

000 001 002 003 004 005 WE-MATH 2.0: A VERSATILE MATHBOOK SYSTEM FOR 006 INCENTIVIZING VISUAL MATHEMATICAL REASONING 007 008 009

010 **Anonymous authors**
011 Paper under double-blind review
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

ABSTRACT

Multimodal Large Language Models (MLLMs) have demonstrated impressive capabilities across various tasks, but still struggle with complex mathematical reasoning. Existing research primarily focuses on dataset construction and method optimization, often overlooking two critical aspects: comprehensive knowledge-driven design and model-centric data space modeling. In this paper, we introduce **WE-MATH 2.0**, a unified system that integrates a structured mathematical knowledge system, model-centric data space modeling, and a reinforcement learning (RL)-based training paradigm to comprehensively enhance the mathematical reasoning abilities of MLLMs. The key contributions of We-Math 2.0 are fourfold: **(1) MathBook Knowledge System:** We construct a five-level hierarchical system encompassing 491 knowledge points and 1,819 fundamental principles. **(2) MathBook-Standard & Pro:** We develop MathBook-Standard, a dataset that ensures broad conceptual coverage and flexibility through dual expansion. Additionally, we define a three-dimensional difficulty space and generate 7 progressive variants per problem to build **MathBook-Pro**, a challenging dataset for robust training. **(3) MathBook-RL:** We propose a two-stage RL framework comprising: (i) Cold-Start Fine-tuning, which aligns the model with knowledge-oriented chain-of-thought reasoning; and (ii) Progressive Alignment RL, leveraging average-reward learning and dynamic data scheduling to achieve progressive alignment across difficulty levels. **(4) MathBookEval:** We introduce a comprehensive benchmark covering all 491 knowledge points with diverse reasoning step distributions. Experimental results show that MathBook-RL performs competitively with existing baselines on four widely-used benchmarks and achieves strong results on MathBookEval, suggesting promising generalization in mathematical reasoning.

1 INTRODUCTION

Large Language models (LLMs) have demonstrated remarkable capabilities across a wide range of tasks (Achiam et al., 2023; DeepSeek-AI, 2025; Jaech et al., 2024; Wan et al., 2024; Trinh et al., 2024; Xin et al., 2024). Building on this foundation, Multimodal Large Language Models (MLLMs) have shown impressive performance in visual question answering (VQA) (Bai et al., 2025b; Zhu et al., 2025; Guo et al., 2025), optical character recognition (OCR) (Ye et al., 2023b;a; Wei et al., 2024), and object detection (Liu et al., 2025b; Ren et al., 2024). However, MLLMs still face difficulties with complex reasoning tasks, particularly in visual mathematical problem-solving, where generalization remains a fundamental challenge (Lu et al., 2023; Zhang et al., 2024a; Wang et al., 2024).

Recent efforts to enhance mathematical reasoning in MLLMs have primarily focused on three directions: dataset construction Lu et al. (2021a); Zhang et al. (2024b); Shi et al. (2024a); Wang et al. (2025), preference optimization Zhuang et al. (2024); Luo et al. (2025), and reinforcement learning (RL) Huang et al. (2025); Meng et al. (2025). Foundation approaches aggregated datasets from diverse mathematical domains Shi et al. (2024b). Subsequent efforts introduce structured supervision formats (e.g., Chain-of-Thought (CoT)) combined with preference optimization to guide step-by-step reasoning Zhuang et al. (2025). More recently, RL-based studies with curriculum-based training have been employed to further improve model performance on complex reasoning tasks Huang et al. (2025); Wan et al. (2025). Despite this progress, several fundamental challenges remain:

054 **Table 1: Comparison of We-Math 2.0 with several representative multimodal mathematical datasets.**

055 Dataset	056 Usage	057 Data Annotation	058 Knowledge-level Annotation	059 Principle-level Annotation	060 Difficulty Levels
057 Geometry3K Lu et al. (2021a)	058 Training Data & Benchmark	059 Manual	060 -	061 -	062 -
057 MathV360K Shi et al. (2024a)	058 Training Data	059 Collection	060 -	061 -	062 4
057 We-Math Qiao et al. (2024a)	058 Benchmark	059 Manual	060 ✓	061 -	062 -
057 GeoSense Xu et al. (2025)	058 Benchmark	059 Manual	060 ✓	061 ✓	062 -
060 We-Math 2.0 (Ours)	061 Training Data & Benchmark	062 Manual	063 ✓	064 ✓	065 8

062 **(1) Lack of a comprehensive knowledge system:** Existing MLLMs show uneven performance
063 across different subfields of math reasoning. Lu et al. (2023); Wang et al. (2024) Unfortunately,
064 current datasets suffer from limited coverage of knowledge points and domain diversity, underscoring
065 the necessity of establishing a more systematic knowledge system.

066 **(2) Lack of model-centric difficulty modeling:** Existing multimodal training datasets primarily
067 perform difficulty annotation based on human learning stages Meng et al. (2025). However, recent
068 studies (Lei et al., 2024; Qiao et al., 2024b; Lu et al., 2023) reveal that MLLMs do not exhibit
069 learning patterns that align well with these human-defined levels. This highlights the need for a more
070 model-centric approach to modeling data difficulty.

071 **(3) Lack of emphasis on reasoning generalization:** MLLMs are capable of solving complex
072 problems, but perform poorly on corresponding subproblems Qiao et al. (2024b) as well as on similar,
073 same-type tasks Zou et al. (2024). This underscores the current training methods' focus on problem
074 memorization rather than fostering reasoning generalization.

075 To address these limitations, we introduce **WE-MATH 2.0**, a versatile framework that combines a
076 structured mathematical knowledge system, model-centric data space modeling, and a reinforcement
077 learning-based training paradigm to comprehensively improve MLLM's reasoning capabilities (see
078 Table 1. In detail, we begin by establishing the **MathBook Knowledge System**, a five-level hierarchy
079 comprising 491 knowledge points and 1,819 fundamental principles (see Figure 1). This structure is
080 systematically derived from sources such as Wikipedia and open-source textbooks, refined through
081 hierarchical clustering, and further revised by human experts.

082 Building on this foundation, we introduce **MathBook-Standard**, a dataset featuring comprehensive
083 annotations at the level of 1,819 knowledge principles, along with carefully curated problems to
084 ensure broad, balanced coverage, particularly in underrepresented mathematical domains. To foster
085 deeper conceptual understanding, MathBook-Standard employs dual expansions: "*multi-images per*
086 *question*" and "*multi-questions per image*", enabling diverse problem sets that achieve conceptual
087 flexibility. Crucially, we propose a pivotal three-dimensional difficulty modeling framework that
088 redefines mathematical problem construction. By explicitly modeling "step complexity", "visual
089 complexity" and "contextual complexity", each problem is systematically expanded into seven
090 difficulty levels to form **MathBook-Pro**. This design enables structured, progressive learning for
091 MLLMs, laying a strong foundation for improved reasoning across difficulty levels.

092 To further enhance MLLMs' general mathematical reasoning ability, we propose **MathBook-RL**, a
093 two-stage reinforcement learning framework for progressive and robust training:

094 **(1) Cold-Start Fine-tuning:** We first adopt a supervised fine-tuning that guides the MLLM to
095 learn knowledge-oriented CoT reasoning, internalizing it to acquire conceptual understanding and
096 structured problem-solving paradigms. **(2) Progressive Alignment RL:** We propose a curriculum-
097 based RL paradigm. Leveraging the "*one-question-multi-image*" and knowledge-point features in
098 MathBook-Standard, we first align the model's analogical reasoning by introducing an average
099 reward mechanism. Building on this foundation, we progressively train the MLLM on MathBook-Pro
100 and further introduce two dynamic scheduling strategies: **i) Knowledge Increment Scheduling:**
101 When errors occur due to complex reasoning steps, the model is adaptively redirected to relevant
102 incremental-step samples in MathBook-Standard. **ii) Modality Increment Scheduling:** When errors
103 stem from increased modality complexity, the model is guided through single-modality incremental
104 problems. This targeted curriculum enables effective knowledge transfer across difficulty levels.

105 To comprehensively evaluate MLLMs' reasoning capability, we introduce **MathBookEval**, a bench-
106 mark covering all 491 knowledge points with diverse step distributions. Experimental results show
107 that MathBook-RL performs competitively with existing baselines on four widely used benchmarks
and substantially improves generalization and robustness. In summary, our contributions are:

- 108 • We propose the **MathBook Knowledge System**, a five-level hierarchical framework with 491
109 knowledge points and 1,819 fundamental principles, enabling systematic and comprehensive
110 mathematical knowledge supervision.
- 111 • We develop **MathBook-Standard** and **MathBook-Pro**, two novel datasets that combine com-
112 prehensive step-wise annotation, dual expansions for conceptual flexibility, and a principled
113 three-dimensional difficulty modeling framework for structured and progressive learning.
- 114 • We introduce **MathBook-RL**, a two-stage RL framework that integrates structured knowledge
115 supervision and dynamic data scheduling, improving the reasoning capabilities of MLLMs.
- 116 • We present **MathBookEval**, a benchmark designed to comprehensively evaluate model reasoning
117 across diverse knowledge points and step distributions. Extensive experiments demonstrate that
118 our approach achieves remarkable performance in both generalization and robustness.

121 2 RELATED WORK

124 **Visual Mathematical Reasoning.** Recently, visual mathematical reasoning has advanced rapidly Shi
125 et al. (2024a); Zhang et al. (2024b); Zhuang et al. (2024); Han et al. (2024); Luo et al. (2025).
126 Benchmarks such as MathVista Lu et al. (2023) and MathVision Wang et al. (2024) assess overall
127 performance, while MathVerse Zhang et al. (2024a) and Dynamath Zou et al. (2024) examine reasoning
128 mechanisms and robustness. Methodologically, progress has been made through visual-textual
129 alignment Shi et al. (2024a); Zhang et al. (2024b); Wang et al. (2025), step-wise reasoning Zhuang
130 et al. (2024); Luo et al. (2025), and RL-based optimization Huang et al. (2025); Zhang et al. (2025);
131 Meng et al. (2025); Chen et al. (2025); Liu et al. (2025a); AI et al. (2025); Wan et al. (2025); Zheng
132 et al. (2025); Yang et al. (2025a); Team et al. (2025a); Hong et al. (2025); Team et al. (2025b), which
133 show promising gains on complex tasks. However, robust and generalizable visual reasoning remains
134 an open challenge. Therefore, we propose a systematic, model-centric knowledge system, integrate it
135 with RL-based alignment and a new dataset, aiming to provide fresh insights for the community.

136 3 WE-MATH 2.0

138 **Overview.** In this section, we introduce **WE-MATH 2.0**, a unified system designed to advance
139 visual mathematical reasoning in MLLMs, developed from three key aspects: (1) We construct a five-
140 level **MathBook Knowledge System** (§3.1), systematically organizing 491 knowledge points and
141 1,819 fundamental principles for comprehensive mathematical supervision. (2) We propose a **Multi-
142 Dimensional data construction** pipeline (§3.2), incorporating seed problem construction, variant
143 expansion, and principled three-dimensional difficulty modeling. (3) We introduce **MathBookEval**
144 (§3.3), a benchmark aligned with our knowledge system for systematic evaluation.

146 3.1 MATHBOOK KNOWLEDGE SYSTEM

148 **System Overview.** We construct a five-level hierarchical **MathBook Knowledge System** orga-
149 nized by the “*Definition-Theorem-Application*” paradigm (Fitzpatrick, 2008). The core is a set of
150 knowledge points $\mathcal{K} = \{k_1, k_2, \dots, k_N\}$, $N = 491$, spanning primary to university mathematics.
151 Each k_i is associated with a set of fundamental principles $\mathcal{P}_i = \{p_{i1}, \dots, p_{im_i}\}$, $m_i \in [1, 7]$, where
152 $\mathcal{P} = \bigcup_{i=1}^N \mathcal{P}_i$ and $|\mathcal{P}| = 1,819$. (Figure 3 illustrate examples of principles within the MathBook
153 Knowledge System, while Figure 8 shows how problems are aligned with the system.)

154 **Hierarchical Construction via Human-AI Collaboration.** We construct \mathcal{K} through a hybrid
155 process. Human experts first design an initial structure $\mathcal{K}^{\text{human}}$ based on authoritative sources,
156 including textbooks, Wikipedia, and national curriculum standards. In parallel, we sample 30K
157 problems from the existing math dataset Lu et al. (2021b); Johnson et al. (2017); Gao et al. (2023);
158 Shi et al. (2024a); Peng et al. (2024); Luo et al. (2025); Zhuang et al. (2024); Zhang et al. (2024b),
159 merging them into a unified dataset. We then use GPT-4o OpenAI (2024) to assign multi-level
160 topic tags $\mathcal{T} = \{t_1, \dots, t_n\}$, followed by hierarchical clustering on the semantic similarity matrix
161 $S \in \mathbb{R}^{n \times n}$ to obtain an AI-generated structure $\mathcal{K}^{\text{auto}}$. The final knowledge point set \mathcal{K} is produced by
162 expert-guided integration of $\mathcal{K}^{\text{human}}$ and $\mathcal{K}^{\text{auto}}$, with independent review for quality assurance.

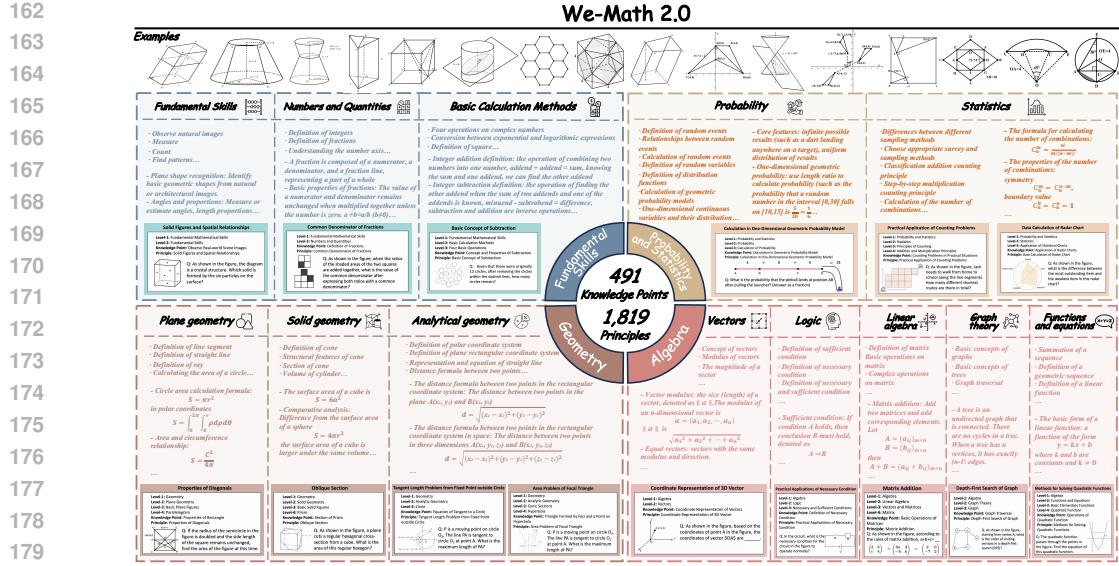


Figure 1: Overview of MathBook, including knowledge points, principles, and sample problems.

Fine-Grained Principle Annotation. Given the constructed \mathcal{K} , we employ GPT-4o to annotate the step-level knowledge points for each problem $q_j \in \mathcal{Q} = \{q_1, \dots, q_M\}$ by mapping each step in its chain-of-thought solution to the corresponding $k_i \in \mathcal{K}$. This yields a mapping $\mathcal{M}_1 : q_j \mapsto (k_{i_1}, k_{i_2}, \dots)$, forming a set of step-level solution paths for each knowledge point. Next, for each k_i , GPT-4o summarizes the set of theorems and principles used across all associated solution paths, resulting in a mapping $\mathcal{M}_2 : k_i \mapsto \{p_{i1}, \dots, p_{im}\}$. Finally, these AI-extracted principles are consolidated and cross-checked with those written by human experts, with iterative refinement to ensure completeness and accuracy of \mathcal{P} . Detailed guideline of our system are listed in Appendix B.1.

3.2 MULTI-DIMENSIONAL DATA CONSTRUCTION

In this section, we introduce our data construction pipeline: **MathBook-Standard** & **MathBook-Pro**.

3.2.1 MATHBOOK-STANDARD: SEED AND VARIANT PROBLEM CONSTRUCTION

Seed Problem Construction. To ensure rich coverage and high-quality design, we construct problems based on the knowledge system following 3 guidelines: (1) All diagrams are rendered with GeoGebra for precise geometric representation; (2) Problems focus on math essence, avoiding reliance on superficial visual cues; (3) Each problem strictly corresponds to its designated principle set \mathcal{P}_i . To achieve these, we adopt a “model-assisted, expert-led” workflow. Given a knowledge point $k_i \in \mathcal{K}$ and its associated principle set \mathcal{P}_i , an LLM first generates a draft problem, including the question, answer, and XML script. We then use GeoGebra, a software that renders diagrams from XML-based scripts, to automate the generation of draft images: $\mathcal{G}_{LM}(k_i, p_{ij}) \rightarrow (q_i^{\text{draft}}, a_i^{\text{draft}}, x_i^{\text{xml draft}})$. The resulting visual drafts serve as references to guide human experts in constructing problems and diagrams via GeoGebra scripting. In practice, almost all drafts were revised or reworked by experts,¹ in order to avoid reliance on superficial visual cues and ensure proper alignment with the underlying mathematical principles. The final seed problem set is $\mathcal{D}_{\text{seed}} = \{(k_i, p_{ij}, q_i, a_i, I_i, x_i^{\text{xml}})\}$, covering all knowledge points and principles. The detailed GeoGebra-based diagram generation guidelines can be found in the Appendix B.2.1.

Variant Problem Expansion. To further enhance the diversity and generalization ability of the dataset, we systematically construct two types of variants based on each seed problem:

(1) **One-Problem-Multi-Image Variants** ($\mathcal{D}_{\text{ImgVar}}$): Given a seed problem $(q_i, a_i, I_i) \in \mathcal{D}_{\text{seed}}$, we fix the problem statement q_i and knowledge annotation (k_i, p_{ij}) , and generate a set of images $\{I_i^{(1)}, I_i^{(2)}, \dots, I_i^{(m)}\}$ by varying the parameters in GeoGebra while maintaining the under-

¹Only 1.2% of the drafts were directly adopted by experts.

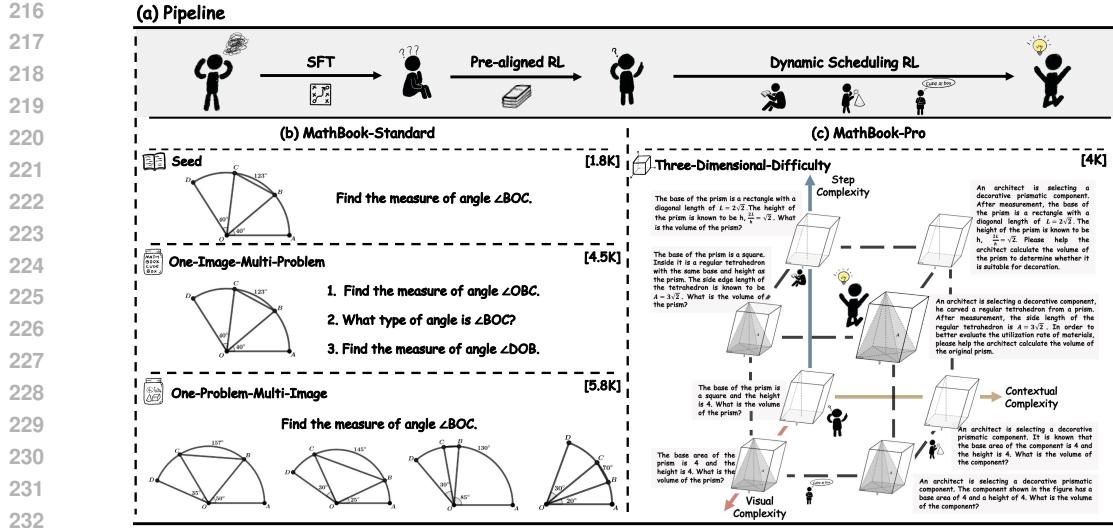


Figure 2: Overview of MathBook dataset and the corresponding training phase.

lying geometric construction. Each image corresponds to a different geometric instantiation (e.g., acute/obtuse/right triangle), resulting in distinct answers $a_i^{(t)}$: $\{(q_i, a_i^{(t)}, I_i^{(t)}) \in \mathcal{D}_{\text{ImgVar}}, t = 1, \dots, m\}$. This approach enriches the visual data diversity while preserving semantic consistency.

(2) One-Image-Multi-Problem Variants ($\mathcal{D}_{\text{QstVar}}$): Given a seed image I_i , we construct multiple new problems $q_i^{(s)}$ targeting different knowledge points $k_i^{(s)}$ and principles $p_{ij}^{(s)}$, curated by experts with language model assistance: $\{((q_i^{(s)}, a_i^{(s)}, I_i) \in \mathcal{D}_{\text{QstVar}}, s = 1, \dots, n\}$. This strategy leverages the reusability of high-quality diagrams to generate diverse problem variants.

By systematically applying these variant construction methods to each seed problem, we build the **MathBook-Standard** dataset with rich semantic and visual diversity.

3.2.2 MATHBOOK-PRO: THREE-DIMENSIONAL DIFFICULTY MODELING

To systematically characterize problem complexity from a model-centric perspective, we define a three-dimensional difficulty space for each seed problem along three orthogonal axes (in Figure 2):

(1) Step Complexity (ϕ_s): Knowledge-oriented reasoning depth is quantified by the number of involved knowledge points, which from the MathBook knowledge system. Given a seed problem with l process-oriented knowledge points, we construct variants requiring $l' > l$ (with at least six).

(2) Visual Complexity (ϕ_v): We increase complexity by adding auxiliary elements (e.g., lines) to the original image via GeoGebra, while preserving the core structure.

(3) Contextual Complexity (ϕ_c): Captures the contextualization of the problem statement. We vary the textual context from concise mathematical descriptions to complex linguistic scenarios.

