CLARO: Controlled Attribute-Driven Reasoning Optimization for Efficient Chain-of-Thought

Oded Schlesinger[1,2] **Young Kyung Kim**[1,2] **J. Matias Di Martino**[1,3] **Guillermo Sapiro**[2,4]

[1]Duke University [2]Princeton University [3]Universidad Católica del Uruguay [4]Apple

Abstract

Large language models exhibit strong reasoning capabilities but often require significant computational resources due to verbose, unstructured Chain-of-Thought outputs. Recent approaches guide reasoning length through token penalties or truncation, risking the omission of necessary steps. We posit that conciseness should be an emergent property of structured thought, rather than a result of artificially forced brevity. To this end, we first demonstrate that *Attribute-Guided Prompting*, a lightweight zero-shot strategy, improves reasoning performance while reducing inference cost. Building on this foundation, we introduce **C**ontro**l**led **A**ttribute-Driven **R**easoning **O**ptimization (*CLARO*), a reinforcement learning framework designed to internalize these benefits. *CLARO* guides models to embed high-quality structural attributes, such as readability, math density, syntactic compression, and low redundancy, within a user-defined token budget. The proposed method outperforms state-of-the-art baselines across diverse benchmarks, yielding accuracy gains of up to 63.6%, demonstrating that guiding generated output language structure enhances reasoning. Overall, our findings establish that optimizing the thought process structure refines reasoning efficacy, with computational efficiency emerging as a derivative benefit of a clearer thought process. Code and models are available at `https://github.com/odedsc/CLARO`.

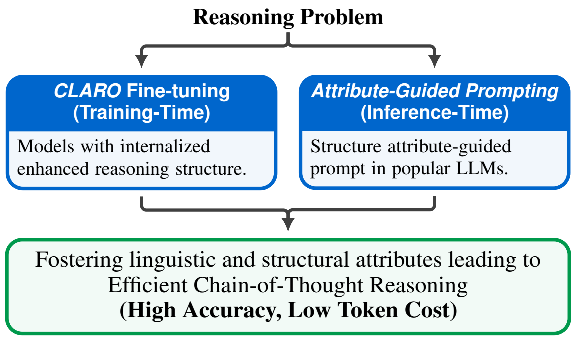

Figure 1: Overview of the proposed *CLARO* framework. The diagram contrasts two proposed approaches for optimizing Chain-of-Thought efficiency: *Attribute-Guided Prompting* (right), which utilizes inference-time instructions to foster linguistic structural attributes (e.g., readability), and *CLARO* fine-tuning (left), which leverages reinforcement learning to permanently internalize these interpretable attributes into the model, promoting high accuracy with reduced token costs.

1 Introduction

Recent advancements in large language models (LLMs) have demonstrated remarkable capabilities in reasoning across a wide array of domains, including question answering (Brown et al., 2020; Zhou et al., 2022) and logical inference (Kojima et al., 2022; Lewkowycz et al., 2022). These models have fundamentally transformed problem-solving frameworks by leveraging natural language explanations, most notably through the Chain-of-Thought (CoT) paradigm (Wei et al., 2022), enabling models to effectively emulate step-by-step thinking and execute inferential processes. However, despite their remarkable performance, these models typically require significant inference-time compute (Snell et al., 2024). Improving reasoning quality within limited inference costs is therefore essential for making reasoning models economically viable, scalable, and sustainable.

A primary factor contributing to this computational burden is the common excessive verbosity of generated outputs, which inflates the number of generated tokens (Xu et al., 2025), increasing inference costs, latency, and overall computational demands (Hassid et al., 2025). Empirical studies indicate that superfluous reasoning trajectories often yield redundant or repetitive steps without delivering accuracy gains, and may even degrade performance (Chen et al., 2024; Yang et al., 2025).

Recent research (Muennighoff et al., 2025; Ag-

Findings of the Association for Computational Linguistics: ACL 2026, pages 26779–26799

garwal and Welleck, 2025) has explored strategies for regulating reasoning length by appending end-of-thinking tokens or incorporating reinforcement learning (RL). However, these approaches focus on quantity without considering content structure, thereby overlooking the linguistic attributes that inherently facilitate concise and clear reasoning. We contend that effective conciseness must reflect the quality of content, ensuring the output budget is used for meaningful inferential progress rather than being artificially extended or shortened.

To operationalize this concept, we draw upon established linguistic and structural attributes associated with efficient communication. Readability, which quantifies text comprehensibility, has been associated with improved clarity in LLM outputs (Markowitz, 2024). Similarly, syntactic compression through shorter sentence structures is often associated with improved communicative efficiency (Siddharthan, 2014). Moreover, natural language explanations can be fragile, a phenomenon where minor changes in phrasing or ambiguity in natural language reasoning can degrade model performance (Mirzadeh et al., 2024). Conversely, mathematical notation offers a compact, unambiguous, and robust alternative for representing intermediate reasoning steps. Finally, redundancy in generated text, characterized by the repetition or unnecessary use of terms, can impair reasoning clarity and introduce inefficiencies in the output (Chiang and Lee, 2024).

We argue that while LLMs implicitly encode these core inferential qualities, they often do not utilize them or their latent capacity for more efficient communication; standard prompting paradigms fail to fully elicit these underlying patterns, resulting in more inefficient writing styles as defaults.

To address this, we define a set of interpretable and easily computable language attribute-inspired metrics and integrate them into a unified optimization framework. Contrary to previous work focused on brevity, our proposed method, *CLARO*, provides more refined control over the generated output, fostering succinct and informative outputs that follow linguistic standards of quality of expression.

As outlined in Figure 1, we demonstrate the efficacy of this approach at both prompting and fine-tuning paradigms: (1) *Attribute-Guided Prompting* (Inference-time), a zero-shot strategy that explicitly instructs models to structure their reasoning, providing an immediate plug-and-play performance enhancement; and (2) *CLARO* (Training-time), an attribute-driven fine-tuning framework that internalizes these behaviors via RL.

Extensive empirical evaluation shows that *CLARO* significantly improves performance, surpassing current state-of-the-art baselines across diverse benchmarks by up to 63.6%. This demonstrates that guiding the structure of generated output refines and enhances reasoning across distinct domains, as illustrated in Figures 2 and A.4. Our analysis disentangles individual attribute contributions through ablations, revealing how each component affects reasoning behavior. Through extensive experimentation and analysis, we demonstrate that explicit attention to reasoning form and properties as facilitated by *CLARO* yields superior efficiency–performance trade-offs.

2 Related Works

Inference-Time Compute. Allocating increased computational resources during inference has been demonstrated to be effective for complex reasoning, a paradigm fundamentally enabled by the CoT framework (Wei et al., 2022). Prominent techniques include evaluating multiple candidate solutions (Cobbe et al., 2021; Lightman et al., 2023; Snell et al., 2024), utilizing larger computational budgets (OpenAI, 2024; Jurayj et al., 2025), and systematically exploring reasoning pathways (Yao et al., 2023; Bi et al., 2024). Additionally, self-feedback mechanisms involving iterative refinement of generated outputs significantly boost reasoning quality, though they inherently increase token generation costs (Zelikman et al., 2022; Madaan et al., 2023). In contrast to these compute-intensive strategies, our approach optimizes the reasoning structure itself to enhance performance within limited budgets.

Efficiency and Conciseness in CoT. Recent literature has addressed the computational cost of reasoning by promoting brief and concise CoT outputs. Certain approaches incentivize brevity by sampling multiple responses and rewarding correct solutions with shorter lengths (Arora and Zanette, 2025; Luo et al., 2025a). Compression-based methods transform extended reasoning into compact forms, as in C3oT (Kang et al., 2025), which trains and infers with compressed CoT, TokenSkip (Xia et al., 2025), which selectively removes tokens deemed as less important, or ThinkPrune (Hou et al., 2025), which trims outputs beyond a predefined length threshold. Alternative strategies (Nayab et al., 2024; Yuan

Problem

A software company employs exactly seven sales representatives—Kim, Mahr, Parra, Quinn, Stuckey, Tiao, and Udall—to work in its three sales zones—Zone 1, Zone 2, and Zone 3. Each sales representative works in exactly one of the sales zones, in accordance with the following conditions: Either Parra or Tiao (but not both) works in Zone 1. Either Tiao or Udall (but not both) works in Zone 2. Parra and Quinn work in the same sales zone as each other. Stuckey and Udall work in the same sales zone as each other. There are more of the sales representatives working in Zone 3 than in Zone 2.
Q: If more sales representatives work in Zone 1 than in Zone 3, then which one of the following could be true?
Answer Choices: (A)Kim works in Zone 2. (B)Mahr works in Zone 2. (C)Parra works in Zone 3. (D)Tiao works in Zone 1. (E)Udall works in Zone 3.

Length-Controlled Baseline

Okay, so we have a problem with seven sales reps and three zones. We need to figure out which option could be true based on the given conditions. The answer choices are A to E, each suggesting who works in a particular zone. The key is to determine the correct option where the conditions are satisfied and the conclusion holds. After analyzing the constraints, I think option A is correct because it fits all the conditions without violating any. </think> The answer is $\boxed{A}$. ✗

***CLARO* (Ours)**

To solve, consider constraints. If Zone 1 > Zone 3, then possible assignments must balance. Given Parra and Quinn in same zone, and others' constraints, E (Udall in Z3) might fit. Answer E. </think> The answer is $\boxed{E}$. ✓

▲ Readability Math notation rate ▲
▲ Syntactic compression Redundancy ▼

Figure 2: Qualitative comparison of general-domain reasoning trajectories under identical settings. The base model (left), before fine-tuning using our proposed method, produces a verbose and unstructured trace that fails to solve the problem. In contrast, *CLARO* (right) generates a significantly more concise and efficient output. By optimizing for structural attributes, *CLARO* fosters a more focused reasoning process that successfully reaches the correct answer.

et al., 2024) introduced prompting methods that explicitly guide the model to limit CoT length. A separate research avenue separates natural language reasoning from calculations and symbolic computation, as in Program of Thoughts (Chen et al., 2022), which offloads arithmetic and logical operations to a program interpreter for a more compact, verifiable CoT. Contrary to these methods that often force brevity, our framework treats conciseness as an emergent property of improved reasoning structure rather than a standalone objective.

Controllable Text Generation. Beyond conciseness, extensive research has focused on guiding models toward controllable or desired output attributes. Foundational efforts include preference-optimization approaches (Ouyang et al., 2022; Rafailov et al., 2023; Shao et al., 2024), discriminator-guided methods such as GeDi (Krause et al., 2020), which guide outputs using lightweight modules, and steering mechanisms to avoid undesired behaviors or terms (Liu et al., 2021; Lu et al., 2022; Schick et al., 2021).

Regarding reasoning specifically, prompting studies (Bsharat et al., 2023; Weston and Sukhbaatar, 2023) demonstrate that enforcing structural directives, such as eliminating filler words or removing irrelevant context, can enhance performance. Complementing these, recent training-based methods extend controllability to CoT optimization; approaches like s1 (Muennighoff et al., 2025) and L1 (Aggarwal and Welleck, 2025) have developed frameworks specifically to regulate the length of the reasoning process. While prior methods focus primarily on length regulation, our work extends controllability to specific linguistic attributes that facilitate the inferential process.

