Federal Technological University of Paraná — UTFPR/CP
Departament of Computer
Cornélio Procópio — Brazil

Experiment: Evaluation of the use of concept map to support the selection of primary studies in the systematic review process — EXECUTION

Instructions: Please read and perform the following task.

Task 1. Evaluate the concept maps in each item considering the affirmation. You must tick ☑ in the item that represents your opinion about each affirmation.

Input: *Likert scale evaluation of Concept Maps generated (Appendix A), Abstracts and Concept maps (Appendix B).*
Output waited: *All Concept Maps must be evaluated in Likert scale.*

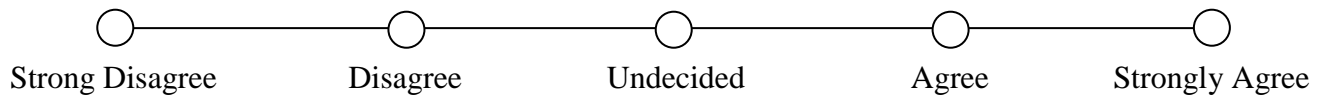*Important:* **Before starting make sure that you have understood how to evaluate the Concept Maps**

Important: **After you have finished the above task, answer the question on Appendix C**
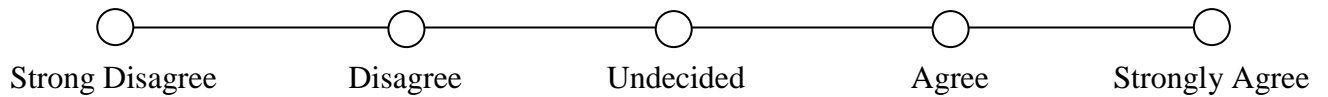

Code: 1401TO

Appendix A – Questionnaire

Evaluation

- **RQ1 –** The concepts presented in the Concept Map represent ideas and information related to the element "Result" of the correspondent abstract.

  ○————————○————————○————————○————————○
  Strong Disagree   Disagree   Undecided   Agree   Strongly Agree

- **RQ2 –** The relationships (links) adequately connect the concepts related to the element "Result" of the correspondent abstract.

  ○————————○————————○————————○————————○
  Strong Disagree   Disagree   Undecided   Agree   Strongly Agree

- **RQ3 –** The Concept Map cover all information related to the element "Result" of the correspondent abstract.

  ○————————○————————○————————○————————○
 Strong Disagree   Disagree   Undecided   Agree   Strongly Agree
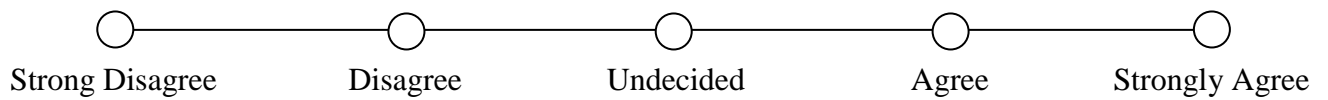
Appendix B – List of papers and their respective concept maps

Paper 1 - Towards Cross-Browser Incompatibilities Detection: a Systematic Literature Review

