FUNC-R1 : Incentivizing Mathematical Function Reasoning in Multimodal Large Language Models

Mingze Yin[1,2,*], Xiaohan Wang[2,*], Dian Li[2,†], Haichao Yao[1,2], Yilin Zhao[2], Youjun Chen[2], Gang Liu[2], Jintai Chen[3], Yiheng Zhu[4,†], Chang-Yu Hsieh[5,†] Aimin Pan[6,7,†]

[1]College of Computer Science and Technology, Zhejiang University
[2]Tencent [3]AI Thrust, HKUST (GZ) [4]Zhongguancun Academy
[5]College of Pharmaceutical Sciences, Zhejiang University [6]VNET Group [7]Zhejiang Lab

[*]Equal Contribution. [†]Corresponding Authors.

Abstract

Performing deliberate mathematical reasoning in visual contexts is a hallmark of advanced Multimodal Large Language Models (MLLMs) and requires a sophisticated synthesis of perceptual grounding and symbolic logic. However, in the realm of mathematical functions, our investigation reveals a critical *modality interference phenomenon*: even advanced models, while performing textual computational reasoning, tend to disregard or misinterpret essential visual cues. To address this challenge, we propose **FUNC-R1**, which synergistically harmonizes precise visual perception and rigorous logical reasoning. Concretely, built upon an explicitly decoupled architecture, we employ a hierarchical post-training framework to progressively identify critical visual evidence and conduct in-depth theoretical reasoning. Furthermore, the *Perception-Aligned Theoretic Optimization (PATO)* strategy is proposed to steer policy updating towards internalizing fundamental theoretical properties while dynamically rectifying heterogeneous visual information throughout the reasoning process. Extensive experiments across diverse benchmarks demonstrate that FUNC-R1 delivers the optimal performance among open-source MLLMs, even surpassing GPT-5 with an 8.4% improvement on MathVerse's function-oriented tasks.

1 Introduction

Mathematical reasoning in visual contexts represents the sophistication of human intelligence, as it hinges on precise visual interpretation, substantive domain knowledge, and rigorous logical inference. Significant advances in Multimodal Large Language Models (MLLMs) have spurred discourse on the prospect of demonstrating precise mathematical reasoning, advancing towards Artificial General Intelligence (AGI) (Trinh et al., 2024; Zhou et al., 2025; Yin et al., 2026; Wang et al., 2025a).

Mathematical functions, grounded in abstraction-centric symbolic formalisms, do not abide by the same constraints that govern conventional problems in geometry, statistics, topology or combinatorics. The function-driven computational deduction demands global coherence and quantitative exactitude, whereas diagram parsing emphasizes localized perception derived from limited and approximate information. Pronounced information asymmetries compel the model to encode bifurcating trajectories, thereby hindering the reconciliation of function-oriented diagram interpretation and computational reasoning (May, 1976).

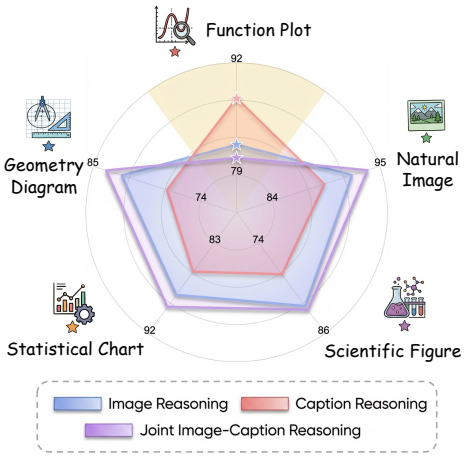


Figure 1: GPT-5's mathematical reasoning performance trends under diverse multimodal conditions vary systematically across distinct visual contexts.

Herein, a simple and intuitive heuristic experiment is conducted to reveal the domain-specific *modality interference phenomenon*. Specifically, we employ the most powerful LLM to date GPT-5

Correspondence to: mzyin256@gmail.com, kimhsieh@zju.edu.cn, pan.aimin@vnet.com

Findings of the Association for Computational Linguistics: EMNLP 2026, pages 40616–40632
October 24-29, 2026 ©2026 Association for Computational Linguistics

to execute reasoning over diverse problem types from MathVista (Lu et al., 2024). Three inference settings are specified: *(i)* supplying the diagram (Image Reasoning), *(ii)* supplying a textual description of the diagram (Caption Reasoning), and *(iii)* supplying both (Joint Image-Caption Reasoning). As illustrated in Figure 1, while performance trends match expectations for most problem types, reasoning with function diagrams is notably inferior to reasoning with textual descriptions. Paradoxically, jointly providing synergistic visual diagrams and textual descriptions yields the worst performance. This anomalous pattern indicates that MLLMs are prone to selectively disregard or partially forget crucial mathematical visual clues when engaging in text-based computational reasoning. Such intrinsic irreconcilability constitutes a major impediment to incentivizing deliberate function reasoning.

To overcome such interference issue, an explicitly decoupled model architecture is designed, consisting of a visual math parser and a textual math reasoner. We curate function-oriented mathematical corpora comprising 3.5k problems, with annotations capturing both visual understanding and logical reasoning, constructed through an LLM-enhanced automated pipeline and validated by human mathematics experts. In the hierarchical training framework, we first perform structured Supervised Fine-Tuning (SFT) for the visual math parser. The systematic captioning trajectory encourages the model to move beyond basic visual element extraction and capture function properties, fostering deeper visual understanding. In the subsequent training stage, we propose the Perception-Aligned Theoretic Optimization (PATO) strategy to systematically enhance the rigor and depth of a textual math reasoner's computational deduction. In contrast to conventional Group Relative Policy Optimization (GRPO), we adopt a smooth and adaptive policy optimization scheme to balance exploitation and exploration, yielding stable and performant Reinforcement Learning (RL) policy updating. Furthermore, the customized reward signals progressively incentivize the model to rectify visual perception anomalies, assimilate sophisticated function-theoretic properties, and execute nuanced computational reasoning. Eventually, the collaboration of visual math parser and textual math reasoner involves revisiting visual evidence and leveraging theoretical knowledge throughout the problem-solving process, therefore catalyzing advanced reasoning over mathematical functions.

Our contributions can be summarized as follows:

- *Modality Interference Issue.* We reveal intrinsic modality interference in the realm of mathematical functions and leverage an explicitly decoupled architecture and a hierarchical post-training strategy to alleviate it.
- *Curated Diverse Corpora.* We construct 3.5k function-centric corpora with LLM-assisted annotation and human validation, supporting the optimization of perception and reasoning.
- *Innovative Policy Optimization.* The tailored PATO is a training-stable, online-decidable policy update mechanism with hybrid rewards, coupling visual interpretation calibration, theoretical knowledge guidance, and verifiable computational reasoning.
- *Substantial Reasoning Enhancements.* FUNC-R1 attains open-source state-of-the-art reasoning performance on five benchmarks, notably outperforming GPT-5 by 8.4% on MathVerse.

2 Method

As illustrated in Figure 2, an explicitly decoupled architecture is incorporated to facilitate the integration of visual perception and logical reasoning. We enhance the visual perception capability of the visual math parser through SFT (Sec. 2.2), and improve the logical reasoning ability of the textual math reasoner via RL (Sec. 2.3).

2.1 Mathematical Corpus Construction

We collect over 10k mathematical function problems from publicly authoritative mathematics websites. The MinerU framework (Wang et al., 2024a) is first employed for data preprocessing, including problem parsing and function plot extraction. To enrich the annotations, we leverage GPT-5 and Gemini-3-Pro to generate image captions, identify relevant knowledge concepts, and refine reasoning trajectories. In addition, Qwen3-VL, Claude Opus 4.1, and GLM-4.1V are adopted as arbitration models to verify the annotations and resolve discrepancies. Finally, five Ph.D. students majoring in mathematics are recruited to perform manual validation, focusing on the visual evidence, key formulas, and logical soundness of the reasoning process. Eventually, two function-centric datasets comprising 3,519 high-quality, well-annotated mathematical problems are constructed: `FuncCaption`, comprising high-quality diagram captions, and

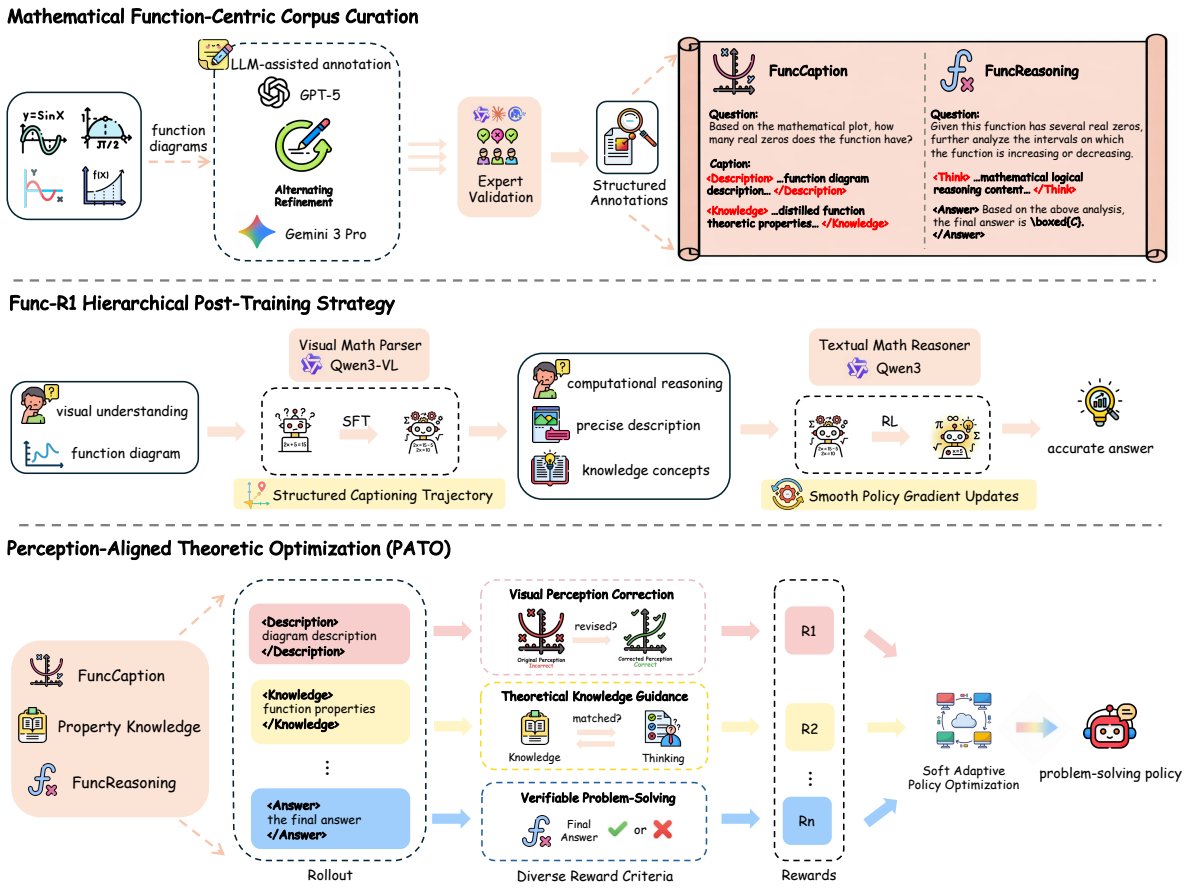


Figure 2: **Overview of the FUNC-R1 framework**. We adopt a progressive optimization approach that combines SFT and RL under an explicitly decoupled architecture to integrate visual perception and logical reasoning capabilities.

FuncReasoning, consisting of function-oriented question–answer pairs.

2.2 Domain Knowledge Enhancement

We begin with the optimization of visual math parser to precisely perceive and interpret mathematical function diagrams through domain-specific Supervised Fine-Tuning (SFT). Based on the problem statement and function diagram, the structured captioning trajectory is organized as: <Description>...</Description> <Property>...</Property>. By elaborating progressive reasoning captions, we empower the MLLM to faithfully comprehend intricate mathematical properties pertinent to subsequent problem-solving. Leveraging the curated FuncCaption for SFT, we employ a standard cross-entropy loss to endow visual math parser with the structural captioning format and function-theoretic knowledge. This SFT phase provides a structured cognitive prior, improving the visual perception and knowledge comprehension for downstream mathematical problem-solving tasks.

2.3 Perception-Aligned Theoretic Optimization (PATO)

2.3.1 Smooth Policy Gradient Updates

Under conventional Group Relative Policy Optimization (GRPO) (Shao et al., 2024), hard gradient clipping induces high-variance advantage estimates, which in turn constrain the attainable convergence performance. As a refinement, Soft Adaptive Policy Optimization (SAPO) (Gao et al., 2025a) adaptively attenuates off-policy updates to maintain a sustained and stable optimization signal, but often at the expense of slower convergence. Inspired by SAPO, our PATO adapts a smooth, temperature-controlled gating mechanism to rich and informative reward learning signals, enabling stable and efficient RL post-training. (See Sec. 3.6, Policy Gradient Updates, for convergence comparisons.)

$$\mathcal{L}_{\mathrm{RL}}(\theta) = \mathbb{E}_{q\sim\mathcal{D},\,\{y_i\}_{i=1}^{G}\sim\pi_{\theta_{\mathrm{old}}}(\cdot|q)} \left[\frac{1}{G}\sum_{i=1}^{G}\frac{1}{|y_i|}\sum_{t=1}^{|y_i|} f_{i,t}(r_{i,t}(\theta))\ \hat{A}_{i,t} \right], \quad (1)$$

$$\widehat{A}_{i,t} = \widehat{A}_i = \frac{R_i - \texttt{mean}(\{R_j\}_{j=1}^G)}{\texttt{std}\left(\{R_j\}_{j=1}^G\right)}, \quad (2)$$

Concretely, a token-level soft trust region is introduced to enhance sequence-level coherence, and an asymmetric temperature scheme is adopted to accommodate the differing dynamics of positive versus negative token updates.

$$f_{i,t}(x) = \sigma(\tau_{i,t}(x-1)) \cdot \frac{4}{\tau_{i,t}}, \quad (3)$$

$$\tau_{i,t} = \begin{cases} \tau_{\text{pos}}, & \text{if} \hat{A}_{i,t} > 0, \\ \tau_{\text{neg}}, & \text{otherwise.} \end{cases} \quad (4)$$

where τ_{pos} and τ_{neg} are the temperatures for positive and negative tokens, respectively.

2.3.2 Reward Signal Design

To effectively guide the model's mathematical reasoning optimization, we design a multi-component reward that integrates *robust visual parsing, profound knowledge understanding, and rigorous computational reasoning*. The proposed mathematical problem-solving policy optimization framework comprehensively incorporates rule-based, LLM-assisted, and directly verifiable rewards.

Visual Perception Correction Reward. We incentivize the problem-solving model to rectify heterogeneous visual perceptual information, given the intrinsic complexity and observational ambiguities of function plots. LLM-as-a-Judge framework is incorporated to determine whether the model successfully avoids, detects, or corrects visual perception errors during the mathematical reasoning process. Concretely, Qwen3-VL-30B-A3B (Yang et al., 2025a) is invoked to produce discrete labels rather than a free-form scalar score. *(i)* $r \in \{0,1\}$ indicates model directly refers to the original caption; *(ii)* $d \in \{0,1\}$ indicates model identifies caption errors and selectively ignores parts; *(iii)* $c \in \{0,1\}$ indicates model precisely detects caption errors and explicitly corrects them. The final reward is deterministically assigned through a fixed rule-based mapping.

$$\mathcal{R}_v = \begin{cases} 2, & c = 1, \\ 1, & c = 0,\ d = 1, \\ -1, & d = 0,\ r = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

This design reduces sensitivity to prompt variations, enhances reproducibility, and provides stable reward signals that incentivize the model to proactively engage in visual perception correction.

Theoretical Knowledge Guidance Reward. We further encourage the model to generate rigorous reasoning processes guided by underlying mathematical principles. Let K_q and K_r denote the knowledge items associated with the question and invoked in the reasoning trajectory, respectively. Intuitively, high-quality reasoning should cover the key theoretical concepts needed for solving current problems, while avoiding irrelevant or spurious knowledge. We therefore measure global knowledge alignment with the F1 score $\mathrm{F}_1(K_q, K_r)$, which captures both the coverage and precision of the theoretical knowledge. Additionally, to prevent the model from merely mentioning relevant concepts without effectively applying them in reasoning, we incorporate the maximum similarity between intermediate reasoning steps and theoretical knowledge items to characterize semantic alignment at a finer granularity.

$$\mathcal{R}_k = \mathrm{F}_1(K_q, K_r) \cdot \mathbb{E}_{k \sim K_q \cap K_r}\left[\texttt{sim}(k_q, k_r)\right]. \quad (6)$$

This reward formulation combines global knowledge matching with local knowledge utilization, jointly promoting overall consistency and logically productive knowledge application throughout intricate reasoning.

Verifiable Problem-Solving Reward. Building on well-calibrated visual parsing information and thorough theoretical knowledge comprehension, we incentivize the model to engage in structured inference and profound computational reasoning to achieve correct solutions. The solving model is required to generate responses that adhere to a strict format containing a reasoning segment and an answer segment (*e.g.*, `<Think>...</Think> <Answer>...</Answer>`). Moreover, the reward for problem-solving correctness is standardized owing to the inherently verifiable nature of mathematics (Lambert et al., 2024).

$$\mathcal{R}_{\text{s}} = \begin{cases} \mathbf{1}(y_i = \hat{y}_i), & \text{valid format,} \\ -1, & \text{invalid format.} \end{cases} \quad (7)$$

where $\hat{y}_i \in \hat{\mathbf{y}} = [\hat{y}_1, \ldots, \hat{y}_n]$ represents the extracted predicted answer, $y_i \in \mathbf{y} = [y_1, \ldots, y_n]$ indicates the reference solution, and $\mathbf{1}(\cdot)$ denotes the indicator function. The textual math reasoner is incentivized to execute rigorous computational derivation grounded in structured reasoning, facilitating mathematical problem-solving efficacy.

Model	MathVista				Math-V	MathVerse				We-Math
	Disc.	Lin.	Nonlin.	Comp.	Func.	Disc.	Lin.	Nonlin.	Comp.	Func.
Closed-Source Multimodal Large Language Models (MLLMs)										
GPT-4V	52.3	50.5	51.8	47.2	36.1	41.2	38.0	35.1	33.6	45.6
GPT-4o	60.3	59.2	64.7	60.3	37.0	40.5	39.9	37.0	35.6	50.1
Qwen-VL-Max	63.2	62.6	60.6	54.8	40.7	29.1	27.8	26.8	22.8	49.8
Gemini 2.5 Flash	83.2	80.4	77.2	73.4	53.2	58.2	56.8	51.0	49.8	56.6
Gemini 3 Flash	84.1	80.0	78.5	75.3	51.8	60.2	58.3	54.4	52.1	59.7
GPT-5	84.7	81.0	**79.9**	76.1	**57.0**	61.6	60.4	56.3	55.6	67.3
Open-Source Multimodal Large Language Models (MLLMs)										
LLaVA-V1.5-7B	22.7	23.6	10.2	27.2	14.7	17.6	10.4	11.6	12.6	27.9
LLaVA-V1.5-13B	27.7	25.6	22.1	21.3	18.0	18.5	14.2	15.9	13.5	23.6
InternVL2-8B	34.2	34.0	35.4	32.5	11.3	23.1	21.7	22.7	18.2	33.0
InternVL2-40B	38.4	37.2	33.4	30.7	19.5	25.1	24.0	21.7	19.2	38.5
DeepSeek-VL2	60.0	57.0	55.7	53.8	30.8	43.8	42.1	43.0	37.5	39.3
GLM-4.1V	63.3	62.7	65.2	68.3	36.2	44.2	40.3	43.7	39.9	40.4
Qwen3-VL-8B	66.8	65.3	60.1	62.7	39.3	47.8	45.3	44.1	42.7	47.3
Qwen3-VL-32B	70.3	68.4	65.9	62.1	42.1	53.3	51.7	50.5	44.8	58.1
Math-LLaVA	49.8	47.6	43.7	42.4	23.9	23.6	28.4	19.7	19.3	37.7
R1-VL	69.6	66.0	65.4	66.4	36.7	54.3	50.5	50.3	45.9	55.7
MathCoder2	73.9	70.4	67.8	66.9	45.8	56.0	54.7	50.9	50.1	60.1
Vision-R1	75.7	72.6	69.9	68.8	46.9	56.8	55.9	51.8	51.0	61.1
CodePercept	77.2	74.1	71.3	70.2	48.2	58.3	57.4	53.5	52.7	62.6
FUNC-R1-Lite: *Qwen3-VL-4B SFT + Qwen3-8B SAPO*										
SFT	71.5	69.7	69.0	63.3	35.7	52.8	47.0	51.2	43.5	57.1
SFT + PATO	77.4	73.6	73.3	70.6	42.4	58.9	58.7	57.0	55.3	64.8
FUNC-R1-Pro: *Qwen3-VL-8B SFT + Qwen3-32B SAPO*										
SFT	76.2	74.1	73.0	66.5	46.9	60.2	60.5	56.8	57.7	61.2
SFT + PATO	**85.3**	**82.6**	79.3	**78.0**	56.4	**66.1**	**65.7**	**61.2**	**60.5**	**70.8**

Table 1: **Out-of-distribution function reasoning performance over diverse mathematical benchmarks.** The optimal and suboptimal open-source models are highlighted in red and blue, respectively. We highlight the best results, including closed-source models, in **bold**.

