Unlocking Software Documentation: Sentiment Classification and On-hold Self-Admitted Technical Debt Identification

Name: Rungroj Maipradit

Laboratory's name: Software Engineering

Supervisor's name: Kenichi Matsumoto

Abstract

In software development, software documentation plays a crucial role in knowledge sharing between software developers. Despite the benefit of software documentation, document creation and maintenance are frequently overlooked. To help developers in maintaining and understanding software documents, two problems are identified (1) help developers identify and removed unnecessary self-admitted technical debt (i.e., situations where a software developer knows that their current implementation is not optimal and indicates this using a source code comment), and (2) support developers in understanding the underlining messages using sentiment analysis.

The first part of this thesis finds what kind of tasks are amenable to automated management. And found one particular class of debt amenable to automated management: On-hold SATD, i.e., debt which contains a condition to indicate that a developer is waiting for a certain event or an updated functionality having been implemented elsewhere. The second part of this thesis investigates the potential of automatically detecting On-hold SATD and identifying its condition. The results show that the proposed design can reliably identify On-hold SATD and also mines the issue tracker to check if the On-hold SATD instances are "superfluous" and can be removed. The third part of this thesis investigates the potential of automatically identifying sentiment analysis on software engineering data sets. The results show that using automated machine learning with n-gram inverse document frequency shows promising results in tackling this problem.