Progressive Problem Complexity Gaps in JaCoText and Pass@k Accuracy Across Model Scales

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: What is the impact of increasing the complexity gap between base and progressive problems in JaCoText on pass@k accuracy across different model sizes, as evaluated on HumanEval Pro and MBPP Pro. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: HumanEval Pro and MBPP Pro: Evaluating Large Language Models on Self-invoking Code Generation. Research question: What is the impact of increasing the complexity gap between base and progressive problems in JaCoText on pass@k accuracy across different model sizes, as evaluated on HumanEval Pro and MBPP Pro benchmarks?.

2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.5/10.

3 Results

15 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.5/10.

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/2606.05868v1
- http://arxiv.org/abs/2412.21199v2
- http://arxiv.org/abs/2605.27000v2