Each seed problem $(q_0, a_0, I_0) \in \mathcal{D}_{\text{seed}}$ serves as the origin in a structured difficulty space. To enable controlled and interpretable expansion, we generate derived problems by varying a single dimension $d \in \{\phi_s, \phi_v, \phi_c\}$ at a time, yielding variants $(q_i^{(d)}, a_i^{(d)}, I_i^{(d)})$. Through multiple rounds of such single-axis transformations, we progressively construct more complex problems by composing changes across multiple dimensions. Formally, the most advanced variant takes the form: $(q^*, a^*, I^*) = \phi_s \circ \phi_v \circ \phi_c(q_0, a_0, I_0)$. In **MathBook-Pro**, the expansion along each dimension is implemented as:

(1) Along the ϕ_s dimension, we introduce intermediate conclusions as new conditions, enabling a knowledge-driven, progressive deepening of reasoning, expressed as $K_{i+1} = K_i + 1$, where K_i denotes the number of knowledge points involved at step i . In MathBook-Pro, the most complex step variants involve at least 6 knowledge points.

(2) Along the ϕ_v , we increase visual complexity by adding auxiliary lines, altering geometric configurations or introducing new spatial constructs via GeoGebra, while preserving the core structure.

(3) Along ϕ_c , we embed the mathematical core into real-world contexts or linguistically abstract scenarios, increasing the semantic and contextual demands of the problem statement.

By expanding along the defined dimensions, we generate a set of difficulty-controlled problem variants for each knowledge point, forming the difficulty modeling subset $\mathcal{D}_{\text{difficulty}}$ of MathBook-Pro.

3.3 MATHBOKEVAL

Design Principles. To ensure the quality and interpretability of annotations in visual math reasoning tasks, MathBookEval is designed based on the following principles: (1) **Comprehensive Knowledge Coverage:** Problems involve 491 knowledge points, spanning primary to university level, demonstrating broad coverage. (2) **Multi-level Reasoning Depth:** Each problem integrates 1–10 knowledge points, compared to 1–3 (Level 1) in existing process-oriented benchmarks, as illustrated in Figure 27. Notably, our annotation adheres to three principles: (1) integrating public and newly constructed problems under a unified guideline; (2) expert step-by-step annotation with explicit knowledge-point mapping; and (3) independent cross-validation, retaining only consistently annotated items.

Data Statistics and Evaluation Protocol. MathBookEval contains **1,000** fully annotated problems, covering all **491** knowledge points in the unified knowledge system \mathcal{K} , with 600 problems collected from existing benchmarks and 400 newly curated (Detailed statistics are presented in Table 8). We provide detailed statistics and splits along two key dimensions: (1) **Reasoning Dimension:** Problems are divided by reasoning steps into three levels: 1-3 (Level 1), 4-6 (Level 2), and 7-10 (Level 3), reflecting different reasoning depths. (2) **Knowledge Dimension:** The 491 knowledge points are grouped into 4 domains and 13 subdomains, covering primary to university level. Figure 27 demonstrates superior coverage of knowledge points and reasoning depth. All problems are in multiple-choice or fill-in-the-blank format.

4 METHODOLOGY

In this section, we introduce MathBook-RL, a two-stage framework that progressively guides MLLMs to develop reasoning capabilities from easy to hard. The first stage is a cold-start fine-tuning phase that establishes a knowledge-driven reasoning paradigm (§4.1); the second is a dynamic reinforcement learning phase that enhances the model’s generalization ability (§4.2).

4.1 COLD-START FINE-TUNING

The cold-start supervised fine-tuning (SFT) stage aims to instill explicit awareness of knowledge system and a knowledge-driven reasoning paradigm, avoiding rote memorization. The initial training set $\mathcal{D}_{\text{init}}$ is built from MathBook-Standard, which fully covers all 491 knowledge points. To improve rationale interpretability, we use GPT-4o OpenAI (2024) to rewrite each sample with natural language explanations that explicitly reference the relevant knowledge. The model is then trained using standard supervised fine-tuning: $\mathcal{L}_{\text{SFT}}(\theta) = \mathbb{E}_{(x,y) \sim \mathcal{D}_{\text{init}}} [-\log P_\theta(y | x)]$. This stage enhances the model’s ability to internalize the knowledge system and follow knowledge-guided reasoning chains.

4.2 PROGRESSIVE ALIGNMENT REINFORCEMENT LEARNING

(1) Pre-aligned RL. Prior to the dynamic scheduling stage, we perform initial RL training on MathBook-Standard dataset to ensure that the model develops genuine understanding of mathematical knowledge. Specifically, we utilize the $\mathcal{D}_{\text{ImgVar}}$ subset, where each group contains multiple variants of the same knowledge principle: $(q_i, a_i^{(t)}, I_i^{(t)}) \in \mathcal{D}_{\text{ImgVar}}$, $t = 1, \dots, m$. To encourage consistent and robust performance across different formulations, we adopt a mean-based reward function: $r = \frac{1}{m} \sum_{t=1}^m r^{(t)}$, where $r^{(t)} = 0.9$ if the answer is correct, 0.1 if only the format is correct, and 0 otherwise. Specifically, for problems corresponding to the same knowledge principle, rollout rewards are first sorted within each problem. Next, the mean reward at each sorted position is calculated across these problems and subsequently employed in the calculation of A_i . Instead of focusing on individual problems, this design integrates rewards across all problems corresponding to the same knowledge principle, thereby providing a more comprehensive critic.

(2) **Dynamic Scheduling RL.** In this section, we introduce a dynamic RL algorithm based on MathBook-Pro. The training process is organized as a curriculum along a main trajectory of increasing difficulty, primarily centered on the knowledge dimension. For each base problem (q_0, a_0, I_0) , denoted as x_0 , we construct a sequence of increasingly challenging variants as follows:

$$x_0 \rightarrow \phi_s(x_0) \rightarrow \phi_s \circ \phi_v(x_0) \rightarrow \phi_s \circ \phi_c(x_0) \rightarrow \phi_s \circ \phi_v \circ \phi_c(x_0) \quad (1)$$

where ϕ_s denotes increasing the number of knowledge points, ϕ_v and ϕ_c denotes increasing visual complexity and contextual abstraction. This forms a progressive path from basic to advanced reasoning for each knowledge anchor.

Incremental Learning Mechanism. At each curriculum transition $x \rightarrow \phi(x)$, if the model fails on $\phi(x)$ after succeeding on x , we introduce an incremental learning step. Specifically, we define the incremental set $\Delta(x, \phi)$ as a collection of samples that isolate the new knowledge or modality introduced by ϕ . The model is first trained on $\Delta(x, \phi)$ to address the incremental challenge, then reattempts $\phi(x)$. Concretely:

- **Knowledge Increment Scheduling:** For $x_0 \rightarrow \phi_s(x_0)$, if the model fails on $\phi_s(x_0)$, we construct $\Delta(x_0, \phi_s)$, comprising auxiliary problems x'_0 that target the new knowledge point(s) from ϕ_s .
- **Modality Increment Scheduling:** For $\phi_s(x_0) \rightarrow \phi_s \circ \phi_v(x_0)$ (or $\phi_s \circ \phi_c(x_0)$), if the model fails on the more complex sample, we construct $\Delta(\phi_s(x_0), \phi_v)$ (or $\Delta(\phi_s(x_0), \phi_c)$), which contains samples isolating the new visual or contextual complexity.

This incremental adaptation, denoted by $\Delta(x, \phi)$ at each step, ensures that the model can efficiently bridge the gap between curriculum stages. Notably, our overall RL objective is optimized using Group Relative Policy Optimization (GRPO) (Shao et al., 2024), which extends PPO by estimating the baseline from group scores instead of a separate critic. The GRPO objective is:

$$\mathcal{J}(\theta) = \mathbb{E}[q \sim P(Q), \{o_i\}_{i=1}^G \sim \pi_{\theta_{old}}(O|q)] \frac{1}{G} \sum_{i=1}^G \frac{1}{|o_i|} \sum_{t=1}^{|o_i|} \left\{ \min \left[\frac{\pi_{\theta}(o_{i,t}|q, o_{i,<t})}{\pi_{\theta_{old}}(o_{i,t}|q, o_{i,<t})} \hat{A}_{i,t}, \text{clip} \left(\frac{\pi_{\theta}(o_{i,t}|q, o_{i,<t})}{\pi_{\theta_{old}}(o_{i,t}|q, o_{i,<t})}, 1 - \epsilon, 1 + \epsilon \right) \hat{A}_{i,t} \right] - \beta \mathbb{D}_{KL} [\pi_{\theta} || \pi_{ref}] \right\}, \quad (2)$$

where ϵ and β are hyperparameters, q denotes the input, $\{o_i\}_{i=1}^G$ are sampled outputs, and r_i is the corresponding reward. $\hat{A}_{i,t}$ is the normalized advantage value for the i -th trajectory in the group. This curriculum-driven RL process, augmented with explicit incremental adaptation at each stage, enables the MLLM to progressively master complex, multi-dimensional reasoning tasks while preserving stability and generalization across knowledge, visual, and contextual variations.

5 EXPERIMENTS

5.1 EXPERIMENTAL SETUP

Datasets. All training data are sourced from WE-MATH 2.0 in compliance with copyright and licensing requirements, and all expert-constructed problems will be released under appropriate CC licenses. We use 1K, 5.8K and 4K samples for SFT, pre-aligned RL, and dynamic scheduling RL stages, respectively. Experiments are conducted on four standard mathematical reasoning benchmarks: MathVista Lu et al. (2023), MathVision Wang et al. (2024), MathVerse Zhang et al. (2024a), and We-Math Qiao et al. (2024a). Detailed evaluation protocols are provided in Appendix C.5.1.

Baselines. We conduct our training based on both Qwen2.5-VL-7B and Qwen2.5-VL-3B, and compare our method with three categories of baselines: (1) **Closed-source models** (e.g., GPT-4o OpenAI (2024)); (2) **Open-source general models** (e.g., InternVL2.5 series Chen et al. (2024a), Qwen2.5-VL series Bai et al. (2025b)); (3) **Open-source reasoning models** (e.g., R1-VL Zhang et al. (2025)). Our evaluation is based on VLMEvalKit Duan et al. (2024). Detailed descriptions of baselines are provided in Appendix C.5.2.

5.2 MAIN RESULTS

Table 2 displays the performance of our MathBook-7B across various benchmarks. Overall, our method achieves remarkable performance, clearly demonstrating its superiority. Further analysis reveals the following observations. (Results for the 3B setting are provided in Appendix C.2.)

378
 379 Table 2: Performance comparison across four widely-used mathematical reasoning benchmarks. Each
 380 benchmark follows its standard evaluation metric: MathVista and MathVision use accuracy, We-Math
 381 reports the strict score, and MathVerse evaluates on the vision-only subset with accuracy. Data sizes
 382 used for SFT and RL are annotated in **blue** and **red**, respectively.

Model	#Data	Avg.	MathVista	MathVision	We-Math	MathVerse
<i>Closed-source</i>						
GPT-4o-latest	-	54.0	71.6	43.8	50.6	49.9
Gemini-1.5-Pro	-	53.6	67.9	41	50.5	54.8
<i>Open-source (General)</i>						
Qwen2.5-VL-7B	-	42.6	68.2	25.1	36.0	41.1
InternVL2.5-8B-BoN-8	-	41.7	68.2	25.6	38.6	34.5
<i>Open-source (Reasoning)</i>						
Math-PUMA-7B	1.88M	-	47.9	-	19.2	26.0
URSA-8B	2.96M	37.8	58.8	28.7	32.8	31.0
R1-OneVision-7B	155K+10K	-	64.1	29.9	30.1	-
R1-VL-7B	260K+10K	-	63.5	24.7	22.7	-
MM-Eureka-7B	15K	45.2	73.0	26.9	34.5	46.2
WeThink-7B	120K+20K	47.5	71.6	26.0	48.0	44.2
VLAAThinker-7B	25K	46.0	68.0	26.4	41.5	48.2
OpenVLThinker-7B	35K+15K	-	72.3	25.9	-	-
MathBook-7B (Ours)	1K+9.8K	48.7	73.0	28.0	48.4	45.2
Δ (vs Qwen2.5-VL-7B)	-	+6.1	+4.8	+2.9	+12.4	+4.1

401
 402 (1) **Overall superiority of MathBook.** Compared to the backbone Qwen2.5-VL-7B, MathBook-7B
 403 achieves over a 6% improvement across all benchmarks, validating the effectiveness of our approach.

404 (2) **Effectiveness of progressive alignment reinforcement learning on knowledge generalization.**
 405 Focusing on the We-Math benchmark, which requires solving both complex multi-step questions and
 406 their corresponding subproblems, MathBook-7B outperforms strong RL baselines. This demonstrates
 407 the effectiveness of progressive alignment reinforcement learning in knowledge generalization.

408 (3) **Less is More: Efficiency with limited training data.** MathBook-7B achieves strong performance
 409 using only 9.8K training samples. We attribute this to the high-quality, structured mathematical
 410 knowledge system we constructed, enabling efficient alignment and generalization with limited data.

413 5.3 RESULTS ON MATHBOOKEVAL

415 To investigate MLLM abilities in reasoning depth
 416 and knowledge coverage breadth, we conduct Math-
 417 BookEval and observe the following (see Table 5).
 418 [Complete results are provided in Appendix B.4.6.](#)

419 (1) **MLLMs performance negatively correlates**
 420 **with the number of required knowledge points.**
 421 As reasoning steps increase, model accuracy de-
 422 clines. In particular, problems requiring 7–10
 423 knowledge points yield the lowest accuracy (be-
 424 low 50%). These results highlight the challenge of
 425 multi-step reasoning and validate knowledge points as a key measure for modeling problem difficulty.

426 (2) **MLLMs perform well on algebra but struggle with geometry.** Along the knowledge dimension,
 427 most MLLMs demonstrate strong performance in algebra, achieving accuracies above 50%. However,
 428 their consistently poor performance in geometry highlights ongoing challenges in spatial reasoning.

429 (3) **Larger models yield more consistent improvements across all dimensions.** Within the
 430 InternVL2.5 and Qwen2.5-VL families, increasing model size leads to consistent gains across all
 431 dimensions and in overall scores, emphasizing the role of scale in enhancing reasoning capabilities.

432 Table 3: Results of the ablation study. **MVt:**
 433 MathVista; **MVs:** MathVision; **WM:** We-Math

Method	SFT	RL-Pre	RL-Dyn	MVt	MVs	WM
M_0	✓	✓	✓	73.0	28.0	48.4
M_1	✓	✓	-	72.4	27.0	47.2
M_2	✓	-	✓	72.0	26.3	43.3
M_3	-	✓	✓	71.5	26.3	46.7
M_4	✓	-	-	65.8	25.7	38.3

Table 5: The performance of different MLLMs on MathBookEval for reasoning evaluation. **Acc.:** Accuracy; **FS.:** Foundational skills; **PS.:** Probability and statistics; **Geo.:** Geometry; **Alg.:** Algebra

Models	Acc.	Reasoning			Knowledge			
		Level1	Level2	Level3	FS.	PS.	Geo.	Alg.
<i>Closed-source MLLMs</i>								
GPT-4o	50.8	52.8	48.9	41.7	33.8	57.6	44.2	67.2
GPT-4V	42.8	44.0	43.0	31.9	36.8	56.6	33.5	59.4
<i>Open-source MLLMs</i>								
InternVL2.5-78B	51.8	52.5	51.8	45.8	50.0	64.2	42.6	67.6
Qwen2.5-VL-72B	57.1	58.3	56.4	50.0	52.9	58.5	52.1	68.8
LLaVA-OneVision-72B	43.0	44.8	42.0	31.9	38.2	52.8	37.0	53.5
InternVL2.5-8B	37.9	40.7	34.5	27.8	33.8	46.2	31.4	50.0
Qwen2.5-VL-7B	46.7	50.1	43.0	33.3	44.1	58.5	38.8	60.2
LLaVA-OneVision-7B	31.6	34.3	28.0	23.6	36.8	41.5	24.9	41.0
R1-VL-7B	38.0	41.9	32.6	27.8	38.2	38.7	32.3	50.4
MM-Eureka-7B	50.0	50.6	51.5	38.9	47.1	52.8	44.0	62.9
MathBook-7B	50.4	52.0	48.2	45.8	57.4	67.9	40.5	63.3

5.4 QUANTITATIVE ANALYSIS

Ablation Study. As shown in Table 3, we conduct ablation studies on the training stages. M0 denotes MathBook-7B, while M1–M4 represent models at different training stages (RL-Pre: Pre-aligned RL; RL-Dyn: Dynamic Scheduling RL). We lead to following two key findings:

(i) Both RL stages contribute significantly. Each RL stage (M0-M3) yields progressive improvements over M4. In particular, pre-aligned reinforcement learning (RL) in the first stage yields impressive results on MathVista and We-Math benchmarks, highlighting the crucial role of knowledge learning in enhancing mathematical reasoning abilities. **(ii) SFT alone offers limited gains, but is crucial for unlocking RL potential.** Comparing M0, M3, M4, we find that SFT alone yields marginal improvements over the Qwen2.5-VL backbone. However, when combined with RL, SFT version substantially boosts overall performance, highlighting its critical role in shifting the model's reasoning paradigm and enabling more effective RL optimization.

Analysis of SFT Data Paradigm and Scale. We explore the impact of data paradigm and scale during the SFT stage. Based on the M0 setting, we consider two variants: (1) Replacing the natural language CoT format with a structured, step-wise CoT format (Zhuang et al., 2024) aligned with \mathcal{K} ; (2) Increasing the SFT data scale with larger datasets (from 1K to 15K).

(i) Natural language CoT outperforms the structured step-wise format in SFT. As shown in Table 4, natural language CoT outperforms the structured format in RL. This highlights the advantage of natural language prompts in cultivating flexible visual mathematical reasoning skills. **(ii) Minimal SFT suffices** up SFT data does not improve performance. Models trained on comparably or even better than those trained on larger datasets. SFT set suffices to establish the reasoning paradigm for effective learning.

Table 4: SFT Data Analysis. **SFT (Str.)** and **SFT (Lar.)** denotes structured and large-scale SFT training.

Setting	MVt	MVs	WM
M_0	73.0	28.0	48.4
SFT(Str.)	71.9	26.0	46.7
SFT(Lar.)	72.8	27.0	49.0

6 CONCLUSION

In this work, we present **WE-MATH 2.0**, a unified framework for multi-modal mathematical reasoning. It comprises: (1) **MathBook Knowledge System**, a five-level hierarchy covering 491 knowledge points and 1,819 fundamental principles for comprehensive supervision; (2) **MathBook-Standard** and **MathBook-Pro**, two richly annotated datasets with conceptual expansions and principled difficulty modeling for structured learning; (3) **MathBook-RL**, a two-stage reinforcement learning framework that leverages knowledge-guided supervision and dynamic data scheduling; and (4) **MathBookEval**, a benchmark for evaluating reasoning across diverse knowledge and step distributions. Extensive experiments validate MathBook’s effectiveness in enhancing generalization of MLLMs.

486 REPRODUCIBILITY STATEMENT
487488 **Data.** We ensure that all datasets developed in this work, including *MathBook Knowledge System*,
489 *MathBook-Standard*, *MathBook-Pro*, and *MathBookEval*, will be fully released upon publication. To
490 support reproducibility during the review phase, we provide representative data samples (up to the
491 maximum size allowed by the submission system, 100MB) in the supplementary material.492 **Experimental Setup.** The experimental protocol, including dataset sizes, training stages, and
493 evaluation benchmarks, is described in Section 5 and Appendix C.1. In addition, our code contains
494 complete environment specifications (e.g., `requirements.txt`), along with scripts for dataset
495 preparation, training, and evaluation. These files ensure that the reported results can be reproduced
496 on standard hardware with minimal configuration.497 **Code and Model Checkpoints.** We include the full code in the supplementary material, together with
498 scripts for data preparation, training, and evaluation. Due to file size limitations, pretrained model
499 checkpoints cannot be submitted at this stage. Upon acceptance, we will release all checkpoints in
500 the camera-ready version to facilitate rapid validation of our results.501 **Methodology.** We detail the proposed framework in Section 3 and Section 4. Specifically, we
502 (i) specify the *MathBook Knowledge System* (five-level hierarchy; **491** knowledge points; **1,819**
503 principles), (ii) describe the *MathBook-Standard* pipeline with two orthogonal expansions (one-
504 question–multi-image; one-image–multi-question), (iii) formalize the three-axis difficulty space for
505 *MathBook-Pro* (step, visual, contextual) yielding seven progressive variants per seed, and (iv) present
506 *MathBook-RL*, a two-stage training paradigm.507
508 ETHICS STATEMENT
509510 **Licensing and Open Access.** For all referenced or incorporated data in WE-MATH 2.0, we only
511 use existing datasets with clear and appropriate licenses. All data curated by our team will be released
512 under the CC BY 4.0 license, ensuring open access for the research community. The entire MathBook
513 dataset, including both external and newly constructed components, will be made publicly available
514 to facilitate further research and development.515 **Data Sources and Privacy.** All data in WE-MATH 2.0 are either sourced from publicly available
516 datasets or generated by our expert team, and do not contain any personal user information. Therefore,
517 there are no privacy concerns related to our dataset.518 **Expert Compensation.** Experts involved in annotation are compensated on a per-task basis, with
519 payment issued only after cross-validation and quality assurance. All compensation meets or exceeds
520 the local minimum wage standards.521
522 REFERENCES
523524 Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman,
525 Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report.
526 *arXiv preprint arXiv:2303.08774*, 2023.527 Shuai Bai, Fudong Wang, Jiajia Liu, Jingdong Chen, Jun Zhou, Kaixiang Ji, Lixiang Ru, Qingpei
528 Guo, Ruobing Zheng, Tianqi Li, et al. M2-reasoning: Empowering mllms with unified general and
529 spatial reasoning. *arXiv preprint arXiv:2507.08306*, 2025.530 Shuai Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Sibo Song, Kai Dang, Peng Wang,
531 Shijie Wang, Jun Tang, Humen Zhong, Yuanzhi Zhu, Mingkun Yang, Zhaohai Li, Jianqiang Wan,
532 Pengfei Wang, Wei Ding, Zheren Fu, Yiheng Xu, Jiabo Ye, Xi Zhang, Tianbao Xie, Zesen Cheng,
533 Hang Zhang, Zhibo Yang, Haiyang Xu, and Junyang Lin. Qwen2.5-v1 technical report. *arXiv
534 preprint arXiv:2502.13923*, 2025a.535 Shuai Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Sibo Song, Kai Dang, Peng Wang,
536 Shijie Wang, Jun Tang, et al. Qwen2. 5-v1 technical report. *arXiv preprint arXiv:2502.13923*,
537 2025b.

540 Hardy Chen, Haoqin Tu, Fali Wang, Hui Liu, Xianfeng Tang, Xinya Du, Yuyin Zhou, and Cihang
 541 Xie. Sft or rl? an early investigation into training rl-like reasoning large vision-language models.
 542 *arXiv preprint arXiv:2504.11468*, 2025.

543

544 Zhe Chen, Jiannan Wu, Wenhui Wang, Weijie Su, Guo Chen, Sen Xing, Muyan Zhong, Qinglong
 545 Zhang, Xizhou Zhu, Lewei Lu, Bin Li, Ping Luo, Tong Lu, Yu Qiao, and Jifeng Dai. Internvl:
 546 Scaling up vision foundation models and aligning for generic visual-linguistic tasks. *arXiv preprint*
 547 *arXiv:2312.14238*, 2023.

548 Zhe Chen, Weiyun Wang, Yue Cao, Yangzhou Liu, Zhangwei Gao, Erfei Cui, Jinguo Zhu, Shenglong
 549 Ye, Hao Tian, Zhaoyang Liu, Lixin Gu, Xuehui Wang, Qingyun Li, Yimin Ren, Zixuan Chen,
 550 Jiapeng Luo, Jiahao Wang, Tan Jiang, Bo Wang, Conghui He, Botian Shi, Xingcheng Zhang, Han
 551 Lv, Yi Wang, Wenqi Shao, Pei Chu, Zhongying Tu, Tong He, Zhiyong Wu, Huipeng Deng, Jiaye
 552 Ge, Kai Chen, Min Dou, Lewei Lu, Xizhou Zhu, Tong Lu, Dahua Lin, Yu Qiao, Jifeng Dai, and
 553 Wenhui Wang. Expanding performance boundaries of open-source multimodal models with model,
 554 data, and test-time scaling. *CoRR*, abs/2412.05271, 2024a. doi: 10.48550/ARXIV.2412.05271.
 555 URL <https://doi.org/10.48550/arXiv.2412.05271>.

556 Zhe Chen, Weiyun Wang, Hao Tian, Shenglong Ye, Zhangwei Gao, Erfei Cui, Wenwen Tong, Kongzhi
 557 Hu, Jiapeng Luo, Zheng Ma, et al. How far are we to gpt-4v? closing the gap to commercial
 558 multimodal models with open-source suites. *arXiv preprint arXiv:2404.16821*, 2024b.

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

562

563 DeepSeek-AI. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning,
 564 2025. URL <https://arxiv.org/abs/2501.12948>.

565

566 Yihe Deng, Hritik Bansal, Fan Yin, Nanyun Peng, Wei Wang, and Kai-Wei Chang. Openvlthinker:
 567 An early exploration to complex vision-language reasoning via iterative self-improvement. *arXiv*
 568 *preprint arXiv:2503.17352*, 2025.

569

570 Haodong Duan, Junming Yang, Yuxuan Qiao, Xinyu Fang, Lin Chen, Yuan Liu, Xiaoyi Dong, Yuhang
 571 Zang, Pan Zhang, Jiaqi Wang, et al. Vlmevalkit: An open-source toolkit for evaluating large
 572 multi-modality models. In *Proceedings of the 32nd ACM international conference on multimedia*,
 573 pp. 11198–11201, 2024.

574

575 Richard Fitzpatrick. Euclid's elements of geometry. 2008.

576

577 Jiahui Gao, Renjie Pi, Jipeng Zhang, Jiacheng Ye, Wanjun Zhong, Yufei Wang, Lanqing Hong,
 578 Jianhua Han, Hang Xu, Zhenguo Li, et al. G-llava: Solving geometric problem with multi-modal
 579 large language model. *arXiv preprint arXiv:2312.11370*, 2023.

580

581 Team GLM, Aohan Zeng, Bin Xu, Bowen Wang, Chenhui Zhang, Da Yin, Dan Zhang, Diego Rojas,
 582 Guanyu Feng, Hanlin Zhao, et al. Chatglm: A family of large language models from glm-130b to
 583 glm-4 all tools. *arXiv preprint arXiv:2406.12793*, 2024.

584

585 Dong Guo, Faming Wu, Feida Zhu, Fuxing Leng, Guang Shi, Haobin Chen, Haoqi Fan, Jian Wang,
 586 Jianyu Jiang, Jiawei Wang, et al. Seed1. 5-vl technical report. *arXiv preprint arXiv:2505.07062*,
 587 2025.

588

589 Xiaotian Han, Yiren Jian, Xuefeng Hu, Haogeng Liu, Yiqi Wang, Qihang Fan, Yuang Ai, Huaibo
 590 Huang, Ran He, Zhenheng Yang, et al. Infimm-webmath-40b: Advancing multimodal pre-training
 591 for enhanced mathematical reasoning. *arXiv preprint arXiv:2409.12568*, 2024.

592

593 Wenyi Hong, Wenmeng Yu, Xiaotao Gu, Guo Wang, Guobing Gan, Haomiao Tang, Jiale Cheng,
 594 Ji Qi, Junhui Ji, Lihang Pan, et al. Glm-4.1 v-thinking: Towards versatile multimodal reasoning
 595 with scalable reinforcement learning. *arXiv preprint arXiv:2507.01006*, 2025.

596

597 Wenxuan Huang, Bohan Jia, Zijie Zhai, Shaosheng Cao, Zheyu Ye, Fei Zhao, Zhe Xu, Yao Hu, and
 598 Shaohui Lin. Vision-r1: Incentivizing reasoning capability in multimodal large language models.
 599 *arXiv preprint arXiv:2503.06749*, 2025.

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

597

598 Justin Johnson, Bharath Hariharan, Laurens Van Der Maaten, Li Fei-Fei, C Lawrence Zitnick, and
 599 Ross Girshick. Clevr: A diagnostic dataset for compositional language and elementary visual
 600 reasoning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.
 601 2901–2910, 2017.

602

603 Zhikai Lei, Tianyi Liang, Hanglei Hu, Jin Zhang, Yunhua Zhou, Yunfan Shao, Linyang Li, Chenchui
 604 Li, Changbo Wang, Hang Yan, et al. Gaokao-eval: Does high scores truly reflect strong capabilities
 605 in llms? *arXiv preprint arXiv:2412.10056*, 2024.

606

607 Bo Li, Yuanhan Zhang, Dong Guo, Renrui Zhang, Feng Li, Hao Zhang, Kaichen Zhang, Peiyuan
 608 Zhang, Yanwei Li, Ziwei Liu, et al. Llava-onevision: Easy visual task transfer. *arXiv preprint*
 609 *arXiv:2408.03326*, 2024.

610

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

613

614 Xiangyan Liu, Jinjie Ni, Zijian Wu, Chao Du, Longxu Dou, Haonan Wang, Tianyu Pang, and
 615 Michael Qizhe Shieh. Noisyrollout: Reinforcing visual reasoning with data augmentation. *arXiv
 preprint arXiv:2504.13055*, 2025a.

616

617 Ziyu Liu, Zeyi Sun, Yuhang Zang, Xiaoyi Dong, Yuhang Cao, Haodong Duan, Dahua Lin, and Jiaqi
 618 Wang. Visual-rft: Visual reinforcement fine-tuning. *arXiv preprint arXiv:2503.01785*, 2025b.

619

620 Pan Lu, Ran Gong, Shubiao Jiang, Liang Qiu, Siyuan Huang, Xiaodan Liang, and Song-Chun Zhu.
 621 Inter-gps: Interpretable geometry problem solving with formal language and symbolic reasoning.
 622 *arXiv preprint arXiv:2105.04165*, 2021a.

623

624 Pan Lu, Liang Qiu, Jiaqi Chen, Tony Xia, Yizhou Zhao, Wei Zhang, Zhou Yu, Xiaodan Liang,
 625 and Song-Chun Zhu. Iconqa: A new benchmark for abstract diagram understanding and visual
 language reasoning. *arXiv preprint arXiv:2110.13214*, 2021b.

626

627 Pan Lu, Hritik Bansal, Tony Xia, Jiacheng Liu, Chunyuan Li, Hannaneh Hajishirzi, Hao Cheng,
 628 Kai-Wei Chang, Michel Galley, and Jianfeng Gao. Mathvista: Evaluating mathematical reasoning
 629 of foundation models in visual contexts. *arXiv preprint arXiv:2310.02255*, 2023.

630

631 Ruilin Luo, Zhuofan Zheng, Yifan Wang, Yiyao Yu, Xinzhe Ni, Zicheng Lin, Jin Zeng, and Yujiu
 632 Yang. Ursu: Understanding and verifying chain-of-thought reasoning in multimodal mathematics.
 633 *arXiv preprint arXiv:2501.04686*, 2025.

634

635 Fanqing Meng, Lingxiao Du, Zongkai Liu, Zhixiang Zhou, Quanfeng Lu, Daocheng Fu, Tiancheng
 636 Han, Botian Shi, Wenhui Wang, Junjun He, et al. Mm-eureka: Exploring the frontiers of multimodal
 637 reasoning with rule-based reinforcement learning. *arXiv preprint arXiv:2503.07365*, 2025.

638

639 OpenAI. Hello gpt-4o, 2024. URL <https://openai.com/index/hello-gpt-4o/>.

640

641 R OpenAI. Gpt-4v (ision) system card. *Citekey: gptvision*, 2023.

642

643 Shuai Peng, Di Fu, Liangcai Gao, Xiuqin Zhong, Hongguang Fu, and Zhi Tang. Multimath: Bridging
 644 visual and mathematical reasoning for large language models. *arXiv preprint arXiv:2409.00147*,
 645 2024.

646

647 Runqi Qiao, Qiuna Tan, Guanting Dong, Minhui Wu, Chong Sun, Xiaoshuai Song, Zhuoma Gongque,
 648 Shanglin Lei, Zhe Wei, MiaoXuan Zhang, Runfeng Qiao, Yifan Zhang, Xiao Zong, Yida Xu, Muxi
 649 Diao, Zhimin Bao, Chen Li, and Honggang Zhang. We-math: Does your large multimodal model
 650 achieve human-like mathematical reasoning? *CoRR*, abs/2407.01284, 2024a. doi: 10.48550/
 651 ARXIV.2407.01284. URL <https://doi.org/10.48550/arXiv.2407.01284>.

648 Runqi Qiao, Qiuna Tan, Guanting Dong, Minhui Wu, Chong Sun, Xiaoshuai Song, Zhuoma GongQue,
 649 Shanglin Lei, Zhe Wei, MiaoXuan Zhang, et al. We-math: Does your large multimodal model
 650 achieve human-like mathematical reasoning? *arXiv preprint arXiv:2407.01284*, 2024b.

651

652 Tianhe Ren, Qing Jiang, Shilong Liu, Zhaoyang Zeng, Wenlong Liu, Han Gao, Hongjie Huang,
 653 Zhengyu Ma, Xiaoke Jiang, Yihao Chen, et al. Grounding dino 1.5: Advance the "edge" of
 654 open-set object detection. *arXiv preprint arXiv:2405.10300*, 2024.

655

656 Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang,
 657 Mingchuan Zhang, YK Li, Y Wu, et al. Deepseekmath: Pushing the limits of mathematical
 658 reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

659

660 Wenhao Shi, Zhiqiang Hu, Yi Bin, Junhua Liu, Yang Yang, See-Kiong Ng, Lidong Bing, and
 661 Roy Ka-Wei Lee. Math-LLaVA: Bootstrapping mathematical reasoning for multimodal large
 662 language models. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pp.
 663 4663–4680, November 2024a.

664

665 Wenhao Shi, Zhiqiang Hu, Yi Bin, Junhua Liu, Yang Yang, See-Kiong Ng, Lidong Bing, and Roy
 666 Ka-Wei Lee. Math-llava: Bootstrapping mathematical reasoning for multimodal large language
 667 models. *arXiv preprint arXiv:2406.17294*, 2024b.

668

669 Core Team, Zihao Yue, Zhenru Lin, Yifan Song, Weikun Wang, Shuhuai Ren, Shuhao Gu, Shicheng
 670 Li, Peidian Li, Liang Zhao, Lei Li, Kainan Bao, Hao Tian, Hailin Zhang, Gang Wang, Dawei Zhu,
 671 Cici, Chenhong He, Bowen Ye, Bowen Shen, Zihan Zhang, Zihan Jiang, Zhixian Zheng, Zhichao
 672 Song, Zhenbo Luo, Yue Yu, Yudong Wang, Yuanyuan Tian, Yu Tu, Yihan Yan, Yi Huang, Xu Wang,
 673 Xinzhe Xu, Xingchen Song, Xing Zhang, Xing Yong, Xin Zhang, Xiangwei Deng, Wenyu Yang,
 674 Wenhan Ma, Weiwei Lv, Weiji Zhuang, Wei Liu, Sirui Deng, Shuo Liu, Shimao Chen, Shihua Yu,
 675 Shaohui Liu, Shande Wang, Rui Ma, Qiantong Wang, Peng Wang, Nuo Chen, Menghang Zhu,
 676 Kangyang Zhou, Kang Zhou, Kai Fang, Jun Shi, Jinhao Dong, Jiebao Xiao, Jiaming Xu, Huaqiu
 677 Liu, Hongshen Xu, Heng Qu, Haochen Zhao, Hanglong Lv, Guoan Wang, Duo Zhang, Dong
 678 Zhang, Di Zhang, Chong Ma, Chang Liu, Can Cai, and Bingquan Xia. Mimo-vl technical report,
 679 2025a. URL <https://arxiv.org/abs/2506.03569>.

680 Gemini Team, Rohan Anil, Sebastian Borgeaud, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut,
 681 Johan Schalkwyk, Andrew M Dai, Anja Hauth, Katie Millican, et al. Gemini: a family of highly
 682 capable multimodal models. *arXiv preprint arXiv:2312.11805*, 2023.

683

684 Kwai Keye Team, Biao Yang, Bin Wen, Changyi Liu, Chenglong Chu, Chengru Song, Chongling
 685 Rao, Chuan Yi, Da Li, Dunju Zang, et al. Kwai keye-vl technical report. *arXiv preprint
 686 arXiv:2507.01949*, 2025b.

687

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

690

691 Zhongwei Wan, Zhihao Dou, Che Liu, Yu Zhang, Dongfei Cui, Qinjian Zhao, Hui Shen, Jing
 692 Xiong, Yi Xin, Yifan Jiang, et al. Srpo: Enhancing multimodal llm reasoning via reflection-aware
 693 reinforcement learning. *arXiv preprint arXiv:2506.01713*, 2025.

694

695 Ziyu Wan, Xidong Feng, Muning Wen, Stephen Marcus McAleer, Ying Wen, Weinan Zhang, and
 696 Jun Wang. Alphazero-like tree-search can guide large language model decoding and training. In
 697 *Forty-first International Conference on Machine Learning*, 2024.

698

699 Ke Wang, Junting Pan, Weikang Shi, Zimu Lu, Houxing Ren, Aojun Zhou, Mingjie Zhan, and
 700 Hongsheng Li. Measuring multimodal mathematical reasoning with math-vision dataset. *Advances
 701 in Neural Information Processing Systems*, 37:95095–95169, 2024.

702

703 Ke Wang, Junting Pan, Linda Wei, Aojun Zhou, Weikang Shi, Zimu Lu, Han Xiao, Yunqiao
 704 Yang, Houxing Ren, Mingjie Zhan, et al. Mathcoder-vl: Bridging vision and code for enhanced
 705 multimodal mathematical reasoning. *arXiv preprint arXiv:2505.10557*, 2025.

706

707 Haoran Wei, Chenglong Liu, Jinyue Chen, Jia Wang, Lingyu Kong, Yanming Xu, Zheng Ge, Liang
 708 Zhao, Jianjian Sun, Yuang Peng, et al. General ocr theory: Towards ocr-2.0 via a unified end-to-end
 709 model. 2024.

702 Huajian Xin, ZZ Ren, Junxiao Song, Zhihong Shao, Wanja Zhao, Haocheng Wang, Bo Liu, Liyue
 703 Zhang, Xuan Lu, Qiushi Du, et al. Deepseek-prover-v1. 5: Harnessing proof assistant feedback for
 704 reinforcement learning and monte-carlo tree search. *arXiv preprint arXiv:2408.08152*, 2024.

705 Liangyu Xu, Yingxiu Zhao, Jingyun Wang, Yingyao Wang, Bu Pi, Chen Wang, Mingliang Zhang,
 706 Jihao Gu, Xiang Li, Xiaoyong Zhu, et al. Geosense: Evaluating identification and application of
 707 geometric principles in multimodal reasoning. *arXiv preprint arXiv:2504.12597*, 2025.

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

711 Yi Yang, Xiaoxuan He, Hongkun Pan, Xiyan Jiang, Yan Deng, Xingtao Yang, Haoyu Lu, Dacheng
 712 Yin, Fengyun Rao, Minfeng Zhu, et al. R1-onevision: Advancing generalized multimodal reasoning
 713 through cross-modal formalization. *arXiv preprint arXiv:2503.10615*, 2025b.

714 Jiabo Ye, Anwen Hu, Haiyang Xu, Qinghao Ye, Ming Yan, Yuhao Dan, Chenlin Zhao, Guohai Xu,
 715 Chenliang Li, Junfeng Tian, et al. mplug-docowl: Modularized multimodal large language model
 716 for document understanding. *arXiv preprint arXiv:2307.02499*, 2023a.

717 Jiabo Ye, Anwen Hu, Haiyang Xu, Qinghao Ye, Ming Yan, Guohai Xu, Chenliang Li, Junfeng Tian,
 718 Qi Qian, Ji Zhang, et al. Ureader: Universal ocr-free visually-situated language understanding
 719 with multimodal large language model. *arXiv preprint arXiv:2310.05126*, 2023b.

720 Jingyi Zhang, Jiaxing Huang, Huanjin Yao, Shunyu Liu, Xikun Zhang, Shijian Lu, and Dacheng Tao.
 721 R1-vl: Learning to reason with multimodal large language models via step-wise group relative
 722 policy optimization. *arXiv preprint arXiv:2503.12937*, 2025.

723 Renrui Zhang, Dongzhi Jiang, Yichi Zhang, Haokun Lin, Ziyu Guo, Pengshuo Qiu, Aojun Zhou, Pan
 724 Lu, Kai-Wei Chang, Yu Qiao, et al. Mathverse: Does your multi-modal llm truly see the diagrams
 725 in visual math problems? In *European Conference on Computer Vision*, pp. 169–186. Springer,
 726 2024a.

727 Renrui Zhang, Xinyu Wei, Dongzhi Jiang, Ziyu Guo, Shicheng Li, Yichi Zhang, Chengzhuo Tong,
 728 Jiaming Liu, Aojun Zhou, Bin Wei, et al. Mavis: Mathematical visual instruction tuning with an
 729 automatic data engine. *arXiv preprint arXiv:2407.08739*, 2024b.

730 Ziwei Zheng, Michael Yang, Jack Hong, Chenxiao Zhao, Guohai Xu, Le Yang, Chao Shen, and Xing
 731 Yu. Deepeyes: Incentivizing" thinking with images" via reinforcement learning. *arXiv preprint
 732 arXiv:2505.14362*, 2025.

733 Jinguo Zhu, Weiyun Wang, Zhe Chen, Zhaoyang Liu, Shenglong Ye, Lixin Gu, Yuchen Duan, Hao
 734 Tian, Weijie Su, Jie Shao, et al. Internvl3: Exploring advanced training and test-time recipes for
 735 open-source multimodal models. *arXiv preprint arXiv:2504.10479*, 2025.

736 Wenwen Zhuang, Xin Huang, Xiantao Zhang, and Jin Zeng. Math-puma: Progressive upward
 737 multimodal alignment to enhance mathematical reasoning. *arXiv preprint arXiv:2408.08640*,
 738 2024.

739 Wenwen Zhuang, Xin Huang, Xiantao Zhang, and Jin Zeng. Math-puma: Progressive upward
 740 multimodal alignment to enhance mathematical reasoning. In *Proceedings of the AAAI Conference
 741 on Artificial Intelligence*, volume 39, pp. 26183–26191, 2025.

742 Chengke Zou, Xingang Guo, Rui Yang, Junyu Zhang, Bin Hu, and Huan Zhang. Dynamath: A
 743 dynamic visual benchmark for evaluating mathematical reasoning robustness of vision language
 744 models. *arXiv preprint arXiv:2411.00836*, 2024.

745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755

756 Appendix

757 CONTENTS

758	1
759	1
760	1
761	1
762	1
763	2
764	2
765	3
766	3
767	3.1 MathBook Knowledge System
768	3.2 Multi-Dimensional Data Construction
769	3.2.1 MathBook-Standard: Seed and Variant Problem Construction
770	3.2.2 MathBook-Pro: Three-Dimensional Difficulty Modeling
771	3.3 MathBookEval
772	6
773	6
774	4
775	4
776	4.1 Cold-Start Fine-tuning
777	4.2 Progressive Alignment Reinforcement Learning
778	6
779	6
780	5
781	5
782	5.1 Experimental Setup
783	5.2 Main Results
784	5.3 Results on MathBookEval
785	8
786	5.4 Quantitative Analysis
787	9
788	6
789	6
790	A
791	A
792	B
793	B
794	B.1 MathBook Knowledge System
795	B.1.1 Hierarchical Structure of Knowledge Points
796	B.1.2 Knowledge Principles
797	17
798	B.2 MathBook-Standard
799	B.2.1 GeoGebra-based Diagram Generation.
800	B.2.2 Dataset Diversity and Variant Construction.
801	26
802	B.3 MathBook-Pro
803	34
804	B.4 MathBookEval
805	B.4.1 Dataset Construction and Annotation Protocol.
806	B.4.2 Task Dimensions.
807	34
808	B.4.3 Dataset Statistics.
809	36
	B.4.4 Evaluation Protocol and Metrics.
	37

810	B.4.5 Experiment Setup	37
811	B.4.6 Additional Results on MathBookEval	38
812		
813		
814	C More Details on MathBook-RL	39
815	C.1 Implementation details	39
816	C.2 Additional Experimental Results for MathBook-3B	40
817	C.3 Generalization to Text-only Mathematical Reasoning Tasks	40
818	C.4 Case Study	40
819	C.5 Experiment Setup	41
820	C.5.1 Details of the Evaluation.	41
821	C.5.2 Details of the Baselines	41
822		
823		
824		
825		
826	D The Use of Large Language Models (LLMs)	45
827		
828		
829		
830		
831		
832		
833		
834		
835		
836		
837		
838		
839		
840		
841		
842		
843		
844		
845		
846		
847		
848		
849		
850		
851		
852		
853		
854		
855		
856		
857		
858		
859		
860		
861		
862		
863		

864

A BROADEN IMPACT

865
 866 **Towards principled and generalizable mathematical model training.** WE-MATH 2.0 provides a
 867 comprehensive and structured mathematical knowledge system, which fills the gap left by previous
 868 works that lack a complete and systematic framework. By introducing a five-level hierarchy with
 869 491 knowledge points and 1,819 fundamental principles, MathBook enables more principled and
 870 interpretable mathematical learning for MLLMs. The dual expansion strategy ("multi-images per
 871 question" and "multi-questions per image") and the three-dimensional difficulty modeling not only
 872 enrich the diversity of training data but also facilitate robust and progressive learning. This system-
 873 atic approach can inspire future research to adopt knowledge-driven and model-centric data space
 874 modeling, leading to more reliable and generalizable mathematical reasoning models. Furthermore,
 875 the fine-grained annotations and progressive difficulty levels provide a valuable resource for bench-
 876 marking and analyzing the strengths and weaknesses of different MLLMs, promoting transparency
 877 and interpretability in model development.