2.3.3 Overall Training Objective

During the RL phase, we aim to optimize the policy model π_θ utilizing the SAPO training paradigm. The crux of this procedure is the design of the reasoning reward signal, which directs the model's learning process. The overall reward $\mathcal{R}_{\text{PATO}}$, used to compute the group-wise relative advantages, is formulated as a weighted combination of the multi-faceted components detailed in Sec. 2.3.2.

$$\mathcal{R}_{\text{PATO}} = \lambda_1 \mathcal{R}_{\text{v}} + \lambda_2 \mathcal{R}_{\text{k}} + \mathcal{R}_{\text{s}}. \tag{8}$$

These non-negative weight coefficients λ_1 and λ_2 serve as critical hyperparameters for calibrating the model, allowing us to modulate the emphasis among dynamically refining visual analysis, promoting understanding of theoretical knowledge, and preserving a coherent reasoning process.

3 Experiments

In the experimental section, we progressively present the following findings: *(i)* FUNC-R1 achieves the best function problem-solving results on the internal test set, with an average accuracy of 91.4% (Sec. 3.2). *(ii)* FUNC-R1 attains outstanding reasoning performance on five mainstream mathematical benchmarks comparable to that of the powerful closed-source GPT-5 (Sec. 3.3). *(iii)* The training approach is further extended to other reasoning domains, such as geometry, to demonstrate strong universality (Sec. 3.4).

3.1 Experimental Setups

We build two model variants, FUNC-R1-Lite and FUNC-R1-Pro, based on the Qwen3 model family. Using a hierarchical post-training strategy, we progressively optimize the visual math parser and the

textual math reasoner. Specifically, `FuncCaption` is used entirely for SFT, while `FuncReasoning` is split into 75% / 10% / 15% partitions to support RL optimization and in-distribution evaluation. Training details are provided in Appendix C, while the LLM baselines and evaluation benchmarks are described in Appendix E. The code and data will be publicly available at https://github.com/zju-ai4s/Func-R1 upon completion of the company's disclosure review process.

3.2 In-Distribution Function Reasoning

To rigorously interrogate the mathematical problem-solving competence of MLLMs involving visual diagrams, we begin with function reasoning assessments on the in-distribution `FuncReasoning` test set. Evaluation data is meticulously classified into four function types: Discrete Function (Disc. Func.), Linear Function (Lin. Func.), Nonlinear Function (Nonlin. Func.), and Composite Function (Comp. Func.). As presented in Table 2, with increasing function complexity, all models exhibit a declining trend in problem-solving efficacy. FUNC-R1 attains state-of-the-art in-distribution function reasoning, with an average accuracy of 91.4, surpassing closed-source LLMs, including Gemini 3 Flash, Claude Opus 4.7, and GPT-5.

Model	Disc. Func.	Lin. Func.	Nonlin. Func.	Comp. Func.
Closed-Source MLLMs				
GPT-4o	77.0	76.4	72.9	69.5
Gemini 2.5 Flash	80.9	82.2	79.4	75.9
Gemini 3 Flash	84.8	84.6	83.3	78.1
Claude Opus 4.1	82.1	77.5	76.4	73.8
Claude Opus 4.7	89.3	91.3	88.2	79.9
GPT-5	91.5	89.0	82.6	83.4
Open-Source MLLMs				
CogVLM2	74.4	70.2	65.4	58.7
WeThink	75.0	71.8	62.0	59.2
Math-LLaVA	74.8	70.5	67.1	64.9
GLM4.1-V	79.3	72.7	71.4	68.8
Qwen3-VL-8B	78.3	74.8	70.5	67.0
Qwen3-VL-32B	82.6	81.1	79.7	75.4
FUNC-R1-Lite	92.7	89.4	84.5	82.9
FUNC-R1-Pro	**94.3**	**92.7**	**90.0**	**88.5**

Table 2: **In-distribution function reasoning performance** covering granular function classes.

3.3 Out-Of-Distribution Function Reasoning

To further validate the proposed paradigm confers broad improvements in mathematical function problem-solving, we holistically conduct more challenging out-of-distribution function reasoning evaluations. As presented in Table 1, FUNC-R1 significantly advances the state-of-the-art performance on 8 of 10 evaluation suites. Notably, FUNC-R1 even outperforms GPT-5 across all MathVerse function subcategory reasoning tasks, yielding an average relative improvement of 8.4%. The exceptional performance across numerous mathematical reasoning benchmarks further validates the effectiveness of the proposed training strategy and optimization algorithm.

3.4 Generalized Multi-Discipline Reasoning

We further explore whether the proposed training framework could enhance MLLMs' reasoning capabilities across diverse disciplinary contexts. Under the proposed training framework, GEO-R1 is trained on GeoReasoning-10K, whereas SYNSCI-R1 is trained on `FuncCaption`, `FuncReasoning` and GeoReasoning-10K, adhering to the identical base models (*i.e.*, Gemma3-4B and Qwen2.5-7B-Instruct) as in Xin et al. (2025). The generalized reasoning performance is thoroughly assessed across seven college-level core disciplines with the MathVista (Lu et al., 2024) and MMMU (Yue et al., 2024) databases. Extensive commonly-used geometric baselines are incorporated, including Gemma3 (Kamath et al., 2025), AutoGeo (Huang et al., 2025), GeoPeP (Sun et al., 2025), GeoGPT4V (Cai et al., 2024), Geo170K (Gao et al., 2025b), and GeoReasoning (Xin et al., 2025) to facilitate comprehensive comparison. Table 3 shows that GEO-R1 delivers outstanding geometric problem-solving performance, leading a 21% improvement over the previous state-of-the-art. SYNSCI-R1 surpasses GEO-R1 on 6 of 7 tasks, indicating the function-driven mathematical knowledge strengthens cross-domain deductive reasoning. Due to the broad-spectrum reasoning capacity afforded by injecting core knowledge of mathematical functions, GEO-R1 and SYNSCI-R1 demonstrate significant academic merit and considerable practical potential for deployment in real-world intelligent education scenarios.

3.5 Qualitative Reasoning Analysis

We also provide a qualitative analysis to demonstrate practical utility of the constructed model. As shown in Figure 3, when confronted with complex function-solving problems, GPT-5, Gemini 3 Pro, and Qwen3-VL demonstrate varying degrees of visual perceptual and logical reasoning hallucinations, leading to suboptimal solutions. In contrast, empowered by the hierarchical training regimen, FUNC-R1 explicitly decouples the function-

Models	MathVista	MMMU					
	Geometry	Art & Design	Business	Science	Health & Medicine	Human. & Social Sci.	Tech. & Eng.
Gemma3	60.7±1.9	57.8±4.0	44.1±0.6	34.3±0.9	46.8±2.2	59.2±2.1	29.0±1.3
AutoGeo	62.3±2.4	59.3±1.4	43.3±1.1	34.9±1.3	47.4±1.1	58.9±1.5	30.7±2.6
GeoPeP	61.0±2.3	59.2±1.1	40.4±1.4	34.0±1.7	45.1±0.9	59.6±1.0	32.6±0.7
GeoGPT4V	60.5±0.7	60.2±1.1	43.1±1.5	34.5±0.7	46.0±0.7	58.3±1.6	30.8±2.0
Geo170K	62.2±1.5	58.5±0.8	43.6±1.4	30.9±2.0	46.8±2.2	59.9±2.2	30.9±1.6
GeoReasoning	62.8±1.3	60.2±2.0	44.5±2.5	36.0±2.0	46.7±1.1	60.0±0.5	32.9±1.3
GEO-R1	75.9±0.8	64.0±1.6	50.5±1.9	42.2±0.6	48.4±1.8	64.7±0.8	37.6±1.4
SYNSCI-R1	81.8±1.1	65.7±1.2	51.3±2.0	48.3±1.0	49.5±1.3	66.6±0.4	35.2±1.1

Table 3: **Generalized deliberate reasoning Performance**. The abbreviations "Human. & Social Sci." and "Tech. & Eng." refer to "Humanities and Social Science" and "Technology and Engineering".