3 Methodology

3.1 Problem Setting

CoT typically involves generating an output y, comprising reasoning r and a final answer a, conditioned on input x. Standard LLMs maximize $P(a|x, r)$, typically ignoring or even aiming to increase the computational cost of r, lacking control over the structural quality of generated reasoning. We approach this problem by learning a policy that maximizes the probability of the correct answer $a = a^*$ while optimizing r using a set of interpretable structural attribute metrics $\mathcal{M}$.

Unlike methods relying on increased compute or post-hoc trajectory analysis, we leverage these attributes to directly shape model behavior, treating conciseness as an emergent property of structural optimization rather than a simple truncation target, thereby maintaining or even improving accuracy.

3.2 Linguistic and Structural Attributes

Building upon prior research that identified key linguistic and structural attributes associated with enhanced text and reasoning (Markowitz, 2024; Siddharthan, 2014; Mirzadeh et al., 2024; Chiang and Lee, 2024), we leverage a set of metrics designed to promote conciseness in generated reasoning.

Analogous to how certain qualities improve human thought organization, we examine how cultivating similar attributes can shape LLMs' reasoning. By explicitly incorporating these attributes, we aim not only to enhance textual clarity but also to influence the inferential process itself, supporting the broader hypothesis that well-structured writing attributes can modulate reasoning quality and dynamics. Each attribute is assigned a corresponding score, collectively defining a suite of interpretable metrics that guide model optimization:

- **Readability**: Evaluated via the Flesch Reading Ease formula (Kincaid et al., 1975), indicating the comprehensibility of the generated text.
- **Mathematical expression rate**: Measures the proportion of mathematical content in the output, which tends to correlate with more compact, precise symbolic notation that conveys computations and relationships efficiently.
- **Syntactic compression (average sentence length)**: Offers insights into the granularity of multi-step solutions, with shorter sentences often supporting clearer, stepwise reasoning.
- **Redundancy**: Computed using TF-IDF (Salton and Buckley, 1988), capturing the frequency of word and term repetition. While moderate redundancy facilitates flow in a solution, excessive repetition indicates inefficiency or verbosity.

Details on preprocessing and text standardization used before computing these metrics are in Appendix A.1.1; computation details for each metric are provided in Appendix A.1.2.

Our objective is to influence model behavior during problem-solving, guiding reasoning outputs toward being less redundant and more readable, mathematically dense, and syntactically concise.

3.3 *Attribute-Guided Prompting*

Using a pretrained LLM, we aim to elicit and promote these attributes in generated reasoning trajectories by explicitly instructing the model to *"Prioritize readability, syntactic compression, and mathematical notation while eliminating words and terms redundancy and verbose explanations."*

Crucially, this method aims not only to avoid incurring additional inference-time computational overhead (e.g., using multi-shot examples) but also to reduce the lengths of generated reasoning trajectories, leading to reduced computational cost in length-unrestrained models.

3.4 *CLARO*

We introduce *CLARO* to internalize these behaviors. Starting from a pretrained reasoning model, LLM_ϕ, capable of limiting output length (subject to deviations), we fine-tune it using RL with two complementary reward functions, R_C and R_L, assigned with equal probability to training samples.

Formally, given an input prompt x and an allocated token budget $\hat{L}$, the model generates a response y of length shorter than $\hat{L}$, while optimizing for a set of metrics $\mathcal{M}$ and answer correctness. Length limitation is achieved by embedding a directive within input x that specifies the desired output length threshold, enabling control over response verbosity. Specifically, consistent with established length-constraint protocols (Aggarwal and Welleck, 2025), we append the instruction *"Think for maximum $\hat{L}$ tokens."* to the end of the input prompt.

We define a set of attribute-motivated metrics, $\mathcal{M}$, derived from interpretable performance criteria that capture essential qualities of effective reasoning: readability, mathematical density, syntactic compression, and redundancy. To this end, R_C incorporates correctness with attribute alignment. To ensure holistic optimization, we incorporate a length contribution term (C_L) alongside the structural attribute metrics,

$$R_C = \mathbf{I}(a = a^*) \cdot (\alpha_L \cdot C_L + \sum_{j=1}^{K} \alpha_j \cdot MC_j), \quad (1)$$

where K denotes the metric count (four in the described setting), and $\mathbf{I}$ is a solution correctness indicator. Each α_j (and α_L) is a weight balancing an attribute contribution, while MC_j represents its normalized, clipped value,

$$MC_j = \text{clip}[\frac{1}{2}(\beta_j \cdot (m_j - \hat{m}_j) + \gamma), 0, 1]. \quad (2)$$

Here, β_j is a weighting coefficient, $\hat{m}_j$ is the target value of metric j, and γ is a constant that ensures partial reward for correct answers with suboptimal metric values. The same formulation applies to C_L. The clipping operation mitigates numerical instability during training. This reward incentivizes accuracy while implicitly favoring concise, attribute-aligned reasoning as a regularizer.

The second reward function, R_L, emphasizes strict adherence to the specified output length, aim-

Model	Prompting Strategy	Accuracy (%)									
		AMC	AIME 2024	AIME 2025	MATH-500	OlympiadBench	GPQA	LSAT	MMLU	**Avg**	**Avg Output Tokens**
Phi-4 (3.8B)	Standard	32.38	8.12	5.00	46.64	21.58	27.05	20.62	67.11	28.56	725.05
	Conciseness Only	28.99	7.50	5.83	45.39	20.54	25.73	20.11	67.22	27.66 ▼	663.75 ▼
	Attribute-Guided (ours)	32.00	8.96	5.21	48.25	21.65	25.95	21.41	68.66	**29.01** ▲	**641.64** ▼
Mathstral (7B)	Standard	15.66	1.67	0.42	36.89	11.94	21.37	15.27	48.63	18.98	631.94
	Conciseness Only	16.96	2.29	0.62	36.91	11.86	22.06	15.11	48.23	19.26 ▲	**585.31** ▼
	Attribute-Guided (ours)	18.98	1.46	0.83	36.50	11.92	22.73	15.52	47.25	**19.40** ▲	613.04 ▼
DeepSeek-R1-Distill-Llama (8B)	Standard	53.84	20.83	19.38	56.42	30.01	23.20	27.58	69.89	37.64	4483.25
	Conciseness Only	57.15	21.88	18.54	55.36	27.48	28.12	30.35	68.77	**38.46** ▲	4228.91 ▼
	Attribute-Guided (ours)	55.27	25.42	17.71	54.05	26.00	26.96	30.27	66.04	37.71 ▲	**4185.92** ▼

Table 1: Performance comparison of prompting strategies. While the conciseness-only baseline occasionally reduces tokens further, it may degrade accuracy. In contrast, the *Attribute-Guided Prompting* proposed here consistently improves upon the standard baseline across diverse benchmarks and offers the most robust trade-off, achieving high efficiency without the instability associated with forced brevity. See Table A.1 for a detailed token count breakdown.

ing to reinforce the model's capacity for length-sensitive reasoning,

$$R_L = \mathbf{I}(a = a^*) - \delta \cdot |L - \hat{L}|, \qquad (3)$$

where δ is a penalty coefficient, $\hat{L}$ is the allocated output length, and L is the output length of y.

In both reward functions (as defined in Equations 1 and 3), the reward is explicitly gated by the correctness indicator $\mathbf{I}(a = a^*)$. Therefore, if the model generates a response that achieves high structural metric alignment but is logically incorrect, the reward is effectively nullified. Consequently, the model cannot sacrifice logical accuracy for stylistic adherence; it must correctly solve the problem to accrue any structural benefit, thereby turning the proposed metrics into regularizers rather than primary objectives.

We optimize LLM_ϕ to produce LLM_C, a model yielding length-aware, attribute-driven reasoning. At inference-time, LLM_C requires specifying only the desired allocated output length as input.

This approach parallels the RLHF paradigm (Ouyang et al., 2022); however, instead of training and relying on a "black box" reward model, *CLARO* leverages a suite of deterministic, interpretable structural objectives. This eliminates the noise and instability often associated with training reward proxies, enabling direct, fine-grained control over reasoning properties without the need for costly human preference data.

4 Experiments

4.1 Experimental Setup

Datasets & Models. We conducted evaluations on mathematical reasoning benchmarks (AIME 2024 and 2025, MATH-500 (Lightman et al., 2023), AMC, and Olympiad-Bench (He et al., 2024)). To validate our method on non-mathematical problems and text, our evaluation also includes substantial non-mathematical, prose-heavy domains, specifically LSAT (Zhong et al., 2022) (logical reasoning and reading comprehension), MMLU (Hendrycks et al., 2020) (general knowledge, including humanities, social sciences, medicine, etc.), and GPQA (Rein et al., 2024) (biology, physics, chemistry, etc.), representing out-of-distribution (OOD) tasks distinct from the primary training distribution.

For fine-tuning, we use the DeepScaleR-Preview-Dataset (Luo et al., 2025b), a comprehensive mathematics collection containing 40,315 question–answer pairs, via a 30-step adaptation procedure instead of large-scale pretraining, resulting in the use of 3,840 randomly selected pairs. Our training setup relies exclusively on problem-answer pairs; no ground-truth reasoning traces or Chain-of-Thought examples were introduced during the fine-tuning process, as the model was provided only with the problem statement and its corresponding final answer (e.g., "42" or "B"). See Appendix A.2 for full details.

To evaluate our proposed prompting strategy, we employed state-of-the-art models, including Phi-4 (Abdin et al., 2024), DeepSeek-R1-Distill-Llama (Guo et al., 2025), and domain-specific Mathstral (Mistral AI, 2024).

As base models, we fine-tune L1-Max (1.5B and 7B) (Aggarwal and Welleck, 2025), which are length-controlled variants of DeepScaleR-Preview (Luo et al., 2025b) (DeepScaleR) and DeepSeek-R1-Distill-Qwen (Guo et al., 2025) (DeepSeek-R1), respectively. The resulting models are hereafter referred to as *CLARO* (1.5B and 7B), representing instances of LLM_C.

Baselines. Evaluation of the proposed fine-tuning method was assessed through comparative analysis against both length-unconstrained models (Deep-

ScaleR and DeepSeek-R1) and length-controlled models (s1 and L1-Max). Detailed baseline configurations are provided in Appendix A.2.

Evaluation. To assess our prompting strategy efficacy, we employed a three-way comparison to isolate specific structural benefits from generic conciseness. To this end, keeping the core reasoning instruction constant across all variants, we compared: (1) a standard baseline using a conventional zero-shot CoT trigger (e.g., *"Please reason step by step..."*); (2) a conciseness-focused baseline appending the instruction *"Prioritize conciseness."* to the CoT trigger, to disentangle the value of grounding conciseness in tangible structural properties versus simply requesting concise outputs; and (3) our *Attribute-Guided Prompting* strategy, described in Section 3.3. Full prompt details are provided in Appendix A.2.

To evaluate the efficacy of the proposed *CLARO* method, we report its performance across diverse established reasoning benchmarks and two model scales, systematically assessing results under varying computational budgets (maximal response lengths). Furthermore, we quantitatively assess the method's generalizability to OOD tasks and analyze the emergent reasoning behaviors. Additional experimental results, including prompting token usage breakdown (Table A.1), $\hat{m}_j$ sensitivity analysis (Figure A.2), and detailed ablation breakdowns (Figure A.3), are provided in the Appendices.