878
 879 **Bridging AI and Education: high-quality datasets for teaching and learning.** WE-MATH 2.0's
 880 datasets are not only designed for model training but also have significant educational value. Each
 881 problem is accompanied by a GeoGebra (GGB) file, which can serve as high-quality teaching material
 882 for educators and students. The hierarchical knowledge system and step-wise annotations make it
 883 easier to design personalized learning paths and targeted exercises, supporting adaptive learning and
 884 formative assessment. The multi-modal and multi-perspective problem sets encourage students to
 885 develop flexible thinking and deepen their conceptual understanding. By bridging the gap between
 886 AI research and educational practice, MathBook has the potential to enhance mathematics education,
 887 facilitate interactive and engaging learning experiences, and support the development of intelligent
 888 tutoring systems.

889
 890 **Enhancing RL generalization through progressive and dynamic training.** WE-MATH 2.0 intro-
 891 duces a novel, model-centric curriculum for RL-based training, where problems are systematically
 892 organized from easy to hard based on explicit difficulty modeling. This approach provides a new
 893 perspective for designing RL curricula, enabling more effective and efficient learning. The "one-
 894 question-multi-image" and "one-image-multi-question" strategies, together with dynamic scheduling
 895 mechanisms, enhance the robustness and generalization of RL-trained models. These innovations
 896 can inspire the broader RL community to explore curriculum learning, dynamic data scheduling, and
 897 multi-modal data augmentation for complex reasoning tasks. Moreover, these hierarchical knowledge
 898 approaches also offers a new solution for tool learning. MathBook thus serves as a valuable testbed
 899 for advancing RL methods in the context of mathematical reasoning and beyond.

900

B DETAILS OF WE-MATH 2.0

901

B.1 MATHBOOK KNOWLEDGE SYSTEM

902

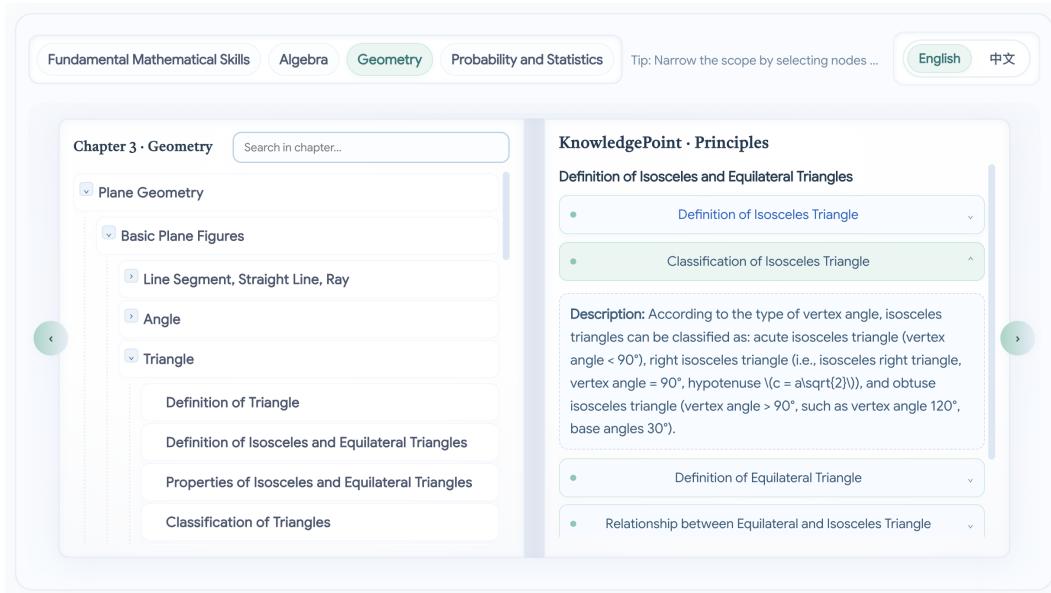
B.1.1 HIERARCHICAL STRUCTURE OF KNOWLEDGE POINTS

903
 904 As illustrated in Figure 3, we provide an overall view of the hierarchical structure of knowledge points
 905 in the **MathBook Knowledge System**. Figures 4–7 further present partial examples of different
 906 substructures at varying depths of the hierarchy. The system is organized as a five-level hierarchical
 907 structure of knowledge points $\mathcal{K} = \{k_1, k_2, \dots, k_N\}$, where $N = 491$ denotes the total number
 908 of knowledge points at the lowest level of the hierarchy. The first level consists of four categories:
 909 *Geometry*, *Fundamental Skills*, *Algebra*, and *Probability and Statistics*.

910
 911 The construction of \mathcal{K} follows a two-track, human-AI collaborative process. First, an initial version
 912 $\mathcal{K}^{\text{human}}$ is constructed by collecting and merging knowledge point lists from authoritative sources,
 913 including Wikipedia, open-source mathematics textbooks, and national curriculum standards. This
 914 initial structure is deduplicated, reorganized, and refined for logical consistency and comprehensive
 915 coverage.

916
 917 In parallel, a large-scale problem set \mathcal{Q} is collected, including 30,000 sampled from existing math
 918 datasets. GPT-4o is used to assign multi-level topic tags $\mathcal{T} = \{t_1, \dots, t_n\}$ to each problem, and a

918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938



939 Figure 3: Overall view of the hierarchical structure of knowledge points in MathBook.
940
941

942 semantic similarity matrix $S \in \mathbb{R}^{n \times n}$ is computed. Hierarchical clustering is then applied to S to
943 generate an AI-assisted hierarchical structure of knowledge points $\mathcal{K}^{\text{auto}}$.
944

945 Finally, the AI-assisted structure $\mathcal{K}^{\text{auto}}$ is integrated with the initial structure $\mathcal{K}^{\text{human}}$ through systematic
946 comparison, merging, and revision. The $\mathcal{K}^{\text{auto}}$ serves as a reference for revising and refining the
947 manually constructed hierarchical structure of knowledge points, resulting in the final knowledge
948 point set \mathcal{K} .
949

950 B.1.2 KNOWLEDGE PRINCIPLES 951

952 As shown in Figures 9–12, we provide several examples of knowledge principles, which include
953 definitions, theorems, and other foundational statements associated with each knowledge point.
954 The annotation of principles $\mathcal{P} = \bigcup_{i=1}^N \mathcal{P}_i$ ($|\mathcal{P}| = 1,819$) also follows a two-track, human-AI
955 collaborative approach.

956 Based on the constructed knowledge hierarchy \mathcal{K} , a set of core principles for each knowledge point
957 k_i is first drafted, referencing authoritative sources such as Wikipedia, textbooks, and international
958 curriculum standards.
959

960 In parallel, for each k_i , a set of representative problems from \mathcal{Q} is selected and their chain-of-thought
961 (CoT) solutions are annotated. Each step in the CoT is mapped to the corresponding knowledge point
962 using GPT-4o, and the relevant CoT steps for each k_i are extracted. The CoT steps associated with
963 each knowledge point are then aggregated and reviewed to supplement, refine, and validate the set of
964 principles for k_i .
965

966 This process is repeated iteratively, consolidating both the expert-written and data-driven principles,
967 cross-checking against original sources and annotated solution paths, until the set of principles \mathcal{P}_i for
968 each knowledge point is comprehensive and precise.
969

970 This process is repeated iteratively, consolidating both the expert-written and data-driven principles,
971 cross-checking against original sources and annotated solution paths, until the set of principles \mathcal{P}_i for
972 each knowledge point is comprehensive and precise. As illustrated in Figure 8, we further provide
973 an example from **We-Math 2.0**, where each problem is explicitly aligned with its corresponding
974 knowledge point and associated principle.
975

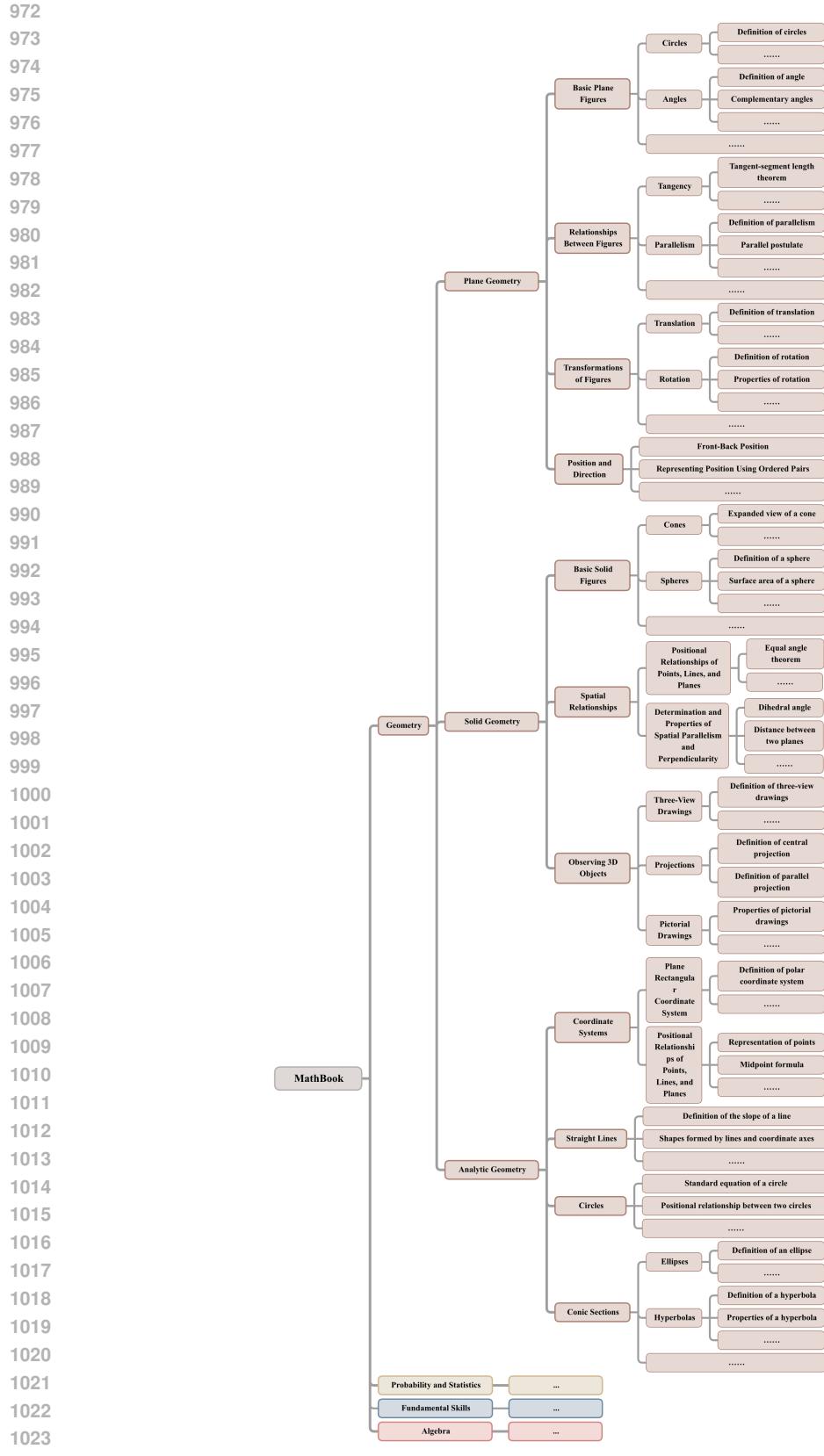


Figure 4: Overview of the hierarchical structure of knowledge points in MathBook (1).

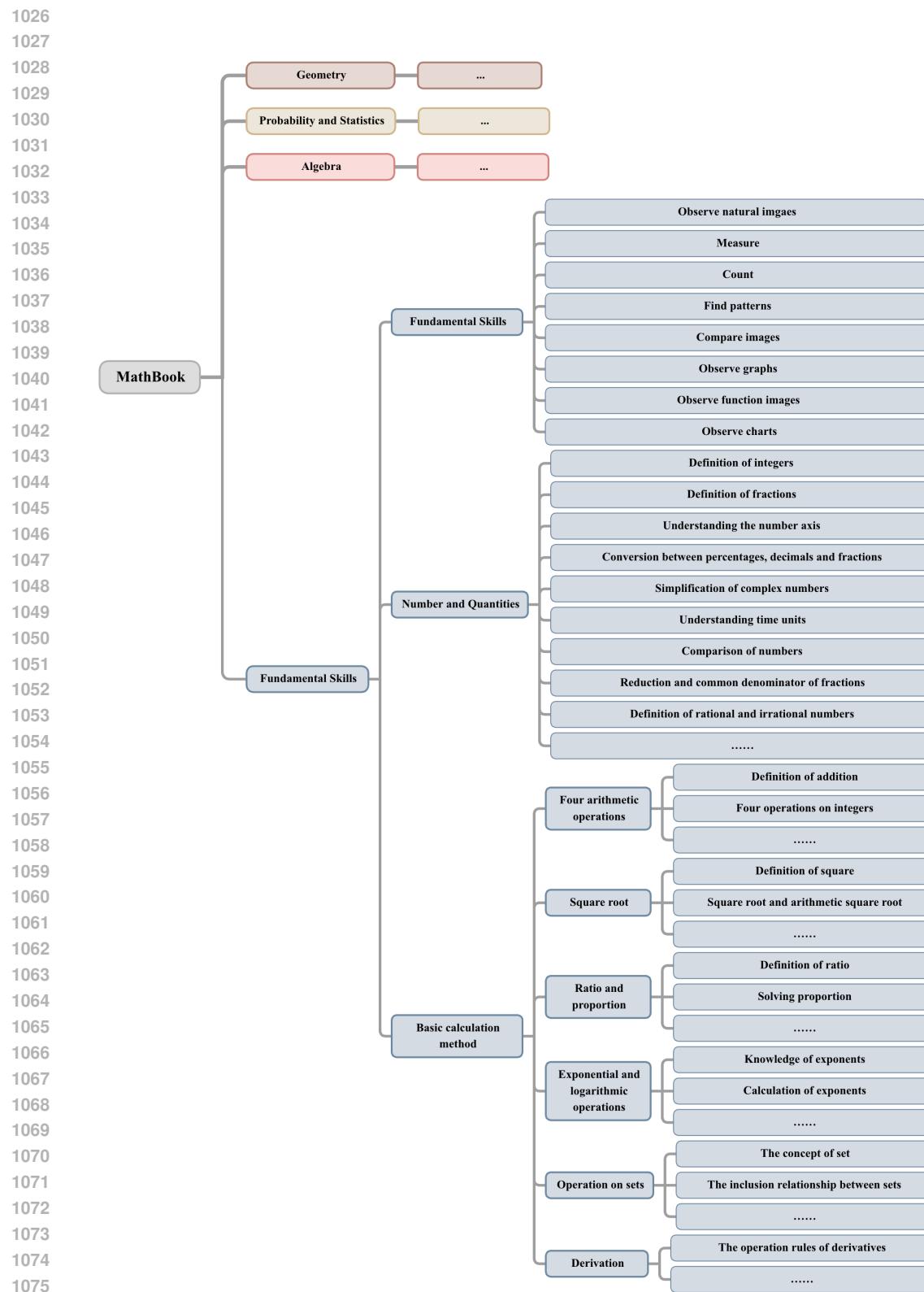


Figure 5: Overview of the hierarchical structure of knowledge points in MathBook (2).

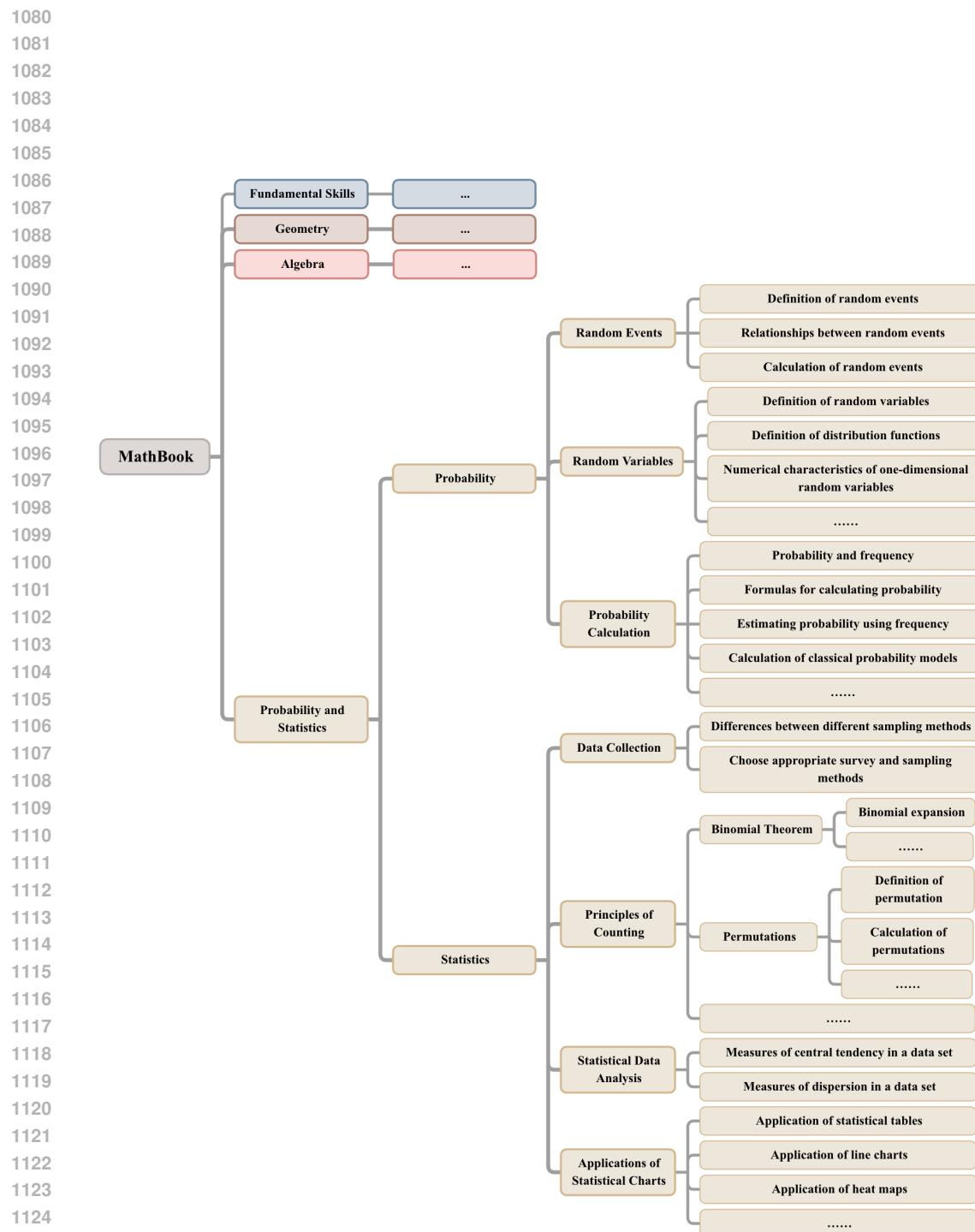


Figure 6: Overview of the hierarchical structure of knowledge points in MathBook (3).

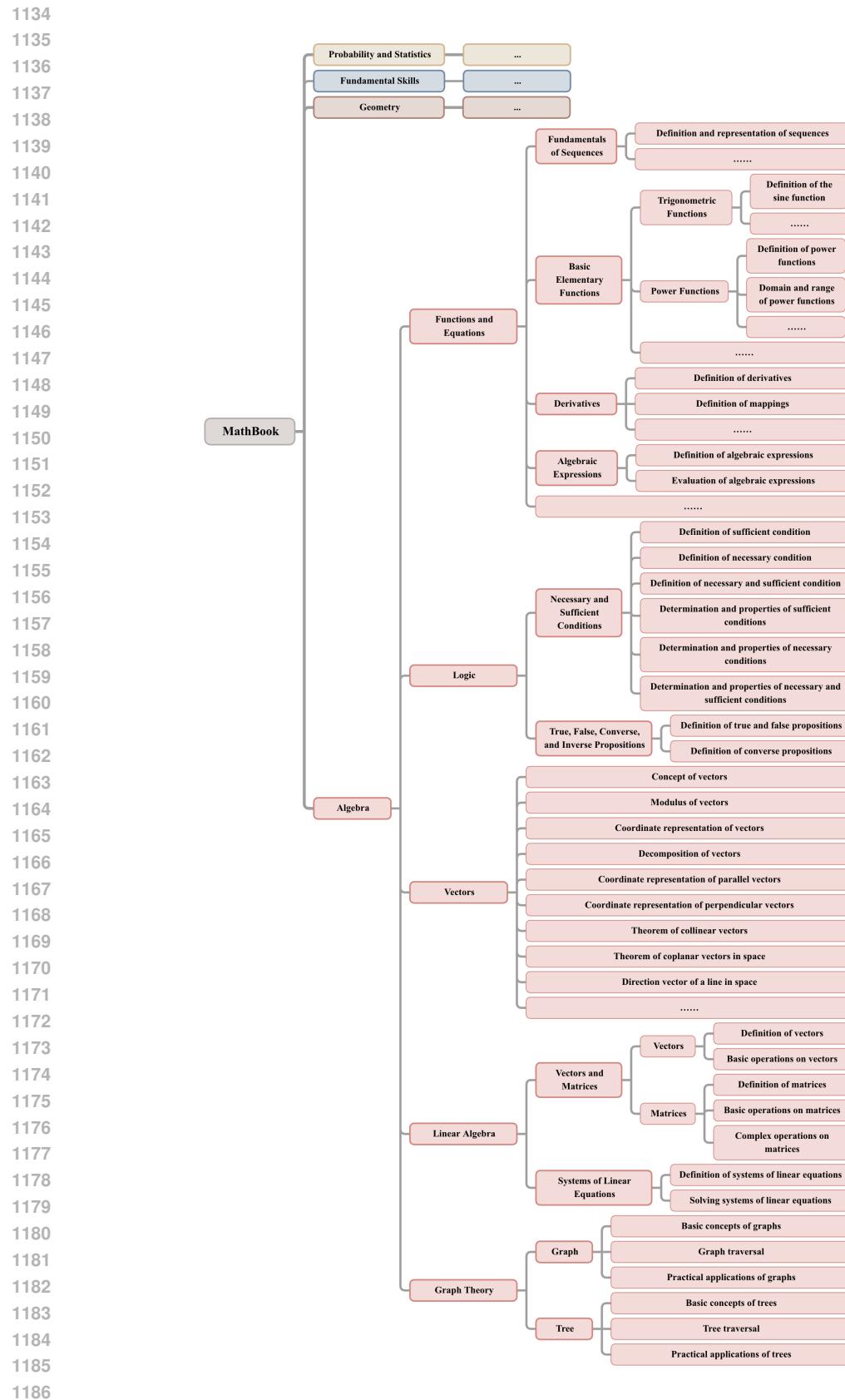


Figure 7: Overview of the hierarchical structure of knowledge points in MathBook (4).

