Quantization Impact on Reasoning Capabilities in Large Language Models

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: How does model quantization affect reasoning capability in large language models v19. 7 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Teaching LLMs to Plan: Logical Chain-of-Thought Instruction Tuning for Symbolic Planning. Research question: How does model quantization affect reasoning capability in large language models v19.

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 4.7/10.

3 Results

16 papers retrieved. 7 claims extracted; 1 independently verified. Quality review score: 4.7/10.

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.

5 Extracted Claims

Claim	Verified	Confidence
PDDL-INSTRUCT achieves planning validity rates of up to 94% in standard planning domains.	✓	0.16
PDDL-INSTRUCT significantly outperforms both baseline models and traditionally instruction-tuned models.	×	0.12
For Llama-3, PDDL-INSTRUCT with detailed feedback and η = 15 achieves validity rates of 94%, 64%, and 79% across the Blo	×	0.04
For GPT-4, PDDL-INSTRUCT with detailed feedback and η = 15 achieves validity rates of 91%, 59%, and 78% across the Block	×	0.03
PDDL-INSTRUCT represents an average absolute improvement of 35%(SD = 20%) over basic instruction tuning, and of 66%(SD =	×	0.09
PDDL-INSTRUCT represents an average absolute improvement of 48%(SD = 5%) over basic instruction tuning, and of 61%(SD =	×	0.08
The explicit reasoning about preconditions, effects, and state transitions enables the models to generate accurate plans	×	0.08

References

- http://arxiv.org/abs/2509.16989v3
- http://arxiv.org/abs/2509.13351v1
- http://arxiv.org/abs/2411.02355v4