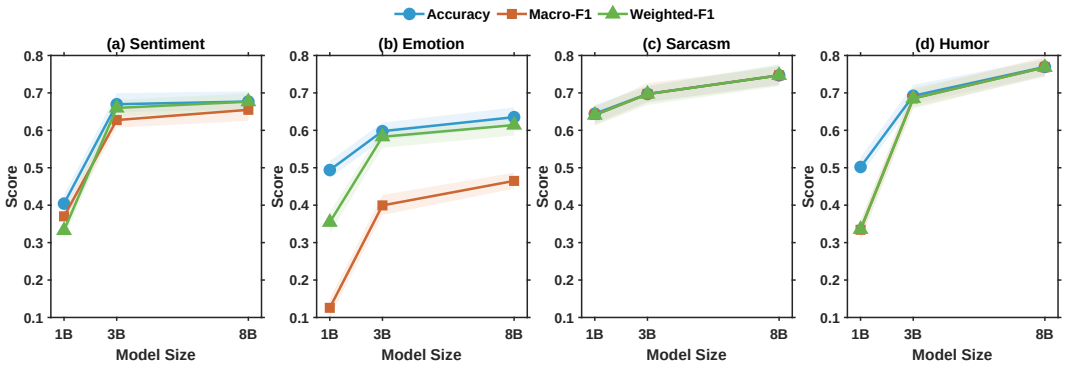

Figure 5: Model Scale Performance Comparison Across Different Tasks.

For the Sentiment, Emotion, and Humor tasks, expanding smaller models from 1B to 3B yields substantial gains, with the F1 score increasing by up to 35%. However, when further scaling from 3B to 8B, the improvements become less pronounced. These results indicate that small-scale models can achieve significant performance gains at relatively low computational cost.

In contrast, the Sarcasm task exhibits a consistent upward trend in performance with increasing model size, achieving about a 5% gain. This steady improvement underscores the potential advantages of larger models for tasks requiring more complex understanding.

Overall, Emotion-o1 strikes an effective balance between cost and performance, making it a viable and efficient option for a range of applications. These findings provide useful guidance for future work in selecting appropriate model scales according to task requirements and resource constraints.

5 Conclusion

This work addresses the limitations of fixed-length chain-of-thought reasoning for emotional understanding in LLMs by proposing Emotion-o1, an adaptive framework that dynamically adjusts reasoning depth based on task complexity. Through multi-stage training with a multi-objective reward (accuracy, brevity, structure, redundancy), our approach achieves significant performance gains: F1 improvements of 11% (sentiment), 14% (emotion), 18% (humor), and 27% (sarcasm) over its backbone. Notably, the 8B model outperforms Grok-3 by 2.1% and within 1% of OpenAI-o1 in critical tasks while reducing reasoning length by 83% versus OpenAI-o1. Emotion-o1 establishes an efficient bridge between structured reasoning and emotional understanding.

REFERENCES

Maciej Besta, Nils Blach, Ales Kubicek, Robert Gerstenberger, Michal Podstawski, Lukas Gianinazzi, Joanna Gajda, Tomasz Lehmann, Hubert Niewiadomski, Piotr Nyczyk, et al. Graph of thoughts: Solving elaborate problems with large language models. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, pp. 17682–17690, 2024.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

Dushyant Singh Chauhan, SR Dhanush, Asif Ekbal, and Pushpak Bhattacharyya. Sentiment and emotion help sarcasm? a multi-task learning framework for multi-modal sarcasm, sentiment and emotion analysis. In *Proceedings of the 58th annual meeting of the association for computational linguistics*, pp. 4351–4360, 2020.

Dushyant Singh Chauhan, Gopendra Vikram Singh, Aseem Arora, Asif Ekbal, and Pushpak Bhattacharyya. A sentiment and emotion aware multimodal multiparty humor recognition in multilingual conversational setting. In *Proceedings of the 29th international conference on computational linguistics*, pp. 6752–6761, 2022.

Qiguang Chen, Libo Qin, Jiaqi Wang, Jingxuan Zhou, and Wanxiang Che. Unlocking the capabilities of thought: A reasoning boundary framework to quantify and optimize chain-of-thought. *Advances in Neural Information Processing Systems*, 37:54872–54904, 2024.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171–4186, 2019.

Dylan Evans. *Emotion: The science of sentiment*. Oxford University Press, USA, 2002.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025.