1188

1189

1190

1191

1192

1193

1194

1195

1196

1197

1198

1199

1200

1201

1202

1203

1204

1205

1206

1207

1208

1209

1210

1211

1212

1213

1214

1215

1216

1217

1218

1219

1220

1221

1222

1223

1224

1225

1226

1227

1228

1229

1230

1231

1232

1233

1234

1235

1236

1237

1238

1239

1240

1241

Problem Retrieval

Tip: Select from the left to narrow to specific knowledge points & principles.

English 中文 Search problems / knowledge / principles..

Filters

Answer:941.192

Level 1: Geometry

Level 2: Solid Geometry

Level 3: Basic Solid Figures

Level 4: Rectangular Prism and Cube

Knowledge: Volume of Cube

Principle: Practical Applications of Cube Volume

Question: What is the maximum volume of a cube that can fit in this box?

Answer:\$16\sqrt{2}

Knowledge - Volume of Cube

Principle - Comparative Analysis of Cube and Rectangular Cuboid Volume Formulas

Question: Find the volume of the rectangular prism based on the cube.

Answer:192

Knowledge - Volume of Cube

Principle - Comparative Analysis of Cube and Rectangular Cuboid Volume Formulas

Question: Find the volume of the rectangular prism based on the cube.

Answer:108

Knowledge - Volume of Cube

Principle - Comparative Analysis of Cube and Rectangular Cuboid Volume Formulas

Question: Find the volume of the cube based on the rectangular prism.

Find the volume of the cube based on the rectangular prism.

Figure 8: An example from We-Math 2.0 showing how each problem is explicitly aligned with a specific knowledge point and its associated principle.

1211

1212

1213

1214

1215

1216

1217

1218

1219

1220

1221

1222

1223

1224

1225

1226

1227

1228

1229

1230

1231

1232

1233

1234

1235

1236

1237

1238

1239

1240

1241

Geometry

Geometry-Plane Geometry-Basic Plane Figures-Triangles -Classification of triangles

Principle:

1. Definition of an acute triangle: A triangle where all three interior angles are less than 90° (e.g., angles of 50° , 60° , 70°), with all altitudes located inside the triangle and the orthocentre inside as well: used in stable structural designs (e.g., Eiffel Tower trusses).
2. Definition of a right triangle: A triangle with one 90° angle (e.g., sides 3, 4, 5 satisfy $3^2 + 4^2 = 5^2$), the other two angles are complementary ($\alpha + \beta = 90^\circ$), and the Pythagorean theorem $a^2 + b^2 = c^2$ is used in GPS triangle positioning for coordinate calculations.
3. Definition of an obtuse triangle: A triangle with one interior angle greater than 90° and less than 180° (e.g., angles of 100° , 40° , 40°), where the side opposite the largest angle is longest; its corresponding altitude lies inside the triangle, and the other two altitudes need to be extended outside the triangle (e.g., for structural stress distribution in engineering blueprints).

Geometry-Plane Geometry-Relationships Between Figures -Intersection and Perpendicularity-Adjacent supplementary angles

Principle:

1. Definition of adjacent supplementary angles: Two adjacent angles (sharing one common side) whose measures sum to 180° are called adjacent supplementary angles (e.g., lines AB and CD intersect at O , then $\angle AOC$ and $\angle COB$ are adjacent supplementary angles).
2. Characteristics of adjacent supplementary angles: (1) A common side is shared and lies between the angles; (2) the other sides are extensions of each other in opposite directions; (3) their sum is a straight angle (mathematical expression: $\angle \alpha + \angle \beta = \pi$). Used in geometric proofs to complete angle relationships.
3. Judgment of adjacent supplementary angles: If two angles share one side and their non-common sides form a straight line, then they are adjacent supplementary angles (e.g., if $\angle 1$ and $\angle 2$ share side OA and OB lie on straight line AB). Applications: constructing interior angles of parallelograms, computing hinge opening angles in doors/windows.
4. Vector expression of adjacent supplementary angles: If the angle between two vectors is θ , then the sum of the angle and its adjacent supplement is $\theta + (180^\circ - \theta) = 180^\circ$.

Geometry-Solid Geometry-Basic Solid Figures-Cones -Section of a cone

Principle:

1. Cross-section of a cone: a cross-section parallel to the base is a circle ($r' = r * h'/h$, where h' is the distance from the apex), used in layered volume calculation.
2. A vertical cross-section along the height is an isosceles triangle (base = $2r$, height = h , area = $\pi * r * h$), used in symmetric cutting of cone sculptures.
3. An oblique section results in a conic section: depending on the angle, the section may be an ellipse (angle < angle between generator and height), a parabola (parallel to generator), or a hyperbola (angle > angle between generator and height). These are used in classifying conic sections and applied in satellite dish design and spotlight reflection paths.

Geometry-Solid Geometry-Basic Solid Figures -Volume of irregular solids

Principle:

1. Object combination: Combine known basic solid figures into a new object and calculate the volume of the new object. Usually, the volume of the basic solid figures as the combination elements can be calculated separately, and finally the volume of the irregular figure can be obtained by adding them up.
2. Object disassembly: Remove a part of the volume from the known basic solid figure and calculate the volume of the new object. Usually, the volume of the original solid figure and the removed volume can be calculated separately, and finally the volume of the irregular figure can be obtained by subtracting them.
3. ...

Geometry-Analytic Geometry-Analytic Geometry -Definition of the slope of a line

Principle:

1. Conditions for two lines to be parallel: When the slopes of the two lines are equal, that is, $k_1 = k_2$, and the intercepts are different (such as $y=2x+3$ and $y=2x-5$ are parallel); Parallel determination questions (verifying whether the two lines are parallel) and equation construction questions (finding equations parallel to the known lines); Easy mistakes: Ignoring the same intercepts leads to the coincidence of the lines
2. Conditions for two lines to be perpendicular: When the product of the slopes of the two lines is -1 , that is, $k_1 \cdot k_2 = -1$, or one line is perpendicular (slope does not exist) and the other is horizontal ($k=0$); vertical determination questions (verifying the orthogonality of lines) and geometric construction questions (finding the equation of the perpendicular line to a given line); easy mistakes: ignoring the special cases of vertical and horizontal lines (such as $x=3$ and $y=5$ are perpendicular)
3. ...

Geometry-Analytic Geometry-Circles -Problem of a circle passing through a fixed point

Principle:

1. Parametric equation processing: For the equation with parameters $x^2+y^2+Ax+By+C=0$, the fixed point satisfies A, B, C and the coefficients are independent of the parameters (for example: the equation $x^2+y^2+1x+1y-5=0$ needs to solve $1x+1y-5=0$ to get $x=1, y=2$)
2. Geometric method: Use the equation of the circle system (such as the intersection line of two circles and the third circle) to determine the fixed point; Easy to make mistakes: Failure to completely eliminate parameters leads to missed solutions, and parameter-related items are mistakenly regarded as independent conditions; Question type: Find all the fixed points that the circle passes through (for example: $kx^2+y^2+1x+1=0$ passes through the fixed point), prove that the circle system passes through the fixed point, and find the equation parameters at a known fixed point
3. ...

Figure 9: Examples of knowledge principles corresponding to specific knowledge points in MathBook (1).

1242

1243

1244

1245

1246

1247

1248

1249

1250

1251

1252

1253

1254

1255

1256

1257

1258

1259

1260

1261

1262

1263

Fundamental Skills

Fundamental Skills-Compare sizes in Images

Principle:

1. Comparison of function values: Compare y -values of different functions at the same x -value.
2. Comparison of rate of change: Compare slopes (rate of increase/decrease) over the same interval.
3. Comparison of extrema: Compare maximum/minimum values of different functions.
4. Comparison of definite integrals: Compare the area enclosed between the curve and the x -axis.

Fundamental Skills-Understanding Numbers and Quantities-Compare sizes in Images-Definition of integers

Principle:

1. Definition of natural numbers: Non-negative integers, used for counting or ordering.
2. Definition of positive integers, negative integers, and zero: Positive integers: Greater than 0, typically used for quantity. Zero: Neither positive nor negative, represents "none" or a reference point. Negative integers: Less than 0, representing opposite quantities.
3. Set notation: The set of integers is denoted by Z . The set of natural numbers is denoted by N .

Fundamental Skills-Basic Calculation Methods-Squaring and Square Rooting-Definition of square

Principle:

1. Definition of square: Square refers to the result of multiplying a number by itself, that is, a number multiplied by itself. For example, the square of 2 is $2 \times 2 = 4$. In mathematics, the symbol " \cdot^2 " is used to represent a square, for example, $2^2 = 4$.
2. Non-negativity of square: The square of any real number is non-negative, that is, the square root does not have a negative solution (such as $(-3)^2 = 9 \geq 0$).
3. Parity of square: If a number is even, then its square is also an even number (such as $4^2 = 16$); if it is odd, its square is also an odd number (such as $3^2 = 9$).
4. Definition of square number: A square number is the square of an integer, which can be written in the form of a perfect square (such as $1 = 1^2$, $4 = 2^2$, $9 = 3^2$, $16 = 4^2$).

Fundamental Skills-Basic Calculation Methods-Four Arithmetic Operations-Definition of division

Principle:

1. The essential definition of division: the inverse operation of multiplication, that is, to find another factor given a product and a factor. If $a \times b = c$, then $c \div b = a$ (where $b \neq 0$), represented by "÷" or fraction line ":", such as $a \div b$ or a/b , the divisor cannot be zero (undefined in mathematics) does not satisfy the commutative law ($6 \div 2 \neq 2 \div 6$).
2. The concept of remainder: when it cannot be divided evenly, the remaining part is called the remainder (such as $7 \div 3 = 2$ remainder 1).
3. Practical application: allocate resources, calculate the average. Easy to make mistakes: the remainder unit is not marked, ignore the divisibility condition.
4. Reciprocal association: $a \div b = a \times (1/b)$, division is converted into multiplication by the reciprocal

Figure 10: Examples of knowledge principles corresponding to specific knowledge points in Math-Book (2).

1264

1265

1266

1267

1268

1269

1270

1271

1272

1273

1274

1275

1276

1277

1278

1279

1280

1281

1282

1283

1284

1285

1286

1287

1288

1289

1290

1291

1292

1293

1294

1295

Probability and Statistics

Probability and Statistics-Probability-Random Events -Definition of random events

Principle:

1. Random event: An experiment satisfies the condition that it can be repeated under the same conditions, and all possible results are clearly known. If the results of each experiment are uncertain, then the experiment is a random experiment, such as shooting a basketball and flipping a coin. Results that may or may not appear in an experiment are called random events. Events that must occur in each experiment are called inevitable events, and events that must not occur are called impossible events, denoted as empty set.
2. Sample space: Each possible result of a random experiment is called a sample point, and the set of all sample points is called the sample space. Random events are always composed of several basic events.
3. ...

Probability and Statistics-Probability-Probability Calculation -Estimating probability using frequency

Principle:

1. The principle of frequency estimation probability: approximate the true probability by the frequency of events in a large number of repeated experiments (the mathematical basis is the law of large numbers: when the number of experiments is $n \rightarrow \infty$, the frequency is $f_n(A) \rightarrow P(A)$)
2. Application scenarios: weather forecast (historical rainfall frequency predicts the probability of precipitation tomorrow), gambling game winning rate calculation (long-term statistics of roulette)
3. Formula example: The frequency of the number 6 appearing in 600 dice tossing is 98 times, and the estimated probability is $\hat{P} = \frac{98}{600} \approx 0.163$

Probability and Statistics-Statistics-Principles of Counting-Addition and Multiplication Principles-Step-by-step multiplication counting principle

Principle:

1. To complete a task with multiple steps, multiply the number of ways for each step (formula: $N = m_1 \times m_2 \times \dots \times m_k$, e.g., 4-digit password, each digit has 10 choices $\rightarrow 10^4 = 10,000$). Key condition: steps must be ordered and independent (e.g., "choose shirt" then "choose pants").
2. Application scenarios: password combinations (letters + numbers), travel routes (3 paths from A to B, 2 from B to C \rightarrow total = $3 \times 2 = 6$)
3. Special cases: repetition allowed (e.g., repeated license plate letters), partial restrictions (e.g., phone number cannot start with 0). Question types: restricted permutations (e.g., 3-digit numbers formed from 1-9 with no repeats)

Probability and Statistics-Statistics-Principles of Counting-Combinations -Combination-based counting problems

Principle:

1. At least/at most problems: "at least m " is total combinations minus combinations less than m (e.g., "at least 3 qualified items" = total - $C(n,0) - C(n,1) - C(n,2)$, formula: $C(n,k) \text{, at least } m = \sum_{i=m}^{n} C(n,i)$, e.g., probability of at least 2 defective items in 10 = $1 - (C(10,0) + C(10,1)) / 2^{10}$)
2. Element distribution problems: same items to different people \rightarrow use divider method (e.g., 10 identical balls to 3 people, each gets ≥ 1 : $C(9,2) = 36$, allowing empty box: $C(n+k-1, k-1)$ (e.g., 7 balls into 4 boxes: $C(10,3) = 120$)
3. Balls into boxes problem: distinguishable balls into distinguishable boxes (n choices per ball $\rightarrow m^n$); indistinguishable balls into distinguishable boxes (divider method); indistinguishable balls into indistinguishable boxes (integer partitions, e.g., Stirling numbers)

Figure 11: Examples of knowledge principles corresponding to specific knowledge points in Math-Book (3).

1296

1297

1298

1299

1300

1301

1302

1303

1304

1305

1306

1307

1308

1309

1310

1311

1312

1313

1314

1315

1316

1317

1318

1319

1320

1321

1322

1323

1324

1325

1326

1327

1328

1329

1330

1331

1332

1333

1334

1335

1336

1337

1338

1339

1340

1341

1342

1343

1344

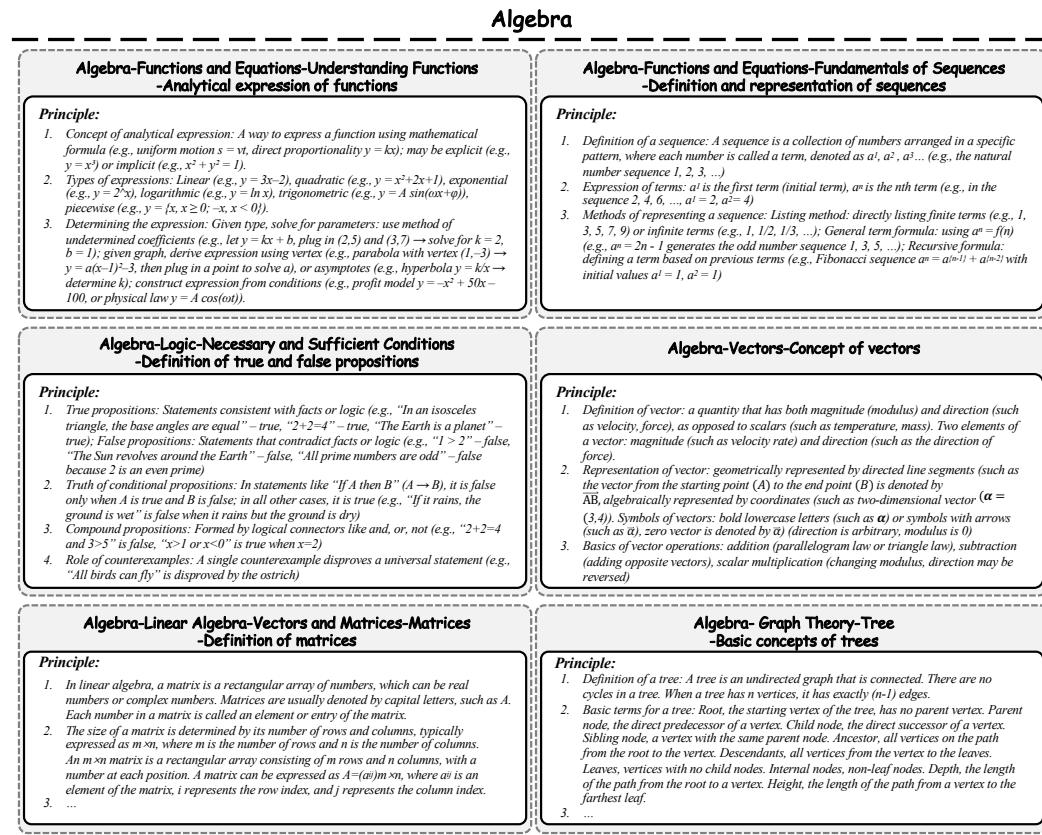
1345

1346

1347

1348

1349



1350 B.2 MATHBOOK-STANDARD
13511352 B.2.1 GEOGEBRA-BASED DIAGRAM GENERATION.
1353

1354 As described in the main text, all diagrams in MathBook-Standard are rendered using **GeoGebra**,
 1355 a dynamic mathematics software that enables precise and reproducible geometric constructions.
 1356 GeoGebra supports both interactive design and programmatic generation of diagrams, making it
 1357 highly suitable for large-scale dataset creation. In our workflow, each problem is paired with a
 1358 high-quality diagram constructed in GeoGebra, ensuring mathematical rigor and visual clarity. For
 1359 automated and scalable generation, we leverage GeoGebra’s ability to encode diagrams as scripts
 1360 (e.g., XML, as shown in Figure 13), which allows for efficient parameter variation and reproducibility,
 1361 but the core advantage lies in GeoGebra’s expressive power and accuracy for mathematical figures.
 1362

1363 Compared to general-purpose plotting libraries such as Python’s `matplotlib`, GeoGebra offers
 1364 richer geometric primitives and more precise control over mathematical relationships, supporting
 1365 a broader range of problem types and visual styles (Table 6). As shown in Figures 14, 15, and 16,
 1366 diagrams generated by GeoGebra exhibit higher fidelity and better alignment with mathematical
 1367 conventions, which is essential for both algorithmic evaluation and educational use. It is evident that
 1368 the complexity and precision of these diagrams would be difficult to achieve using Python-based
 1369 plotting tools alone.
 1370

Table 6: Comparison between GeoGebra and Python Plotting tools.

Tool	Command Line Plotting	Interactive Graphic Editing	Image Re-editing	Precise Parameter Control
GeoGebra	✓	✓	✓	✓
Python Plotting	✓	✗	✗	Limited

1376 B.2.2 DATASET DIVERSITY AND VARIANT CONSTRUCTION.
1377

1378 Building on the GeoGebra-based pipeline, we systematically construct a diverse dataset as detailed
 1379 in the main text. For each knowledge point and principle, a **seed problem** is designed with a
 1380 corresponding diagram. To further enhance diversity, we introduce two types of variants: **one-
 1381 problem-multi-image** (generating multiple diagram instances for the same problem by varying
 1382 parameters in GeoGebra) and **one-image-multi-problem** (curating multiple questions for a single
 1383 diagram, each derived from different knowledge points or mathematical principles). Representative
 1384 examples of seed problems and their variants are shown in Figure 17–24, demonstrating the flexibility
 1385 and extensibility of our approach.
 1386

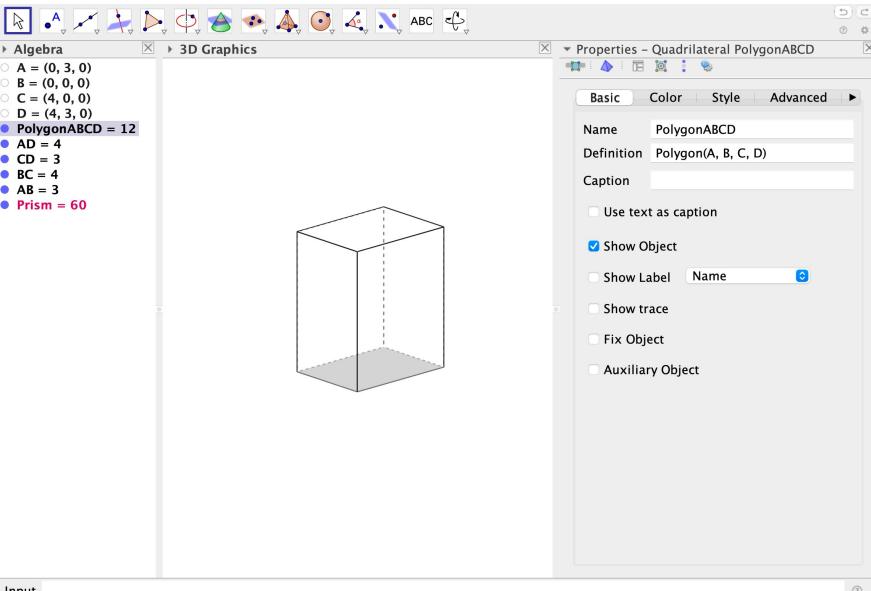
1387 By leveraging GeoGebra’s capabilities, MathBook-Standard achieves both high-fidelity geometric
 1388 representation and systematic dataset expansion, ensuring rich semantic and visual diversity for
 1389 mathematical reasoning tasks.
 1390

1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403

```

1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424 GGB
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439 XML
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457

```



The screenshot shows the GeoGebra interface with the following components:

- Algebra View:** Displays points A, B, C, D and a polygon ABCD with area 12, and a prism with volume 60.
- 3D Graphics View:** Shows a 3D prism with a base quadrilateral ABCD and a top quadrilateral EFGH.
- Properties View:** Shows the properties of the polygon ABCD, including its name, definition, and various checkboxes for visibility and labeling.

```

<command name="Polygon">
    <input a0="A" a1="B" a2="C" a3="D" />
    <output a0="PolygonABCD" a1="AB" a2="BC" a3="CD" a4="AD" />
</command>

<command name="Prism">
    <input a0="PolygonABCD" a1="5" />
    <output a0="Prism" a1="E" a2="F" a3="G" a4="H" a5="faceABFE" a6="faceBCGF"
           a7="faceCDHG" a8="faceADHE" a9="faceEFGH" a10="AE" a11="edgeBF" a12="CG" a13="edgeDH"
           a14="edgeEF" a15="edgeFG" a16="edgeGH" a17="edgeEH" />
</command>

```

Figure 13: Overview of the GeoGebra interface and part of the corresponding XML script, showing core commands for defining geometric objects.