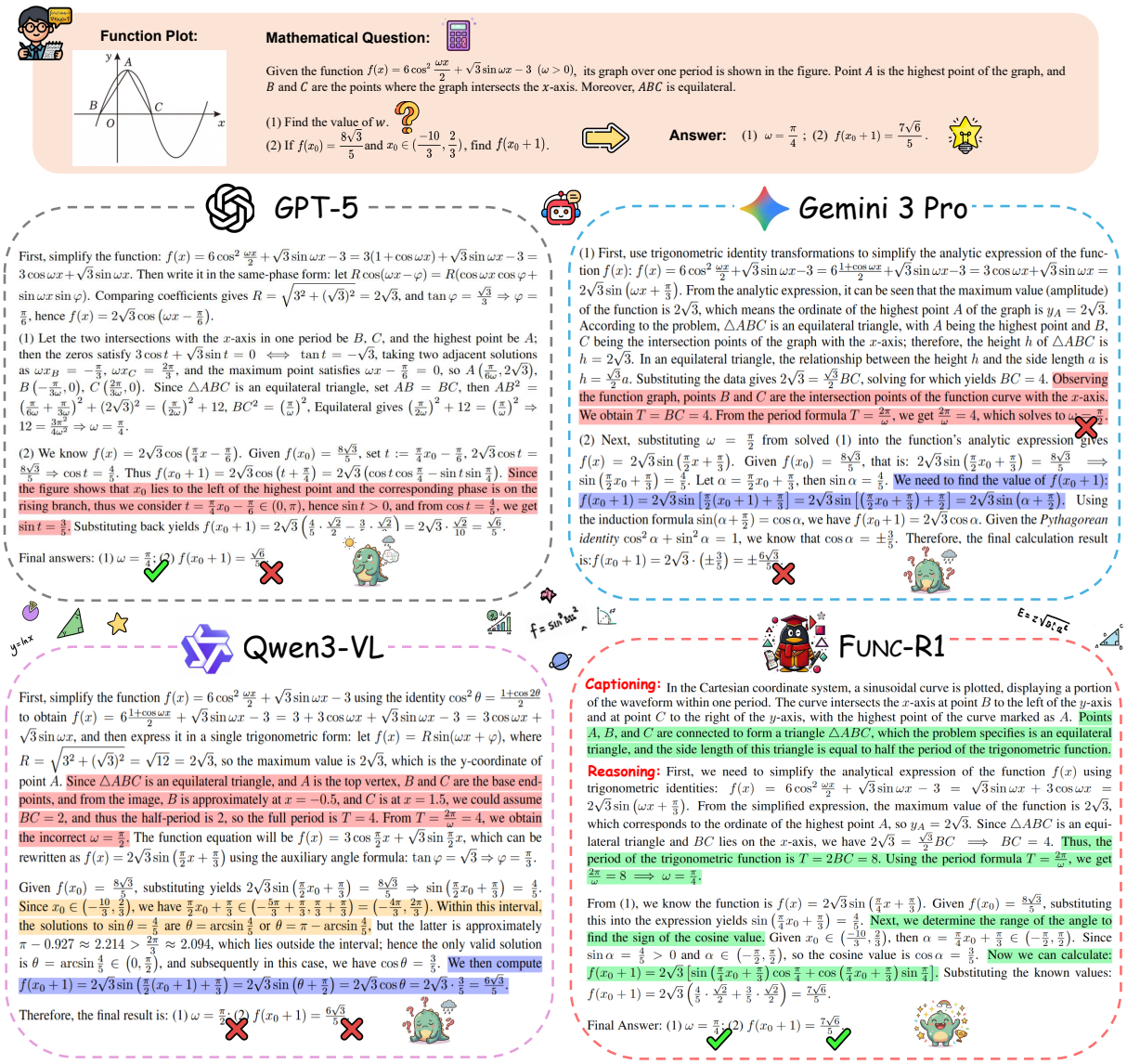


Figure 3: Qualitative evaluation of deductive reasoning in mathematical function problems. The **precise solving procedure** is highlighted, and we delineate discrete categories of causative factors that culminated in erroneous solutions: **inaccurate visual perception**, **theoretical knowledge deficit**, and **defective logical reasoning**.

solving pipeline and accomplishes rigorous mathematical reasoning grounded in meticulous visual interpretation. This demonstrates that FUNC-R1 is a versatile AI assistant, coupling numerically robust symbolic derivation with automated reasoning over mathematical functions, enabling reliable decision support in engineering design, quantitative finance, and scientific computing.

3.6 Ablation Studies

Single-MLLM Bottlenecks. Herein, we leverage Qwen3-VL-8B to investigate the upper bound of reasoning performance achievable by a single MLLM in MathVerse. We comprehensively apply CoT prompting, SFT, and RL strategies to improve the MLLM's abilities in computational reasoning. Figure 4 reveals the performance ceiling of a single open-source MLLM still remains substantially below that of GPT-5, achieving 54.7 compared to GPT-5's 61.6. FUNC-R1 employs a decoupled two-stage problem-solving pipeline that mitigates the performance bottleneck of a single MLLM architecture, surpassing GPT-5 by 8.4% in mathematical reasoning performance.

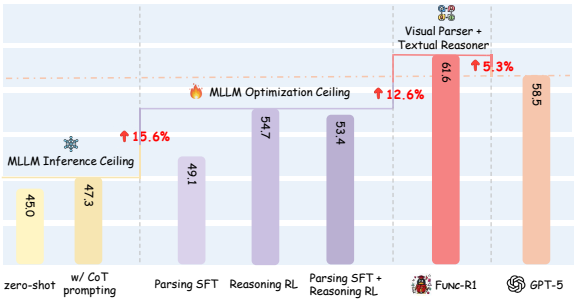


Figure 4: Reasoning bottlenecks in single-MLLMs.

Policy Gradient Updates. During the RL training phase, we replace hard gradient clipping in GRPO with a smooth, temperature-controlled trust region and incorporate comprehensive reward signals, thereby enabling more stable and efficient policy optimization. We conduct a comprehensive comparison among widely used gradient update algorithms, with the reward function, decoding configuration, and evaluation protocol kept identical. On the `FuncReasoning` validation set, Figure 5 demonstrates that PATO converges faster and achieves superior solve rates compared to vanilla GRPO, SAPO, and DAPO.

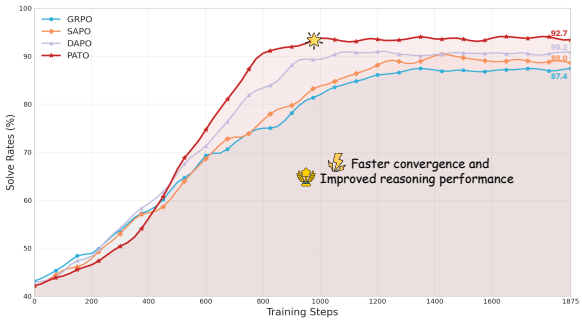


Figure 5: Policy optimization algorithm comparison.

Reward Signal Design. The customized reward signals are calibrated to specifically address function-oriented reasoning challenges. We present comprehensive ablation results under full and partial reward signals in Table 4. Metrics indicate that the absence of any reward leads to degraded performance. $\mathcal{R}_s$ serves as the cornerstone for reasoning optimization, while $\mathcal{R}_v$ and $\mathcal{R}_k$ further advance visual perception correction and theoretical knowledge assimilation.

Rewards $\mathcal{R}_s$	$\mathcal{R}_v$	$\mathcal{R}_k$	MathVista	Math-Vision	MathVerse
✓			65.7	32.1	51.4
✓	✗	✓	69.3	36.0	54.6
✓	✓	✗	70.5	37.4	53.2
✓	✓	✓	73.1	42.4	57.5

Table 4: Ablation study of reward signal components.

4 Related Work

4.1 Multimodal Mathematical Benchmarks

Rigorous assessment of the mathematical reasoning capabilities of multimodal foundation models requires comprehensive and systematic mathematical benchmarks. MMMU (Yue et al., 2024), MathVista (Lu et al., 2024), and Math-Vision (Wang et al., 2024b) innovatively incorporate numerous multimodal mathematical problems to comprehensively evaluate the reasoning performance. Through an in-depth evaluation of mathematical reasoning, MathVerse (Zhang et al., 2024) reveals persistent limitations in MLLMs' diagram perception. LiveK12Bench (Wang et al., 2026) targets key problem-solving challenges arising from exam parsing, process verification, and efficiency constraints. Additionally, We-Math (Qiao et al., 2025) explores the progressive mathematical reasoning principles beyond holistic performance.

4.2 Mathematical Reasoning Optimization

Accurate mathematical problem-solving in multimodal scenarios depends on the integration of robust visual perception with meticulous mathematical reasoning. Substantial efforts have been devoted to monolithically improve MLLMs' mathematical problem-solving capabilities. Specifically, Math-LLaVA (Shi et al., 2024), MathCoder-VL (Wang et al., 2025a), G-LLaVA (Gao et al., 2025b), and Math-LMM (Liu et al., 2025) employ instruction tuning to enhance the cross-modal alignment between visual diagrams and textual content.

Furthermore, Vision-R1 (Huang et al., 2026) leverage reinforcement learning techniques to advance stepwise mathematical problem-solving.

5 Conclusion

This work takes mathematical function reasoning as a testbed to reveal that contemporary MLLMs still exhibit a fundamental tension between precise visual perception and rigorous intricate reasoning. The curated `FuncCaption` and `FuncReasoning` corpora constitute robust empirical resources that underpin an emerging research direction focused on mathematical function problem-solving under multimodal reasoning scenarios. We present FUNC-R1, a multimodal reasoning LLM with advanced mathematical proficiency, which delivers sophisticated reasoning grounded in profound visual parsing. The proposed hierarchical training strategy integrates domain-specific supervised fine-tuning and perception-aligned theoretically grounded policy optimization, furnishing broad guidance for future multimodal reasoning optimization studies. FUNC-R1 attains outstanding sophisticated function reasoning capability commensurate with GPT-5, and the proposed techniques effectively foster cross-disciplinary generalized reasoning.

Limitations

In this research work, the amount of high-quality data for function-oriented multimodal mathematical reasoning remains relatively limited. Building such data requires careful alignment among graph perception, mathematical property annotation, and solution-level reasoning, which makes large-scale collection costly and challenging. Although our current dataset is sufficient to support the main findings, broader coverage of function types, visual forms, and problem difficulty would likely further improve robustness and generalization. Moreover, while the decoupled architecture design helps alleviate modality interference, it may still suffer from error propagation from visual parsing to downstream reasoning. Future work will further explore more interactive perception-reasoning mechanisms to enhance reliability across diverse multimodal reasoning settings.

Acknowledgments

This research is supported by the National Natural Science Foundation of China under Grant No. U22A6001 and No. 62606436.

References

Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, Irwan Bello, and 2 others. 2024. GPT-4 technical report. *arXiv preprint arXiv:2303.08774*.

