Project Description: V&V of an Open Source Product

Kai Petersen

January 27, 2015

PRODUCT

The static and dynamic analysis is every year conducted on a different product. This years product is http://sourceforge.net/projects/sudokuki/files/?source=directory.

PART I: STATIC ANALYSIS

DESCRIPTION OF THE TASK

Scenario: You have taken over the project for the system specified above and your product manager asked you to first conduct a static analysis for the product to assure good quality prior to further extending and modernizing it.

You are to use static automated code analysis (automated inspection of the source code) to find defects in the code. A variety of tools exist and they have different capabilities. When using them, they provide you with a list of warnings that need to be reviewed, and you have to decide whether they are a false positive (wrong warning that does not actually constitute a fault) or true positive (real fault).

Take the following steps:

1. Before starting the inspection with the tool, as a group decide on the inspection goals that you consider most relevant for the given scenario (do not make it too trivial, e.g. just checking coding standards, etc., what does really matter?).

2. Identify static analysis tools that allow you to achieve your defined goals (this could be one, or more than one depending on your goals).

3. Now each of you **individually reads** the aggregated lists of warnings given by the tool(s) and classify the warnings into: True positive, false positive, uncertain (if you are not

able to decide). Each individual should clearly document that, and it should be traceable (Excel files for each individual and their rating should be delivered to me together with the report; use the Excel file we prepared for you[1]). If there are too many warnings to inspect, how do you filter them to come to a list to review in more detail, could you do some pre-selection of warnings? **Please also log time time** each of you **individually** spends on inspecting the warnings, and the time you spend on **discussing them** together. This will later help you to reflect on cost efficiency (see Section Reporting).

4. **After you have done that individually** come together and discuss the warnings with each other, discuss agreements and disagreements and come to a conclusion together. Decide which warnings you would prioritized highest in terms of what needs to be fixed, and relate that to your goals. **Please do only discuss within your group, and not between groups**.

REPORTING

In your report, the following should be highlighted and very clear:

1. **Motivate the goals that you defined for your automated analysis.** Why did you prioritize those goals over others that could have been relevant? That is, first present a set of potentially relevant goals (use literature for help), and then motivate your selection.

2. **Motivate the selection of the static analysis tool(s) chosen.** why did you select a specific tool? What is the advantage of those tools over other ones, and based on which criteria did you determine the advantage?

3. **Evaluate the outcome of the static analysis approach.** Were the tools helpful in achieving your goals (be very explicit why this was the case or it was not the case)? When you individually decided on the warnings, how well were you in agreement when comparing your individual answers? Please explain the reasons why there were differences in the interpretation (e.g. were specific types of faults easier to judge than others, why? any other reasons?). Did the tools perform well enough in accurately identifying faults from a practical point of view? If so, explain why or why not. Would you consider automated static code analysis to be cost-effective (explain why or why not)? Please compare your findings and reflections with what researchers report in literature.

PART II: DYNAMIC ANALYSIS

DESCRIPTION OF THE TASK

In software testing companies aim at a high level of automation, which includes testing tools. Different tools implement a variety of approaches to testing, and are designed for different test levels (unit testing, system testing, acceptance testing, etc.). Academia is intensively

[1] The Excel file has several Tabs. The first five Tabs are for your individual logging of each individual team member. The last Tab shows all your decisions side by side. Now in a group discuss and reach a consensus.

working on creating easy to use and effective testing tools and frameworks, such as FitNesse for acceptance testing, Behaviour driven testing tools for acceptance testing, EvoSuite for unit testing, etc.

In this assignment you should conduct a dynamic analysis using testing tools. First you select two tools. The tools should be non-commercial, look for academic and/or open source tools. The focus of the test tools should be on test design, i.e. they should automatically generate the test cases. The tools should be in-line with the goals you have chosen for your testing.

It is important that you motivate the choice of your tool, and which techniques/approaches you are choosing and why. The why is always very important, never state: *"we are doing this"*; instead always say: *"we are doing this because..."*. It also helps not just saying why you do a certain thing, but also why you do not use something else (e.g. discard a tool). In particular, it should come across that you understand which test techniques and approaches are implemented in the tools, and what the strengths and weaknesses of the tools are. That is, you are expected to look at different tools and also try different tools.

After you have chosen two tools you should run both of them on the system. Thereafter, you should compare their findings with each other and discuss: What is different in the findings? Why is it different?

FORMALITIES

LENGTH

The maximum number of pages is 15 pages in IEEE format (6 pages for Part I and 9 for Part II; references and Appendices, such as tables with defect lists, do not count). Do not submit a higher number of pages, as you should convey and carefully choose the most important information relevant to your report. A report longer than the page limit receives the grade F and needs to be resubmitted in revised form.

GROUP SIZE

The task is to be done in a group size of 4 to 5 students. Please form the groups yourself.