1458

1459

1460

1461

1462

1463

1464

1465

1466

1467

1468

1469

1470

1471

1472

1473

1474

1475

1476

1477

1478

1479

1480

1481

1482

1483

1484

1485

1486

1487

1488

1489

1490

1491

1492

1493

1494

1495

1496

1497

1498

1499

1500

1501

1502

1503

1504

1505

1506

1507

1508

1509

1510

1511

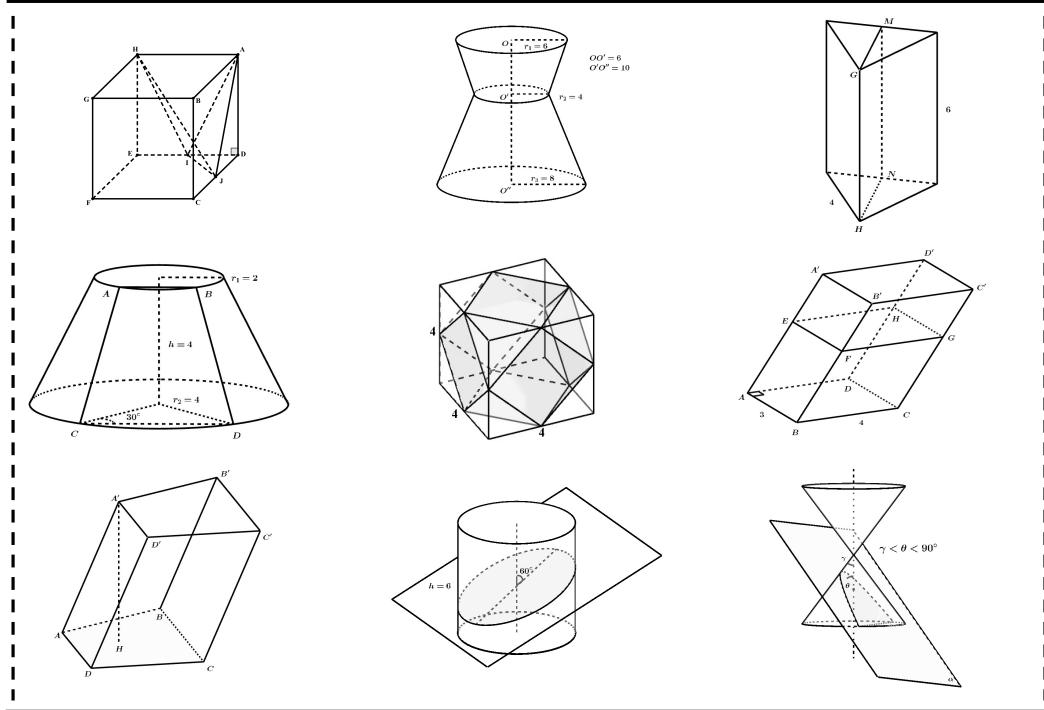
Examples (1)

Figure 14: Overview of a group of GeoGebra-generated images in MathBook (1).

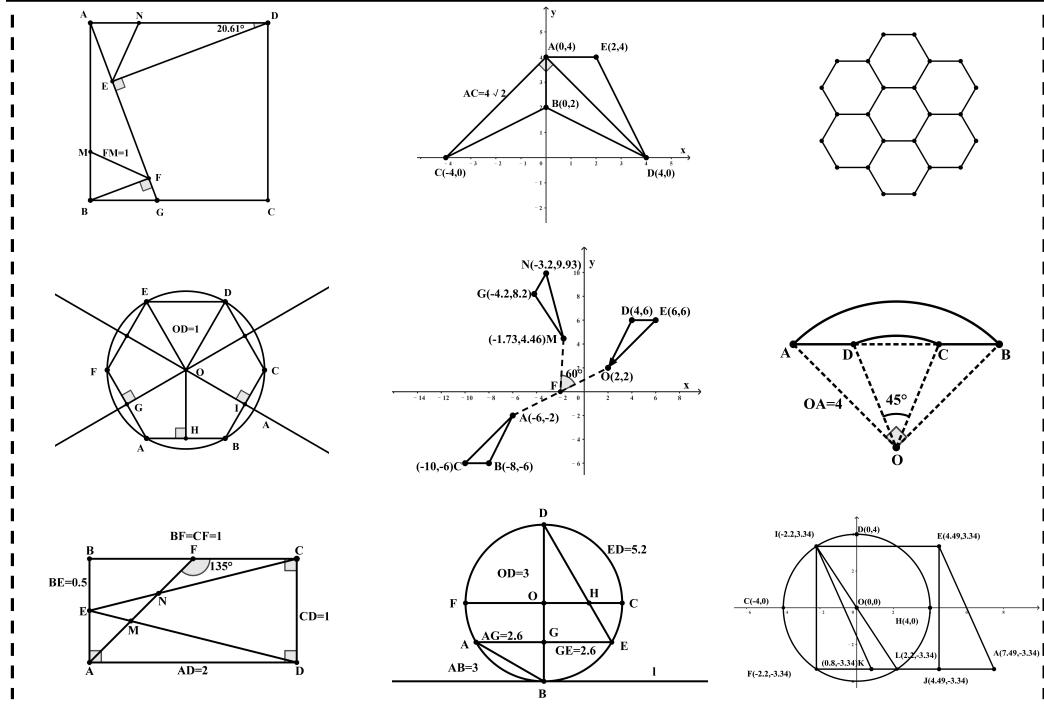
Examples (2)

Figure 15: Overview of a group of GeoGebra-generated images in MathBook (2).

1512

1513

1514

1515

1516

1517

1518

1519

1520

1521

1522

1523

1524

1525

1526

1527

1528

1529

1530

1531

1532

1533

1534

1535

1536

1537

1538

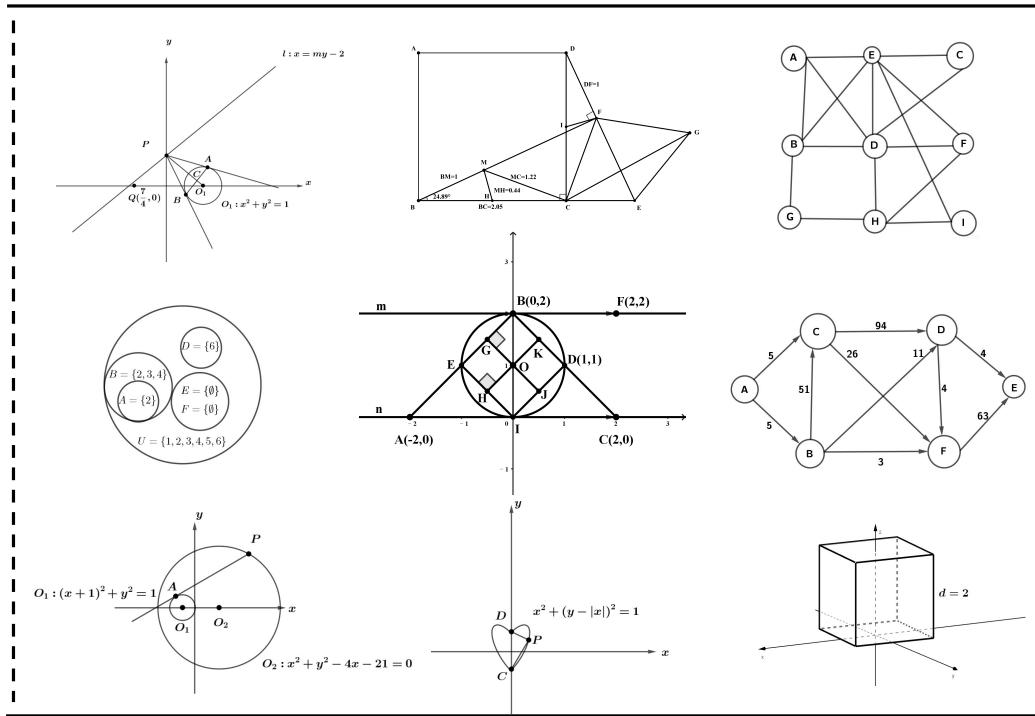
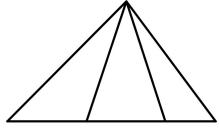
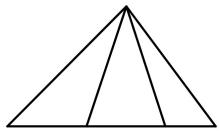
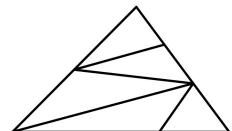
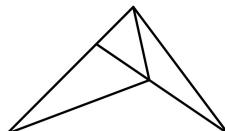
Examples (3)

Figure 16: Overview of a group of GeoGebra-generated images in MathBook (3).

Example 1**Seed****How many triangles are there in the picture?****One-Image-Multi-Problem****1. How many triangles are there in the picture?****2. How many line segments are there in the picture?****3. How many angles are there in the picture?****One-Problem-Multi-Image****How many triangles are there in the picture?**

1564

1565

Figure 17: An example of MathBook-Standard data instance (1).

1566

1567

1568

1569

1570

1571

1572

1573

1574

1575

1576

1577

1578

1579

1580

1581

1582

1583

1584

1585

1586

1587

1588

1589

1590

1591

1592

1593

1594

1595

1596

1597

1598

1599

1600

1601

1602

1603

1604

1605

1606

1607

1608

1609

1610

1611

1612

1613

1614

1615

1616

1617

1618

1619

Example 2

Seed

Compare the magnitudes of the fractions represented by the shaded parts in circles A and B.

One-Image-Multi-Problem

1. What is the difference between the fractions indicated by the shaded areas of Circle B and Circle A?
2. What is the sum of the fractions indicated by the shaded areas of Circle B and Circle A?
3. Compare the size of the fractions represented by the blank parts in circle A and circle B.

One-Problem-Multi-Image

Compare the magnitudes of the fractions represented by the shaded parts in circles A and B.

Figure 18: An example of MathBook-Standard data instance (2).

1595

1596

1597

1598

1599

1600

1601

1602

1603

1604

1605

1606

1607

1608

1609

1610

1611

1612

1613

1614

1615

1616

1617

1618

1619

Example 3

Seed

What three-dimensional geometric solid is generated by rotating the figure about the dotted line?

One-Image-Multi-Problem

1. What three-dimensional geometric solid is generated by rotating the figure about the dotted line?
2. What type of figure represents the side profile of a solid generated by rotation?
3. What is the shape of the base of the solid generated by rotation?

One-Problem-Multi-Image

What three-dimensional geometric solid is generated by rotating the figure about the dotted line?

Figure 19: An example of MathBook-Standard data instance (3).

1620

1621

1622

1623

1624

1625

1626

1627

1628

1629

1630

1631

1632

1633

1634

1635

1636

1637

1638

1639

1640

1641

1642

1643

1644

1645

1646

1647

1648

1649

1650

1651

1652

1653

1654

1655

1656

1657

1658

1659

1660

1661

1662

1663

1664

1665

1666

1667

1668

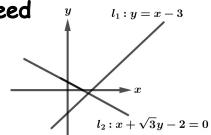
1669

1670

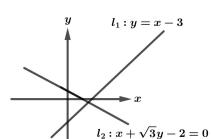
1671

1672

1673

Example 4**Seed**

What is the Angle between the two straight lines in the figure?

**One-Image-Multi-Problem**

1. What is the slope of the straight line l1?
2. What type of triangle is formed by the two lines and the y-axis in the figure?
3. What is the intercept of the straight line l1?

**One-Problem-Multi-Image**

What is the Angle between the two straight lines in the figure?

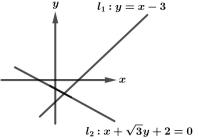
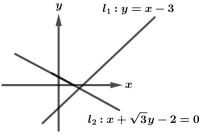
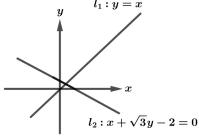
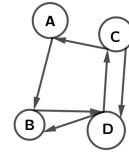
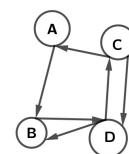


Figure 20: An example of MathBook-Standard data instance (4).

Example 5**Seed**

What are the in-degree and out-degree of vertex B respectively?

**One-Image-Multi-Problem**

1. What are the in-degree and out-degree of vertex A respectively?
2. What is the length of the shortest path from C to B?
3. Does a cycle exist in the figure?

**One-Problem-Multi-Image**

What are the in-degree and out-degree of vertex B respectively?

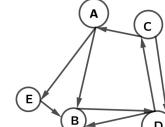
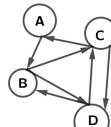
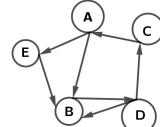


Figure 21: An example of MathBook-Standard data instance (5).

1674

1675

1676

1677

1678

1679

1680

1681

1682

1683

1684

1685

1686

1687

1688

1689

1690

1691

1692

1693

1694

1695

1696

1697

1698

1699

1700

1701

1702

1703

1704

1705

1706

1707

1708

1709

1710

1711

1712

1713

1714

1715

1716

1717

1718

1719

1720

1721

1722

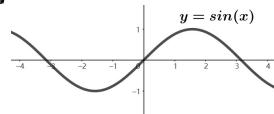
1723

1724

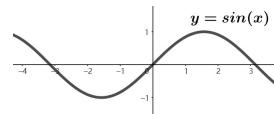
1725

1726

1727

Example 6**Seed**

Determine the monotonically increasing interval of the function in the graph.

**One-Image-Multi-Problem**

1. Determine the monotonically decreasing interval of the function in the graph.
2. Determine the zero point of the function in the graph.
3. Determine the symmetry points of the functions in the graph.

**One-Problem-Multi-Image**

Determine the monotonically increasing interval of the function in the graph.

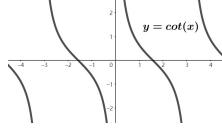
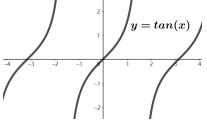
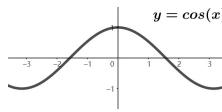


Figure 22: An example of MathBook-Standard data instance (6).

Example 7**Seed**

Divide the circles in the following picture into 3 equal parts. How many are there in each part?

**One-Image-Multi-Problem**

1. Divide the circles in the following picture into 2 equal parts. How many are there in each part?
2. If the circles are arranged in a rectangular grid with 3 rows, how many columns are there?
3. If all circles are tightly packed into a square shape, how many circles are on the outermost layer?

**One-Problem-Multi-Image**

Divide the circles in the following picture into 3 equal parts. How many are there in each part?

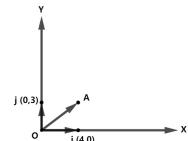


Figure 23: An example of MathBook-Standard data instance (7).

1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742

1743 **Example 8**

1744  **Seed**

1745  The vector OA can be decomposed into the vectors i and j shown in the figure. What are the coordinates of vector OA ?

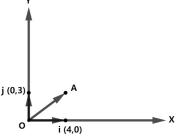
1746

1747

1748

1749

1750  **One-Image-Multi-Problem**

1751  1. What is the abscissa of vector OA ?
2. What is the ordinate of vector OA ?
3. What is the modulus of vector OA ?

1752

1753

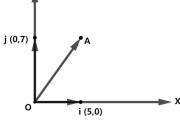
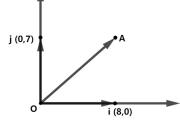
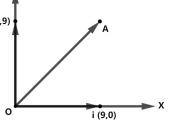
1754

1755

1756

1757  **One-Problem-Multi-Image**

1758 The vector OA can be decomposed into the vectors i and j shown in the figure.
1759 What are the coordinates of vector OA ?

1760   

1761

1762

1763

1764

1765

1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781

Figure 24: An example of MathBook-Standard data instance (8).

1782 B.3 MATHBOOK-PRO
1783

1784 As shown in Figure 25, we present a concrete example from MathBook-Pro to illustrate the
1785 construction and expansion of problem variants within the three-dimensional difficulty space. The seed
1786 problem, positioned at the origin, focuses on the *arc length formula* in plane geometry, and involves
1787 knowledge points such as *definition of angle*, *definition of circles*, *four arithmetic operations of*
1788 *integers* and *four arithmetic operations of fractions*.

1789 We first demonstrate how the seed problem is expanded along each individual dimension:
1790

1791 **Step Complexity:** The number of required knowledge points is increased by introducing new
1792 intermediate conclusions as conditions. In this example, the extended variant requires not only the
1793 *arc length formula*, but also incorporates *circumference of a circle* and *area of a circle* as additional
1794 knowledge points. The solution to the new problem depends on the answer to the seed problem,
1795 reflecting a progressive deepening of reasoning.

1796 **Visual Complexity:** The original diagram is enhanced by adding shaded regions, which increases
1797 the visual and interpretive demands while keeping the core mathematical focus unchanged.

1798 **Contextual Complexity:** The problem statement is recontextualized from a direct geometric descrip-
1799 tion to a real-world application scenario. Although the narrative becomes more complex, the essential
1800 assessment of the *arc length formula* remains at the core.

1801 By systematically combining these single-dimension expansions, we further generate multi-
1802 dimensional variants that integrate increased step, visual, and contextual complexity. In total,
1803 starting from the seed problem, we construct 7 variants corresponding to all possible combinations
1804 of the three dimensions. Each new variant is constructed through progressive modifications to both
1805 the problem statement and the accompanying image, resulting in a diverse and interpretable set
1806 of difficulty-controlled problems. The full set of variants and their corresponding dimensions are
1807 summarized in Table 7.

1808 As shown in Figure 26, MathBook-Pro supports the Dynamic Scheduling RL framework by pro-
1809 viding difficulty-controlled problem variants. These structured training samples enable progressive
1810 adaptation across different difficulty levels in a dynamic training process.

1811
1812
1813 Table 7: Difficulty-controlled variants constructed from the seed problem in MathBook-Pro. Each
1814 variant corresponds to a unique combination of step, visual, and contextual complexity.

1815 Variant	1816 Step	1817 Visual	1818 Contextual
1819 Seed	1820 -	1821 -	1822 -
1823 Variant 1	1824 ✓	1825 -	1826 -
1827 Variant 2	1828 -	1829 ✓	1830 -
1831 Variant 3	1832 -	1833 -	1834 ✓
1835 Variant 4	1836 ✓	1837 ✓	1838 -
1839 Variant 5	1840 ✓	1841 -	1842 ✓
1843 Variant 6	1844 -	1845 ✓	1846 ✓
1847 Variant 7	1848 ✓	1849 ✓	1850 ✓

1825 B.4 MATHBOOKEVAL
1826

1827 B.4.1 DATASET CONSTRUCTION AND ANNOTATION PROTOCOL.

1828 To ensure both comprehensive knowledge coverage and rigorous, interpretable annotations for
1829 visual mathematical reasoning, MathBookEval is constructed through a multi-stage, process-oriented
1830 pipeline. We begin by integrating representative samples collected from five open-source benchmarks:
1831 MathVista Lu et al. (2023), MathVerse Zhang et al. (2024a), MathVision Wang et al. (2024), We-
1832 Math Qiao et al. (2024a) and DynaMath Zou et al. (2024), systematically filtering out redundant or
1833 highly similar items to maximize diversity in knowledge point combinations and reasoning patterns.
1834 All problems are re-annotated under unified guidelines, ensuring consistency in annotation style and
1835

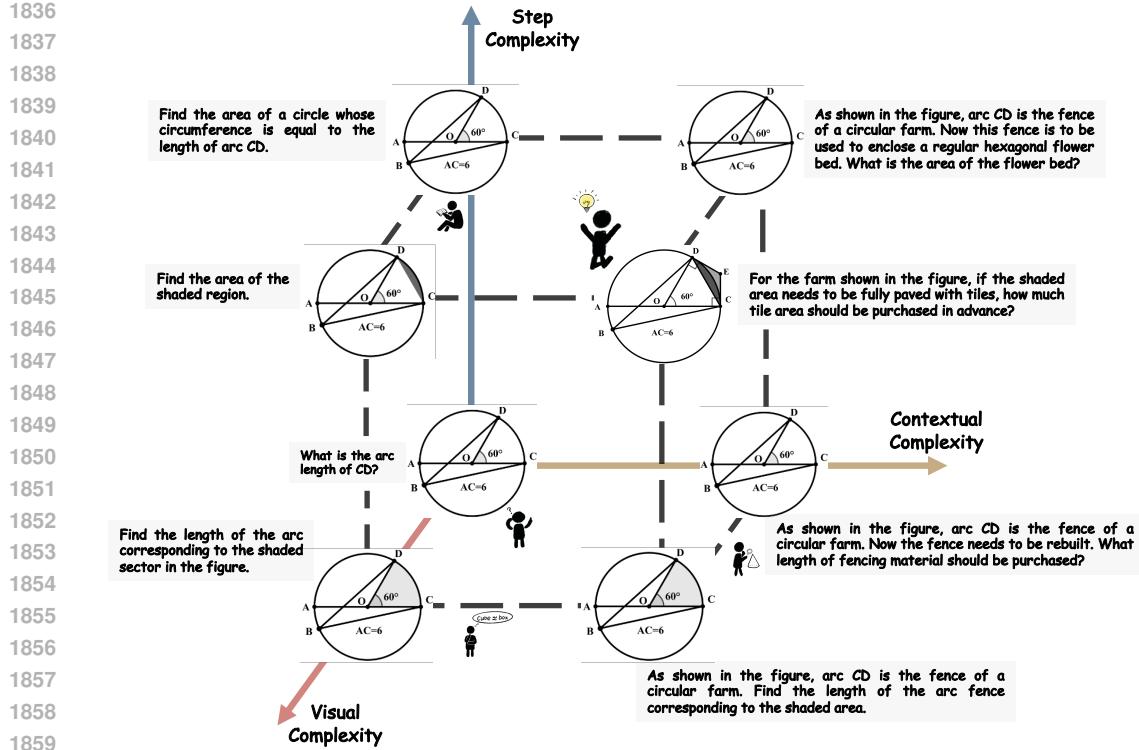


Figure 25: An example from MathBook-Pro in the difficulty space.