Anthropic. 2025a. Claude Opus 4.1.

Anthropic. 2025b. Introducing Claude Opus 4.5.

Jinze Bai, Shuai Bai, Shusheng Yang, Shijie Wang, Sinan Tan, Peng Wang, Junyang Lin, Chang Zhou, and Jingren Zhou. 2023. Qwen-VL: A versatile vision-language model for understanding, localization, text reading, and beyond. *arXiv preprint arXiv:2308.12966*.

Shuai Bai, Yuxuan Cai, Ruizhe Chen, Keqin Chen, Xionghui Chen, Zesen Cheng, Lianghao Deng, Wei Ding, Chang Gao, Chunjiang Ge, Wenbin Ge, Zhifang Guo, Qidong Huang, Jie Huang, Fei Huang, Binyuan Hui, Shutong Jiang, Zhaohai Li, Mingsheng Li, and 45 others. 2025. Qwen3-VL technical report. *arXiv preprint arXiv:2511.21631*.

Shihao Cai, Keqin Bao, Hangyu Guo, Jizhi Zhang, Jun Song, and Bo Zheng. 2024. GeoGPT4V: Towards geometric multi-modal large language models with geometric image generation. In *Conference on Empirical Methods in Natural Language Processing*, pages 750–766.

Google DeepMind. 2025. Gemini 3 Pro.

Chang Gao, Chujie Zheng, Xiong-Hui Chen, Kai Dang, Shixuan Liu, Bowen Yu, An Yang, Shuai Bai, Jingren Zhou, and Junyang Lin. 2025a. Soft adaptive policy optimization. *arXiv preprint arXiv:2511.20347*.

Jiahui Gao, Renjie Pi, Jipeng Zhang, Jiacheng Ye, Wanjun Zhong, Yufei Wang, Lanqing HONG, Jianhua Han, Hang Xu, Zhenguo Li, and Lingpeng Kong. 2025b. G-LLaVA: Solving geometric problem with multi-modal large language model. In *International Conference on Learning Representations*.

Wenyi Hong, Weihan Wang, Ming Ding, Wenmeng Yu, Qingsong Lv, Yan Wang, Yean Cheng, Shiyu Huang, Junhui Ji, Zhao Xue, Lei Zhao, Zhuoyi Yang, Xiaotao Gu, Xiaohan Zhang, Guanyu Feng, Da Yin, Zihan Wang, Ji Qi, Xixuan Song, and 6 others. 2024. CogVLM2: Visual language models for image and video understanding. *arXiv preprint arXiv:2408.16500*.

Wenyi Hong, Wenmeng Yu, Xiaotao Gu, Guo Wang, Guobing Gan, Haomiao Tang, Jiale Cheng, Ji Qi, Junhui Ji, Lihang Pan, Shuaiqi Duan, Weihan Wang, Yan Wang, Yean Cheng, Zehai He, Zhe Su, Zhen Yang, Ziyang Pan, Aohan Zeng, and 13 others. 2025. GLM-4.5V and GLM-4.1V-Thinking: Towards versatile

multimodal reasoning with scalable reinforcement learning. *arXiv preprint arXiv:2507.01006*.

Wenxuan Huang, Bohan Jia, Zijie Zhai, Shaosheng Cao, Zheyu Ye, Fei Zhao, Zhe Xu, Yao Hu, and Shaohui Lin. 2026. Vision-R1: Incentivizing reasoning capability in multimodal large language models. In *International Conference on Learning Representations*.

Zihan Huang, Tao Wu, Wang Lin, Shengyu Zhang, Jingyuan Chen, and Fei Wu. 2025. AutoGeo: Automating geometric image dataset creation for enhanced geometry understanding. *IEEE Transactions on Multimedia*.

Aishwarya Kamath, Johan Ferret, Shreya Pathak, Nino Vieillard, Ramona Merhej, Sarah Perrin, Tatiana Matejovicova, Alexandre Ramé, Morgane Rivière, Louis Rouillard, Thomas Mesnard, Geoffrey Cideron, Jean bastien Grill, Sabela Ramos, Edouard Yvinec, Michelle Casbon, Etienne Pot, Ivo Penchev, Gaël Liu, and 54 others. 2025. Gemma 3 technical report. *arXiv preprint arXiv:2503.19786*.

Nathan Lambert, Jacob Morrison, Valentina Pyatkin, Shengyi Huang, Hamish Ivison, Faeze Brahman, Lester James V Miranda, Alisa Liu, Nouha Dziri, Shane Lyu, and 1 others. 2024. Tulu 3: Pushing frontiers in open language model post-training. *arXiv preprint arXiv:2411.15124*.

Haotian Liu, Chunyuan Li, Yuheng Li, and Yong Jae Lee. 2024. Improved baselines with visual instruction tuning. *Conference on Computer Vision and Pattern Recognition*.

Wentao Liu, Qianjun Pan, Yi Zhang, Zhuo Liu, Ji Wu, Jie Zhou, Aimin Zhou, Qin Chen, Bo Jiang, and Liang He. 2025. CMM-Math: A chinese multimodal math dataset to evaluate and enhance the mathematics reasoning of large multimodal models. In *ACM International Conference on Multimedia*, page 12585–12591.

Pan Lu, Hritik Bansal, Tony Xia, Jiacheng Liu, Chunyuan Li, Hannaneh Hajishirzi, Hao Cheng, Kai-Wei Chang, Michel Galley, and Jianfeng Gao. 2024. MathVista: Evaluating mathematical reasoning of foundation models in visual contexts. In *International Conference on Learning Representations*.

Robert M May. 1976. Simple mathematical models with very complicated dynamics. *Nature*, 261(5560):459–467.

OpenAI. 2025. Introducing gpt-5.

Runqi Qiao, Qiuna Tan, Guanting Dong, MinhuiWu MinhuiWu, Chong Sun, Xiaoshuai Song, Jiapeng Wang, Zhuoma GongQue, Shanglin Lei, Yifan Zhang, and 1 others. 2025. We-Math: Does your large multimodal model achieve human-like mathematical reasoning? In *Annual Meeting of the Association for Computational Linguistics*, pages 20023–20070.

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*.

Wenhao Shi, Zhiqiang Hu, Yi Bin, Junhua Liu, Yang Yang, See Kiong Ng, Lidong Bing, and Roy Ka-Wei Lee. 2024. Math-LLaVA: Bootstrapping mathematical reasoning for multimodal large language models. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 4663–4680.

Yanpeng Sun, Shan Zhang, Wei Tang, Aotian Chen, Piotr Koniusz, Kai Zou, Yuan Xue, and Anton van den Hengel. 2025. MATHGLANCE: Multimodal large language models do not know where to look in mathematical diagrams. *arXiv preprint arXiv:2503.20745*.

Trieu H Trinh, Yuhuai Wu, Quoc V Le, He He, and Thang Luong. 2024. Solving olympiad geometry without human demonstrations. *Nature*, 625(7995):476–482.

Bin Wang, Chao Xu, Xiaomeng Zhao, Linke Ouyang, Fan Wu, Zhiyuan Zhao, Rui Xu, Kaiwen Liu, Yuan Qu, Fukai Shang, and 1 others. 2024a. MinerU: An open-source solution for precise document content extraction. *arXiv preprint arXiv:2409.18839*.

Ke Wang, Junting Pan, Weikang Shi, Zimu Lu, Houxing Ren, Aojun Zhou, Mingjie Zhan, and Hongsheng Li. 2024b. Measuring multimodal mathematical reasoning with MATH-Vision dataset. In *Conference on Neural Information Processing Systems Datasets and Benchmarks Track*.

Ke Wang, Junting Pan, Linda Wei, Aojun Zhou, Weikang Shi, Zimu Lu, Han Xiao, Yunqiao Yang, Houxing Ren, Mingjie Zhan, and Hongsheng Li. 2025a. MathCoder-VL: Bridging vision and code for enhanced multimodal mathematical reasoning. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 2505–2534.

Weiyun Wang, Zhangwei Gao, Lixin Gu, Hengjun Pu, Long Cui, Xingguang Wei, Zhaoyang Liu, Linglin Jing, Shenglong Ye, Jie Shao, Zhaokai Wang, Zhe Chen, Hongjie Zhang, Ganlin Yang, Haomin Wang, Qi Wei, Jinhui Yin, Wenhao Li, Erfei Cui, and 56 others. 2025b. InternVL3.5: Advancing open-source multimodal models in versatility, reasoning, and efficiency. *arXiv preprint arXiv:2508.18265*.

Xiaohan Wang, Mingze Yin, Yilin Zhao, Gang Liu, and Dian Li. 2026. LiveK12Bench: Have large multimodal models truly conquered high school-level examinations? *arXiv preprint arXiv:2605.26781*.

Zhiyu Wu, Xiaokang Chen, Zizheng Pan, Xingchao Liu, Wen Liu, Damai Dai, Huazuo Gao, Yiyang Ma, Chengyue Wu, Bingxuan Wang, Zhenda Xie, Yu Wu, Kai Hu, Jiawei Wang, Yaofeng Sun, Yukun Li, Yishi Piao, Kang Guan, Aixin Liu, and 8 others.

2024. DeepSeek-VL2: Mixture-of-experts vision-language models for advanced multimodal understanding. *arXiv preprint arXiv:2412.10302*.

Yue Xin, Wenyuan Wang, Rui Pan, Ruida Wang, Howard Meng, Renjie Pi, Shizhe Diao, and Tong Zhang. 2025. Generalizable geometric image caption synthesis. *arXiv preprint arXiv:2509.15217*.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, Chujie Zheng, Dayiheng Liu, Fan Zhou, Fei Huang, Feng Hu, Hao Ge, Haoran Wei, Huan Lin, Jialong Tang, and 41 others. 2025a. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*.

Jie Yang, Feipeng Ma, Zitian Wang, Dacheng Yin, Kang Rong, Fengyun Rao, and Ruimao Zhang. 2025b. WeThink: Toward general-purpose vision-language reasoning via reinforcement learning. *arXiv preprint arXiv:2506.07905*.

Mingze Yin, Yiheng Zhu, Jialu Wu, Jian Ma, Hanjing Zhou, Mingyang Li, Yuhua Zhou, Jintai Chen, Tingjun Hou, Jieping Ye, and Aimin Pan. 2026. Caduceus: Moe foundation models for unifying biological and natural language. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*.

Xiang Yue, Yuansheng Ni, Kai Zhang, Tianyu Zheng, Ruoqi Liu, Ge Zhang, Samuel Stevens, Dongfu Jiang, Weiming Ren, Yuxuan Sun, and 1 others. 2024. MMMU: A massive multi-discipline multimodal understanding and reasoning benchmark for expert AGI. In *Conference on Computer Vision and Pattern Recognition*.

Jingyi Zhang, Jiaxing Huang, Huanjin Yao, Shunyu Liu, Xikun Zhang, Shijian Lu, and Dacheng Tao. 2025. R1-VL: Learning to reason with multimodal large language models via step-wise group relative policy optimization. In *International Conference on Computer Vision*.

Renrui Zhang, Dongzhi Jiang, Yichi Zhang, Haokun Lin, Ziyu Guo, Pengshuo Qiu, Aojun Zhou, Pan Lu, Kai-Wei Chang, Yu Qiao, and 1 others. 2024. MathVerse: Does your multi-modal llm truly see the diagrams in visual math problems? In *European Conference on Computer Vision*, pages 169–186.

Hanjing Zhou, Mingze Yin, Danny Chen, Jian Wu, and Jintai Chen. 2025. Group-On: Boosting one-shot segmentation with supportive query. In *IEEE International Conference on Multimedia and Expo*.

A Large Language Model Usage

In this research, large language models are utilized as generation and proofreading tools to produce mathematical function reasoning data and refine academic English expression. The model development relies exclusively on publicly available, open-source foundation models from the Qwen series (Yang et al., 2025a; Bai et al., 2025). It is worth noting that the conceptual innovation and motivation underlying the study are conceived independently by the human author team, without any conceptual or cognitive inspiration from large language models. Multiple large language models, including GPT-5 (OpenAI, 2025), Gemini 3 Pro (DeepMind, 2025), Claude Opus 4.5 (Anthropic, 2025b), QWen3-VL-235B-A22B-Thinking (Bai et al., 2025), and GLM-4.1V-9B-Thinking (Hong et al., 2025), are used within the data generation engine to annotate function plots and design reasoning-oriented problem solving data. We employ large language models exclusively through the official API endpoints provided by respective vendors. All human–AI interactions are conducted in full compliance with all applicable terms of service and licensing conditions. We solemnly affirm that the constructed dataset is used exclusively for academic research and entails no commercial purpose. The manuscript is originally written by the human authors, and large language models are engaged only at the final stage to assist in polishing key sections of the language and word choice.

B Ethical Considerations

- **Copyright and Licensing:** Adherence to all applicable copyright and licensing requirements is mandatory.
- **Data Privacy:** Compliance with data protection laws and ethical standards in data handling is paramount. Annotators must refrain from collecting questions or content containing personal or sensitive information.
- **Model Development:** The developed model is built upon publicly available pre-trained Qwen series foundation models (Yang et al., 2025a; Bai et al., 2025). We disclose the exact model versions, licensing terms, and training details, and we comply with all license terms and usage restrictions.

C Reproducibility Statement

To alleviate modality interference in mathematical function reasoning, we explicitly decouple the architecture into two specialized components: a visual math parser and a textual math reasoner. An

ablation study examining the performance ceiling of a unified multimodal architecture is presented in Sec. 3.6, Single-MLLM Bottlenecks. Specifically, FUNC-R1-Lite utilizes Qwen3-VL-4B as the visual math parser, and adopts Qwen3-8B to serve as the textual math reasoner for logical reasoning computation. FUNC-R1-Pro employs Qwen3-VL-8B as the visual math parser and Qwen3-32B as the textual math reasoner. For the visual math parser, the SFT stage runs for 3 epochs with batch size 8 and learning rate $1e{-}5$. For the textual math reasoner, the RL stage lasts for 5 epochs with batch size 64, learning rate $1e{-}6$, and a rollout number $N = 8$.

D Detailed Prompt Template

Problem-Solving Prompt

You are a mathematics expert. Your task is to combine the given function expression with the attached graph, carry out rigorous mathematical reasoning, and provide a complete solution. In your reasoning and calculations, explicitly use standard theoretical properties of mathematical function. Present your step-by-step solution process and mark your final answer according to the following rules:
- Single-choice: place the correct option letter inside `\boxed{}`.
- Multiple-choice: place all correct option letters inside one `\boxed{}`.
- Fill-in-the-blank: enclose each blank's answer in a separate `\boxed{}`.
- Open-ended: enclose each sub-question's answer in a separate `\boxed{}`.

Answer Extraction Prompt

You are an expert grader. Determine whether the student's answer is consistent with the standard answer. In accordance with the following grading criteria:
- Judge each sub-question independently.
- Accept equivalent expressions (mathematical or textual). If equivalence cannot be determined, mark the answer incorrect.
- Do not re-derive the answer; for questions with definitive results, compare the final answer only. For non-definitive questions (*e.g.*, proofs), evaluate whether the reasoning approach is sound.
- Output the number of correctly answered sub-questions, enclosed in `\boxed{}`.

E Additional Experimental Details

LLM baselines. To facilitate comprehensive comparison, we diversely incorporate the current state-of-the-art open-source and closed-source MLLMs. Particularly, we select LLaVA series (Liu et al., 2024), InternVL series (Wang et al., 2025b), Qwen-VL series (Bai et al., 2025), DeepSeek-VL2 (Wu et al., 2024), GLM-4.1V-Thinking (Hong et al., 2025), CogVLM2 (Hong et al., 2024), We-Think (Yang et al., 2025b), R1-VL (Zhang et al., 2025), and Math-LLaVA (Shi et al., 2024) for open-source models. Additionally, for closed-source models, we employ Qwen-VL-Max (Bai et al., 2023), Gemini 2.5 Flash, Gemini 3 Flash (DeepMind, 2025), Claude Opus 4.1 (Anthropic, 2025a), GPT-4o, GPT-4V (Achiam et al., 2024), and GPT-5 (OpenAI, 2025).

Evaluation benchmarks. Experiments are holistically conducted on in-distribution benchmark `FuncReasoning`, as well as diverse out-of-distribution multimodal mathematical reasoning benchmarks, including MathVista (Lu et al., 2024), Math-Vision (Wang et al., 2024b), MathVerse (Zhang et al., 2024), and We-Math (Qiao et al., 2025). We consistently employ an open-ended evaluation protocol whereby the final answer is extracted from the LLM-based solver's output and its correctness is established through comparison with the ground-truth solution. Implementation details are provided in the Appendix C.

F Data Generative Design Principles

When executing the knowledge concept sedimentation, LLMs (*i.e.*, GPT-5 (OpenAI, 2025), Gemini 3 Pro (DeepMind, 2025)) receive the sample-level function annotations and the knowledge concept panel from the previous iteration. Instead of relying on downstream reasoning feedback, the thinkers directly optimize the knowledge content based on the three following steps:

(i) *Parsing:* Thinker parses the sample-level concept into a structured key-value format.

(ii) *Optimization:* Thinker scrutinizes the extracted knowledge concepts against the fundamental requirements. In the initial round, it directly condenses raw descriptions; in subsequent rounds, it filters out superficial, redundant, or misleading keys that fail to reflect intrinsic mathematical properties.

(iii) *Integration:* The remaining high-value knowledge items are aggregated into a consolidated knowledge concept panel, marking the completion of the current iteration.

To ensure rich coverage and high-quality design throughout the self-evolving knowledge consolidation engine, we construct mathematical function-oriented question-answer pairs following three fundamental requirements:

(i) *Logical Reachability*: The questions must be

framed within the constraints of the given function plots and captions, allowing for final answers to be reached logically.

(ii) *Intrinsic Focus*: Generative LLMs is guided to focus on the intrinsic properties of mathematical functions, avoiding dependence on superficial visual cues.

(iii) *Concept Assessment*: Designed function problems tend to assess reasoning abilities related to at least one relevant knowledge concept.

This engine offers a rigorous, reliable pipeline for generating high-quality multimodal mathematical reasoning data. The resulting `FuncCaption` and `FuncReasoning` comprise 3,519 entries, providing a solid foundation for downstream multimodal mathematical function reasoning applications.

G Empirical Analysis of Mathematical Functions

Herein, we elucidate our research motivation by highlighting the unique challenges of function reasoning and the necessity of task-specific reward signals. The mathematical function's explicit parameterization, computational complexity, and robust measurability pose reasoning obstacles that diverge from traditional problem formats. We conduct a rigorous, systematic comparative analysis to characterize GPT-5's problem-solving error distribution across Multiple problem categories from MathVerse (*i.e.*, plane geometry, solid geometry, mathematical function). We also integrate the novel `FuncReasoning` dataset into the analysis, and detailed composition statistics are shown in Table 5.

Problem Category	Number & Percentage
MathVerse (Zhang et al., 2024)	
- Plane geometry	1,746 (66.8%)
- Solid geometry	332 (12.7%)
- **Mathematical function**	**534 (20.5%)**
`FuncReasoning`	
- **Mathematical function**	**3,519 (100%)**

Table 5: Key statistics of the mathematical database.

A comprehensive depiction of error-type proportions for distinct problem categories is shown in Figure 6. With increasing problem difficulty (*i.e.*, from plane geometry through solid geometry to function), the predominant error types transition from computation deviation to visual misperception and knowledge deficit. This further reinforces

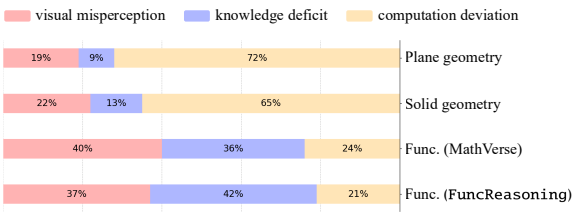


Figure 6: Attribution distribution of mathematical problem-solving errors.

the limitations of existing approaches that optimize exclusively for rigorous computational reasoning. As detailed in Sec. 2.3.2, we tailored comprehensive reward signals within the PATO strategy (*i.e.*, visual perception correction, theoretical knowledge guidance, and verifiable problem-solving) to deliver targeted interventions for challenges encountered in function-oriented complex reasoning.