Aaron Jaech, Adam Kalai, Adam Lerer, Adam Richardson, Ahmed El-Kishky, Aiden Low, Alec Helyar, Aleksander Madry, Alex Beutel, Alex Carney, et al. Openai o1 system card. *arXiv preprint arXiv:2412.16720*, 2024.

Rui Mao, Qian Liu, Kai He, Wei Li, and Erik Cambria. The biases of pre-trained language models: An empirical study on prompt-based sentiment analysis and emotion detection. *IEEE transactions on affective computing*, 14(3):1743–1753, 2022.

Iman Mirzadeh, Keivan Alizadeh, Hooman Shahrokhi, Oncel Tuzel, Samy Bengio, and Mehrdad Farajtabar. Gsm-symbolic: Understanding the limitations of mathematical reasoning in large language models. *arXiv preprint arXiv:2410.05229*, 2024.

Guangtao Nie and Yibing Zhan. A review of affective generation models. *arXiv preprint arXiv:2202.10763*, 2022.

Shereen Oraby, Vrindavan Harrison, Lena Reed, Ernesto Hernandez, Ellen Riloff, and Marilyn Walker. Creating and characterizing a diverse corpus of sarcasm in dialogue. *arXiv preprint arXiv:1709.05404*, 2017.

Soujanya Poria, Devamanyu Hazarika, Navonil Majumder, Gautam Naik, Erik Cambria, and Rada Mihalcea. Meld: A multimodal multi-party dataset for emotion recognition in conversations. *arXiv preprint arXiv:1810.02508*, 2018.

Qwen Team. Qwq-32b: Embracing the power of reinforcement learning, March 2025. URL `https://qwenlm.github.io/blog/qwq-32b/`.

Palak Verma, Neha Shukla, and AP Shukla. Techniques of sarcasm detection: A review. In *2021 international conference on advance computing and innovative technologies in engineering (ICACITE)*, pp. 968–972. IEEE, 2021.

Orion Weller and Kevin Seppi. Humor detection: A transformer gets the last laugh. *"Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing"*, November 2019.

Heming Xia, Chak Tou Leong, Wenjie Wang, Yongqi Li, and Wenjie Li. Tokenskip: Controllable chain-of-thought compression in llms, 2025. URL https://arxiv.org/abs/2502.12067.

Ben Yao, Yazhou Zhang, Qiuchi Li, and Jing Qin. Is sarcasm detection a step-by-step reasoning process in large language models? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 25651–25659, 2025.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models, 2023. *URL https://arxiv. org/abs/2305.10601*, 3:1, 2023.

Yazhou Zhang, Yang Yu, Qing Guo, Benyou Wang, Dongming Zhao, Sagar Uprety, Dawei Song, Qiuchi Li, and Jing Qin. Cmma: benchmarking multi-affection detection in chinese multi-modal conversations. *Advances in Neural Information Processing Systems*, 36:18794–18805, 2023.

Yazhou Zhang, Chunwang Zou, Zheng Lian, Prayag Tiwari, and Jing Qin. Sarcasmbench: Towards evaluating large language models on sarcasm understanding, 2024. URL https://arxiv.org/abs/2408.11319.

Yongchao Zhou, Andrei Ioan Muresanu, Ziwen Han, Keiran Paster, Silviu Pitis, Harris Chan, and Jimmy Ba. Large language models are human-level prompt engineers. In *The eleventh international conference on learning representations*, 2022.

A Theoretical Justification of Task Difficulty

In this section, we present a theoretical analysis to justify why sarcasm detection is intrinsically more difficult than sentiment classification. We formalize the two tasks, analyze their Bayes error bounds, provide a task reduction argument, and compare their statistical learning complexity using VC-dimension theory. Together, these results show that sarcasm detection requires strictly more information and model capacity to achieve comparable performance.

A.1 Problem Formalization

Let x denote a text utterance and c denote its context, which may include dialogue history, speaker identity, social background, or cultural knowledge. We define the following:

- $S(x) \in \{0, 1\}$: the *literal sentiment polarity* of x, where 0 indicates negative and 1 indicates positive sentiment.
- $C(x, c) \in \{0, 1\}$: the *contextual polarity*, i.e., the true stance once context c is taken into account.
- $Z(x, c) = \mathbb{1}[S(x) \neq C(x, c)]$: the *sarcasm indicator*, which equals 1 if the literal sentiment disagrees with the contextual stance, and 0 otherwise.

