



AutoCodeRover: Autonomous Program Improvement

Yuntong Zhang, Haifeng Ruan, Zhiyu Fan, Abhik
Roychoudhury

Presenter: Nhat Minh Le – 40323412

Motivation

- Large Language Models can assist human in solving problems especially in automated program repair and coding.
- AutoCodeRover is an open-source project.
- Clear setup instructions

Agenda

I. Tool Information

II. Evaluation

III. Limitation

IV. Demo

I. Tool Information

- AutoCodeRover: a novel LLM-driven multi-agent framework combining with code search capabilities for automating bug detection, structuring bug reporting, and program repair for Python Projects (SWE-bench).
- In February, 2025, It is acquired by Sonar

I. Tool Information

Table 1: List of Context Retrieval APIs.

API name	Description	Output
<code>search_class (cls)</code>	Search for class <code>cls</code> in the codebase.	Signature of the searched class.
<code>search_class_in_file (cls, f)</code>	Search for class <code>cls</code> in file <code>f</code> .	Signature of the searched class.
<code>search_method (m)</code>	Search for method <code>m</code> in the codebase.	Implementation of the searched method.
<code>search_method_in_class (m, cls)</code>	Search for method <code>m</code> in class <code>cls</code> .	Implementation of the searched method.
<code>search_method_in_file (m, f)</code>	Search for method <code>m</code> in file <code>f</code> .	Implementation of the searched method.
<code>search_code (c)</code>	Search for code snippet <code>c</code> in the codebase.	+/- 3 lines of the searched snippet <code>c</code> .
<code>search_code_in_file (c, f)</code>	Search for code snippet <code>c</code> in file <code>f</code> .	+/- 3 lines of the searched snippet <code>c</code> .

I. Tool Information

ISSTA '24, September 16–20, 2024, Vienna, Austria

Yuntong Zhang, Haifeng Ruan, Zhiyu Fan, and Abhik Roychoudhury

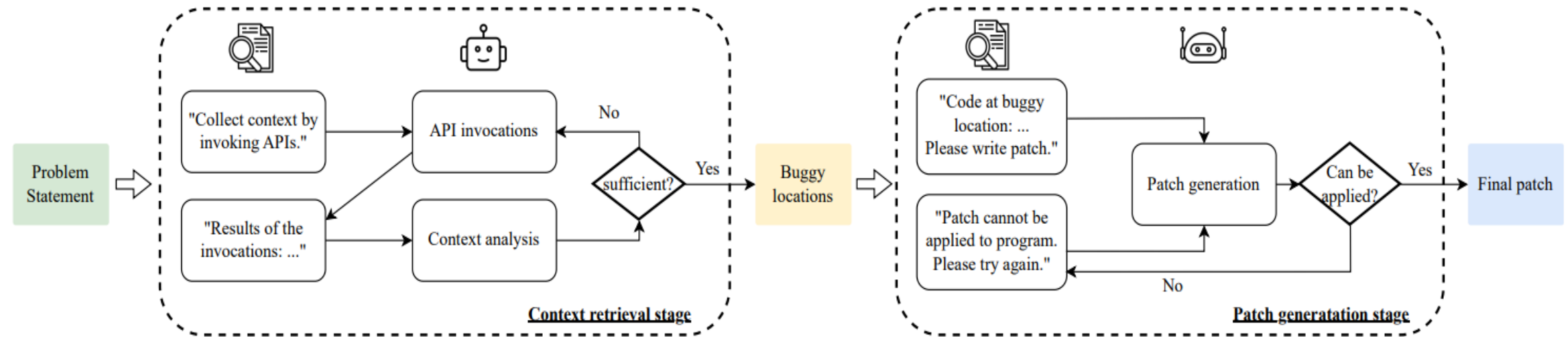


Figure 2: Overall workflow of AUTOCODEROVER.

II. Evaluation

AutoCodeRover achieves lower cost (on average \$0.43 USD), compared to SWE-Agent

AutoCodeRover can resolve more tasks with less cost (31) compared to SWE-Agent (23)

Limitation:

1. Rely on LLM to decide bug and generate code
2. Overfittting patch

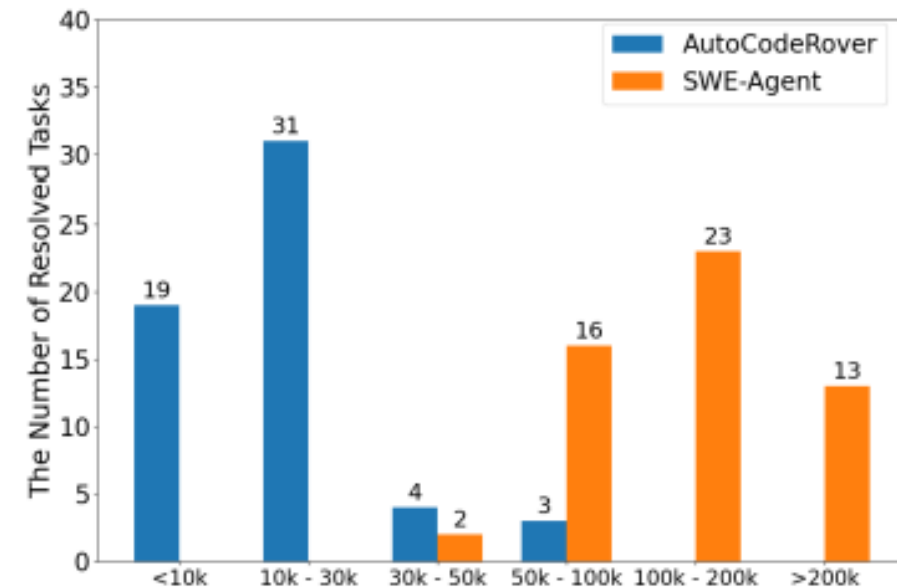


Figure 8: The number of resolved tasks and token cost distributions of AUTOCODEROVER and SWE-AGENT.

III. Demo

- **Environment setup**
 - Docker
 - OPENAI_API_KEY, ANTHROPIC_API_KEY, GROQ_API_KEY
 - Link setting up AutoCodeRover: <https://youtu.be/ZwRGiMduXh4>
 - Link setting up SWE-bench: <https://youtu.be/BVswfxKhkO8>
- **Demo for one task**
 - https://youtu.be/Co_TxGwZNrk
- **Demo for multiple tasks and SWE-bench-Lite (300 tasks)**
 - https://youtu.be/9MoMlz_Vcp0

Thank you!

Questions?

References

- AutoCodeRover: <https://arxiv.org/abs/2404.05427>
- AutoCodeRover's Github:
<https://github.com/AutoCodeRoverSG/auto-code-rover>
- SWE-bench: <https://www.swebench.com/>