Master Thesis Topic
Combining LLMs for Automated Program Repair

Motivation and Background:
Automated program repair (APR) [1] describes the process of automatically fixes software bugs, and hence, promises great support for software developers in practice. Recent approaches in this field, like AlphaRepair [2] and ChatRepair [3], as well as recently proposed SE agents [4] already made significant progress towards this goal. This thesis project will explore the combination of various LLMs with different tasks within the APR context. Therefore, we will implement our own SE agent, which autonomously repairs software issues.

Student Task and Responsibilities:
- Make yourself familiar with the state-of-the-art LLM-based APR techniques and the current LLM-based SE agents.
- Systematically explore the usage of multiple LLMs that act as a team of experts to repair software bugs and build a tool that performs APR in this context.
- Apply your tool on the Defects4J benchmark (https://github.com/rjust/defects4j).
- Explore more sophisticated SE tasks with your approach (i.e., beyond just repairing a known bug), e.g., by investigating the issues in SWE-bench (https://www.swebench.com/).
- Design/select evaluation metrics and conduct a thorough evaluation of your approach.
- Analyze the results and document your findings.

Deliverables:
- LLM-based APR tool that combines multiple LLMs
- Evaluation artifacts (dataset, tools, etc.)
- Documented findings of the conducted experiments

Pre-Requisites: (Programming Languages, OS, Skills, Papers, etc)
Knowledge in Java, Python, and LLM prompt engineering is helpful for this project.

[1] C. Le Goues, M. Pradel, A. Roychoudhury and S. Chandra, "Automatic Program Repair," in *IEEE Software*, vol. 38, no. 4, pp. 22-27, July-Aug. 2021. https://doi.org/10.1109/MS.2021.3072577

[2] C. S. Xia and L. Zhang. 2022. "Less training, more repairing please: revisiting automated program repair via zero-shot learning". In ESEC/FSE 2022. ACM, NY, USA, 959–971. https://doi.org/10.1145/3540250.3549101

[3] C. S. Xia and L. Zhang, "Keep the Conversation Going: Fixing 162 out of 337 bugs for $0.42 each using ChatGPT," Apr. 01, 2023. http://arxiv.org/abs/2304.00385

[4] I. Bouzenia, P. Devanbu, and M. Pradel, "RepairAgent: An Autonomous, LLM-Based Agent for Program Repair," Mar. 25, 2024. http://arxiv.org/abs/2403.17134

Contacts
Prof. Dr. Yannic Noller (`sq-office@rub.de`)
Software Quality group, Faculty of Computer Science, Ruhr University of Bochum