

Language Models for Competition-Level Software Engineering Problem Solving

Assignee Research

June 6, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: What techniques enable language models to solve competition-level software engineering problems v10. 8 claims were extracted from source literature; 6 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 7.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: MLFC: From 10 to 50 Planners in the Multi-Agent Programming Contest. Research question: What techniques enable language models to solve competition-level software engineering problems v10.

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 7.3/10.

3 Results

16 papers retrieved. 8 claims extracted; 6 independently verified. Quality review score: 7.3/10.

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 MLFC team achieved 2nd place in the 15th edition of the Multi-Agent Programming Contest.	✓	0.29
The scenario used in the 15th edition of the contest is an extension of the 14th edition scenario titled 'Agents Assembl	✓	0.26
In the contest scenario, two teams of agents move around a 2D grid to compete in assembling complex block structures.	✓	0.25
The MLFC team developed a cartography strategy to determine the size of the map.	×	0.12
The MLFC team performed formal verification of the map merging protocol to provide assurances that it works when increas	✓	0.30
The MLFC team implemented a plan cache strategy to efficiently scale the number of planners.	×	0.15
The MLFC team implemented a task achievement strategy involving forming groups of agents to achieve tasks.	✓	0.20
The MLFC team implemented a 'bullies' strategy where agents focus on stopping agents from the opposing team.	✓	0.20

References

- <http://arxiv.org/abs/2406.04710v2>
- <http://arxiv.org/abs/2110.08172v2>
- <http://arxiv.org/abs/2006.02739v1>