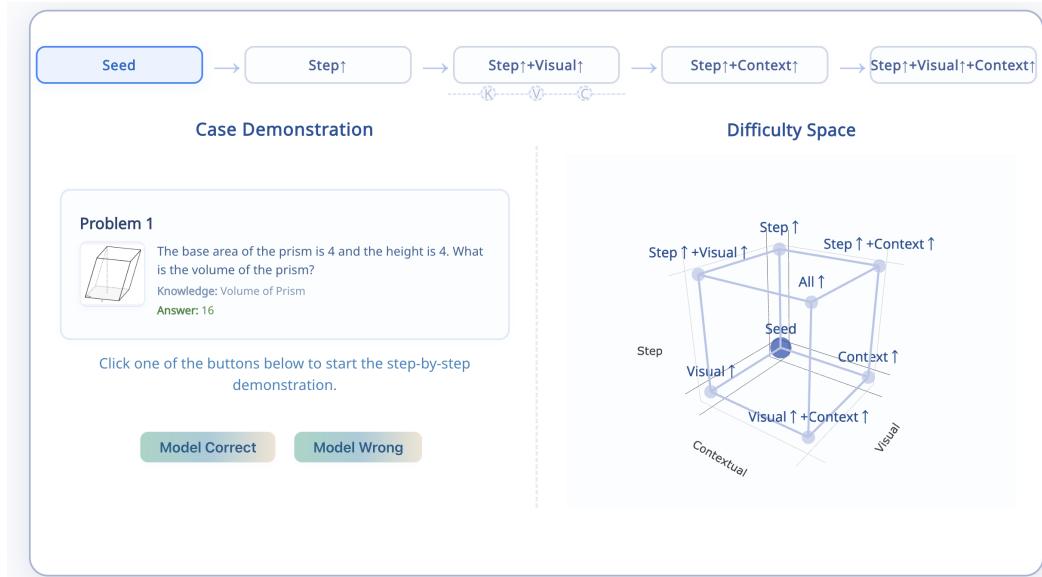


Figure 26: Overview of the Dynamic Scheduling RL training stage with MathBook-Pro.

1884
1885
1886
1887
1888
1889

granularity. For each problem, at least two human experts independently provide a complete, step-by-step solution, where each step is explicitly decomposed according to the underlying knowledge point(s) from the unified knowledge system \mathcal{K} . This knowledge-point-based decomposition is fundamental: it enables precise and systematic quantification of reasoning depth, as each reasoning

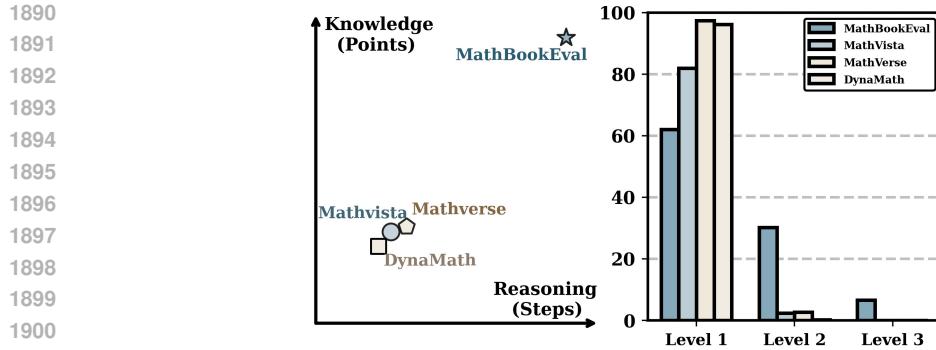


Figure 27: Comparison of MathBookEval and open-source benchmarks

step directly corresponds to a specific knowledge point. Only those problems for which the set of annotated knowledge points at each step is exactly consistent across expert annotations are retained in the final dataset, ensuring high reliability and objectivity. To address knowledge points and reasoning depths insufficiently covered by existing benchmarks, additional samples are newly constructed by human experts following the same process-oriented, knowledge-point-level annotation protocol. As a result, MathBookEval comprises both collected samples from open-source benchmarks and newly constructed samples, achieving balanced and comprehensive coverage across mathematical domains and reasoning complexities, with every annotation step tightly aligned to the relevant knowledge points for robust reasoning depth modeling.

B.4.2 TASK DIMENSIONS.

MathBookEval is organized along two dimensions, capturing both the diversity of mathematical knowledge and the depth of reasoning required for problem solving.

(1) Reasoning Dimension: Problems are categorized by the number of reasoning steps required to reach the solution. Each step is explicitly defined and directly mapped to a specific knowledge point in the unified knowledge system \mathcal{K} . We define three levels: Level 1 (1–3 steps, basic reasoning), Level 2 (4–6 steps, intermediate reasoning), and Level 3 (7–10 steps, complex reasoning). This step-by-step annotation based on knowledge points allows for clear and objective quantification of reasoning depth, enabling detailed analysis of model performance across different levels of reasoning complexity.

(2) Knowledge Dimension: The 491 knowledge points in \mathcal{K} are distributed across 4 top-level domains and 13 subdomains, all of which are covered in the benchmark. Each problem is annotated with all the knowledge points involved in its solution, and every reasoning step is aligned with a specific knowledge point. This enables fine-grained evaluation of model capabilities across various mathematical topics and educational stages.

B.4.3 DATASET STATISTICS.

Table 8 summarizes the key statistics of MathBookEval. The benchmark contains 1,000 fully annotated problems, covering all 491 knowledge points in the unified knowledge system. Of these, 600 are sourced from open-source benchmarks and 400 are newly constructed to address coverage gaps and increase diversity. Problems are distributed across multiple formats, including multiple-choice and fill-in-the-blank, and span a wide range of reasoning depths and knowledge domains.

As shown in Figure 27, we compare MathBookEval and existing benchmarks. In the right panel, the y-axis represents the percentage of problems at each reasoning level. Note that for some benchmarks, the total does not reach 100% because we exclude problems that experts annotate as belonging to other subjects such as physics, chemistry, or biology, in order to ensure a rigorous comparison. It is evident from the figure that existing benchmarks contain less than 3% of problems at Level 2 (4–6 steps), and none at Level 3 (7–10 steps). In contrast, MathBookEval substantially supplements these two categories, providing a more comprehensive evaluation of multi-step reasoning abilities.

1944

1945

1946

1947

1948

1949

1950

1951

1952

1953

1954

1955

1956

1957

1958

1959

1960

1961

1962

1963

1964

1965

B.4.4 EVALUATION PROTOCOL AND METRICS.

MathBookEval incorporates existing evaluation protocols, including both LLM-as-a-judge and rule-based approaches. To ensure consistency, MathBookEval adopts the LLM-as-a-judge protocol as the evaluation rule. Specifically, we adopt the LLM-as-a-judge protocol, following MathVista and MathVerse by employing GPT-4o as the judge model and reporting overall accuracy as the evaluation metric. The detailed prompt used for evaluation is shown in Table 9.

Table 8: Key statistics of MathBookEval.

Statistics	Number
Total Problems	1,000
- Open-source Benchmarks	600
- MathVista	150
- MathVerse	150
- MathVision	100
- We-Math	100
- DynaMath	100
- Newly Constructed	400
Knowledge Points Covered	491
- Domains	4
- Subdomains	13
Reasoning Depth	
- Level 1 (1–3 steps)	62.0%
- Level 2 (4–6 steps)	30.2%
- Level 3 (7–10 steps)	7.8%

Table 9: Prompt templates for evaluation on MathBookEval.

Type	Prompt Template
Evaluation Prompt	<p>Now, we require you to solve a math question. Please briefly describe your thought process and provide the final answer.</p> <p>For multiple-choice questions, return the selected option and its content. For direct answer selection, return only the chosen result. For fill-in-the-blank questions, answer directly.</p> <p>Question: <Question></p> <p>Regarding the format, please answer following the template below, and be sure to include two <> symbols:</p> <p><Thought process>: <<your thought process>></p> <p><Answer>: <<your answer>></p>

B.4.5 EXPERIMENT SETUP

Details of the Evaluated Models. To evaluate the performance of various multimodal large models on mathematical tasks, we include a diverse set of recent models in our benchmark. Table 10 lists the release dates and official sources of all evaluated models. Additionally, Table 11 provides an overview of their architectural designs to support a comprehensive comparison.

Details of the Model Hyperparameters. For all closed-source models accessed via API, we adopt the standard generation settings and perform inference on CPUs, with the process typically completing within a day. For open-source models, inference is conducted on a cluster equipped with 8 NVIDIA A800-SXM4-80GB GPUs, using the hyperparameter configurations provided in the official inference examples. If no specific instructions are available, default settings are applied. The detailed generation parameters are summarized in Table 12 and Table 13.

1998

1999

2000

2001

2002

2003

2004

2005

2006

2007

2008

2009

2010

2011

2012

2013

2014

2015

2016

2017

2018

2019

2020

2021

2022

2023

2024

2025

2026

2027

2028

2029

2030

2031

2032

2033

2034

2035

2036

2037

2038

2039

2040

2041

2042

2043

2044

2045

Table 10: The release time and model source of MLLMs used in MathBookEval

Model	Release Time	Source
GPT-4o OpenAI (2024)	2024-05	https://gpt4o.ai/
GPT-4V OpenAI (2023)	2024-04	https://openai.com/index/gpt-4v-system-card/
InternVL2.5-78B (Chen et al., 2024b)	2024-12	https://huggingface.co/OpenGVLab/InternVL2_5-78B
InternVL2.5-8B (Chen et al., 2024b)	2024-12	https://huggingface.co/OpenGVLab/InternVL2_5-8B
Qwen2.5-VL-72B (Bai et al., 2025a)	2025-01	https://huggingface.co/Qwen/Qwen2.5-VL-72B-Instruct
Qwen2.5-VL-7B (Bai et al., 2025a)	2025-01	https://huggingface.co/Qwen/Qwen2.5-VL-7B-Instruct
Qwen2.5-VL-3B (Bai et al., 2025a)	2025-01	https://huggingface.co/Qwen/Qwen2.5-VL-3B-Instruct
LLaVA-OneVision-72B Li et al. (2024)	2024-08	https://huggingface.co/lmms-lab/llava-onevision-qwen2-72b-ov-chat
LLaVA-OneVision-7B (Li et al., 2024)	2024-08	https://huggingface.co/lmms-lab/llava-onevision-qwen2-7b-ov
GLM-4V-9B (GLM et al., 2024)	2024-06	https://huggingface.co/THUDM/glm-4v-9b

Table 11: Model architecture of 10 MLLMs evaluated on MathBookEval.

Models	LLM	Vision Encoder
GPT-4o	-	-
GPT-4V	-	-
InternVL2.5-78B	Qwen2.5-72B-Instruct	InternViT-6B-448px-V2_5
InternVL2.5-8B	internlm2_5-7b-chat	InternViT-6B-448px-V2_5
Qwen2.5-VL-72B	Qwen2.5-72B-Instruct	CLIP ViT-bigG-P14
Qwen2.5-VL-7B	Qwen2.5-7B-Instruct	CLIP ViT-bigG-P14
Qwen2.5-VL-3B	Qwen2.5-3B-Instruct	CLIP ViT-bigG-P14
LLaVA-OneVision-72B	Qwen2-72B	SigLip-so400m-P14-384
LLaVA-OneVision-7B	Qwen2-7B	SigLip-so400m-P14-384
GLM-4V-9B	GLM-9B	EVA_02_CLIP-E-P14

B.4.6 ADDITIONAL RESULTS ON MATHBOOKEVAL

Table 14 reports the complete results of all evaluated models on MathBookEval, providing a more comprehensive perspective on model behavior across both reasoning depth and knowledge-coverage dimensions. Overall, general-purpose open-source MLLMs exhibit stable performance trends, with larger models such as Qwen2.5-VL-72B and InternVL2.5-78B demonstrating strong accuracy across multiple reasoning levels and knowledge categories. In contrast, reasoning-oriented systems (for example, MM-Eureka-7B) show performance patterns influenced by their reasoning-focused training pipelines, excelling in certain multi-step inference settings while presenting varied outcomes across different knowledge domains. These results collectively illustrate how different training paradigms and objectives shape mathematical reasoning capabilities in complementary ways. Within this broader landscape, MathBook-7B achieves competitive performance across both reasoning and knowledge dimensions, reflecting the effectiveness of the structured knowledge design and progressive alignment strategies introduced in We-Math 2.0.

Furthermore, as shown in Figure 31 to Figure 40, we present the performance of different models on various subdomains in MathBookEval, where subdomains belonging to the same domain are indicated with the same color. It can be observed that models generally perform worse on geometry-related problems, especially in subdomains such as solid geometry and analytic geometry, which involve higher visual complexity and require more advanced reasoning. In contrast, models tend to achieve better results on algebra and fundamental skills, particularly in subdomains related to computational methods.

2052

2053

Table 12: Generating parameters for Open-Source MLLMs.

Model	Generation Setup
InternVL2.5-78B	do_sample = False, max_new_tokens = 1024
InternVL2.5-8B	do_sample = False, max_new_tokens = 1024
Qwen2.5-VL-72B	do_sample = False, max_new_tokens = 1024
Qwen2.5-VL-7B	do_sample = False, max_new_tokens = 1024
Qwen2.5-VL-3B	do_sample = False, max_new_tokens = 1024
LLaVA-OneVision-72B	do_sample = False, max_new_tokens = 1024
LLaVA-OneVision-7B	do_sample = False, max_new_tokens = 1024
GLM-4V-9B	do_sample = False

2065

2066

2067

Table 13: Generating parameters for Closed-Source MLLMs.

Model	Generation Setup
GPT-4o	"model" : "gpt-4o", "temperature" : 0, "max_tokens" : 1024
GPT-4V	"model" : "gpt-4-turbo", "temperature" : 0, "max_tokens" : 1024

2073

2074

Table 14: Full results of different MLLMs on MathBookEval. **Acc.:** Accuracy; **FS.:** Foundational skills; **PS.:** Probability and statistics; **Geo.:** Geometry; **Alg.:** Algebra.

Models	Acc.	Reasoning			Knowledge			
		Level1	Level2	Level3	FS.	PS.	Geo.	Alg.
<i>Closed-source MLLMs</i>								
GPT-4o	50.8	52.8	48.9	41.7	33.8	57.6	44.2	67.2
GPT-4V	42.8	44.0	43.0	31.9	36.8	56.6	33.5	59.4
<i>Open-source (General)</i>								
InternVL2.5-78B	51.8	52.5	51.8	45.8	50.0	64.2	42.6	67.6
Qwen2.5-VL-72B	57.1	58.3	56.4	50.0	52.9	58.5	52.1	68.8
LLaVA-OneVision-72B	43.0	44.8	42.0	31.9	38.2	52.8	37.0	53.5
InternVL2.5-8B	37.9	40.7	34.5	27.8	33.8	46.2	31.4	50.0
Qwen2.5-VL-7B	46.7	50.1	43.0	33.3	44.1	58.5	38.8	60.2
LLaVA-OneVision-7B	31.6	34.3	28.0	23.6	36.8	41.5	24.9	41.0
Qwen2.5-VL-3B	36.9	38.7	34.2	33.3	35.3	49.1	29.1	49.6
GLM-4V-9B	22.2	23.7	20.5	16.7	26.5	23.6	18.4	28.9
<i>Open-source (Reasoning)</i>								
MM-Eureka-7B	50.0	50.6	51.5	38.9	47.1	52.8	44.0	62.9
R1-VL-7B	38.0	41.9	32.6	27.8	38.2	38.7	32.3	50.4
OpenVLThinker-7B	32.6	33.0	32.6	29.2	45.6	38.7	24.9	43.4
VLAAThinker-7B	35.7	38.3	31.9	29.2	46.6	42.5	30.2	41.6
MathBook-7B	50.4	52.0	48.2	45.8	57.4	67.9	40.5	63.3

2099

2100

C MORE DETAILS ON MATHBOOK-RL

2102

2103

C.1 IMPLEMENTATION DETAILS

2104

2105

We use Qwen2.5-VL-7B-Instruct as the base model and conduct all experiments on 8×A800 GPUs. The training process consists of two stages.

2106 In the first stage, we perform cold-start supervised fine-tuning (SFT) to help the model develop
 2107 explicit awareness of knowledge system and a knowledge-driven reasoning paradigm. The SFT stage
 2108 uses a learning rate of 1.0×10^{-5} , is trained for 1 epoch, and adopts a warmup ratio of 0.1.
 2109

2110 In the second stage, we apply dynamic reinforcement learning (RL) to further improve the model’s
 2111 generalization ability. For RL, we set the rollout temperature to 1.0 and generate 8 rollouts per sample.
 2112 The learning rate is set to 1×10^{-6} , and the maximum completion length is 1024 tokens. The system
 2113 prompt used in this stage is illustrated in Table 15. The reward function combines answer accuracy
 2114 (weight 0.9) and response format compliance (weight 0.1). Specifically, we use MathVerify to extract
 2115 and compare answers for accuracy, while format compliance ensures the output follows the required
 2116 structure.
 2117

2118 **Table 15: The system prompt template for response generation in the RL stage.**

Type	Prompt Template
System Prompt	A conversation between User and Assistant. The user asks a question, and the Assistant solves it. The assistant first outputs the thinking process in <think> </think> and then provides the final answer(number, option, phrase, or LaTeX expression as appropriate) in <answer> </answer> tags. User: {question} Assistant:

2125 C.2 ADDITIONAL EXPERIMENTAL RESULTS FOR MATHBOOK-3B

2128 Table 16 summarizes the performance of MathBook-3B across multiple benchmarks. Our analysis
 2129 yields the following observations.
 2130

2132 **Table 16: Performance comparison between Qwen2.5-VL-3B and MathBook-3B across benchmarks.**

Model	MathVista	MathVision	WeMath	MathVerse	MathBookEval
Qwen2.5-VL-3B	60.8	21.1	22.9	29.1	36.9
MathBook-3B	63.2	24.3	28.1	32.6	40.2

2137 **(1) Consistent improvements across benchmarks.** MathBook-3B surpasses the Qwen2.5-VL-3B
 2138 backbone on all evaluated datasets, with notable gains on WeMath (+5.2), MathBookEval (+3.3), and
 2139 MathVerse (+3.5). These results indicate that the structured data design and training methodology
 2140 introduced in We-Math 2.0 remain effective even at smaller model scales, enhancing both multimodal
 2141 mathematical reasoning and generalization to diverse problem formats.
 2142

2143 **(2) Performance gains independent of model size.** Despite using a significantly smaller backbone,
 2144 MathBook-3B achieves improvements comparable in trend to those observed with MathBook-7B.
 2145 This suggests that the benefits of We-Math 2.0 arise from its principled knowledge structure and
 2146 training pipeline, rather than relying on large model capacity.
 2147

2148 C.3 GENERALIZATION TO TEXT-ONLY MATHEMATICAL REASONING TASKS

2149 To further examine whether the structured knowledge design in We-Math 2.0 transfers beyond the
 2150 multimodal setting, we additionally evaluate MathBook-7B on two widely used text-only mathemat-
 2151 ical reasoning benchmarks: GSM8K (Cobbe et al., 2021) and MATH-500 Lightman et al. (2023).
 2152 As shown in Table 17, MathBook-7B achieves consistent improvements over the Qwen2.5-VL-7B
 2153 backbone. Although these datasets differ substantially from the multimodal training distribution, the
 2154 gains observed here suggest that the structured knowledge supervision introduced in We-Math 2.0
 2155 can support more generalizable reasoning behavior across modalities.
 2156

2157 C.4 CASE STUDY

2158 To further illustrate the strengths of our approach, we present several representative case studies
 2159 comparing our model with the Qwen2.5-VL-7B baseline across different benchmarks.
 2160

2160
2161
2162
2163
2164
2165
2166

2167 Table 17: Results on text-only reasoning benchmarks.

Model	MATH-500	GSM8K
Qwen2.5-VL-7B	66.0	86.2
MathBook-7B	67.2	87.6

Case 1: Conciseness and Reasoning Process. Figure 28 compares the response patterns of our model and the Qwen2.5-VL-7B baseline on We-Math. Our model produces more concise answers, with a reduced average response length, while still retaining all necessary formulas and reasoning steps. This demonstrates that our approach effectively mitigates the issue of overthinking, resulting in more focused and efficient solutions without sacrificing completeness or mathematical rigor.

Case 2: Spatial Reasoning Enhancement. Figure 29 highlights the performance of our model on MathVision. Compared to the baseline, which often fails to correctly interpret or solve such problems, our model (MathBook-7B) shows clear improvements in spatial reasoning. This is particularly evident in questions requiring the understanding of geometric relationships or positional logic, indicating that our training strategy significantly enhances the model’s ability to handle spatially complex scenarios.

Case 3: Knowledge-Oriented and Context-Aware Mathematical Reasoning. In Figure 30, we compare the responses of our model and the Qwen2.5-VL-7B baseline on MathVista. Our model not only applies the relevant mathematical concepts correctly but also demonstrates improved integration of mathematical knowledge with real-world problem contexts. This case exemplifies the model’s strengthened ability to bridge abstract mathematical reasoning and practical application, reflecting the benefits of our knowledge-oriented and context-aware training paradigm.

Overall, these case studies provide qualitative evidence that our model achieves more concise, accurate, and contextually appropriate reasoning compared to strong baselines, particularly in scenarios demanding spatial understanding and knowledge-oriented problem solving.

2187 C.5 EXPERIMENT SETUP

2188 C.5.1 DETAILS OF THE EVALUATION.

2189 We evaluate our model on four representative benchmarks: MathVista Lu et al. (2023), Math-
2190 Verse Zhang et al. (2024a), MathVision Wang et al. (2024), and We-Math Qiao et al. (2024b). During
2191 evaluation, we strictly follow the official scoring protocols provided in the respective benchmark
2192 GitHub repositories to ensure fair and consistent comparison. Specifically, we report results on the
2193 testmini split for MathVista, MathVerse, MathVision, and We-Math. For We-Math, we adopt the
2194 main evaluation metric as defined in the original paper, reporting “Score (Strict)” as the primary
2195 metric. For the other benchmarks, we report the average accuracy as the main evaluation result.

