Training Strategies for Language Model Generalization in Mathematical Reasoning

Assignee Research

June 6, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: What training strategies improve language model generalization to novel mathematical reasoning problems v11. 12 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: ARS: Adaptive Reasoning Suppression for Efficient Large Reasoning Language Models. Research question: What training strategies improve language model generalization to novel mathematical reasoning problems v11.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.3/10.

3 Results

16 papers retrieved. 12 claims extracted; 1 independently verified. Quality review score: 4.3/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
The study evaluates Qwen2.5-Math-1.5B-Instruct, Qwen2.5-Math-7B-Instruct, and DeepSeek-R1-Distill-Qwen-7B models.	×	0.04
The ARS algorithm utilizes difficulty thresholds (d1, d2) and confidence thresholds (c1, c2, c3) as inputs.	×	0.03
In FAST mode, the ARS policy is configured with 2 drafts and 10 tokens per draft.	×	0.01
In MOD mode, the ARS policy uses a budget of 64 tokens.	×	0.02
In the default mode, the ARS policy uses a sc_k parameter of 3.	×	0.02
The generation process in the ARS algorithm terminates if the text length reaches 1200 tokens.	×	0.02
ARS consistently achieves superior length reduction while maintaining competitive accuracy across all tested model scale	×	0.09
Reasoning behaviors in Large Reasoning Language Models are triggered by specific keywords including 'Wait', 'But', and '	×	0.11
The standard generation process produces output tokens where each token is sampled from the model distribution condition	×	0.02
The objective of the ARS framework is to minimize expected output length while keeping accuracy degradation below a thre	×	0.03
The ARS framework consists of three core components: Multi-checkpoint certainty estimation, Progressive threshold adapta	✓	0.16
ARS establishes multiple checkpoints at regular intervals during generation to estimate model certainty through tentativ	×	0.04

References

- http://arxiv.org/abs/2510.00071v2
- http://arxiv.org/abs/2509.25160v1
- http://arxiv.org/abs/2604.25926v1