The sentiment classification task is to predict $S(x)$ from x, while sarcasm detection requires predicting $Z(x, c)$ from (x, c). Intuitively, sarcasm arises when the surface form of an utterance is positive but the intended meaning is negative (or vice versa). This naturally suggests that sarcasm detection involves reasoning about both literal semantics and pragmatic intent.

A.2 Irreducible Error Bound

We first compare the Bayes error of the two tasks. If only x is observable, the optimal classifier for sarcasm detection must rely on the conditional probability

$$p(x) = \Pr(Z = 1 \mid x) = \Pr(S(x) \neq C(x, c) \mid x). \tag{13}$$

The corresponding Bayes error is

$$\mathcal{E}_{\text{sar}}^{(x)} = \mathbb{E}_x\big[\min\{p(x), 1-p(x)\}\big]. \tag{14}$$

If there exists a subset

$$A_\varepsilon = \{x : \varepsilon \le p(x) \le 1-\varepsilon\}, \quad \varepsilon \in (0, 0.5], \tag{15}$$

then

$$\mathcal{E}_{\text{sar}}^{(x)} \ge \varepsilon \cdot \Pr(A_\varepsilon) > 0. \tag{16}$$

Interpretation. Whenever context c is essential for disambiguating whether an utterance is sarcastic, the Bayes error given only x is strictly positive. This irreducible error reflects the inherent ambiguity of sarcasm. In contrast, sentiment classification typically assumes that $S(x)$ is a deterministic function of x, yielding $H(S \mid x) = 0$, where $H(S \mid x)$ is the *conditional entropy*:

$$H(S \mid x) = -\sum_{s \in \mathcal{S}} \Pr(S = s \mid x) \log \Pr(S = s \mid x) \tag{17}$$

measuring the uncertainty of sentiment label S given text x, and consequently zero irreducible error.

A.3 Task Reduction Argument

Sarcasm detection can be formalized as an XOR composition:

$$Z(x, c) = S(x) \oplus C(x, c), \tag{18}$$

where $\oplus$ denotes logical exclusive-or. This representation shows that detecting sarcasm requires distinguishing two sources of evidence: the literal polarity $S(x)$ derived from the surface text, and the contextual polarity $C(x, c)$ inferred from background knowledge.

Formally, if $\mathcal{H}_S$ and $\mathcal{H}_C$ are hypothesis classes that approximate S and C, then sarcasm detection requires approximating functions of the form

$$\mathcal{H}_{\text{sar}} = \big\{\, h(x, c) = h_S(x) \oplus h_C(x, c) : h_S \in \mathcal{H}_S,\, h_C \in \mathcal{H}_C \,\big\}. \tag{19}$$

Implication. Any model that solves sarcasm detection must have the representational capacity to jointly encode both literal sentiment and contextual stance. Thus, sarcasm detection is at least as hard as solving both subtasks simultaneously.

A.4 Statistical Learning Complexity

We now analyze the VC-dimension of sarcasm detection. By standard results for composite hypothesis classes, we obtain the following bounds:

$$\max\{\text{VCdim}(\mathcal{H}_S), \text{VCdim}(\mathcal{H}_C)\} \;\le\; \text{VCdim}(\mathcal{H}_{\text{sar}}) \;\le\; \text{VCdim}(\mathcal{H}_S) + \text{VCdim}(\mathcal{H}_C). \tag{20}$$

The lower bound shows that sarcasm detection is at least as complex as the harder of the two subtasks, while the upper bound suggests that under mild independence assumptions, the complexity may approach the sum of both. In practice, this implies that sarcasm detection requires significantly larger sample sizes and stronger inductive biases to achieve the same generalization performance as sentiment classification.

A.5 Discussion and Conclusion

The above analysis yields three important insights:

1. Sarcasm detection has a strictly positive irreducible error when context is missing, while sentiment classification does not.
2. The sarcasm label is definable only as an XOR of sentiment and contextual stance, making the task intrinsically compositional.

3. The VC-dimension of sarcasm detection is strictly larger than that of sentiment classification, often close to the sum of both subtasks.

Taken together, these results demonstrate that sarcasm detection is provably more difficult than sentiment classification. This conclusion aligns with empirical findings that sarcasm models require richer contextual modeling, more complex architectures, and substantially more data to achieve robust performance.

