Bachelor Thesis Topic

A Systematic Analysis of Faults in the Defects4J Benchmark

Motivation and Background
Automated software fault localization helps developers to save precious time. The invention of new techniques for fault localization requires ample and well-maintained benchmarks on which the techniques can be tested and compared with each other. One of the biggest and most recent benchmarks is the Defects4J benchmark, containing 357 buggy versions of five open source Java projects.

Goals
Due to the lack of proper detailed information on each of the bugs, it is relatively hard to evaluate the results of conducted fault localization experiments in a precise manner. Therefore, the goal of this thesis is to evaluate the given data and to generate parseable information for future experiments.

Description of the Task
The specific tasks are:
- Obtain a general understanding of automated fault localization techniques.
- Identify important properties of the given bugs that are not provided by the benchmark and are valuable for automated fault localization.
- Find a proper way to store the identified properties with special attention on the ability of automated processing.
- Generate the respecting information for each bug in the Defects4J benchmark.

Research Type
Theoretical Aspects: ★★★★★
Industrial Relevance: ★★★★★
Implementation ★★★★★

Prerequisite
The student should be enrolled in the bachelor/master of software engineering/informatics program, and has completed the required course modules to start a bachelor/master thesis.

Skills required
Programming skills in Java. Understanding of, or willingness to learn, the software engineering methods and the statistical techniques needed for the project.

References

Contacts
Simon Heiden, Humboldt-Universität zu Berlin, Institut für Informatik, Lehrstuhl Software Engineering, Unter den Linden 6, 10099 Berlin, Germany

Application
Please contact me during my office hours or send me an email with the title: “[ThesisProject]-SAFDB” to se-career@informatik.hu-berlin.de