H Supplementary Baseline Details

Additional introductory details on the included baseline models are provided in the form of model cards, as illustrated in Table 6.

I Mathematical Reasoning Examples

We assemble `FuncCaption` and `FuncReasoning` as comprehensive corpora as possible of diverse function diagrams to facilitate FUNC-R1 's thorough acquisition of foundational mathematical theoretical knowledge. As illustrated in Figure 7, we provide a detailed presentation of high-quality mathematical visualizations that encompass the discrete, linear, nonlinear, and composite functions.

LLaVA series (Liu et al., 2024)

By leveraging high-quality multi-turn visual instruction data and synthetic augmentation, LLaVA scales effectively and improves performance on complex visual question answering, fine-grained image description, and multimodal reasoning. This framework features a lightweight, efficient training pipeline and an open data/code ecosystem that facilitate rapid iteration.

InternVL series (Wang et al., 2025b)

A family of general-purpose vision-language models for industrial and research applications, emphasizing large-scale pretraining and fine-grained region alignment. Integrating high-resolution feature aggregation with dynamic receptive fields, the models enhance dense-scene and OCR-intensive comprehension and enable robust long-context multi-image reasoning and cross-page document QA.

CogVLM2 (Hong et al., 2024)

CogVLM2 emphasizes stronger visual encoding and language alignment. Trained with larger-scale multimodal instruction data and high-resolution inputs, it exhibits robustness on web page screenshots, GUIs, charts, and scene text.

Qwen-VL series (Bai et al., 2025)

Developed by the Qwen multimodal team, this system integrates strong Chinese language proficiency with deep visual understanding, focusing on OCR, chart analysis, practical QA, and tool use across mobile and server deployments. It leverages high-resolution image encoding and region-level supervision, supports multi-image, multi-turn, and multi-format inputs.

DeepSeek-VL2 (Wu et al., 2024)

DeepSeek-VL2 prioritizes efficient alignment and robust reasoning. Through refined visual pretraining and instruction distillation, it reduces hallucinations and improves complex QA performance.

GLM-4.1V-Thinking (Hong et al., 2025)

The thinking variant of the GLM-4 series emphasizes explicit multi-step chain-of-thought and visual reasoning, supporting coordinated image–text planning and intermediate reasoning traces to enhance interpretability on complex tasks.

Math-LLaVA (Shi et al., 2024)

Strengthened by math-domain instruction data and structured annotations, Math-LLaVA improves symbol recognition, quantitative derivation, while improving visual–semantic alignment to reduce formula recognition errors.

WeThink (Yang et al., 2025b)

By integrating process rewards and preference alignment, WeThink improves robustness in visual reasoning and decision-making. The framework supports multi-image, multi-evidence integration to mitigate hallucination and inconsistency.

R1-VL (Zhang et al., 2025)

R1-VL emphasizes stepwise reasoning, error correction, and reflective alignment to improve performance on long-horizon tasks. The approach yields notable gains on geometric math, scientific chart interpretation, and quantitative visual analysis.

Table 6: Detailed model introduction card for incorporated baselines.

GPT-5

First, simplify the function: $f(x) = 6\cos^2\frac{\omega x}{2} + \sqrt{3}\sin\omega x - 3 = 3(1+\cos\omega x) + \sqrt{3}\sin\omega x - 3 = 3\cos\omega x + \sqrt{3}\sin\omega x$. Then write it in the same-phase form: let $R\cos(\omega x - \varphi) = R(\cos\omega x\cos\varphi + \sin\omega x\sin\varphi)$. Comparing coefficients gives $R = \sqrt{3^2 + (\sqrt{3})^2} = 2\sqrt{3}$, and $\tan\varphi = \frac{\sqrt{3}}{3} \Rightarrow \varphi = \frac{\pi}{6}$, hence $f(x) = 2\sqrt{3}\cos\left(\omega x - \frac{\pi}{6}\right)$. (1) Let the two intersections with the x-axis in one period be B, C, and the highest point be A; then the zeros satisfy $3\cos t + \sqrt{3}\sin t = 0 \iff \tan t = -\sqrt{3}$, taking two adjacent solutions as $\omega x_B = -\frac{\pi}{3}$, $\omega x_C = \frac{2\pi}{3}$, and the maximum point satisfies $\omega x - \frac{\pi}{6} = 0$, so $A\left(\frac{\pi}{6\omega}, 2\sqrt{3}\right)$, $B\left(-\frac{\pi}{3\omega}, 0\right)$, $C\left(\frac{2\pi}{3\omega}, 0\right)$. Since $\triangle ABC$ is an equilateral triangle, set $AB = BC$, then $AB^2 = \left(\frac{\pi}{6\omega} + \frac{\pi}{3\omega}\right)^2 + (2\sqrt{3})^2 = \left(\frac{\pi}{2\omega}\right)^2 + 12$, $BC^2 = \left(\frac{\pi}{\omega}\right)^2$, Equilateral gives $\left(\frac{\pi}{2\omega}\right)^2 + 12 = \left(\frac{\pi}{\omega}\right)^2 \Rightarrow 12 = \frac{3\pi^2}{4\omega^2} \Rightarrow \omega = \frac{\pi}{4}$. (2) We know $f(x) = 2\sqrt{3}\cos\left(\frac{\pi}{4}x - \frac{\pi}{6}\right)$. Given $f(x_0) = \frac{8\sqrt{3}}{5}$, set $t := \frac{\pi}{4}x_0 - \frac{\pi}{6}$, $2\sqrt{3}\cos t = \frac{8\sqrt{3}}{5} \Rightarrow \cos t = \frac{4}{5}$. Thus $f(x_0+1) = 2\sqrt{3}\cos\left(t + \frac{\pi}{4}\right) = 2\sqrt{3}\left(\cos t\cos\frac{\pi}{4} - \sin t\sin\frac{\pi}{4}\right)$. Since the figure shows that x_0 lies to the left of the highest point and the corresponding phase is on the rising branch, thus we consider $t = \frac{\pi}{4}x_0 - \frac{\pi}{6} \in (0, \pi)$, hence $\sin t > 0$, and from $\cos t = \frac{4}{5}$, we get $\sin t = \frac{3}{5}$. Substituting back yields $f(x_0+1) = 2\sqrt{3}\left(\frac{4}{5}\cdot\frac{\sqrt{2}}{2} - \frac{3}{5}\cdot\frac{\sqrt{2}}{2}\right) = 2\sqrt{3}\cdot\frac{\sqrt{2}}{10} = \frac{\sqrt{6}}{5}$. Final answers: (1) $\omega = \frac{\pi}{4}$; (2) $f(x_0+1) = \frac{\sqrt{6}}{5}$.

Gemini-3-Pro

(1) First, use trigonometric identity transformations to simplify the analytic expression of the function $f(x)$: $f(x) = 6\cos^2\frac{\omega x}{2} + \sqrt{3}\sin\omega x - 3 = 6\frac{1+\cos\omega x}{2} + \sqrt{3}\sin\omega x - 3 = 3\cos\omega x + \sqrt{3}\sin\omega x = 2\sqrt{3}\sin\left(\omega x + \frac{\pi}{3}\right)$. From the analytic expression, it can be seen that the maximum value (amplitude) of the function is $2\sqrt{3}$, which means the ordinate of the highest point A of the graph is $y_A = 2\sqrt{3}$. According to the problem, $\triangle ABC$ is an equilateral triangle, with A being the highest point and B, C being the intersection points of the graph with the x-axis; therefore, the height h of $\triangle ABC$ is $h = 2\sqrt{3}$. In an equilateral triangle, the relationship between the height h and the side length a is $h = \frac{\sqrt{3}}{2}a$. Substituting the data gives $2\sqrt{3} = \frac{\sqrt{3}}{2}BC$, solving for which yields $BC = 4$. Observing the function graph, points B and C are the intersection points of the function curve with the x-axis. We obtain $T = BC = 4$. From the period formula $T = \frac{2\pi}{\omega}$, we get $\frac{2\pi}{\omega} = 4$, which solves to $\omega = \frac{\pi}{2}$. (2) Next, substituting $\omega = \frac{\pi}{2}$ from solved (1) into the function's analytic expression gives $f(x) = 2\sqrt{3}\sin\left(\frac{\pi}{2}x + \frac{\pi}{3}\right)$. Given $f(x_0) = \frac{8\sqrt{3}}{5}$, that is: $2\sqrt{3}\sin\left(\frac{\pi}{2}x_0 + \frac{\pi}{3}\right) = \frac{8\sqrt{3}}{5} \implies \sin\left(\frac{\pi}{2}x_0 + \frac{\pi}{3}\right) = \frac{4}{5}$. Let $\alpha = \frac{\pi}{2}x_0 + \frac{\pi}{3}$, then $\sin\alpha = \frac{4}{5}$. We need to find the value of $f(x_0+1)$: $f(x_0+1) = 2\sqrt{3}\sin\left[\frac{\pi}{2}(x_0+1) + \frac{\pi}{3}\right] = 2\sqrt{3}\sin\left[\left(\frac{\pi}{2}x_0 + \frac{\pi}{3}\right) + \frac{\pi}{2}\right] = 2\sqrt{3}\sin\left(\alpha + \frac{\pi}{2}\right)$. Using the induction formula $\sin(\alpha + \frac{\pi}{2}) = \cos\alpha$, we have $f(x_0+1) = 2\sqrt{3}\cos\alpha$. Given the *Pythagorean identity* $\cos^2\alpha + \sin^2\alpha = 1$, we know that $\cos\alpha = \pm\frac{3}{5}$. Therefore, the final calculation result is:$f(x_0+1) = 2\sqrt{3}\cdot\left(\pm\frac{3}{5}\right) = \pm\frac{6\sqrt{3}}{5}$.

Table 7: The detailed solution provided by leading closed-source models.

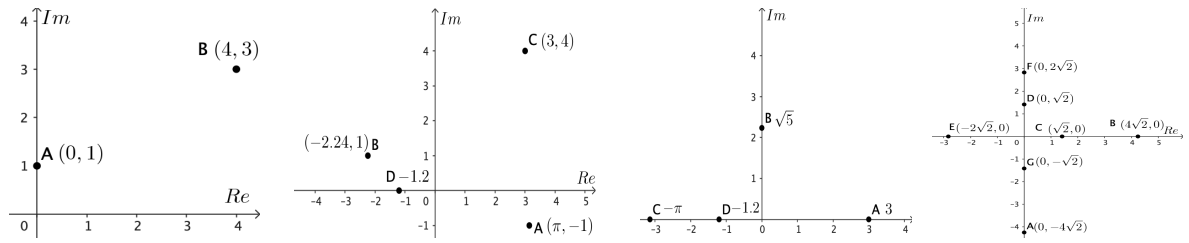
Qwen3-VL-235B-A22B

First, simplify the function $f(x) = 6\cos^2\frac{\omega x}{2} + \sqrt{3}\sin\omega x - 3$ using the identity $\cos^2\theta = \frac{1+\cos 2\theta}{2}$ to obtain $f(x) = 6\frac{1+\cos\omega x}{2} + \sqrt{3}\sin\omega x - 3 = 3 + 3\cos\omega x + \sqrt{3}\sin\omega x - 3 = 3\cos\omega x + \sqrt{3}\sin\omega x$, and then express it in a single trigonometric form: let $f(x) = R\sin(\omega x + \varphi)$, where $R = \sqrt{3^2 + (\sqrt{3})^2} = \sqrt{12} = 2\sqrt{3}$, so the maximum value is $2\sqrt{3}$, which is the y-coordinate of point A. Since $\triangle ABC$ is an equilateral triangle, and A is the top vertex, B and C are the base endpoints, and from the image, B is approximately at $x = -0.5$, and C is at $x = 1.5$, we could assume $BC = 2$, and thus the half-period is 2, so the full period is $T = 4$. From $T = \frac{2\pi}{\omega} = 4$, we obtain the incorrect $\omega = \frac{\pi}{2}$. The function equation will be $f(x) = 3\cos\frac{\pi}{2}x + \sqrt{3}\sin\frac{\pi}{2}x$, which can be rewritten as $f(x) = 2\sqrt{3}\sin\left(\frac{\pi}{2}x + \frac{\pi}{3}\right)$ using the auxiliary angle formula: $\tan\varphi = \sqrt{3} \Rightarrow \varphi = \frac{\pi}{3}$. Given $f(x_0) = \frac{8\sqrt{3}}{5}$, substituting yields $2\sqrt{3}\sin\left(\frac{\pi}{2}x_0 + \frac{\pi}{3}\right) = \frac{8\sqrt{3}}{5} \Rightarrow \sin\left(\frac{\pi}{2}x_0 + \frac{\pi}{3}\right) = \frac{4}{5}$. Since $x_0 \in \left(-\frac{10}{3}, \frac{2}{3}\right)$, we have $\frac{\pi}{2}x_0 + \frac{\pi}{3} \in \left(-\frac{5\pi}{3} + \frac{\pi}{3}, \frac{\pi}{3} + \frac{\pi}{3}\right) = \left(-\frac{4\pi}{3}, \frac{2\pi}{3}\right)$. Within this interval, the solutions to $\sin\theta = \frac{4}{5}$ are $\theta = \arcsin\frac{4}{5}$ or $\theta = \pi - \arcsin\frac{4}{5}$, but the latter is approximately $\pi - 0.927 \approx 2.214 > \frac{2\pi}{3} \approx 2.094$, which lies outside the interval; hence the only valid solution is $\theta = \arcsin\frac{4}{5} \in \left(0, \frac{\pi}{2}\right)$, and subsequently in this case, we have $\cos\theta = \frac{3}{5}$. We then compute $f(x_0 + 1) = 2\sqrt{3}\sin\left(\frac{\pi}{2}(x_0 + 1) + \frac{\pi}{3}\right) = 2\sqrt{3}\sin\left(\theta + \frac{\pi}{2}\right) = 2\sqrt{3}\cos\theta = 2\sqrt{3} \cdot \frac{3}{5} = \frac{6\sqrt{3}}{5}$. Therefore, the final result is: (1) $\omega = \frac{\pi}{2}$; (2) $f(x_0 + 1) = \frac{6\sqrt{3}}{5}$.

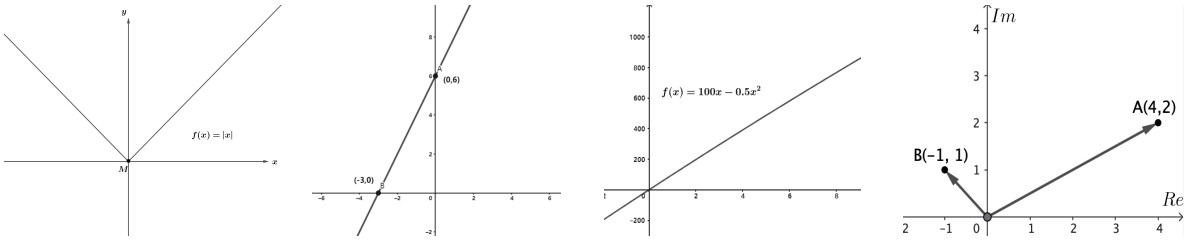
FUNC-R1

In the Cartesian coordinate system, a sinusoidal curve is plotted, displaying a portion of the waveform within one period. The curve intersects the x-axis at point B to the left of the y-axis and at point C to the right of the y-axis, with the highest point of the curve marked as A. Points A, B, and C are connected to form a triangle $\triangle ABC$, which the problem specifies is an equilateral triangle, and the side length of this triangle is equal to half the period of the trigonometric function. First, we need to simplify the analytical expression of the function $f(x)$ using trigonometric identities: $f(x) = 6\cos^2\frac{\omega x}{2} + \sqrt{3}\sin\omega x - 3 = \sqrt{3}\sin\omega x + 3\cos\omega x = 2\sqrt{3}\sin\left(\omega x + \frac{\pi}{3}\right)$. From the simplified expression, the maximum value of the function is $2\sqrt{3}$, which corresponds to the ordinate of the highest point A, so $y_A = 2\sqrt{3}$. Since $\triangle ABC$ is an equilateral triangle and BC lies on the x-axis, we have $2\sqrt{3} = \frac{\sqrt{3}}{2}BC \implies BC = 4$. Thus, the period of the trigonometric function is $T = 2BC = 8$. Using the period formula $T = \frac{2\pi}{\omega}$, we get $\frac{2\pi}{\omega} = 8 \implies \omega = \frac{\pi}{4}$. From (1), we know the function is $f(x) = 2\sqrt{3}\sin\left(\frac{\pi}{4}x + \frac{\pi}{3}\right)$. Given $f(x_0) = \frac{8\sqrt{3}}{5}$, substituting this into the expression yields $\sin\left(\frac{\pi}{4}x_0 + \frac{\pi}{3}\right) = \frac{4}{5}$. Next, we determine the range of the angle to find the sign of the cosine value. Given $x_0 \in \left(\frac{-10}{3}, \frac{2}{3}\right)$, then $\alpha = \frac{\pi}{4}x_0 + \frac{\pi}{3} \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Since $\sin\alpha = \frac{4}{5} > 0$ and $\alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, so the cosine value is $\cos\alpha = \frac{3}{5}$. Now we can calculate: $f(x_0 + 1) = 2\sqrt{3}\left[\sin\left(\frac{\pi}{4}x_0 + \frac{\pi}{3}\right)\cos\frac{\pi}{4} + \cos\left(\frac{\pi}{4}x_0 + \frac{\pi}{3}\right)\sin\frac{\pi}{4}\right]$. Substituting the known values: $f(x_0 + 1) = 2\sqrt{3}\left(\frac{4}{5} \cdot \frac{\sqrt{2}}{2} + \frac{3}{5} \cdot \frac{\sqrt{2}}{2}\right) = \frac{7\sqrt{6}}{5}$. Final Answer: (1) $\omega = \frac{\pi}{4}$; (2) $f(x_0 + 1) = \frac{7\sqrt{6}}{5}$.

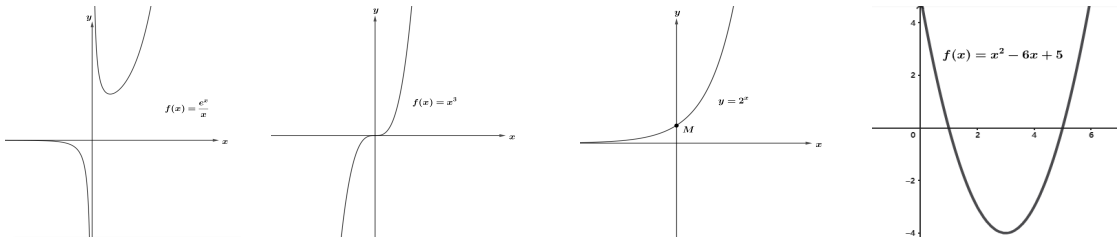
Table 8: The elaborate response generated by prominent open-source models.



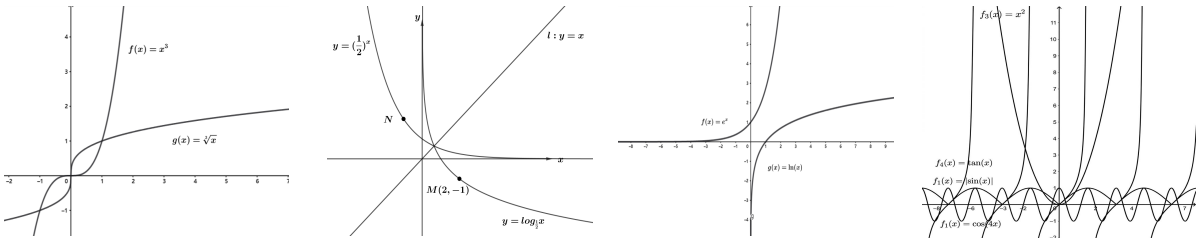
(a) Examples of the mathematical diagram for discrete function.



(b) Examples of the mathematical diagram for linear function.



(c) Examples of the mathematical diagram for nonlinear function.



(d) Examples of the mathematical diagram for composite function.

Figure 7: Function-oriented mathematical visualization of Synthetic Datasets.