B Error Analysis

Fig. 6 shows the error analysis for the four tasks. Sarcasm and Humor detection show notable false negatives, indicating difficulties in identifying complex contextual cues. Sentiment analysis exhibits polarity interpretation challenges, particularly in distinguishing neutral from negative expressions. Emotion recognition demonstrates a tendency to default to neutral classifications, with authentic emotions like joy and anger frequently misclassified, alongside cross-category confusion. The observed false positives in sentiment and sarcasm tasks suggest occasional oversensitivity to certain linguistic signals. Collectively, these patterns highlight room to refine contextual understanding and develop task-specific approaches to capture linguistic nuances more effectively.

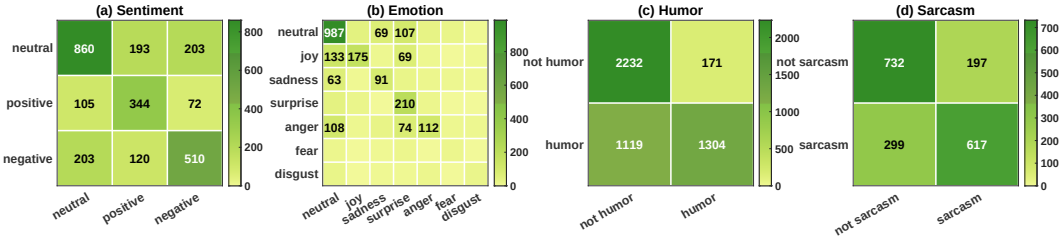

Figure 6: The confusion matrices of the four tasks.

C Instruction Study

This section presents the prompt design used in constructing part of the CoT dataset. For each emotion task, we designed prompts tailored to the specific characteristics of the task(see task definitions in Table 3). To extract long CoT from LLMs, as shown in Table 4, we first defined several distinct reasoning structures. These structures served as guidelines to steer the model in flexibly selecting and applying different reasoning approaches, without restrictions on their order or frequency of use. After completing the reasoning process, the model was required to perform conclusion verification and error checking to ensure completeness and logical soundness.

Table 3 Task definitions for sentiment, emotion, sarcasm, and humor classification

Task	**Task Definition**
Sentiment	0=neutral: no clear emotional cues 1=positive: features like positive lexicon, uplifting emojis, achievement expressions 2=negative: contains negative elements, expressions of unpleasant events
Emotion	0=neutral: no clear emotional cues 1=joy: features like positive lexicon, uplifting emojis, achievement expressions 2=sadness: contains loss/grief elements, negative event descriptions 3=surprise: unexpected events or cognitive dissonance 4=anger: aggressive language, confrontational rhetoric 5=fear: threat-related content, anxiety indicators 6=disgust: expressions of revulsion, descriptions of unpleasant events

Table 3 Task definitions for sentiment, emotion, sarcasm, and humor classification

Task	Task Definition
Sarcasm	1=sarcasm: contains features like surface praise with underlying criticism, contextual incongruity, exaggerated contrast, etc. 0=not sarcasm
Humor	1=humor: contains features like wordplay/puns, exaggerated scenarios, unexpected twists, contextual incongruity, absurd juxtapositions, etc. 0=not humor

Specifically, for each dataset sample *text*, we generated four long CoTs and one short CoT. The short CoT followed the conventional step-by-step prompting approach and is not further discussed here. Among the four long CoTs, two were required to employ non-linear reasoning patterns, such as Tree-of-Thought or Graph-of-Thought structures. One was constrained to follow a multi-path reasoning strategy, in which the model was encouraged to explore multiple reasoning trajectories; in our work, the trajectories were not split into separate CoTs but were instead treated as multi-step reasoning within a single CoT. This design preserves structural diversity in the reasoning process.

Table 4 Task-Specific prompt designs for sentiment, emotion, sarcasm, and humor classification

