Review report of a final thesis

Czech Technical University in Prague **Faculty of Information Technology**

Student:	Filippo Ghibellini
Reviewer:	Paley Guangping Li, Ph.D.
Thesis title:	Dynamic test generation for R packages
Branch of the study:	Computer Science

Date: 11. 6. 2017

Evaluation criterion:	*The evaluation scale: 1 to 5.*
1. Difficulty and other comments on the assignment	*1 = extremely challenging assignment,* ***2 = rather difficult assignment,*** *3 = assignment of average difficulty,* *4 = easier, but still sufficient assignment,* *5 = insufficient assignment*

Criteria description:
Characterize this final thesis in detail and its relationships to previous or current projects. Comment what is difficult about this thesis (in case of a more difficult thesis, you may overlook some shortcomings that you would not in case of an easy assignment, and on the contrary, with an easy assignment those shortcomings should be evaluated more strictly.)

Comments:
The framework presented in this thesis is based upon a previous project that contains very limited documentation. The framework was not initially designed for the extension the student has been tasked to incorporate, which made the extension non-trivial in how it would fit into the overall framework.

Evaluation criterion:	*The evaluation scale: 1 to 4.*
2. Fulfilment of the assignment	*1 = assignment fulfilled,* ***2 = assignment fulfilled with minor objections,*** *3 = assignment fulfilled with major objections,* *4 = assignment not fulfilled*

Criteria description:
Assess whether the thesis meets the assignment statement. In Comments indicate parts of the assignment that have not been fulfilled, completely or partially, or extensions of the thesis beyond the original assignment. If the assignment was not completely fulfilled, try to assess the importance, impact, and possibly also the reason of the insufficiencies.

Comments:
The primary goal of the thesis is to extend the functionality of the Genthat framework. The extension includes generating tests for methods which takes environments as their arguments and to prevent the requirement of forcing the arguments of a function to be evaluated upon test generation.

A part of the requirement for this thesis is to evaluate the advantages of tracing based test generation. I had hoped to see the thesis mention some simply tests being generated for some R packages. However, the lack of such evidence is understandable, as even with the extensions provided, Genthat is still very limited and cannot presently cover real world packages. This thesis provided a significant step forward for Genthat to incorporated into mainstream use.

Evaluation criterion:	*The evaluation scale: 1 to 4.*
3. Size of the main written part	***1 = meets the criteria,*** *2 = meets the criteria with minor objections,* *3 = meets the criteria with major objections,* *4 = does not meet the criteria*

Criteria description:
Evaluate the adequacy of the extent of the final thesis, considering its content and the size of the written part, i.e. that all parts of the thesis are rich on information and the text does not contain unnecessary parts.

Comments:
The length of the thesis is satisfactory. The thesis include a rich and deep introduction to the complex nature of the R programming language, as well as an understanding of the benefits of tracing based test generation.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
4. Factual and logical level of the thesis	*90 (A)*

Criteria description:
Assess whether the thesis is correct as to the facts or if there are factual errors and inaccuracies. Evaluate further the logical structure of the thesis, links among the chapters, and the comprehensibility of the text for a reader.

Comments:
The thesis contains no major mistakes or errors. The chapters was structured nicely, and flowed well throughout the thesis. The writing covered well the intent of the thesis.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
5. Formal level of the thesis	*85 (B)*

Criteria description:
Assess the correctness of formalisms used in the thesis, the typographical and linguistic aspect s, see Dean's Directive No. 14/2015, Article 3.

Comments:
The thesis contained some minor typos and minor grammatical errors.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
6. Bibliography	*90 (A)*

Criteria description:
Evaluate the student's activity in acquisition and use of studying materials in his thesis. Characterize the choice of the sources. Discuss whether the student used all relevant sources, or whether he tried to solve problems that were already solved. Verify that all elements taken from other sources are properly differentiated from his own results and contributions. Comment if there was a possible violation of the citation ethics and if the bibliographical references are complete and in compliance with citation standards.

Comments:
The thesis contains a complete bibliography of the background involving Genthat and the R programming language.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
7. Evaluation of results, publication outputs and awards	*85 (B)*

Criteria description:
Comment on the achieved level of major results of the thesis and indicate whether the main results of the thesis extend published state-of-the-art results and/or bring completely new findings. Assess the quality and functionality of hardware or software solutions. Alternatively, evaluate whether the software or source code that was not created by the student himself was used in accordance with the license terms and copyright. Comment on possible publication output or awards related to the thesis.

Comments:
The overall result of the thesis is very encouraging, and is on par with the expectation of the task. The lack of an evaluation for the extension does cause some concerns over its validity.

Any future publication of the Genthat framework would be built on top of the work presented in this thesis. There is a very realistic possibility for Genthat to be present as an industrial grade test generation framework for real world application.

Evaluation criterion:	*No evaluation scale.*
8. Applicability of the results	

Criteria description:
Indicate the potential of using the results of the thesis in practice.

Comments:
The extensions presented in this thesis for test generation has great applicability. Active research is currently being conducted over the Genthat framework, directly on top of the extensions presented in this thesis.

Evaluation criterion:	*No evaluation scale.*
9. Questions for the defence	

Criteria description:
Formulate any question(s) that the student should answer to the committee during the defence (use a bullet list).

Questions:
* If you could, please explain why was there was no evaluation of the extension. How difficult would it be to produce tests for real world R packages? What other features would Genthat require?

* What did you find to to be the most challenging aspect of this project?

* What would you have done differently, if you had to redo this project again?

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
10. The overall evaluation	*88 (B)*

Criteria description:
Summarize the parts of the thesis that had major impact on your evaluation. The overall evaluation **does not** have to be the arithmetic mean or any other formula with the values from the previous evaluation criteria 1 to 9.

Comments:
The thesis presents a comprehensive background of the R programming languages, the challenges of working in R, the aim of the Genthat framework, and the target of it's extension for covering environments and lazy arguments. Overall, the work produced in the thesis and the implementation is of a high standard.

Signature of the reviewer: