

Evolutionary Search Strategies Enhance LLM Reasoning in Competition-Level Code Generation

Assignee Research

June 4, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: What is the impact of incorporating evolutionary search strategies on the reasoning accuracy of LLMs when evaluated on competition-level software engineering datasets like CodeContests. 5 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: What Does a Software Engineer Look Like? Exploring Societal Stereotypes in LLMs. Research question: What is the impact of incorporating evolutionary search strategies on the reasoning accuracy of LLMs when evaluated on competition-level software engineering datasets like CodeContests?.

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 3.2/10.

3 Results

13 papers retrieved. 5 claims extracted; 0 independently verified. Quality review score: 3.2/10.

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
Google’s initiatives like the Inclusive Images Competition [58] aim to recognize diverse cultural backgrounds.	×	0.03
The Draw a Scientist experiment [21] found that children depicted scientists as older white men, reflecting societal stereotypes.	×	0.06
Cutrupi et al. conducted the Draw a Software Engineer Test and found that gender stereotypes increase with age, and university students.	×	0.07
Baltes et al.’s research on ageism in software development shows that LLMs reinforce stereotypes of software engineers.	×	0.11
A study on 278 motion-capture-related works demonstrated that AI systems are often trained on biased data, predominantly from Western sources.	×	0.04

References

- <http://arxiv.org/abs/2407.15026v2>
- <http://arxiv.org/abs/2501.03569v1>
- <http://arxiv.org/abs/2605.28353v1>