Task	Prompt Example
Sentiment	Perform rigorous sentiment analysis by dynamically applying selected reasoning methods. Use the following framework (choose steps, order, and iterations as needed): **[Reasoning Framework]** 1.Decomposition: Break down text elements (semantics/context/rhetoric) 2.Reflection: Question initial assumptions and verify their rationality 3.Verification: Cross-check logical consistency 4.Transition: Handle contradictory information (using "however" - like analysis) 5.Retry: Correct the reasoning path when errors are found **[Process Requirements]** 1.Must include $\geq$ 5 reasoning steps, freely combining the above components, without limitation on the number of times or order, and also free to explore other reasoning methods. 2.Each step must clearly indicate the type of reasoning used (e.g., Step 1 - Decomposition) 3.At least two verification stages must be included: Preliminary conclusion verification and Final decision verification 4.Contradictions in the text must be addressed (demonstrating the use of "however" - like analysis). 5.Error correction must show the complete adjustment of the reasoning path. 6.Final conclusion must align with *sentiment_definition* **[Error Checkpoints]** 1.Sentiment intensity validation 2.Context-text consistency check 3.Emoji-semantic alignment verification Tweet content: *text*, conclude with "Therefore, the sentiment label is: "(0=neutral,1=positive,2=negative)
Emotion	Perform rigorous multi-dimensional emotion analysis by dynamically applying selected reasoning methods. Use the following framework (choose steps, order, and iterations as needed): **[Reasoning Framework]** 1.Decomposition: Break down text elements (semantics/context/rhetoric) 2.Reflection: Question initial assumptions and verify their rationality 3.Verification: Cross-check logical consistency

Table 4 Task-Specific prompt designs for sentiment, emotion, sarcasm, and humor classification

Task	Prompt Example
	4.Transition: Handle contradictory information (using ”however” - like analysis) 5.Retry: Correct the reasoning path when errors are found **[Process Requirements]** 1.Must include $\geq$ 5 reasoning steps, freely combining the above components, without limitation on the number of times or order, and also free to explore other reasoning methods. 2.Each step must clearly indicate the type of reasoning used (e.g., Step 1 - Decomposition) 3.At least two verification stages must be included: Preliminary conclusion verification and Final decision verification 4.Contradictions in the text must be addressed (demonstrating the use of ”however” - like analysis) 5.Error correction must show the complete adjustment of the reasoning path 6.Final conclusion must align with *emotion_definition* **[Error Checkpoints]** 1.Emotional intensity validation 2.Trigger event analysis 3.Emoji/textual consistency verification 4.Cultural context alignment check Tweet content: *text*, conclude with ”Therefore, the emotion label is: ”(0=neutral,1=joy,2=sadness,3=surprise,4=anger,5=fear,6=disgust)
Sarcasm	Perform rigorous sentiment analysis reasoning please strictly follow the structured reasoning process. The reasoning framework includes the following optional components: **[Reasoning Framework]** 1.Decomposition: Break down text elements (semantics/context/rhetoric) 2.Reflection: Question initial assumptions and verify their rationality 3.Verification: Cross-check logical consistency 4.Transition: Handle contradictory information (using ”however” - like analysis) 5.Retry: Correct the reasoning path when errors are found **[Process Requirements]** 1.Must include $\geq$ 5 reasoning steps, freely combining the above components, without limitation on the number of times or order, and also free to explore other reasoning methods. 2.Each step must clearly indicate the type of reasoning used (e.g., Step 1 - Decomposition). 3.At least two verification stages must be included: Preliminary conclusion verification and Final decision verification 4.Contradictions in the text must be addressed (demonstrating the use of ”however” - like analysis). 5.Error correction must show the complete adjustment of the reasoning path. 6.Final conclusion must align with *sarcasm_definition* **[Error Checkpoints]** 1.Rhetorical analysis completeness check 2.Contextual factor weight validation 3.Counterfactual outcome consistency verification Tweet content: *text*, conclude with ”Therefore, the sarcasm label is: ”(1=sarcasm, 0=none)
Humor	Perform rigorous sentiment analysis reasoning please strictly follow the structured reasoning process. The reasoning framework includes the following optional components: **[Reasoning Framework]** 1.Decomposition: Break down text elements (semantics/context/rhetoric)

Table 4 Task-Specific prompt designs for sentiment, emotion, sarcasm, and humor classification

Task	Prompt Example
	2.Reflection: Question initial assumptions and verify their rationality 3.Verification: Cross-check logical consistency 4.Transition: Handle contradictory information (using ”however” - like analysis) 5.Retry: Correct the reasoning path when errors are found **[Process Requirements]** 1.Must include $\geq$ 5 reasoning steps, freely combining the above components, without limitation on the number of times or order, and also free to explore other reasoning methods. 2.Each step must clearly indicate the type of reasoning used (e.g., Step 1 - Decomposition). 3.At least two verification stages must be included: Preliminary conclusion verification and Final decision verification 4.Contradictions in the text must be addressed (demonstrating the use of ”however” - like analysis). 5.Error correction must show the complete adjustment of the reasoning path. 6.Final conclusion must align with *humor_definition* **[Error Checkpoints]** 1.Rhetorical analysis completeness check (wordplay/puns detection) 2.Contextual absurdity validation 3.Expectation-subversion consistency verification Tweet content: *text*, conclude with ”Therefore, the humor label is: ”(1=humor, 0=none)