4.2 *Attribute-Guided Prompting*

We first investigate the efficacy of the proposed attributes on widely adopted pretrained models. To isolate the specific contribution of the proposed linguistic structural attributes, we compare our method not only against standard prompting but also against the generic conciseness-focused baseline defined in Section 4.1. This three-way comparison, detailed in Table 1, reveals that while generic directives for brevity can improve efficiency, they are often unstable and insufficient, whereas instructing models to adhere to structural attributes yields robust performance improvements.

As hypothesized, simply demanding conciseness can compromise reasoning depth. For example, with Phi-4, the conciseness-focused baseline reduces token count but degrades average accuracy compared to standard prompting. In contrast, our attribute-guided strategy is robust, not only preventing this degradation but consistently improving accuracy and reducing token consumption. This trend of superior efficiency holds across scales. For DeepSeek-R1-Distill-Llama, our method achieves the lowest token consumption. Similarly, on Mathstral, while the conciseness-focused baseline yields fewer tokens, our approach achieves the highest accuracy, offering the most favorable trade-off between cost and performance.

These results confirm that while conciseness is a desirable property for efficient inference, the mechanism by which it is achieved is paramount. LLMs possess latent capabilities for efficient reasoning that are often underutilized by standard prompts. By fostering our proposed attributes, we guide models to achieve conciseness as an emergent property of better structure rather than forced brevity, resulting in safer and more efficient reasoning.

4.3 *CLARO*

We fine-tune using the GRPO algorithm (Shao et al., 2024) with a learning rate of 4×10^{-6} and a batch size of 128, with 16 samples generated per input via the VeRL framework (Sheng et al., 2024). Training was conducted for 30 steps on four NVIDIA A100 80GB GPUs. Further implementation details, including the parameter values for equations 1, 2, and 3, are provided in Appendix A.1.3.

Figure 3 presents a comparative analysis of *CLARO* models' performance against baselines across different tasks. Given that other baselines consistently established a clear performance gap, s1 is evaluated on a representative subset of token budgets, without affecting comparative conclusions. Our findings demonstrate that the proposed fine-tuning robustly yields significant performance enhancements across different datasets, domains, and token budgets, emphasizing *CLARO*'s versatility. While the most substantial gains appear in math benchmarks, *CLARO* maintains robust performance and demonstrates strong generalization in non-mathematical, general reasoning tasks. This adaptability suggests that the internalized structural attributes, such as symbolic notation, generalize robustly across domains, without imposing rigid formats that would degrade general reasoning (as qualitatively illustrated in Figure 2).

In restricted token regimes, *CLARO* achieves substantial gains over the baselines across various benchmarks, as *CLARO* 1.5B and 7B models yield average relative improvements of up to 56.8% and 63.6% (4.7% and 6.9% absolute), respectively, under identical constraints. This sub-

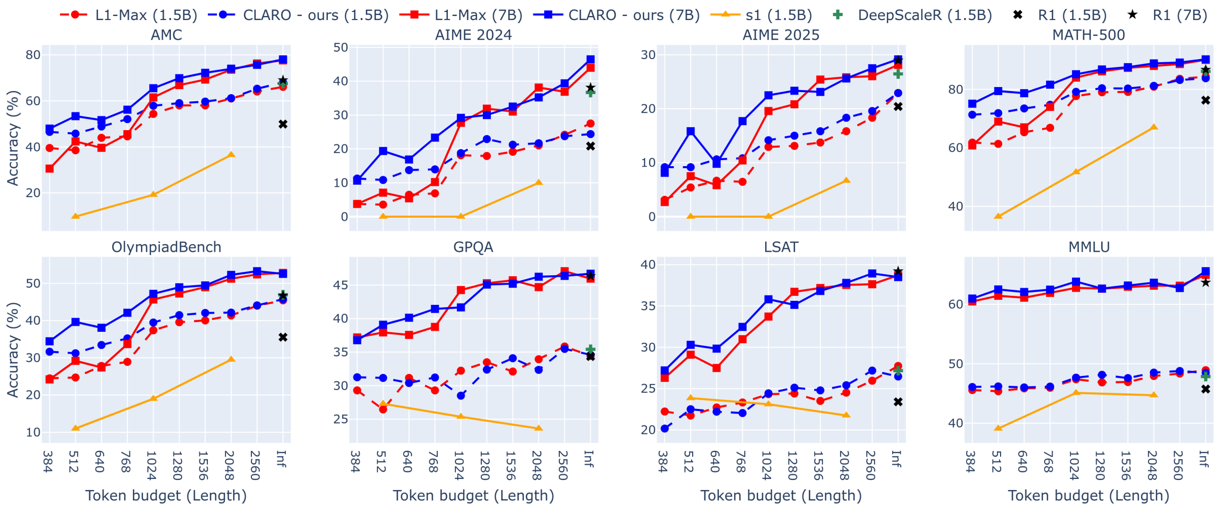

Figure 3: Accuracy vs. allocated length trade-offs across diverse domains. We compare *CLARO* (blue) against length-controlled and unconstrained models across eight distinct benchmarks. While baselines struggle under tight budgets (left side of plots), *CLARO* maintains robust accuracy, significantly outperforming the length-controlled baseline in low-resource regimes. Even as constraints relax (right side), *CLARO* matches or exceeds baselines performance, demonstrating that structural optimization yields superior effective capacity usage.

stantial gap stems from superior effective capacity usage; by fundamentally shifting reasoning dynamics, *CLARO* regulates the structural quality of the output, balancing depth with constraints to ensure that generated tokens are dedicated to inferential progress rather than expansive phrasing. These gains are evident across nearly the entire span of token budgets, slowly converging as constraints relax, while outperforming comparable baselines on average, indicating that improvements are not confined to constrained settings only. By incentivizing the model to embed high-quality structural attributes, our framework ensures that even limited output windows are filled with substantive CoT.

Crucially, these gains do not come at the expense of long-context performance. When given relaxed token budgets, a regime where models often approach budget saturation, *CLARO* maintains parity with or marginally outperforms baselines. Furthermore, as illustrated in Figure A.1, *CLARO* exhibits greater alignment with token budgets compared to both the s1 and L1-Max baselines, demonstrating enhanced control over the requested token budget. Importantly, this trend reinforces that structural optimization is most impactful where efficient reasoning is vital, ensuring that limited space is utilized for substantive inferential progress rather than superficial verbosity. Given this focus on efficiency under constraints, subsequent analyses will center on token-budgeted regimes.

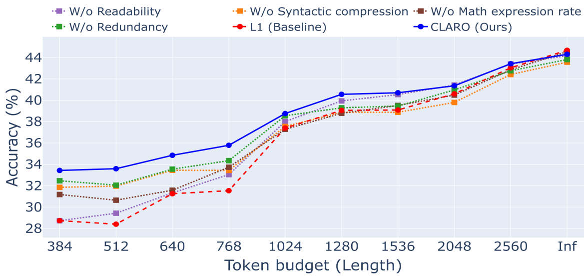

Figure 4: Disentangling the impact of the linguistic attributes incorporated in *CLARO*. We report the accuracy across budgets averaged over all evaluated benchmarks. The *CLARO* model (blue) outperforms all leave-one-out variants, highlighting the synergistic value of the full attribute set. Notably, each variant still outperforms the baseline (red), demonstrating that structural guidance, even when partial, improves reasoning efficacy.

4.4 Attribute Contribution and Analysis

All structure attributes contribute. We conducted a comprehensive analysis to assess and disentangle the individual impact of each of the targeted attribute components $\mathcal{M}$. Figure 4 illustrates different attribute-excluded variants' performance via controlled leave-one-out experiments on *CLARO* (1.5B), where specific attribute coefficients (α_j) in R_C were independently set to 0. A detailed breakdown and analysis of dataset-specific performances is provided in Figure A.3.

The results indicate that while the complete *CLARO* framework, with all attributes optimized synergistically, consistently outperforms all abla-

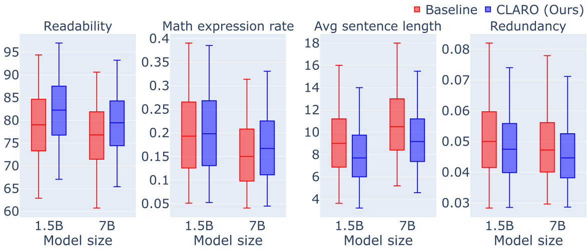

Figure 5: Behavioral shift in structural attributes before and after fine-tuning. We compare four key metric distributions—Readability (higher is better), Math density (higher is better), Average sentence length (lower is better), and Redundancy (lower is better)—between the baseline L1-Max (red) and *CLARO* (blue). *CLARO* effectively shifts model outputs toward target objectives. Crucially, this alignment correlates with the accuracy gains shown in Figure 3, validating our hypothesis that clearer reasoning structure enhances performance.

tion variants, ablated variants still surpass the L1-Max baseline. This suggests that while the combination of all attributes yields the optimal reasoning structure, the imposition of even a subset of these attributes provides tangible benefits. The different attributes exhibit synergistic rather than merely additive effects, suggesting inherent interactions. For instance, readability may sometimes conflict with direct metrics of sentence length; however, these tensions are effectively balanced and mitigated by the relative weighting of the metric components within R_C, highlighting the complementary nature in shaping reasoning behavior. Qualitative examples of reasoning trajectories reflecting these shifts are provided in Figures A.4 and A.5.

Fine-tuning improves language structure and reasoning. To better understand the effect of *CLARO* on generated reasoning properties, we further examined how target structural attributes, quantified by $\mathcal{M}$, shift after fine-tuning. Specifically, we measured their metrics across benchmark outputs derived from the textual generations in Figure 3, spanning the allocated token budgets. The results, illustrated in Figure 5, showcase improvements in targeted attributes across both model scales.

Correct answers have improved structure. We investigate the relation between structural attribute alignment and solution accuracy by analyzing the distribution of structural metric scores for correct versus incorrect solutions across four evaluated models: L1-Max (1.5B and 7B) and *CLARO* (1.5B and 7B). Our analysis results, illustrated in Figure 6, reveal that successful reasoning trajectories consistently exhibit stronger convergence toward our defined structural targets ($\hat{m}_j$) compared to incorrect ones. This suggests that the proposed attributes are not merely stylistic preferences but are intrinsic markers of successful reasoning chains.

Specifically, we observe that correct solutions are statistically more likely to exhibit higher readability (m_1), employing standard English structures rather than fragmented or convoluted phrasing. Additionally, they demonstrate a higher mathematical expression rate (m_2), leveraging symbolic notation to represent logic compactly, whereas incorrect traces often revert to verbose natural language descriptions of arithmetic or calculations. Furthermore, correct outputs display greater syntactic compression (m_3), characterized by shorter, more granular sentences that indicate distinct logical steps.

Interestingly, the Redundancy metric (m_4) exhibits a mixed trend, where correct solutions can display slightly higher redundancy scores compared to incorrect ones. This correlation may stem from two primary factors. First, incorrect solutions often suffer from semantic drift, oscillating between unrelated concepts or hallucinations, which artificially deflates similarity scores. Conversely, correct solutions tend to remain focused on specific variables and constraints, naturally increasing their usage and repetition. Second, the TF-IDF metric is sensitive to text length. Since correct solutions are typically more concise, the recurrence of necessary problem terms has a high impact on similarity scores compared to verbose incorrect traces.

Finally, attribute target values $\hat{m}_j$ sensitivity ablation analyses (see Figure A.2) confirm *CLARO*'s robustness, showing stable performance while modifying metric target values.