FORMAT

You have to use the IEEE paper template for the submission.
`http://www.ieee.org/conferences_events/conferences/publishing/templates.html`

SUBMISSION

Submit your assignment through it's learning before the specified deadline, not via e-mail.

EVALUATION

Table 1: Rubric for Part I

Criteria	4 – Superior Command	3 – Good Control	2 – Fair/some Control	1 – Minimal/No Control
Identifying inspection goals	- Several alternative goals have been specified. - Goals are non-trivial (e.g. aim is not just to check coding conventions) - It is visible that literature has been used as input to identifying the goals - The importance of goals chosen to focus on is documented and rationales are provided	- Several alternative goals have been specified. - Goals are non-trivial - The goals are not based on literature - The importance of goals chosen to focus on is documented and rationales are provided	- Goals are clearly stated and non-trivial, but not motivated, not based on literature	Trivial goals are stated and they are not motivated
Motivate alternatives for selecting static analysis tools	- Several alternatives and their properties are described - Alternatives relate to the goals defined for the inspection meeting, which is made explicit in the report - Then the selection is motivated based on experimentation and analysis based on actual facts	- Several alternatives and their properties are described. - Alternatives relate to the goals defined for the inspection meeting, which is made explicit in the report - The choices are well motivated based on the properties	- Few alternatives are properties and their properties party described. - Alternatives relate to the goals defined for the inspection - A choice from the alternatives is made, but not well motivated.	- No alternatives discussed. - No choices discussed.
Reflection on outcome	The performance of the tools has been critically evaluated with their goals (i.e. it is explicit why or why not they are suitable in achieving the goals). - Warnings have been reviewed following the guidelines (Excel files are filled in as instructed) - Classifications of individuals are compared and it is clear from the report that the reasons for them are explored, as well as what the findings mean for industrial practice - Literature is used to support the reflections made.	- The performance of the tools has been critically evaluated with their goals (i.e. it is explicit why or why not they are suitable in achieving the goals). - Warnings have been reviewed following the guidelines (Excel files are filled in as instructed) - Classifications of individuals are compared and it is clear from the report that the reasons for them are explored, as well as what the findings mean for industrial practice	- The performance of the tools has been described, but not evaluated (e.g. how well it corresponds to the goals, suitability, etc.) - Warnings have been reviewed following the guidelines (Excel files are filled in as instructed) - Classifications of individual reviewers are compared, but reasons for differences and meaning for industry are not explored.	- The description of the performance of the tools is not traceable, warnings are not reviewed and reported following the instructions, no comparative analysis done

Table 2: Rubric for Part II

Criteria	4 – Superior Command	3 – Good Control	2 – Fair/some Control	1 – Minimal/No Control
Identifying inspection goals	- Several alternative goals have been specified. - Goals are non-trivial - It is visible that literature has been used as input to identifying the goals - The importance of goals chosen to focus on is documented and rationales are provided	- Several alternative goals have been specified. - Goals are non-trivial - The goals are not based on literature - The importance of goals chosen to focus on is documented and rationales are provided	- Goals are clearly stated and non-trivial, but not motivated, not based on literature	Trivial goals are stated and they are not motivated
Motivate alternatives for selecting static analysis tools	- Several alternatives and their properties are described - It is clear from the report which testing technique the tool implements and how - Alternatives relate to the goals defined, which is made explicit in the report - Then the selection is motivated based on experimentation and analysis based on actual facts (tried several tools and discusses them)	- Several alternatives and their properties are described. - It is clear from the report which testing technique the tool implements and how - Alternatives relate to the goals defined, which is made explicit in the report - The choices are well motivated based on the properties (tried several tools and discusses them)	- Few alternatives are properties and their properties party described. - the testing technique used in the tool is mentioned - Alternatives relate to the goals defined for the test activity - A choice from the alternatives is made, but not well motivated.	- No alternatives discussed. - No choices discussed.
Reflection on outcome	The performance of the tools has been critically evaluated with their goals (i.e. it is explicit why or why not they are suitable in achieving the goals). - The output of the two testing tools chosen is compared. - The reasons for the differences have been explored between the tools - Literature is used to support the reflections made.	- The performance of the tools has been critically evaluated with their goals (i.e. it is explicit why or why not they are suitable in achieving the goals). - Warnings have been reviewed following the guidelines (Excel files are filled in as instructed) - Classifications of individuals are compared and it is clear from the report that the reasons for them are explored, as well as what the findings mean for industrial practice	- The performance of the tools has been described, but not evaluated (e.g. how well it corresponds to the goals, suitability, etc.) - Test outcomes of the two tools are compared, but reasons for differences and meaning for industry are not explored.	- The description of the performance of the tools is not traceable, defects are not reported/describedü, no comparative analysis done