D Dataset Statistics

The emotion CoT dataset is shown in Fig. 7. After filtering, we compile **21,266**, **22,966**, **21,169**, and **22,341** samples for the *sentiment*, *emotion*, *sarcasm*, and *humor* tasks, respectively. The ratio of linear to non-linear reasoning is approximately **7:3** across all tasks. The ratio of long to short CoT is roughly **1:1** for Sentiment and Emotion, and about **8:2** for Sarcasm and Humor.

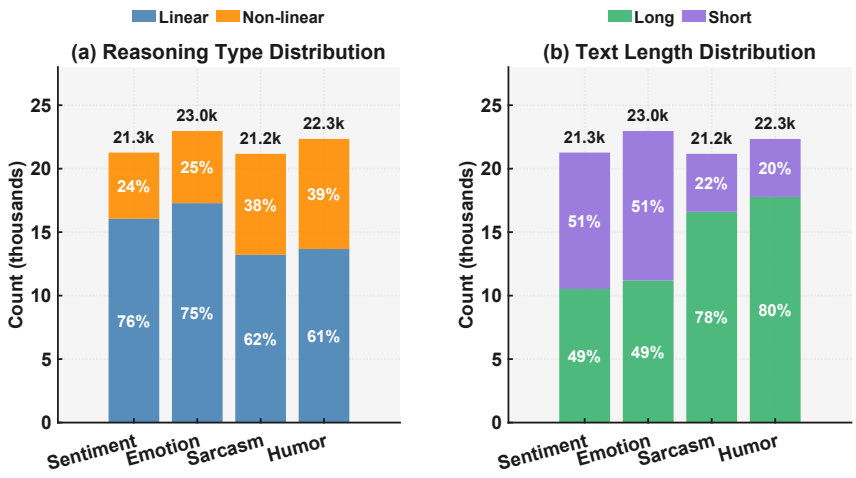

Figure 7: Overview of the CoT Dataset

E CASE STUDY

In this section, we selected the most representative sentiment classification and sarcasm detection tasks from simple and complex tasks to demonstrate their reasoning details, as shown in Fig. 8. For tasks of different difficulties, Emotion-o1 can adaptively select the reasoning mode for reasoning.