- **Abstract**

Cross Browser Incompatibilities (XBI) stands for compatibility issues which can be observed while rendering the same web application in different browsers. Users can interact with the Web through distinct web browsers implementations, such as: Internet Explorer, Microsoft Edge, Mozilla Firefox, Opera, Google Chrome, among others. However, with the increasing number of browser implementations and the continually evolving characteristic of web technologies lead to differences in how browsers behave and render web applications. In order to overcome this issue during the software development process, web developers must detect and fix XBIs before deploying web applications, regardless of the effort and cost required to conduct these inspections. Many developers rely on manual inspection of every web page of their applications rendered in various configuration environments (considering multiple OS platforms, browser implementations and versions) to detect XBIs, independently of the effort and cost required to conduct these inspections. This paper presents a Systematic Literature Review (SLR) on XBI automatic detection strategies published as primary studies. The SLR process was conducted with the goal of identifying distinct techniques which have been used to identify XBI, present the evolution of the developed tools, and guide future research on the topic. **In accordance to our findings, the strategies identified through the primary studies range from DOM (Document Object Model) Structure analysis, screenshot comparison, graph's isomorphism algorithms, machine learning, adaptive random testing, relative layout comparison, and static analysis. The SLR found 17 articles and 7 developed tools.**
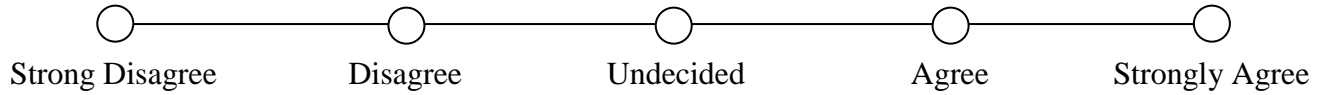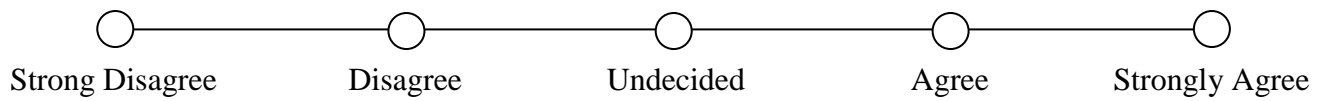
- **Concept Map**
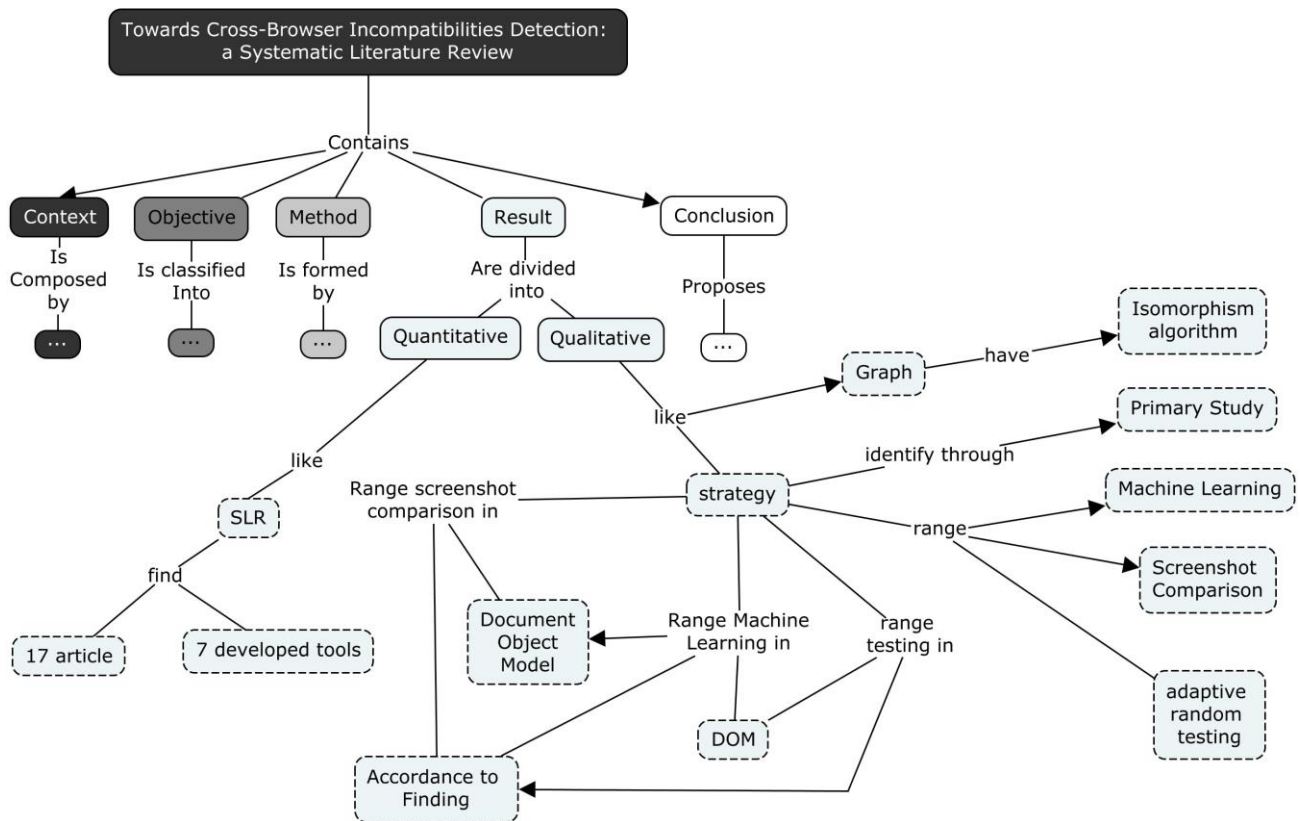
Classification:

- **RQ1 –** The concepts presented in the Concept Map represent ideas and information related to the element "Result" of the correspondent abstract.

| ○———————○———————○———————○———————○ |
| Strong Disagree   Disagree   Undecided   Agree   Strongly Agree |

- **RQ2 –** The relationships (links) adequately connect the concepts related to the element "Result" of the correspondent abstract.

| ○———————○———————○———————○———————○ |
| Strong Disagree   Disagree   Undecided   Agree   Strongly Agree |

- **RQ3 –** The Concept Map cover all information related to the element "Result" of the correspondent abstract.

| ○———————○———————○———————○———————○ |
| Strong Disagree   Disagree   Undecided   Agree   Strongly Agree |
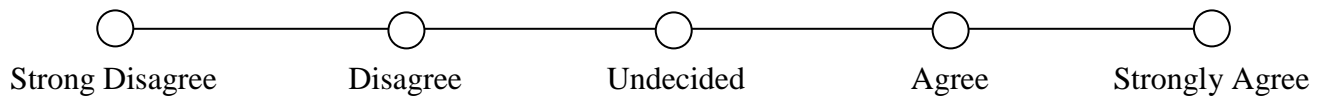
Paper 2 - Evaluating test suite characteristics, cost, and effectiveness of FSM-based testing methods
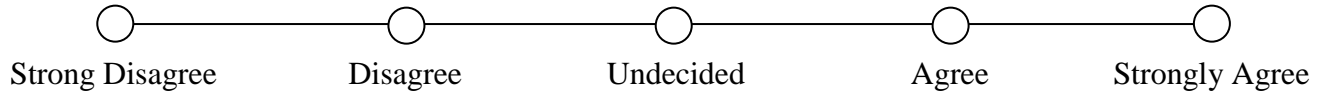
- **Abstract**

