Tree-Based vs. Linear Retrieval Overhead in Multi-Hop Reasoning Benchmarks

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How does the computational overhead of tree-based dynamic iterative retrieval compare to linear chain-of-thought retrieval on multi-hop reasoning benchmarks like HotpotQA when scaling to larger. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: When to Retrieve During Reasoning: Adaptive Retrieval for Large Reasoning Models. Research question: How does the computational overhead of tree-based dynamic iterative retrieval compare to linear chain-of-thought retrieval on multi-hop reasoning benchmarks like HotpotQA when scaling to larger context windows?.

2 Methodology

Systematic literature search across multiple databases yielded 13 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.7/10.

3 Results

13 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.7/10.

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.

References

- http://arxiv.org/abs/2404.14464v1
- http://arxiv.org/abs/2604.26649v1
- http://arxiv.org/abs/2604.18234v1