Conclusion. Taken together, these findings demonstrate that promoting the proposed structural and linguistic attributes, often associated with human communicative efficiency, serves as an effective proxy for improving machine reasoning.

5 Discussion and Conclusion

In this work, we demonstrate that leveraging linguistic structural attributes associated with efficient human communication enhances the efficacy of Chain-of-Thought reasoning. Our main contributions are twofold. First, we establish via prompting widely adopted pretrained models that structural directives, such as increasing mathematical density and syntactic compression, can promote efficient reasoning capabilities. This straightforward, low-

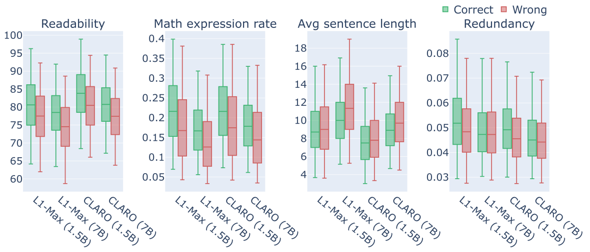

Figure 6: Metric distributions stratified by correctness. Correct solutions (green) show distinct structural properties compared to wrong ones (red), with correct answers consistently exhibiting higher readability, greater mathematical density, and tighter syntactic compression.

overhead *Attribute-Guided Prompting* strategy offers immediate gains and validates the efficacy of our proposed structural constraints. Second, we introduce **C**ontro**l**led **A**ttribute-Driven **R**easoning **O**ptimization (*CLARO*), an RL framework that internalizes these behaviors within models, promoting more concise and constructive CoT reasoning.

CLARO optimizes models not only for brevity, with enhanced control over response length limits, but also for substantive content via increased readability, mathematical density, syntactic simplicity, and lower redundancy. This dual focus distinguishes our approach from prior methods that largely regulate or constrain reasoning length without addressing the quality of the content itself.

We validate the robustness of our framework across distinct model scales, domains, and ablations. Our fine-tuned models and prompting strategy surpass comparable baselines across diverse benchmarks and token budgets, establishing that structural control is model-agnostic and effective under both length-constrained and unconstrained settings, thereby helping reduce inference-time costs while maintaining or even improving accuracy. By integrating linguistic attribute conditioning into the learning process, *CLARO* balances accuracy and conciseness, demonstrating that optimizing such attributes does not artificially shorten responses but actively refines reasoning dynamics to produce more effective trajectories. This capability is critical for real-world deployment where computational cost and latency are tightly bound.

Experimental results confirm that *CLARO* yields robust correctness under length constraints compared to prior state-of-the-art techniques. While these gains are most pronounced in restrictive regimes where efficiency is paramount, the performance advantage is sustained under generous constraints and in unconstrained settings, underscoring *CLARO*'s versatility across both tight and relaxed token budgets. Notably, widely adopted pretrained models exhibit efficiency and performance gains at their preferred lengths when prompted with structural attributes via *Attribute-Guided Prompting*.

Beyond performance gains, *CLARO* illuminates the underlying attributes contributing to effective machine reasoning. By leveraging interpretable metrics, our framework highlights mechanisms that govern generated output quality and efficiency. Importantly, *CLARO* circumvents the risk of Goodhart's Law, where optimization for a proxy metric undermines the original goal, by conditioning structural rewards on answer correctness via the correctness indicator introduced in Equations 1 and 3, thereby fostering substantive logic. This approach eliminates rewards for erroneous answers and effectively prevents metric gaming. Empirical results demonstrate a strong alignment between correct solutions and the desired linguistic properties using *CLARO*, confirming that the targeted attributes serve as functional mechanisms of clear expression rather than mere proxies, with accuracy consistently improving alongside structural adherence.

Ultimately, our findings suggest a fundamental decoupling of reasoning depth from verbosity. Rather than compelling models to "think less" by truncating trajectories, or "think more" by increasing compute, *CLARO* incentivizes them to "think more clearly," compressing effective logical trajectories into concise formats that preserve inferential fidelity. Crucially, our analysis underscores that conciseness must be understood as a property of both form and content; optimizing one without the other risks either verbosity or loss of substantive reasoning. Moreover, by grounding the framework in interpretable linguistic metrics, *CLARO* offers a transparent template for reasoning calibration that bypasses the need for opaque reward models. Thus, this study also provides a granular toolset for reasoning analysis and controllability, offering a practical pathway toward balancing inference cost and problem-solving performance, making CoT reasoning more economically and operationally viable.

Acknowledgments This work was partially supported by ONR, Simons Foundation, NSF, and gifts from Apple and Google. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the funding and gifting agencies and companies.

6 Limitations

While *CLARO* demonstrates significant improvements in reasoning efficiency and accuracy, several limitations warrant consideration.

First, our metric definitions are optimized for English. Extending *CLARO* to other languages would necessitate redefining these metrics to align with different linguistic typologies. However, while the specific readability formulas are language-dependent, the *CLARO* framework, either using *Attribute-Guided Prompting* or Attribute-Driven Reinforcement Learning, is language-agnostic.

Second, the framework's efficacy is partly contingent on the model's adherence to specific instructions and length constraints. While *CLARO* fine-tuning significantly improves budget compliance, models with weaker instruction-following tendencies may exhibit occasional deviations, potentially leading to truncated reasoning or budget violations. Similarly, our zero-shot *Attribute-Guided Prompting* strategy also relies on the model's instruction-following capabilities.

Finally, while metric target values calibrated on a 1.5B model transferred effectively to a 7B model (representing what is available for this type of work in academic settings), this robustness may not extend indefinitely. Deploying *CLARO* on significantly larger architectures or distinct model families may require recalibrating attribute targets ($\hat{m}_j$) or reward weights to accommodate the different baseline distributions.

References

Marah Abdin, Jyoti Aneja, Harkirat Behl, Sébastien Bubeck, Ronen Eldan, Suriya Gunasekar, Michael Harrison, Russell J Hewett, Mojan Javaheripi, Piero Kauffmann, and 1 others. 2024. Phi-4 technical report. *arXiv preprint arXiv:2412.08905*.

Pranjal Aggarwal and Sean Welleck. 2025. L1: Controlling how long a reasoning model thinks with reinforcement learning. *arXiv preprint arXiv:2503.04697*.

Daman Arora and Andrea Zanette. 2025. Training language models to reason efficiently. *arXiv preprint arXiv:2502.04463*.

Zhenni Bi, Kai Han, Chuanjian Liu, Yehui Tang, and Yunhe Wang. 2024. Forest-of-thought: Scaling test-time compute for enhancing LLM reasoning. *arXiv preprint arXiv:2412.09078*.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, and 1 others. 2020. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33:1877–1901.

Sondos Mahmoud Bsharat, Aidar Myrzakhan, and Zhiqiang Shen. 2023. Principled instructions are all you need for questioning Llama-1/2, GPT-3.5/4. *arXiv preprint arXiv:2312.16171*.

Wenhu Chen, Xueguang Ma, Xinyi Wang, and William W Cohen. 2022. Program of thoughts prompting: Disentangling computation from reasoning for numerical reasoning tasks. *arXiv preprint arXiv:2211.12588*.

Xingyu Chen, Jiahao Xu, Tian Liang, Zhiwei He, Jianhui Pang, Dian Yu, Linfeng Song, Qiuzhi Liu, Mengfei Zhou, Zhuosheng Zhang, and 1 others. 2024. Do not think that much for 2+3=? on the overthinking of o1-like LLMs. *arXiv preprint arXiv:2412.21187*.

Cheng-Han Chiang and Hung-yi Lee. 2024. Over-reasoning and redundant calculation of large language models. *arXiv preprint arXiv:2401.11467*.

Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, and 1 others. 2021. Training verifiers to solve math word problems. *arXiv preprint arXiv:2110.14168*.

Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Alex Vaughan, and 1 others. 2024. The Llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, and 1 others. 2025. Deepseek-r1: Incentivizing reasoning capability in LLMs via reinforcement learning. *arXiv preprint arXiv:2501.12948*.

Michael Hassid, Gabriel Synnaeve, Yossi Adi, and Roy Schwartz. 2025. Don't overthink it. Preferring shorter thinking chains for improved LLM reasoning. *arXiv preprint arXiv:2505.17813*.

Chaoqun He, Renjie Luo, Yuzhuo Bai, Shengding Hu, Zhen Leng Thai, Junhao Shen, Jinyi Hu, Xu Han, Yujie Huang, Yuxiang Zhang, and 1 others. 2024. Olympiadbench: A challenging benchmark for promoting agi with olympiad-level bilingual multimodal scientific problems. *arXiv preprint arXiv:2402.14008*.

Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. 2020. Measuring massive multitask language understanding. *arXiv preprint arXiv:2009.03300*.

Bairu Hou, Yang Zhang, Jiabao Ji, Yujian Liu, Kaizhi Qian, Jacob Andreas, and Shiyu Chang. 2025. Thinkprune: Pruning long Chain-of-Thought of LLMs via reinforcement learning. *arXiv preprint arXiv:2504.01296*.

William Jurayj, Jeffrey Cheng, and Benjamin Van Durme. 2025. Is that your final answer? Test-time scaling improves selective question answering. *arXiv preprint arXiv:2502.13962*.

Yu Kang, Xianghui Sun, Liangyu Chen, and Wei Zou. 2025. C3ot: Generating shorter Chain-of-Thought without compromising effectiveness. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 24312–24320.

J Peter Kincaid, Robert P Fishburne Jr, Richard L Rogers, and Brad S Chissom. 1975. Derivation of new readability formulas (automated readability index, fog count and Flesch reading ease formula) for navy enlisted personnel.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2022. Large language models are zero-shot reasoners. *Advances in Neural Information Processing Systems*, 35:22199–22213.

Ben Krause, Akhilesh Deepak Gotmare, Bryan McCann, Nitish Shirish Keskar, Shafiq Joty, Richard Socher, and Nazneen Fatema Rajani. 2020. Gedi: Generative discriminator guided sequence generation. *arXiv preprint arXiv:2009.06367*.

Aitor Lewkowycz, Anders Andreassen, David Dohan, Ethan Dyer, Henryk Michalewski, Vinay Ramasesh, Ambrose Slone, Cem Anil, Imanol Schlag, Theo Gutman-Solo, and 1 others. 2022. Solving quantitative reasoning problems with language models. *Advances in Neural Information Processing Systems*, 35:3843–3857.

Hunter Lightman, Vineet Kosaraju, Yuri Burda, Harrison Edwards, Bowen Baker, Teddy Lee, Jan Leike, John Schulman, Ilya Sutskever, and Karl Cobbe. 2023. Let's verify step by step. In *The Twelfth International Conference on Learning Representations*.

Alisa Liu, Maarten Sap, Ximing Lu, Swabha Swayamdipta, Chandra Bhagavatula, Noah A Smith, and Yejin Choi. 2021. Dexperts: Decoding-time controlled text generation with experts and anti-experts. *arXiv preprint arXiv:2105.03023*.

Ximing Lu, Sean Welleck, Jack Hessel, Liwei Jiang, Lianhui Qin, Peter West, Prithviraj Ammanabrolu, and Yejin Choi. 2022. Quark: Controllable text generation with reinforced unlearning. *Advances in Neural Information Processing Systems*, 35:27591–27609.

Haotian Luo, Li Shen, Haiying He, Yibo Wang, Shiwei Liu, Wei Li, Naiqiang Tan, Xiaochun Cao, and Dacheng Tao. 2025a. O1-pruner: Length-harmonizing fine-tuning for o1-like reasoning pruning. *arXiv preprint arXiv:2501.12570*.

Michael Luo, Sijun Tan, Justin Wong, Xiaoxiang Shi, William Y Tang, Manan Roongta, Colin Cai, Jeffrey Luo, Tianjun Zhang, Li Erran Li, and 1 others. 2025b. Deepscaler: Surpassing o1-preview with a 1.5 b model by scaling rl. *Notion Blog*.

Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegreffe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, Yiming Yang, and 1 others. 2023. Self-refine: Iterative refinement with self-feedback. *Advances in Neural Information Processing Systems*, 36:46534–46594.

David M Markowitz. 2024. From complexity to clarity: How ai enhances perceptions of scientists and the public's understanding of science. *PNAS Nexus*, 3(9):pgae387.

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

Mistral AI. 2024. Mathstral. Accessed: 2025-12-04.

Niklas Muennighoff, Zitong Yang, Weijia Shi, Xiang Lisa Li, Li Fei-Fei, Hannaneh Hajishirzi, Luke Zettlemoyer, Percy Liang, Emmanuel Candès, and Tatsunori Hashimoto. 2025. s1: Simple test-time scaling. *arXiv preprint arXiv:2501.19393*.

Sania Nayab, Giulio Rossolini, Marco Simoni, Andrea Saracino, Giorgio Buttazzo, Nicolamaria Manes, and Fabrizio Giacomelli. 2024. Concise thoughts: Impact of output length on LLM reasoning and cost. *arXiv preprint arXiv:2407.19825*.

OpenAI. 2024. Learning to Reason with LLMs. Accessed: 2025-12-04.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, and 1 others. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744.

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.

David Rein, Betty Li Hou, Asa Cooper Stickland, Jackson Petty, Richard Yuanzhe Pang, Julien Dirani, Julian Michael, and Samuel R Bowman. 2024. Gpqa: A graduate-level Google-proof Q&A benchmark. In *First Conference on Language Modeling*.

Gerard Salton and Christopher Buckley. 1988. Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5):513–523.

Timo Schick, Sahana Udupa, and Hinrich Schütze. 2021. Self-diagnosis and self-debiasing: A proposal for reducing corpus-based bias in NLP. *Transactions of the Association for Computational Linguistics*, 9:1408–1424.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Yang Wu, and 1 others. 2024. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*.

Guangming Sheng, Chi Zhang, Zilingfeng Ye, Xibin Wu, Wang Zhang, Ru Zhang, Yanghua Peng, Haibin Lin, and Chuan Wu. 2024. Hybridflow: A flexible and efficient RLHF framework. *arXiv preprint arXiv: 2409.19256*.

Advaith Siddharthan. 2014. A survey of research on text simplification. *ITL-International Journal of Applied Linguistics*, 165(2):259–298.

Charlie Snell, Jaehoon Lee, Kelvin Xu, and Aviral Kumar. 2024. Scaling LLM test-time compute optimally can be more effective than scaling model parameters. *arXiv preprint arXiv:2408.03314*.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, and 1 others. 2022. Chain-of-Thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35:24824–24837.

Jason Weston and Sainbayar Sukhbaatar. 2023. System 2 attention (is something you might need too). *arXiv preprint arXiv:2311.11829*.

Heming Xia, Chak Tou Leong, Wenjie Wang, Yongqi Li, and Wenjie Li. 2025. Tokenskip: Controllable Chain-of-Thought compression in LLMs. *arXiv preprint arXiv:2502.12067*.

Silei Xu, Wenhao Xie, Lingxiao Zhao, and Pengcheng He. 2025. Chain of draft: Thinking faster by writing less. *arXiv preprint arXiv:2502.18600*.

An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li, Chengyuan Li, Dayiheng Liu, Fei Huang, Guanting Dong, Haoran Wei, Huan Lin, Jialong Tang, Jialin Wang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Ma, and 40 others. 2024. Qwen2 technical report. *arXiv preprint arXiv:2407.10671*.

Wenkai Yang, Shuming Ma, Yankai Lin, and Furu Wei. 2025. Towards thinking-optimal scaling of test-time compute for LLM reasoning. *arXiv preprint arXiv:2502.18080*.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Tom Griffiths, Yuan Cao, and Karthik Narasimhan. 2023. Tree of thoughts: Deliberate problem solving with large language models. *Advances in Neural Information Processing Systems*, 36:11809–11822.

Weizhe Yuan, Ilia Kulikov, Ping Yu, Kyunghyun Cho, Sainbayar Sukhbaatar, Jason Weston, and Jing Xu. 2024. Following length constraints in instructions. *arXiv preprint arXiv:2406.17744*.

Eric Zelikman, Yuhuai Wu, Jesse Mu, and Noah Goodman. 2022. Star: Bootstrapping reasoning with reasoning. *Advances in Neural Information Processing Systems*, 35:15476–15488.

Wanjun Zhong, Siyuan Wang, Duyu Tang, Zenan Xu, Daya Guo, Yining Chen, Jiahai Wang, Jian Yin, Ming Zhou, and Nan Duan. 2022. Analytical reasoning of text. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 2306–2319.

Denny Zhou, Nathanael Schärli, Le Hou, Jason Wei, Nathan Scales, Xuezhi Wang, Dale Schuurmans, Claire Cui, Olivier Bousquet, Quoc Le, and 1 others. 2022. Least-to-most prompting enables complex reasoning in large language models. *arXiv preprint arXiv:2205.10625*.

A Appendix

A.1 Implementation Details

A.1.1 Text Preprocessing

To quantitatively assess distinct linguistic and structural attributes of generated outputs, we employ a multifaceted approach with specific computations for each metric. The detailed procedures are elaborated as follows.

Let the generated output be denoted as a text sequence,

$$y = (t_1, t_2, \ldots, t_T),$$

where $|y| = T$ is the total number of tokens. Let $\mathcal{S} = \{s_1, s_2, \ldots, s_{|\mathcal{S}|}\}$ denote the set of sentences extracted from y.

The generated output is preprocessed to standardize the text and facilitate subsequent analysis. This involves the removal of redundant whitespace and the identification of mathematical expressions and symbolic notation within the text.

A.1.2 Metric Formulations and Computation

The length metric is defined as the total number of tokens contained in the generated output sequence; hence,

$$L = |y|. \tag{A.1}$$

The readability of the textual content was quantitatively assessed utilizing the Flesch Reading Ease formula (Kincaid et al., 1975), as defined in Equation A.2. To maintain the integrity of the analysis and mitigate potential confounds, mathematical expressions were excluded from this process. This exclusion stems from the fundamental mismatch between the design of the Flesch Reading Ease formula for textual sentences and the inherent structural and semantic properties of mathematical notation, which, if included, could potentially introduce artefactual errors and compromise the validity of the readability score. It is formally defined:

$$m_1 = 206.835 \;-\; 1.015 \cdot \frac{W_{\text{text}}}{S_{\text{text}}} \;-\; 84.6 \cdot \frac{Sy_{\text{text}}}{W_{\text{text}}}, \tag{A.2}$$

where W_{text} is the total number of tokens outside of math expressions, S_{text} is the total number of sentences containing at least one non-math token, and Sy_{text} is the total number of syllables in W_{text}. A higher m_1 indicates greater comprehensibility.

The mathematical expression rate metric is determined by computing the ratio of the cumulative length of all identified mathematical expressions, J, to the total text length.

$$m_2 = \frac{\sum_{i=1}^{J} \text{len}(e_i)}{|y|}, \tag{A.3}$$

where $\text{len}(e_i)$ is the token length of math expression e_i. A higher m_2 means a more compact logical representation via math notation.

For the syntactic compression metric, sentences containing mathematical expressions are omitted from this analysis as their structure often deviates significantly from that of natural language. The average length of the remaining sentences, which are representative of standard textual prose, is then calculated.

$$m_3 = \frac{1}{|\mathcal{S}'|} \sum_{s \in \mathcal{S}'} \text{len}(s), \tag{A.4}$$

where $\mathcal{S}' \subseteq \mathcal{S}$ is the subset of sentences containing no math expressions, and $\text{len}(s)$ counts the tokens in sentence s. A smaller m_3 indicates more syntactic compression, i.e., shorter sentences.

Finally, a vectorized representation of the text sentences is generated for the redundancy metric. To this end, each sentence is considered an individual document in the output corpus, and TF-IDF vectorization is applied to these sentences. Given a text comprising $|\mathcal{S}|$ sentences, this process yields $|\mathcal{S}|$ corresponding vectors. A pairwise similarity matrix of dimensions $|\mathcal{S}| \times |\mathcal{S}|$ is then constructed using the cosine similarity metric between these vectors. To avoid duplicates and self-comparison, the redundancy score is then defined as the mean similarity over all unique sentence pairs,

$$m_4 = \frac{2}{|\mathcal{S}|(|\mathcal{S}| - 1)} \sum_{1 \le i < j \le |\mathcal{S}|} \text{sim}(i, j). \tag{A.5}$$

A higher m_4 value indicates more repetition (less efficient expression).

A.1.3 Reward Parameter Configuration

The reward weights in Equation 1, corresponding to Readability ($j = 1$), Mathematical expression rate ($j = 2$), Syntactic compression ($j = 3$), and Redundancy ($j = 4$), are set as follows: $\alpha_1 = 0.2$, $\alpha_2 = \alpha_3 = \alpha_4 = 0.1$, and $\alpha_L = 0.5$, with Equation 2 target values $\hat{m}_1 = 81$, $\hat{m}_2 = 0.275$, $\hat{m}_3 = 7$, and $\hat{m}_4 = 0.04$, and a constant $\gamma = 1$. The allocated length, $\hat{L}$, is sampled from a uniform distribution $U(100, 4000)$. The scalar coefficient in Equation 3 is set to be $\delta = \frac{1}{3000}$.

Model	Prompting Strategy	Avg Output Tokens							
		AMC	AIME 2024	AIME 2025	MATH-500	OlympiadBench	GPQA	LSAT	MMLU
Phi-4 (3.8B)	Standard	836.28	1048.71	904.90	531.56	837.78	708.52	628.20	304.46
	Conciseness Only	748.67	1069.76	857.11	474.37	781.54	619.85	529.98	228.69
	Attribute-Guided (ours)	762.12	939.39	868.24	476.06	755.51	599.70	520.64	211.47
Mathstral (7B)	Standard	663.98	910.89	862.42	518.89	793.25	557.01	504.40	244.72
	Conciseness Only	679.96	808.45	804.72	473.20	729.61	493.41	477.28	215.88
	Attribute-Guided (ours)	663.24	886.57	803.79	501.58	743.31	566.00	496.00	243.83
DeepSeek-R1-Distill-Llama (8B)	Standard	4524.00	5656.94	5527.20	3147.80	4546.34	5279.65	5526.50	1657.57
	Conciseness Only	4293.49	5548.16	5440.71	2664.61	4346.34	4997.98	5305.07	1234.93
	Attribute-Guided (ours)	4340.06	5574.05	5489.81	2567.76	4351.40	4897.89	5215.92	1050.44

Table A.1: Token usage breakdown by benchmark. Average output token counts are detailed across individual datasets. The data confirms that *Attribute-Guided Prompting* yields robust efficiency gains and token reductions across diverse reasoning tasks compared to standard prompting or focusing solely on conciseness.