2196 C.5.2 DETAILS OF THE BASELINES

2197 We compare our method with a comprehensive set of baselines from three perspectives: closed-
2198 source models (GPT-4o OpenAI (2024), Gemini-1.5-Pro Team et al. (2023)), open-source general
2199 models (Qwen2.5-VL-7B Bai et al. (2025a), InternVL2.5-8B Chen et al. (2023)), GLM-4V-9B GLM
2200 et al. (2024), LLaVA-OneVision-7B Li et al. (2024), and open-source reasoning models (Math-
2201 PUMA-7B Zhuang et al. (2024), URSA-8B Luo et al. (2025), R1-Onevision-7B Yang et al. (2025b),
2202 R1-VL-7B Zhang et al. (2025), MM-Eureka-7B Meng et al. (2025), WeThink-7B Yang et al. (2025a),
2203 VLAA-Thinker-7B Chen et al. (2025), OpenVLThinker-7B Deng et al. (2025)), which enables a
2204 thorough and multi-faceted comparison to highlight the advantages of our approach.

2205 **GPT-4o OpenAI (2024):** GPT-4o (“o” for omni) is OpenAI’s 2024 flagship multilingual, multi-
2206 modal large language model that accepts text, images, and audio as input for unified cross-modal
2207 understanding. It is designed for broad adaptability and seamless integration of visual and linguistic
2208 information, supporting complex reasoning across modalities and languages.

2209 **Gemini-1.5-Pro Team et al. (2023):** Gemini 1.5 Pro, developed by Google DeepMind, is a
2210 multimodal model capable of processing text, images, audio, and video inputs, with an extremely

2214
 2215
 2216
 2217
 2218
 2219
 2220
 2221
 2222
 2223
 2224
 2225
 2226
 2227
 2228
 2229
 2230
 2231
 2232
 2233
 2234
 2235
 2236
 2237
 2238
 2239
 2240
 2241
 2242
 2243
 2244
 2245
 2246
 2247
 2248
 2249
 2250
 2251
 2252
 2253
 2254
 2255
 2256
 2257
 2258
 2259
 2260
 2261
 2262
 2263
 2264
 2265
 2266
 2267

Comparison of MathBook-7B and Qwen2.5-VL-7B on We-Math

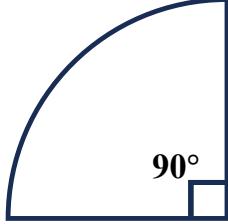
Question:

Now, we require you to solve a multiple-choice math question. Please briefly describe your thought process and provide the final answer(option). Question: As shown in the diagram, what is the arc length of this sector?

Options:

A. 2π ; B. 4π ; C. 8π ; D. No correct answer

Regarding the format, please answer following the template below, and be sure to include two <> symbols: <Thought process>> <Answer>: <<your option>>

4 

Response from Qwen2.5-VL-7B:

<Thought process>: The diagram shows a quarter circle (90 degrees) with a radius of 4 units. The formula for the arc length of a sector is $(\theta/360) * 2\pi r$, where θ is the angle in degrees and r is the radius. For a 90-degree sector, the arc length is $(90/360) * 2\pi * 4 = (1/4) * 2\pi * 4 = \pi * 4 = 4\pi$. Therefore, the arc length of this sector is 4π .

<Answer>: B. 4π 

Response from MathBook-7B:

<Thought process>: The arc length of a sector is calculated using the formula $(\frac{\theta}{360} \times 2\pi r)$, where θ is the angle in degrees and r is the radius. Here, $\theta = 90^\circ$ and $r = 4$. Plugging in the values, we get $(\frac{90}{360} \times 2\pi \times 4) = \frac{1}{4} \times 8\pi = 2\pi$. Thus, the arc length is 2π .

<Answer>: A. 2π 

Figure 28: Comparison of MathBook-7B and Qwen2.5-VL-7B on We-Math.

long context window (up to 1–2 million tokens). It is optimized for multi-task proficiency and structured tool integration, enabling analysis of lengthy and diverse content.

Qwen2.5-VL-7B Bai et al. (2025a): Qwen2.5-VL-7B is an open-source vision-language model with 7 billion parameters, designed to generate both free-form text and structured outputs such as bounding boxes and JSON. It emphasizes fine-grained visual understanding and multi-modal alignment, supporting tasks like document analysis and event detection.

InternVL2.5-8B Chen et al. (2023): InternVL2.5-8B is an 8B-parameter open-source multimodal model that employs progressive scaling and co-training strategies to align its vision and language components. It incorporates training optimizations and a curated dataset to enhance cross-modal reasoning and reduce hallucinations.

GLM-4V-9B GLM et al. (2024): GLM-4V-9B is a 9B-parameter multimodal model built on the GLM series architecture. It integrates visual understanding and language processing capabilities, and is optimized using large-scale multimodal data. With specialized architectural design and training strategies, it demonstrates strong performance in multimodal reasoning tasks and is capable of effectively handling a variety of vision-centric tasks.

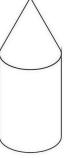
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321

Comparison of MathBook-7B and Qwen2.5-VL-7B on MathVision

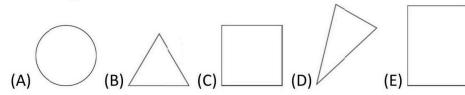
Question:

What do you see if you look at the tower, which is made up of two building blocks, exactly from above?

<image1>



<image2>



Response from Qwen2.5-VL-7B:

To determine what you would see if you looked at the tower from above, we need to analyze the shapes of the building blocks that make up the tower.

1. The tower consists of two building blocks: a cylinder and a cone.
2. When viewed from above, the cylinder and the cone will appear as their respective two-dimensional shapes.

- The cylinder, when viewed from above, appears as a circle.
- The cone, when viewed from above, appears as a triangle.

Since the question asks for the shape you would see from above, we need to identify the shape that corresponds to the top view of the cone.

The options provided are:

(A) Circle (B) Triangle (C) Square (D) Triangle (repeated) (E) Square

The correct answer is the shape that represents the top view of the cone, which is a triangle.

Therefore, the answer is \boxed{B}.



Response from MathBook-7B:

<think>The top view of the tower, which consists of a cone and a cylinder, is a circle. This is because the top surface of the cone is a circle, and the top surface of the cylinder is also a circle. Therefore, the top view is a circle.</think><answer>: A</answer>



Figure 29: Comparison of MathBook-7B and Qwen2.5-VL-7B on MathVision.

LLaVA-OneVision-7B Li et al. (2024): LLaVA-OneVision-7B is an open-source model designed for unified single-image, multi-image, and video understanding. It adopts an AnyRes visual representation strategy to enable cross-scenario capability transfer. Trained via a three-stage curriculum, it performs well across various benchmarks and exhibits emerging capabilities like video-to-video difference analysis and multi-camera self-driving video understanding via task transfer from images.

Math-PUMA-7B Zhuang et al. (2024): Math-PUMA-7B is a vision-language model focused on mathematical reasoning with visual inputs, introducing a three-stage curriculum for aligning textual and visual modalities. The model is optimized for visual math benchmarks and aims to ensure consistent problem-solving across formats such as text and diagrams.

URSA-8B Luo et al. (2025): URSA-8B is an 8B-parameter multimodal model targeting chain-of-thought reasoning in visual mathematical problems, trained on large-scale multimodal CoT datasets. It employs a reward model for stepwise verification, emphasizing both the generation and validation of reasoning chains for reliable solutions.

R1-Onevision-7B Yang et al. (2025b): R1-Onevision-7B is a 7B-parameter multimodal reasoning model that converts images into structured textual representations for symbolic reasoning. It is trained

2322

2323

2324

2325

2326

2327

2328

2329

2330

2331

2332

2333

2334

2335

2336

2337

2338

2339

2340

2341

2342

2343

2344

2345

2346

2347

2348

2349

2350

2351

2352

2353

2354

2355

2356

2357

2358

2359

2360

2361

2362

2363

2364

2365

2366

2367

2368

2369

2370

2371

2372

2373

2374

2375

Comparison of MathBook-7B and Qwen2.5-VL-7B on MathVista

Average waiting time at the DMV	
Month	Waiting time (minutes)
August	17
September	14
October	26
November	17
December	25

Question:

Please answer the question requiring an integer answer and provide the final value, e.g., 1, 2, 3, at the end.

Question: An administrator at the Department of Motor Vehicles (DMV) tracked the average wait time from month to month. According to the table, what was the rate of change between August and September? (Unit: minutes per month)

Response from Qwen2.5-VL-7B:

<Thought process>: To find the rate of change between August and September, we need to calculate the difference in waiting time and divide it by the number of months between August and September.

The waiting time in August is 17 minutes.

The waiting time in September is 14 minutes.

The difference in waiting time is:

17 minutes (August) - 14 minutes (September) = 3 minutes

Since there is only one month between August and September, the rate of change is:

3 minutes / 1 month = 3 minutes per month

Therefore, the rate of change between August and September is 3 minutes per month. 

Response from MathBook-7B:

<think>To find the rate of change between August and September, we calculate the difference in waiting times and divide by the number of months between them. The waiting time in August is 17 minutes, and in September it is 14 minutes. The change is $17 - 14 = 3$ minutes. Since the time decreased, the rate of change is -3 minutes per month.</think> <answer>-3</answer> 

Figure 30: Comparison of MathBook-7B and Qwen2.5-VL-7B on MathVista.

on step-by-step multimodal reasoning annotations and refined with reinforcement learning, enabling precise multi-hop visual-textual inference.

R1-VL-7B Zhang et al. (2025): R1-VL-7B is an open-source 7B vision-language model designed for stepwise reasoning, applying Step-wise Group Relative Policy Optimization (StepGRPO) for dense intermediate rewards. This approach improves logical coherence and multi-step problem-solving, especially in mathematical and logical tasks.

MM-Eureka-7B Meng et al. (2025): MM-Eureka-7B is a vision-language model based on Qwen2.5-VL-7B, fine-tuned with the MMK12 dataset and a rule-based reinforcement learning strategy. It is designed for multidisciplinary visual reasoning in math and science, using rule-based rewards to guide the learning of complex reasoning steps.

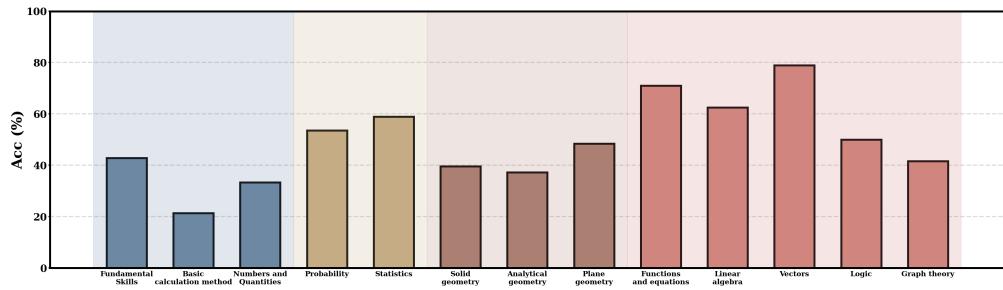
WeThink-VL-7B Yang et al. (2025a): WeThink-VL-7B is a 7B-parameter general-purpose vision-language reasoning model fine-tuned on Qwen2.5-VL-7B via reinforcement learning. It is trained on the WeThink dataset and adopts a hybrid reward mechanism combining rule-based verification. The model enhances performance across both mathematical reasoning and general multimodal tasks by leveraging a scalable multimodal QA synthesis pipeline for diverse data generation and GRPO.

2376
 2377 **OpenVLThinker-7B Deng et al. (2025):** OpenVLThinker-7B is an open-source model tailored for
 2378 complex vision-language reasoning, built by iterating between lightweight supervised fine-tuning
 2379 and curriculum reinforcement learning. SFT initially distills chain-of-thought traces from text-
 2380 based reasoning models, while RL refines these behaviors via a two-stage curriculum. It achieves
 2381 improvements across six benchmarks, outperforming concurrent models with smaller training data.
 2382

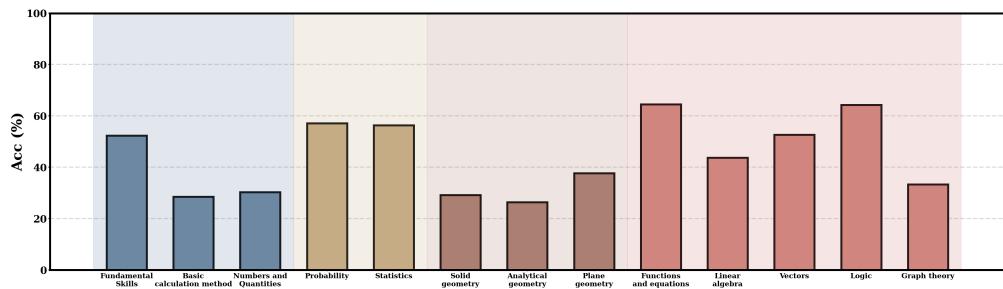
2383 **VLAA-Thinker-7B Chen et al. (2025):** VLAA-Thinker-7B is an open-source large vision-
 2384 language model optimized for multimodal reasoning, trained via Group Relative Policy Optimization
 2385 with a novel mixed reward module. The mixed reward integrates 4 types of rule-based rewards and 1
 2386 open-ended reward, avoiding "pseudo reasoning paths" induced by supervised fine-tuning. It achieves
 2387 good performance on the open LMM reasoning leaderboard.

2388 D THE USE OF LARGE LANGUAGE MODELS (LLMs)

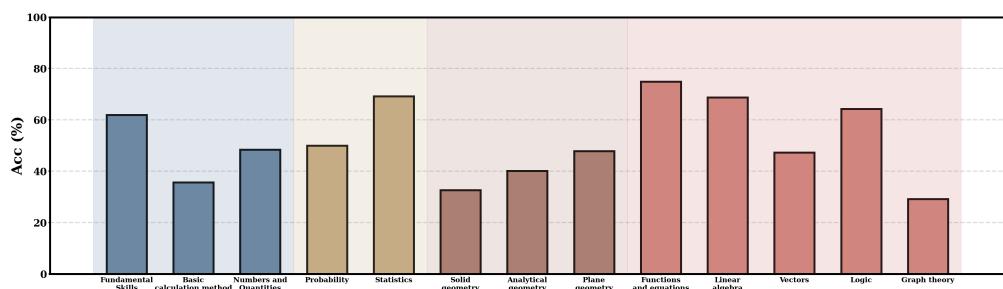
2389 LLMs were used solely for grammar correction and minor language polishing. The conception,
 2390 methodology, experiments, and analysis presented in this paper were entirely designed and imple-
 2391 mented by the authors without relying on LLMs.



2403 Figure 31: Detailed performance of GPT-4o across 13 subdomains.
 2404



2415 Figure 32: Detailed performance of GPT-4V across 13 subdomains.
 2416



2428 Figure 33: Detailed performance of InternVL2.5-78B across 13 subdomains.
 2429

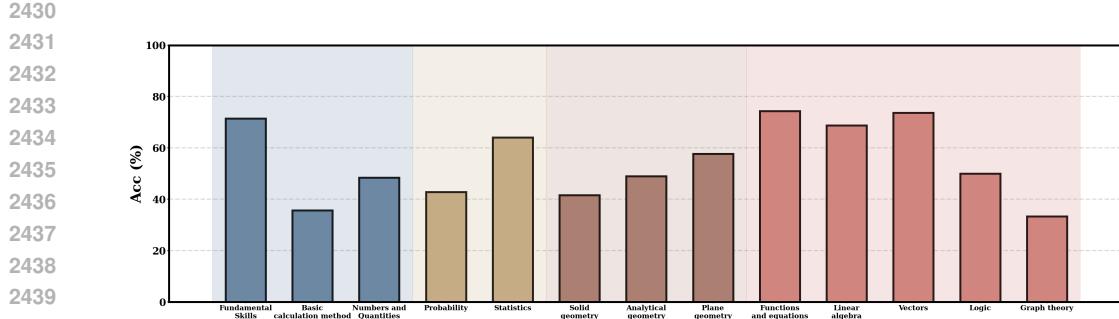


Figure 34: Detailed performance of Qwen2.5-VL-72B across 13 subdomains.

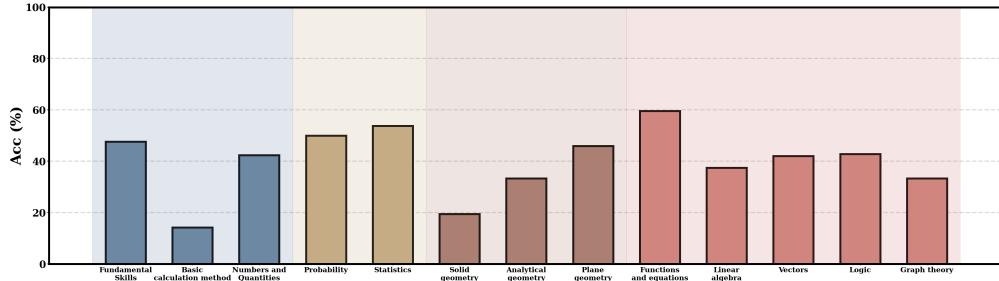


Figure 35: Detailed performance of LLaVA-OneVision-72B across 13 subdomains.

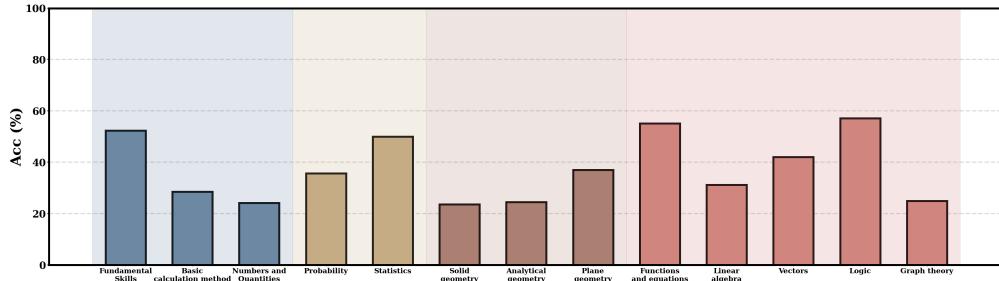


Figure 36: Detailed performance of InternVL2.5-8B across 13 subdomains.

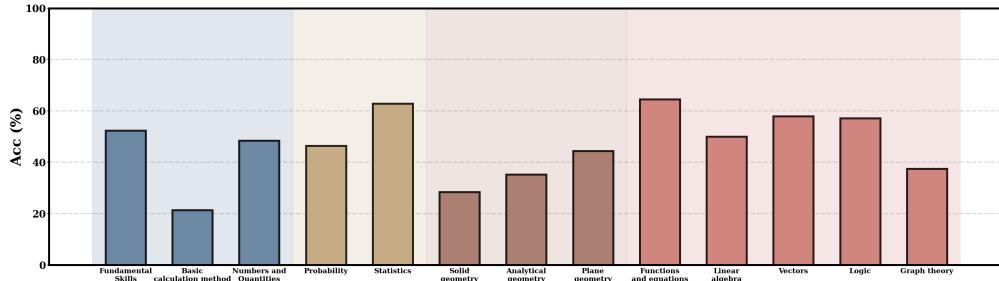


Figure 37: Detailed performance of Qwen2.5-VL-7B across 13 subdomains.

2484

2485

2486

2487

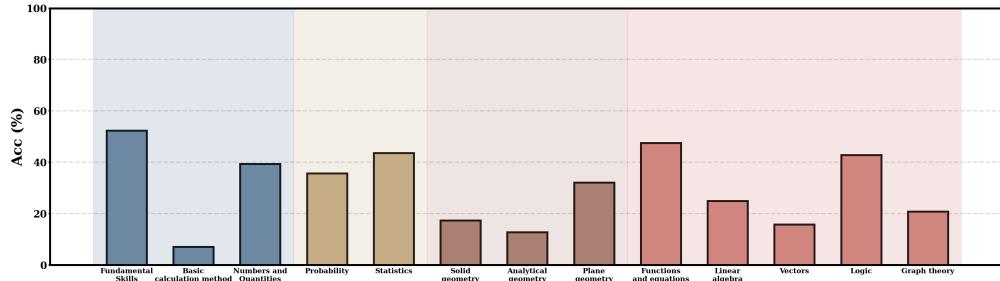


Figure 38: Detailed performance of LLaVA-OneVision-7B across 13 subdomains.

2497

2498

2499

2500

2501

2502

2503

2504

2505

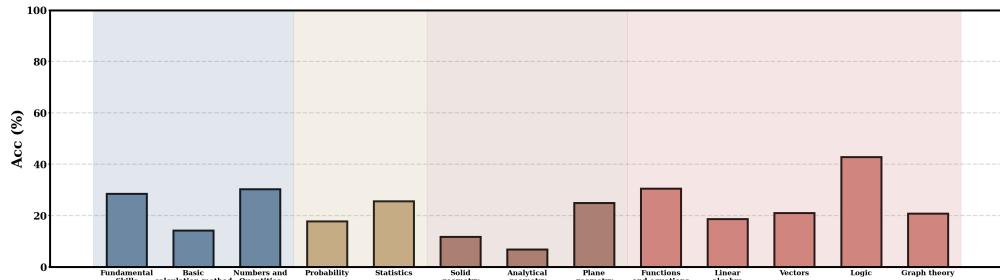


Figure 39: Detailed performance of GLM-4V-9B across 13 subdomains.

2515

2516

2517

2518

2519

2520

2521

2522

2523

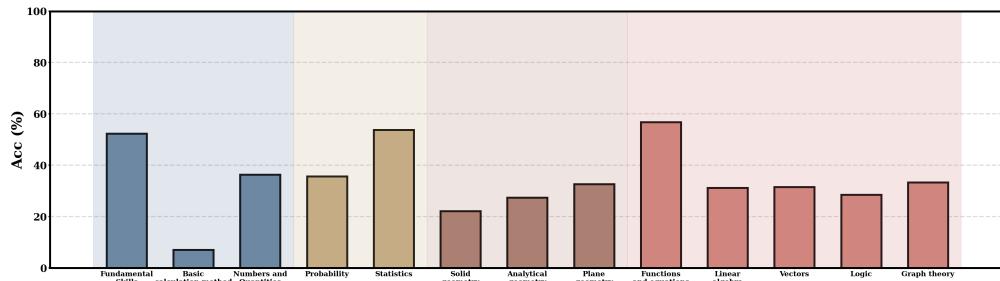


Figure 40: Detailed performance of Qwen2.5-VL-3B across 13 subdomains.

2534

2535

2536

2537