Grading Code Quality of Programming Assignments Based on Bad Smells

Woei-Kae Chen and Pin-Ying Tu
Department of Computer Science and Information Engineering
National Taipei University of Technology
{wkchen, t5599004}@ntut.edu.tw

Abstract

Programming assignments (PAs) are very important to many computer science courses. Traditionally, the grading of a programming assignment is based mainly on the correctness of the code. However, from the view point of software engineering education, such a grading does not encourage students to develop code that is easy to read and maintain. Thus, the authors created a grading policy that considers not only the correctness but also the quality of the code, expecting students to follow the most important discipline — the source code should be written in a way that is readable and maintainable. Instead of using pure subjective code-quality ratings, bad smells are used to assess the code quality of PAs. When a PA is graded by the teaching assistant, a list of bad smells is identified and given to the student so that the student can use refactoring methods to improve the code.

1. Summary

The authors conducted an experiment on the programming assignments of a junior-year Windows Programming course (Fall 2010) to evaluate the new grading policy. The course required students to implement a C# rich-client application in six incremental PAs. The average code size grows from 165 lines of code (LOC) to around 2000. The authors compiled a list of 12 bad smells [1] (long method, duplicated code, etc.) most frequently found in students’ code. The list is explained to students at the beginning of the course and used as the code-quality grading standard. The code-quality rating, which worth about 30% of total grade, is determined by counting the number of bad smells. When the grading of a PA was completed, the student received a code-quality rating, along with a list of bad smells. Since the PAs were incremental, students were motivated to keep removing bad smells to get better ratings in the following PAs. We record the smell density (number of smells per thousand LOC), SD, of students’ code to indicate code quality (the lower the better). The results showed that as the PAs progressed, the average SD decreased from around 52.6 to 18.8, indicating a significant improvement (about 2.8 times of improvement). After the completion of the sixth PA, the authors administered a survey to the students. The students reported that they felt a strong sense of code-quality improvement by removing bad smells (an average rating of 4.3 on a 5-point Likert scale). In the future, the authors expect to give a deeper analysis of the results and study the method for more courses.

References


This research was partially supported by the National Science Council and Ministry of Economic Affair, Taiwan, under contract number NSC 99-2220-E-027-005 and MOEA 99-EC-17-A-02-S1-135, which are gratefully acknowledged.