

The importance of comments and code rules in development procedure

181250090 刘育麟

group: 15 year: 18

2021 年 7 月 4 日

1 Introduction

A project in computer science usually includes a lot of lines codes. A programmer in one project doesn't have enough ability to finish all of codes by himself so that they only code for a part of a project that cause the inevitability to read other people codes to find the bug. Furthermore, the people who work for computer science usually change their work and the new work is fixing or maintaining the existing codes that were coded by the person who leave this project development team. This is the reason for using comments and code rules because these two things can help people read other people codes easily. Comments is the descriptions or the explanation about codes, and code rules are like naming scheme and code indentation. These two things can't improve the coding ability for programmers but they help programmers be familiar with the code when they read some codes which they don't read them before so that they are important in a big project.

2 Motivation

Because of the reason in introduction, we want to get some experience about the comments or code rules from the workers in some projects to help the project which need a lot of people cooperating improve and perfect the rule about comment and code rule. The people in a big project which includes more than one million lines codes can provide some effective suggestions than the students in the school or some small case participants with few codes. Furthermore, we want to know some rules in one big project about the comments and code rules such as the position of comments and the naming scheme of functions and take some experiences about the influence for a programmer read a code file first and the help for a new programmer to know the structure and logic about a project who just get the original code.

3 Research description

3.1 Research questions

1. What are methods for a programmer understanding a new project code structure and logic?
2. Where the code position in the code file can help programmers know the code file logic and function?
3. What naming scheme is good for a programmer to know the method function they call?
4. Which cases need these limitations to help programmers understand codes?
5. How importance about the comments and code rules is in a develop procedure?
6. What limitations you prefer about comments and code rules?

In these six questions, we can know the usage about limitations and explanations which can't improve the performance like comments and naming scheme. These questions can be separated by two parts. The first parts focus on the experience about programmers developing procedure. The second parts focus on the preference about the limitations for the programmers and ask some suggestion about the importance and necessity about these things.

3.2 Method

We conduct a survey to get the answer about these questions. We want to collect a lot of information from many company so that we will post this survey in a public website and allow all of programmers to finish it. To avoid revealing some business security, our question will design focus on the limitations which based on some books mentioning these rules. We design some closed questions and some

opening question to get the experiences by closed questions and the suggestions by opening questions. Furthermore, we will interview some programmers who have a lot of experiences in multiple companies to ask their opinion about the differences about different companies style.

3.3 Methodological characteristics

In table 1.

表 1: Methodological characteristics

Characteristics	study
type	Cross-sectional
format	Self-administrated questionnaires
questions	closed questions with few open questions
type of answer	yes/no answers + response categories + textual answer + numeric values
instrument evaluation	not mention
sampling	Cluster-base sampling
length	finish in 10 minutes
study	quantitative analysis + qualitative analysis

3.4 Survey design

All of the chosen categories are multiple choose.

1. What is your company you have worked?
2. What is your position in companies?
 - (a) back-end engineer
 - (b) front-end engineer
 - (c) project manager
 - (d) structure engineer

(e) algorithm engineer

3. Which language you usually use?

(a) java

(b) c/c++

(c) python

(d) go

(e) R

(f) HTML

(g) javascript

(h) c#

4. Where you add your comments?

(a) file top

(b) class top

(c) method top

(d) line back

5. Which kind of naming scheme you used?

(a) underscore

(b) camelCase

(c) upper flat case

(d) PascalCase

(e) hyphen

6. Numeric values, 1 is most disagreeable and 5 is most agreeable.

(a) How importance you think to comments and code rule like naming scheme?

- (b) The comments help you read codes.
- (c) Good coding style help new co-workers understand project structure easily.
- (d) Too many comments is good than simple and clean codes.
- (e) Many naming scheme may cause no useable names in developing procedure.
- (f) Do you think the naming scheme help you read codes?

7. What is you reasons to choose the agreement level in above?
8. What is your experience about the code rules to help you easily read code files?
9. Which code style you prefer? Maybe you can write some codes, or tell your opinion.
10. If you first read a codes? which things you will first focus on to help you understand?
11. Where the comments you prefer adding?
12. In the companies you have worked, which of the code rules you like most?
13. How old are you?
 - (a) 0-18
 - (b) 18-25
 - (c) 25-35
 - (d) 35-65
 - (e) 65+
14. What is your gender?
 - (a) male
 - (b) female

4 Date collection

We will use a public website to collect the data from anyone. We will use some conversation applications such as QQ, WeChat, Mai-Mai, Weibo and so on to make more people see this survey, and we will add some key words about computer science to attract more people who work in this feild. We plan to open this survey for 2-3 weeks, and we will collect all the questionnaires and prepared to analyze the data. We estimate to collect more than 100 questionnaires which includes more than 10 companies to guarantee the quantitative analysis.

Furthermore, we want to choose some volunteers to be interviewees. We will ask the questions on survey to get their answers person-to-person. Besides the questions on survey, we will ask some questions generated by their answers. These survey answers can guarantee the qualitative analysis because we will record all of the answers and specific ask some deeper questions which can' t get from the survey that written by the programmers in network.

5 Validity

5.1 Internal validity

The survey questions focus on the experiences by programmers, and we will focus on the programmers who work in the companies. Furthermore, because we collect the interviewees work companies, we can know the different companies code style and according to the companies scale to judge the weight about the answers. These methods can guarantee the effectiveness for the answers about the public survey. Then, because we have other interviews which ask the same questions in person. We can analyze the accuracy about the answers in public survey and the interviews help us collect qualitive data. Then, we will delete the outlier in the

public survey to avoid the bias in results.

5.2 External validity

We will collect data for 2-3 weeks. This is a long term to collect enough data to finish our research, and we estimate to collect more than 100 questionnaires which includes more than 10 companies so that we can guarantee the generalization for our survey. Then, to attract the programmers to finish this survey, we will add some key words in our survey names and the post in public network, and we hope to know every level programmers opinion so that we have a chosen questions to ask people position in survey.