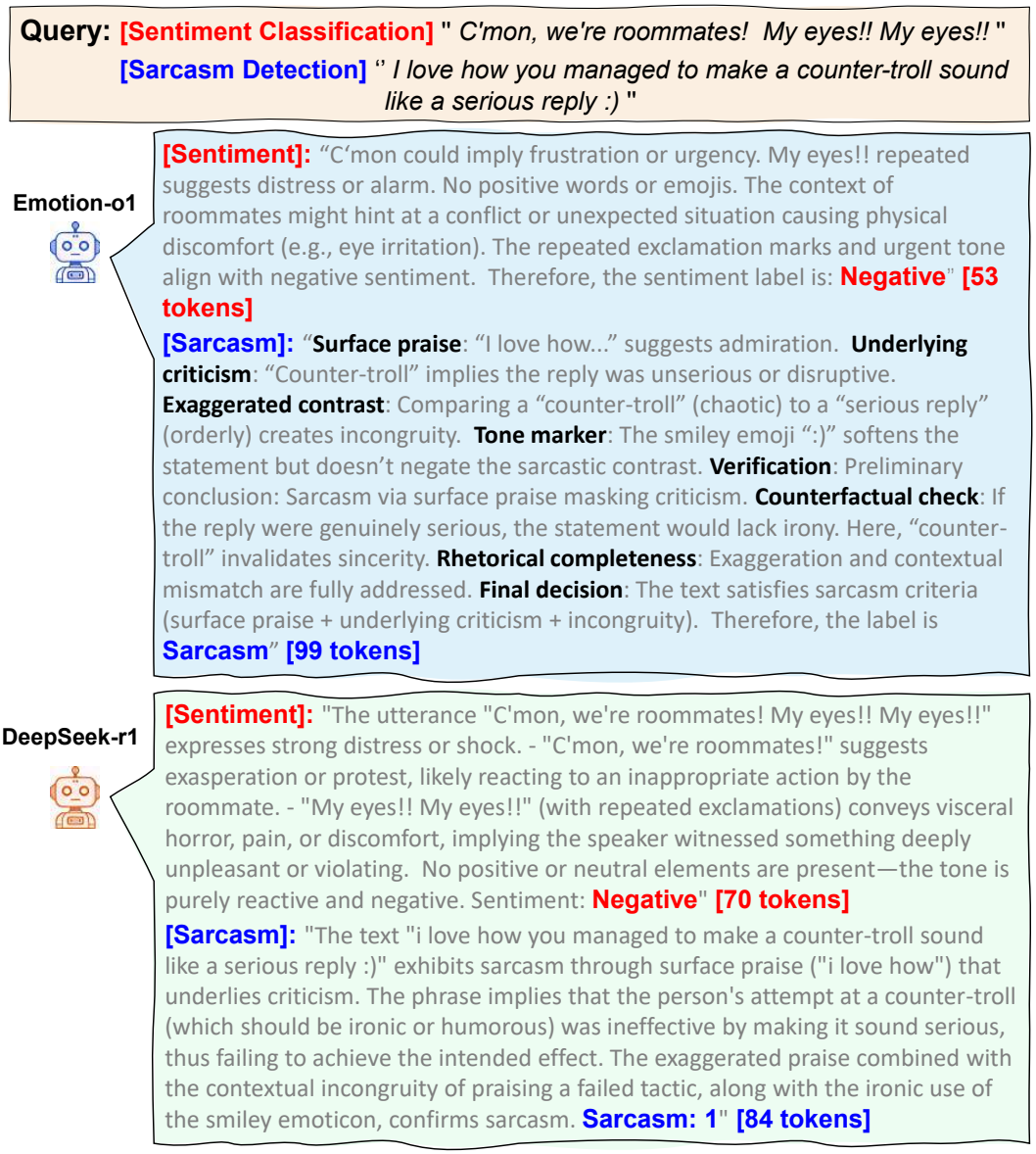

Figure 8: Reasoning details for simple task(Sentiment) and complex task(Sarcasm)

For the sentiment classification task, using a short CoT allows the sentiment to be inferred directly from salient keywords in the text, requiring only a concise response (53 tokens) to achieve the target task, thereby making the reasoning notably more succinct than that of DeepSeek-r1 (70 tokens) and effectively avoiding any unnecessary output. In contrast, for the sarcasm detection task, a long CoT—incorporating verification and reflection—enables the exploration of implicit meanings in the context. This process produces a more detailed response (99 tokens) and, when compared with DeepSeek-r1, presents a noticeably clearer and more organized reasoning structure, thereby ensuring comprehensive coverage of the entire inference process.

Through task-oriented adaptive length reasoning, Emotion-o1 can achieve more reasonable use of computing resources, thereby maximizing model efficiency under limited costs.

F Parameter Settings

Table 5 summarizes the detailed parameter configurations for the various stages of the process. In order to address the specific demands of the Long CoT task, the encoding length was set to relatively large values in each stage. For the PPO stage, initial weight values were assigned based on prior empirical experience, and the optimal parameters were subsequently determined from the best-performing settings across 5 recorded experimental iterations.

Table 5 Summary of experimental parameter settings

Stage	**Parameter**
Distillation	temperance = 0.7 max_tokens = 8192
SFT	max_length = 2048 per_device_train_batch_size = 2 gradient_accumulation_steps = 8 num_train_epochs = 3 learning_rate = 2e-5 bf16 = True gradient_checkpointing = True
PPO	ϵ_{acc} = 0.1 s_{min} = 1 s_{max} = 4 s_{base} = 0.4 N_A = 4 N_C = 5 τ = 0.75 w_{acc} = 0.7 if CoT is Short else 0.6 w_{length} = 0.25 if CoT is Short else 0.15 w_{struct} = 0 if CoT is Short else 0.2 w_{repeat} = 0.05 seed = 42 max_length = 4096 learning_rate = 1e-5 batch_size = 4 ppo_epochs = 4

G Limitations

First, our method is evaluated on four curated emotion-related tasks, which, while diverse, may not cover the full spectrum of affective reasoning challenges in real-world applications. Second, our framework focuses solely on textual input, excluding multimodal signals (e.g., visual or acoustic cues), which are often crucial for understanding emotions in human communication.