Extensive hyperparameter tuning was not performed in this study. Crucially, the limited calibration of metric values was conducted exclusively on the 1.5B model and subsequently applied directly to the 7B model without modification. This successful transfer across model scales attests to the robustness and generalizability of the proposed structural attributes, though it suggests that additional performance improvements may still be possible with model-specific or further optimization.

It is worth noting that the target values $\hat{m}_j$ do not strictly enforce a reduction or increase in the relevant metric for 100% of instances. Rather, they are intended to establish a directional trend, given that baseline models typically exhibit metric distributions opposing these targets (e.g., naturally displaying lower readability levels than desired). Consequently, models are not explicitly optimized to match each target metric value.

In Equation 2, we set $\beta_1 = 1/15$, $\beta_2 = 5$, $\beta_3 = -0.1$, $\beta_4 = -15$, and $\beta_L = -2 \times 10^{-3}$. Note that several β_j values are negative, as these coefficients determine how each attribute is adjusted relative to its target value. Negative values promote the model to decrease attributes that are above their desired targets (e.g., encouraging shorter sentences), while positive values foster increasing attributes that are below their targets (e.g., enhancing readability).

A.2 Expanded Experimental Setup Details

Datasets & Models. For the evaluated benchmarks, we specifically utilized the AMC23 split for AMC, the Diamond split for GPQA, and the AR split for LSAT. For MMLU, we randomly sampled 1,000 questions across all its subsets.

DeepScaleR (Luo et al., 2025b) is originally RL fine-tuned from DeepSeek-R1-Distill-Qwen (1.5B) (Guo et al., 2025). DeepSeek-R1-Distill-Llama is a distilled model based on Llama-3.1-8B (Grattafiori et al., 2024), produced via supervised fine-tuning on the reasoning traces generated by the DeepSeek-R1 model (Guo et al., 2025).

Baselines. Evaluation of the proposed fine-tuning method was assessed through comparative analysis against several state-of-the-art baseline models with a comparable number of parameters. These baselines span the spectrum from unconstrained reasoning solutions to length-aware ones, allowing for a rigorous evaluation of efficiency. This selection facilitates a robust and directly comparable assessment of efficient reasoning capabilities against both length-unconstrained and competitive length-controlled approaches. The considered baselines are:

1. DeepScaleR-Preview (1.5B) is a fine-tuned variant of Deepseek-R1-Distilled-Qwen (1.5B) (Guo et al., 2025), produced using RL.
2. DeepSeek-R1-Distill-Qwen (1.5B and 7B) are distilled models based on Qwen2.5-Math (Yang et al., 2024), produced via supervised fine-tuning on the reasoning traces generated by the DeepSeek-R1 model.
3. s1 (Muennighoff et al., 2025) (1.5B) incorporates a budget-forcing mechanism that regulates maximal reasoning length using inference-time instructions and end-of-thinking tokens injection during inferential processes.
4. L1-Max (Aggarwal and Welleck, 2025) employs a direct input prompt instruction mechanism to control the length of the generated reasoning brevity by incentivizing outputs shorter than an allocated token budget.

Evaluation. Overall, we adopt the evaluation protocols and settings used in (Aggarwal and Welleck, 2025) to ensure a consistent comparison with the

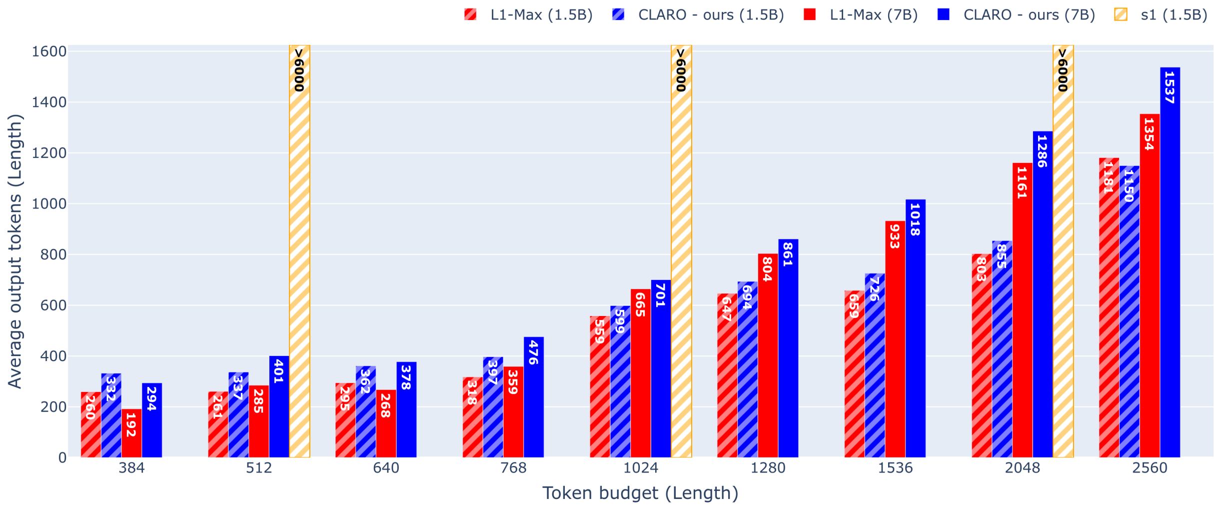

Figure A.1: Bar charts illustrating average output lengths versus allocated token budgets across varying constraints. While *CLARO* exhibits slightly higher mean lengths than the L1-Max baseline, it consistently remains shorter than the allocated budget. This demonstrates enhanced control over the generated output, as *CLARO* better utilizes the allocated computational allowance to support robust reasoning, avoiding the tendency of baselines to under-utilize available capacity. Conversely, s1 fails to enforce constraints effectively, resulting in large deviations and excessive output generation.

baselines. To ensure robust evaluation across different model architectures, standard zero-shot Chain-of-Thought triggers were utilized.

For the *Attribute-Guided Prompting* experiments (see Section 4.2), we utilized the instruction: *"Please reason step by step, and put your final answer within \boxed{}."* Our primary objective is to isolate the independent contributions of the proposed structural attributes under identical conditions. By holding the reasoning instruction constant and exclusively modulating the directives governing output structure, our findings (Table 1) demonstrate that explicitly eliciting these structural properties, rather than enforcing generic brevity, improves downstream reasoning performance.

For fine-tuning and evaluation of *CLARO* and L1 models (see Section 4.3), we used *"Let's think step by step and output the final answer within \boxed{}."* to maintain consistency with the base model's training distribution. For a rigorous comparison, we evaluated models (*CLARO* and baselines) across varying lengths (including unconstrained ones) using a unified protocol, with the specific exception of s1. Given that s1 inherently requires a specialized inference setup, we evaluated it using its original pipeline and associated prompting strategy. However, we ensured that the assessment criteria remained consistent to preserve the validity of the comparative analysis.

For multiple-choice datasets (GPQA, LSAT, and MMLU), the answer directive was modified to elicit the specific multiple-choice label (e.g., 'B') rather than the answer value (e.g., '5'), in order to facilitate automated parsing. To this end, instead of *"...final answer within \boxed{}."*, we ask for *"...final answer (eg, A, B, C, D) within \boxed{}."*.

Finally, we report results based on single experimental runs; however, the consistency of results observed across the extensive array of models, scales, datasets, lengths, and ablation studies provides sufficient evidence for the robustness of our claims.

A.3 Token Usage Analysis

Table A.1 presents granular average output token counts for each prompting strategy across all evaluated benchmarks. This data complements the aggregated averages reported in Table 1, illustrating how *Attribute-Guided Prompting* optimizes token usage across different reasoning tasks.

Figure A.1 illustrates the average response lengths relative to the allocated token budgets, comparing *CLARO* against the length-constrained baselines s1 and L1-Max. The data complements the findings in Section 4.3, revealing that *CLARO* consistently generates outputs shorter on average than the allocated token budget. Although *CLARO* tends to utilize a larger portion of this budget compared to L1-Max, this increased length should not be conflated with the sole driver of performance. Notably, within the L1 model family, L1-Exact (Aggarwal and Welleck, 2025)—a similar variant targeting specific lengths rather than lower bounded by them—produces longer outputs than L1-Max yet yields inferior results, demonstrating that token count does not monotonically correlate with performance. Thus, *CLARO*'s advantage lies not in raw length, but in its effective budget utilization, allowing it to leverage the allocated budget to complete substantive, attribute-aligned reasoning chains. Overall, this demonstrates that our method successfully aligns the model's output with the intended scope of the constraint, allowing for deeper inferential quality while adhering to the defined upper bounds.

The s1 model frequently exhibits divergent behavior and inconsistent length control. Beyond repeating generation loops where identical sentences are reiterated without convergence to a solution, leading to prolonged, out-of-budget outputs, reasoning trajectories are also often truncated by the insertion of the end-of-thinking token, risking the omission of crucial steps required for reaching a solution. Given these pronounced deviations and significant budget violations, s1 is presented for representative budgets only, without affecting comparative conclusions.

A.4 Sensitivity Analysis

To better understand the effects of different target values of each attribute component and assess the stability of the proposed framework, we conducted an ablation study by varying the target parameters $\hat{m}_j$ for each metric. To this end, we adjusted these specific values to either strictly constrain or loosely relax the structural attributes, analyzing the resulting impact on reasoning accuracy.

Specifically, relative to the calibrated defaults utilized in our main experiments (detailed in Appendix A.1.3), we independently evaluated the following alternative target values, while maintaining all other variables at their default values: for Readability, we tested $\hat{m}_1 = 78$ and $\hat{m}_1 = 84$; for Mathematical expression rate, $\hat{m}_2 = 0.25$ and $\hat{m}_2 = 0.30$; for Syntactic compression, $\hat{m}_3 = 6$ and $\hat{m}_3 = 8$; and for Redundancy, $\hat{m}_4 = 0.035$ and $\hat{m}_4 = 0.045$.

As illustrated in Figure A.2, while our specific calibration generally yields the peak performance, *CLARO* demonstrates remarkable robustness to hyperparameter deviations, consistently maintaining high accuracy even with uncalibrated values. Across all evaluated attributes, deviations from the optimal targets resulted in only marginal performance fluctuations rather than catastrophic drops. Consequently, *CLARO* can be effectively deployed without exhaustive calibration.

A.5 Detailed Ablation Results

We provide a comprehensive breakdown of the ablation study results to substantiate the effectiveness of the proposed attributes. Figure A.3 presents the leave-one-out 1.5B variants' performance for each dataset individually, compared to the full *CLARO* model and the L1-Max baseline across all eight evaluated datasets at varying token budgets.

CLARO achieves the highest accuracy across nearly all token budgets. At the tightest constraint (384 tokens), *CLARO* holds a significant lead over the baseline and outperforms the best-performing ablation variant by approximately 1% in absolute terms. The excluded readability variant (purple) exhibits the sharpest performance decline in low-token settings, effectively regressing to the baseline performance; this indicates that as reasoning becomes more compressed, maintaining high readability is essential for preserving logic. On mathematically intensive benchmarks like MATH-500 and AIME, removing the mathematical expression rate attribute excluded the math expression rate variant (brown) results in notable drops. Similarly, the ablation of syntactic compression (orange) reveals that shorter sentence structures are critical for decomposing complex reasoning into manageable steps; without this constraint, performance degrades across budgets, indicating that enforcing granular sentence boundaries aids in error prevention during multi-step inference. Finally, while the excluded redundancy variant (green) performs closest to the full model overall, its removal still results in a consistent lag behind *CLARO* across the entire length spectrum, confirming that enforcing low redundancy contributes to the overall efficiency of the reasoning trace.

A.6 Qualitative Examples

To complement the quantitative results presented in Section 4.4, Figures A.4 and A.5 provide a qualitative comparison of reasoning trajectories generated by the baseline model (L1-Max), *CLARO* model, and its different ablation variants. By examining the same problem instance across these models, we can illuminate the behavioral shifts resulting from the exclusion of individual attributes. These examples illustrate how the proposed method influ-

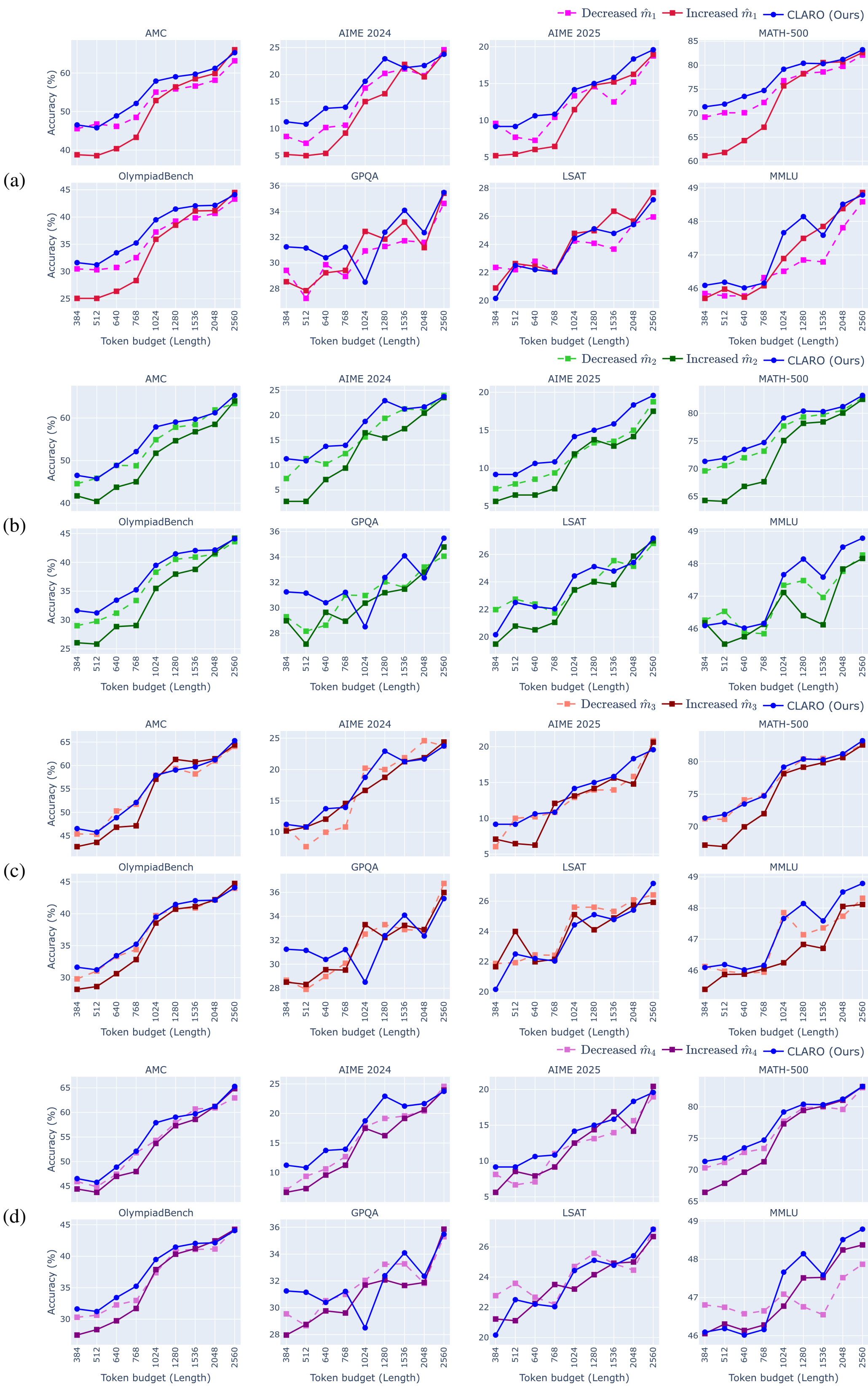

Figure A.2: Sensitivity analysis of reasoning performance across structural attribute targets. We evaluate the impact of increasing or decreasing metric target values for (a) Readability, (b) Mathematical expression rate, (c) Syntactic compression, and (d) Redundancy on model accuracy across diverse benchmarks. While the default *CLARO* calibration (blue) generally yields optimal results, the framework exhibits stability across variants, maintaining strong performance even when constraints are relaxed or tightened. This demonstrates that the efficacy of *CLARO* is driven by the general imposition of structural attributes rather than the precise tuning of their hyperparameters.

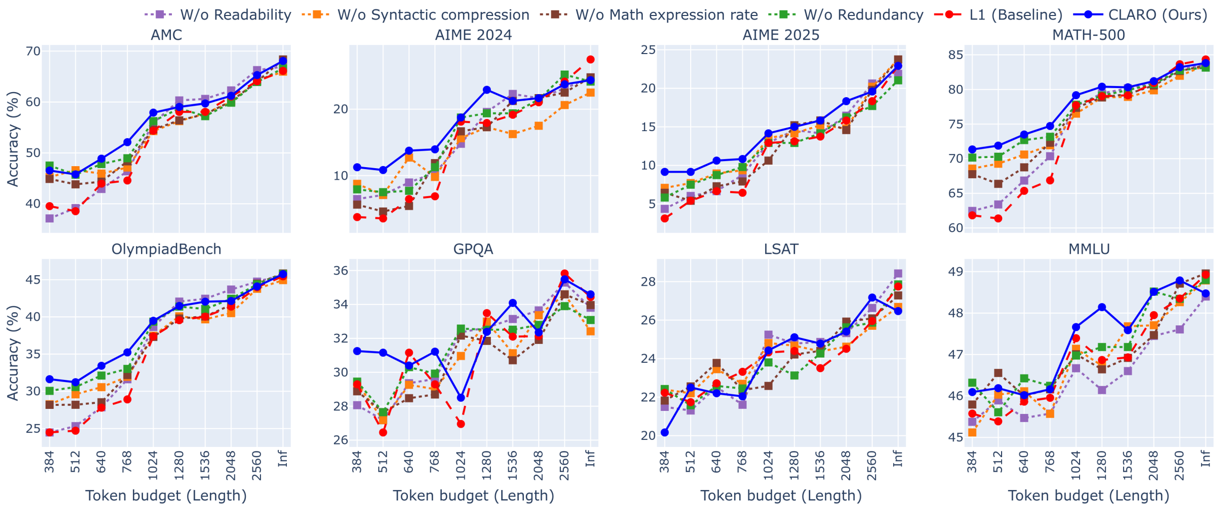

Figure A.3: Detailed breakdown of model performance across benchmarks under varying allocated token budgets. *CLARO* (blue) and all its ablated variants (squares) robustly outperform the L1-Max baseline (red). While guiding reasoning structure using all attributes yields optimal results, partial subsets also improve performance.

ences reasoning trajectories across key structural attributes. By contrasting these side-by-side, the examples help clarify how the attribute-driven fine-tuning alters not only output length but also the substantive qualities of reasoning, thereby offering additional insight into the mechanisms underlying the observed performance improvements.

Problem

Find the largest possible real part of $[(75+117i)z+\frac{96+144i}{z}]$ where z is a complex number with $|z|=4$.

Length-Controlled Baseline

Okay, so I need to find the largest possible real part of $(75+117i)z+(96+144i)/z$, where $|z|=4$. Hmm, let me see. Since $|z|=4$, z can be written as $4e^{i\theta}$. Then, perhaps substitute $z=4e^{i\theta}$ and compute the real part of the expression. Let me try that.
So, $(75+117i)z=(75+117i)4e^{i\theta}=4(75+117i)e^{i\theta}$. Similarly, $(96+144i)/z=(96+144i)/4e^{i\theta}=(24+36i)e^{-i\theta}$. Then, the real part would be the real part of $[4(75+117i)e^{i\theta}+(24+36i)e^{-i\theta}]$. Maybe simplify this expression.
Let me compute each term separately. Let me denote $e^{i\theta}=cos\theta+isin\theta$. Then, $4(75+117i)(cos\theta+isin\theta)+(24+36i)(cos\theta-isin\theta)$. Expanding this, we get real parts and imaginary parts. Then, take the real part of the sum.
This seems a bit involved. Maybe there's a better approach. Let me consider that the expression is linear in z and $1/z$. Since $|z|=4$, $1/z$ has magnitude 1/4. Perhaps use calculus to maximize the real part. Let me set $z=4x$ where |x|=1. Then, the expression becomes $(75+117i)4x+(96+144i)/(4x)=4(75+117i)x+(24+36i)/x$. Then, the real part is $4*75*x+24/x-(117*4)x*(1/4)$??? Wait, maybe this is too messy.
Perhaps the maximum occurs when z is aligned in a certain way. Maybe when z is real and positive. Let me try z=4. Then, compute the real part: $(75+117i)*4+(96+144i)/4=(300+468i)+(24+36i)=324+504i$. Real part is 324.
Alternatively, $z=-4$: $(75+117i)(-4)+(96+144i)/(-4)=(-300-468i)+(-24-36i)=-324-504i$. Real part is -324.
So, perhaps the maximum is 324 when $z=4$. </think> The largest possible real part is $\boxed{324}$. ✗

***CLARO* (Ours)**

Okay, so I need to find the largest possible real part of $(75+117i)z+(96+144i)/z$, where $|z|=4$. Hmm, z is a complex number with magnitude 4. Maybe I can write z as $4e^{i\theta}$. Then, substitute that into the expression and find the real part.
So, let $z=4e^{i\theta}$. Then, $(75+117i)z=(75+117i)*4e^{i\theta}$. And $(96+144i)/z=(96+144i)/4e^{i\theta}=(24+36i)e^{-i\theta}$. I need to find the real part of the sum: $(75+117i)4e^{i\theta}+(24+36i)e^{-i\theta}$. Maybe expand this.
Let me compute each term's real part. For the first term, $(75+117i)4e^{i\theta}=4(75+117i)(cos\theta+isin\theta)$. The real part is $4[75cos\theta-117sin\theta]$.
For the second term, $(24+36i)e^{-i\theta}=(24+36i)(cos\theta-isin\theta)$. The real part is $24cos\theta+36sin\theta$.
So total real part is $4(75cos\theta-117sin\theta)+24cos\theta+36sin\theta$. Simplify: $(300cos\theta-468sin\theta)+24cos\theta+36sin\theta=(324cos\theta-432sin\theta)$.
We need to maximize $324cos\theta-432sin\theta$. This is of the form $Acos\theta+Bsin\theta$, which can be written as $Ccos(\theta+\Phi)$. The maximum is $sqrt(A^2+B^2)$.
Here, $A=324, B=-432$. So $sqrt(324^2+432^2)=sqrt(104976+186624)=sqrt(291600)=540$.
Thus, the largest possible real part is 540. </think> The largest possible real part is $\boxed{540}$. ✓

▲ Readability
▲ Syntactic compression
Math notation rate ▲
Redundancy ▼

Figure A.4: Qualitative comparison of reasoning trajectories on a mathematical reasoning task under identical settings. Consistent with the example in Figure 2, the baseline model (left) produces a verbose and unstructured trace that fails to solve the problem. In contrast, *CLARO* (right) demonstrates superior structural attributes leading to the correct solution. This trajectory serves as the reference standard for the ablation analysis in Figure A.5, where the impact of removing attributes is examined on this same problem instance.

Problem

Find the largest possible real part of $[(75+117i)z+\frac{96+144i}{z}]$ where z is a complex number with $|z|=4$.

Readability metric excluded ($\alpha_1 = 0$)

... Let me see, z is a complex number with magnitude 4. Maybe I can write z as $4e^{i\theta}$. Then, substitute and simplify.
First, $(75+117i)z = (75+117i)4e^{i\theta} = 4(75+117i)e^{i\theta}$. Similarly, $(96+144i)/z = (96+144i)/4e^{i\theta} = (24+36i)e^{-i\theta}$.
Now, the expression becomes $4(75+117i)e^{i\theta} + (24+36i)e^{-i\theta}$. I need the real part of this.
So, $\mathrm{Re}[4(75+117i)e^{i\theta} + (24+36i)e^{-i\theta}]$. Let's compute each term's real part.
For the first term: $4(75\cos\theta - 117\sin\theta + i(75\sin\theta + 117\cos\theta))$. So real part is $4(75\cos\theta - 117\sin\theta)$.
For the second term: $(24\cos\theta + 36\sin\theta) + i(-24\sin\theta + 36\cos\theta)$. So real part is $24\cos\theta + 36\sin\theta$.
Adding both real parts: $4(75\cos\theta - 117\sin\theta) + 24\cos\theta + 36\sin\theta$.
Simplify: $300\cos\theta - 468\sin\theta + 24\cos\theta + 36\sin\theta = (300+24)\cos\theta + (-468+36)\sin\theta = 324\cos\theta - 432\sin\theta$.
We need to maximize $324\cos\theta - 432\sin\theta$. This is of the form $A\cos\theta + B\sin\theta$, which can be written as $C\cos(\theta+\phi)$. The maximum is $\sqrt{A^2+B^2}$.
So $C = \sqrt{324^2+432^2} = \sqrt{104,976+186,624} = \sqrt{291,600} = 540$.
Thus, the largest possible real part is 540. </think> The largest possible real part is $\boxed{540}$. ✓

Syntactic compression metric excluded ($\alpha_3 = 0$)

... Hmm, let's see. Since $|z| = 4$, z can be written as $4e^{i\theta}$. Then, $(75+117i)z$ becomes $(75+117i)4e^{i\theta}$.
Similarly, $(96+144i)/z$ is $(96+144i)/4e^{i\theta}$. I need to find the real part of the sum. Maybe I can write both terms in terms of cosine and sine. Let's see, the real part of $(75+117i)z$ is $\mathrm{Re}[(75+117i)4e^{i\theta}] = 4\mathrm{Re}[(75+117i)e^{i\theta}]$. Similarly, the real part of $(96+144i)/z$ is $\mathrm{Re}[(96+144i)/4e^{i\theta}] = (1/4)\mathrm{Re}[(96+144i)/e^{i\theta}]$. Maybe simplify $(96+144i)$ as $48(2+3i)$. So, $(96+144i)/e^{i\theta} = 48(2+3i)e^{-i\theta}$. Then, $\mathrm{Re}[(96+144i)/z] = (1/4)48\mathrm{Re}[(2+3i)e^{-i\theta}] = 12\mathrm{Re}[(2+3i)e^{-i\theta}]$. So, overall, the real part is $4\mathrm{Re}[(75+117i)e^{i\theta}] + 12\mathrm{Re}[(2+3i)e^{-i\theta}]$. Maybe compute $\mathrm{Re}[(75+117i)e^{i\theta}]$.
Let's compute $(75+117i)e^{i\theta} = (75+117i)(\cos\theta + i\sin\theta) = (75\cos\theta - 117\sin\theta) + i(75\sin\theta + 117\cos\theta)$. So Re part is $75\cos\theta - 117\sin\theta$. Similarly, $\mathrm{Re}[(2+3i)e^{-i\theta}] = 2\cos\theta + 3\sin\theta$. Therefore, total real part is $4(75\cos\theta - 117\sin\theta) + 12(2\cos\theta + 3\sin\theta) = 300\cos\theta - 468\sin\theta + 24\cos\theta + 36\sin\theta = (300+24)\cos\theta + (-468+36)\sin\theta = 324\cos\theta - 432\sin\theta$.
Now, we need to maximize $324\cos\theta - 432\sin\theta$. This is of the form $A\cos\theta + B\sin\theta$, which can be written as $C\cos(\theta+\phi)$. The maximum is $\sqrt{A^2+B^2}$. So $\sqrt{324^2+432^2}$. Compute $324^2 = 104,976$; $432^2 = 186,624$. Sum=291,600. $\sqrt{291,600} = 540$. So the maximum real part is 540. </think> The largest possible real part is $\boxed{540}$. ✓

Mathematical expression rate excluded ($\alpha_2 = 0$)

... Hmm, let's see. Maybe I can write z as $4e^{i\theta}$ since $|z| = 4$. Then, substitute that into the expression and separate real and imaginary parts. But I'm not sure how to proceed from here. Maybe I can express the real part as a function of θ and then find its maximum. Let me try that.
First, $z = 4e^{i\theta}$. Then, $(75+117i)z = (75+117i)4e^{i\theta}$. Similarly, $(96+144i)/z = (96+144i)/4e^{i\theta}$. So, the expression becomes $(75+117i)4e^{i\theta} + (96+144i)/(4e^{i\theta})$. Then, the real part is $\mathrm{Re}[(75+117i)4e^{i\theta}] + \mathrm{Re}[(96+144i)/(4e^{i\theta})]$.
Calculating $\mathrm{Re}[(75+117i)4e^{i\theta}]$ is $4(75\cos\theta - 117\sin\theta)$. And $\mathrm{Re}[(96+144i)/(4e^{i\theta})]$ is $(96-144i)/(4e^{i\theta})$'s real part, which is $(96-144i)/(4e^{i\theta}) = (96-144i)/(4\cos\theta + i4\sin\theta)$. Hmm, this seems complicated. Maybe there's a better way.
Alternatively, perhaps use calculus to maximize the real part as a function of θ. Let me denote $\mathrm{Re}[\text{expression}] = 4(75\cos\theta - 117\sin\theta) + \mathrm{Re}[(96+144i)/(4e^{i\theta})]$. Maybe simplify $(96+144i)/(4e^{i\theta}) = (96+144i)/(4)(\cos\theta + i\sin\theta) = (24+36i)/(\cos\theta + i\sin\theta)$. Then, multiply numerator and denominator by conjugate: $(24+36i)(\cos\theta - i\sin\theta)/(\cos^2\theta + \sin^2\theta) = (24\cos\theta + 36\sin\theta) + i(-24\sin\theta + 36\cos\theta)$. So, Re part is $24\cos\theta + 36\sin\theta$.
Therefore, total Re part is $4(75\cos\theta - 117\sin\theta) + 24\cos\theta + 36\sin\theta = (300\cos\theta + 24\cos\theta) + (-468\sin\theta + 36\sin\theta) = 324\cos\theta - 432\sin\theta$.
Now, to maximize $324\cos\theta - 432\sin\theta$. This is of the form $A\cos\theta + B\sin\theta$, which has maximum $\sqrt{A^2+B^2}$. So, $\sqrt{324^2+432^2} = \sqrt{104976+186624} = \sqrt{291600} = 540$.
Therefore, the largest possible real part is 540. </think> The largest possible real part is $\boxed{540}$. ✓

Redundancy metric excluded ($\alpha_4 = 0$)

... Hmm, z is a complex number with magnitude 4. Let me think about how to approach this.
First, since $|z| = 4$, I can write z as $4e^{i\theta}$, where θ is the argument of z. Then, $(75+117i)z$ would be $(75+117i)4e^{i\theta}$. Similarly, $(96+144i)/z$ would be $(96+144i)/4e^{i\theta}$.
I need to find the real part of this expression. So, let's denote $z = 4e^{i\theta}$.
Then, the expression becomes: $(75+117i)4e^{i\theta} + (96+144i)/(4e^{i\theta})$
Simplify each term: First term: $(75+117i)4e^{i\theta} = 4(75+117i)e^{i\theta}$
Second term: $(96+144i)/(4e^{i\theta}) = (96+144i)/(4)e^{-i\theta} = (24+36i)e^{-i\theta}$
So, the expression is $4(75+117i)e^{i\theta} + (24+36i)e^{-i\theta}$
Now, to find the real part, we take $\mathrm{Re}[4(75+117i)e^{i\theta} + (24+36i)e^{-i\theta}]$
Let's compute each part: $\mathrm{Re}[4(75+117i)e^{i\theta}] = 4\mathrm{Re}[(75+117i)e^{i\theta}]$
Similarly, $\mathrm{Re}[(24+36i)e^{-i\theta}] = \mathrm{Re}[(24+36i)e^{-i\theta}]$
But $e^{i\theta} = \cos\theta + i\sin\theta$, so: $\mathrm{Re}[(75+117i)(\cos\theta + i\sin\theta)] = 75\cos\theta - 117\sin\theta$
Thus, the first part becomes $4(75\cos\theta - 117\sin\theta)$
The second part: $\mathrm{Re}[(24+36i)(\cos\theta - i\sin\theta)] = 24\cos\theta + 36\sin\theta$
Therefore, the total real part is: $4(75\cos\theta - 117\sin\theta) + 24\cos\theta + 36\sin\theta$
Simplify: $(300\cos\theta - 468\sin\theta) + 24\cos\theta + 36\sin\theta = (324\cos\theta - 432\sin\theta)$
So, we need to maximize $324\cos\theta - 432\sin\theta$ This is of the form $A\cos\theta + B\sin\theta$, which can be written as $C\cos(\theta+\phi)$, where $C = \sqrt{A^2+B^2}$
Here, $A = 324, B = -432$ Thus, $C = \sqrt{324^2+432^2} = \sqrt{(324)^2+(432)^2}$ Compute $324^2 = 104,976; 432^2 = 186,624$; sum $= 291,600$; $\sqrt{} = 540$
Thus, the maximum value is 540. Therefore, the largest possible real part is 540. </think> The largest possible real part is $\boxed{540}$. ✓

Figure A.5: Qualitative examples of reasoning traces from *CLARO* ablation variants. Each box illustrates the specific reasoning style that emerges when a target attribute is excluded from the optimization process. Readability metric excluded (Top-Left): The model produces more segmented sentences with lower linguistic flow. Mathematical expression rate excluded (Bottom-Left): Although the typical goal of this metric is to increase mathematical density, in this case, removing it results in unregulated usage—the model generates a high volume of math but embeds it within a disproportionately long and verbose output, failing to achieve efficient density. Syntactic compression metric excluded (Top-Right): The output lacks granular sentence segmentation, leading to long and dense blocks of text that may obscure logical steps. Redundancy metric excluded (Bottom-Right): The model reverts to a repetitive style, frequently restating tasks and intermediate steps (e.g., "simplify", "compute"), inflating token count without adding informational value.