Testing from finite state machines has been investigated due to its well-founded and sound theory as well as its practical application. There has been a recurrent interest in developing methods capable of generating test suites that detect all faults in a given fault domain. However, the proposal of new methods motivates the comparison with traditional methods. We compare the methods that generate complete test suites from finite states machines. The test suites produced by the W, HSI, H, SPY, and P methods are analyzed in different configurations. Complete and partial machines were randomly generated varying numbers of states, inputs, outputs, and transitions. These different configurations were used to compare test suite characteristics (number of resets, test case length) and the test suite length (i.e., the sum of the length of its test cases). The fault detection ratio was evaluated using mutation testing to produce faulty implementations with an extra state. **On average, the recent methods (H, SPY, and P) produced longer test cases but smaller test suites than the traditional methods (W, HSI). The recent methods generated test suites of similar length, though P produced slightly smaller test suites. The SPY and P methods had the highest fault detection ratios and HSI had the lowest. For all methods, there was a positive correlation between the number of resets and the test suite length and between the test case length and the fault detection ratio**. The recent methods rely on fewer and longer test cases to reduce the overall test suite length, while the traditional methods produce more and shorter test cases. Longer test cases are correlated to fault detection ratio which favored SPY, though all methods have a ratio of over 92%.
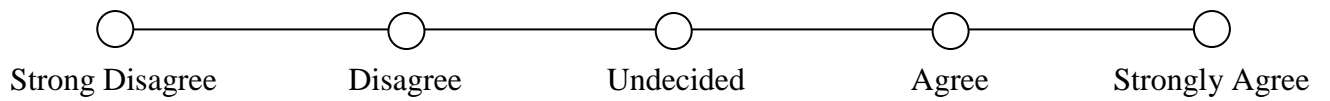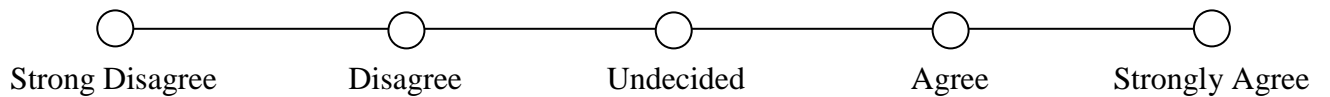
- **Concept Map**
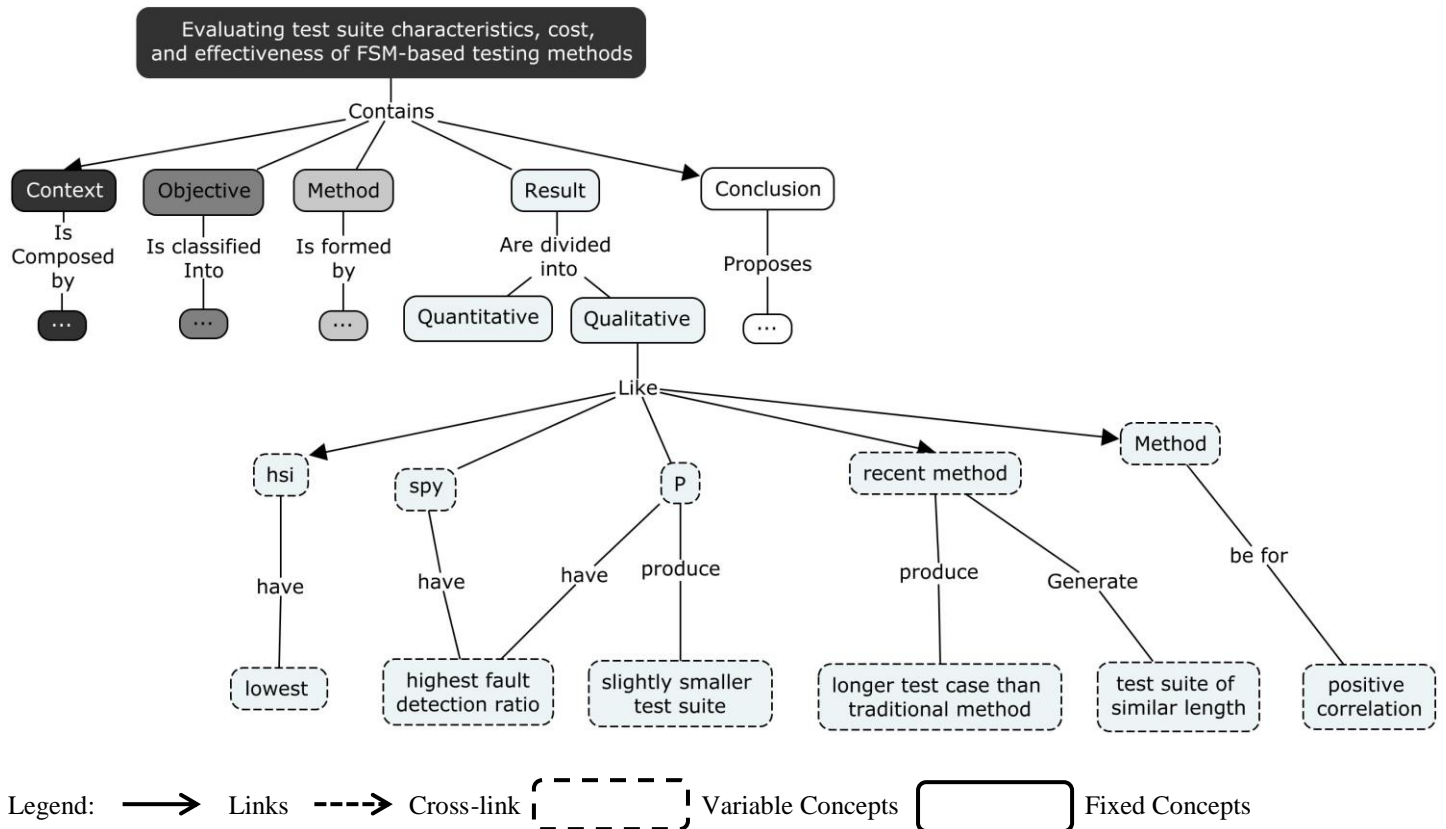
5

Classification:

- **RQ1 –** The concepts presented in the Concept Map represent ideas and information related to the element "Result" of the correspondent abstract.

| Strong Disagree | Disagree | Undecided | Agree | Strongly Agree |

- **RQ2 –** The relationships (links) adequately connect the concepts related to the element "Result" of the correspondent abstract.

| Strong Disagree | Disagree | Undecided | Agree | Strongly Agree |

- **RQ3 –** The Concept Map cover all information related to the element "Result" of the correspondent abstract.

| Strong Disagree | Disagree | Undecided | Agree | Strongly Agree |

Federal Technological University of Paraná — UTFPR/CP
Departament of Computer
Cornélio Procópio — Brazil

Paper 3 - An empirical study to quantify the characteristics of Java programs that may influence symbolic execution from a unit testing perspective

- **Abstract**

In software testing, a program is executed in hopes of revealing faults. Over the years, specific testing criteria have been proposed to help testers to devise test cases that cover the most relevant faulty sce- narios. Symbolic execution has been used as an effective way of automatically generating test data that meet those criteria. Although this technique has been used for over three decades, several challenges re- main and there is a lack of research on how often they appear in real-world applications. In this paper, we analyzed two samples of open source Java projects in order to understand the characteristics that may hinder the generation of unit test data using symbolic execution. The first sample, named SF100, is a third party corpus of classes obtained from 100 projects hosted by SourceForge. The second sample, called R47, is a set of 47 well-known and mature projects we selected from different repositories. Both samples are compared with respect to four dimensions that influence symbolic execution: path explosion, constraint complexity, dependency, and exception-dependent paths. **The results provide valuable insight into how researchers and practitioners can tailor symbolic execution techniques and tools to better suit the needs of different Java applications.**
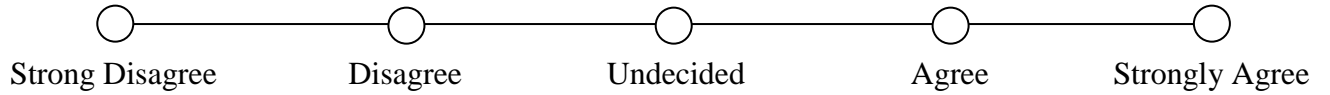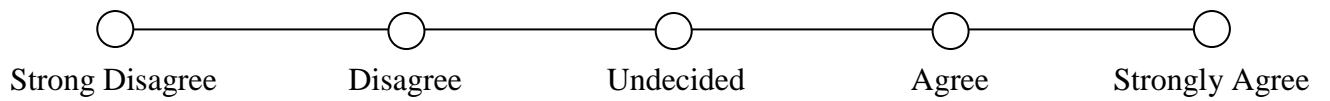
Concept Map
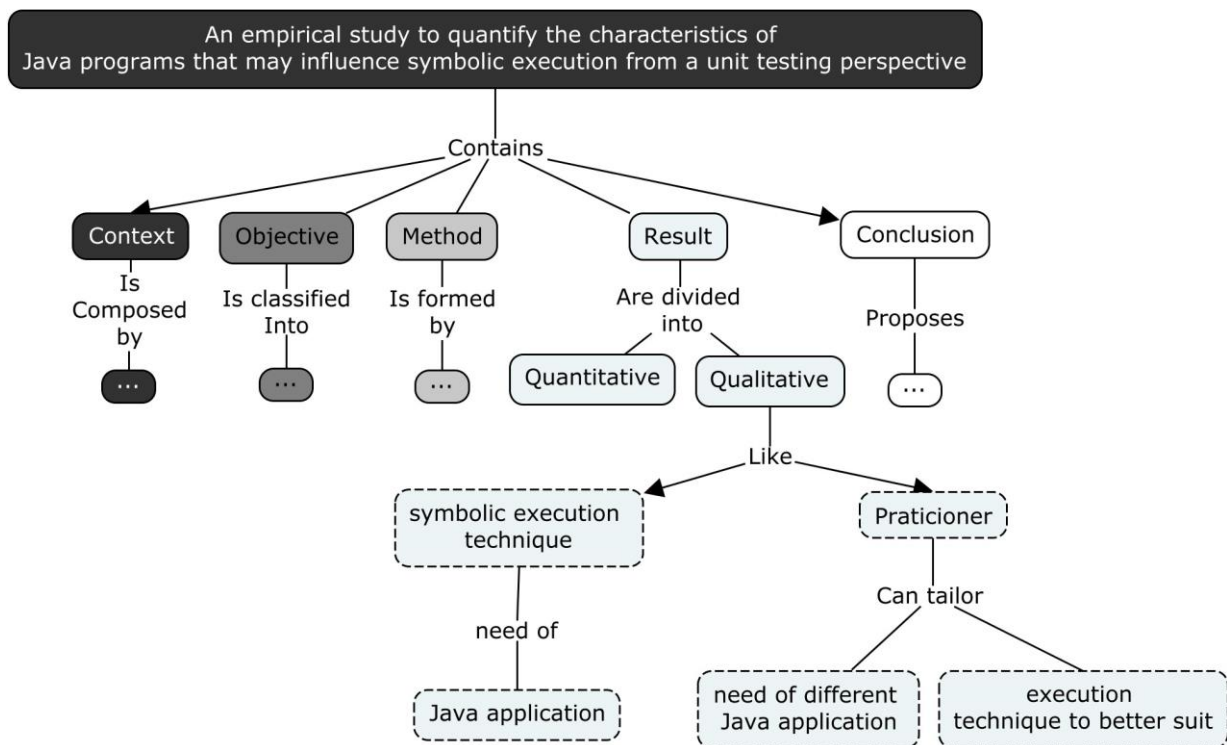
Classification:

- **RQ1 –** The concepts presented in the Concept Map represent ideas and information related to the element "Result" of the correspondent abstract.

| Strong Disagree | Disagree | Undecided | Agree | Strongly Agree |

- **RQ2 –** The relationships (links) adequately connect the concepts related to the element "Result" of the correspondent abstract.

| Strong Disagree | Disagree | Undecided | Agree | Strongly Agree |

- **RQ3 –** The Concept Map cover all information related to the element "Result" of the correspondent abstract.

| Strong Disagree | Disagree | Undecided | Agree | Strongly Agree |
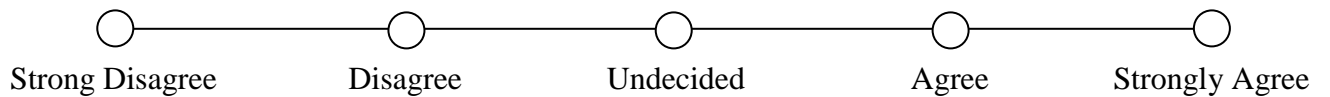
Paper 4 - Test oracles for simulink-like models

- **Abstract**

The design of embedded systems is often supported by the definition of executable models for tools like Matlab/Simulink or Scilab/Xcos. These models play a pivotal role in the development process and their correctness is thus extremely important. Many different solutions exist for the definition of suitable tests to "exercise" these models, but only a few (partial) solutions exist for assessing the quality of execution (simulation) results, that is, for defining suitable oracles. This paper addresses the problem and proposes a formal language for specifying the oracles and relating them to existing models. **It also presents Apolom, a prototype tool for checking simulation results against stated oracles. The empirical assessment we conducted to assess the viability of the proposed solution is organized around four case studies and witnesses interesting results in terms of effectiveness, efficiency, and required resources.**
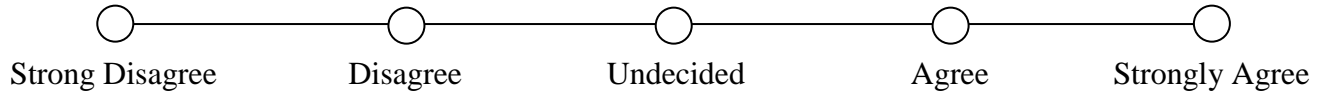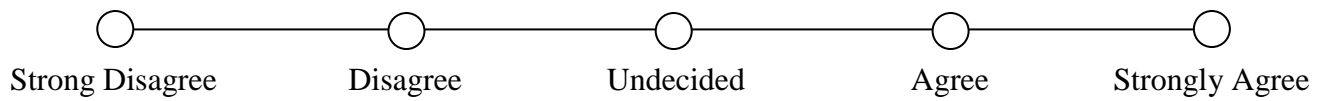
- **Concept Map**
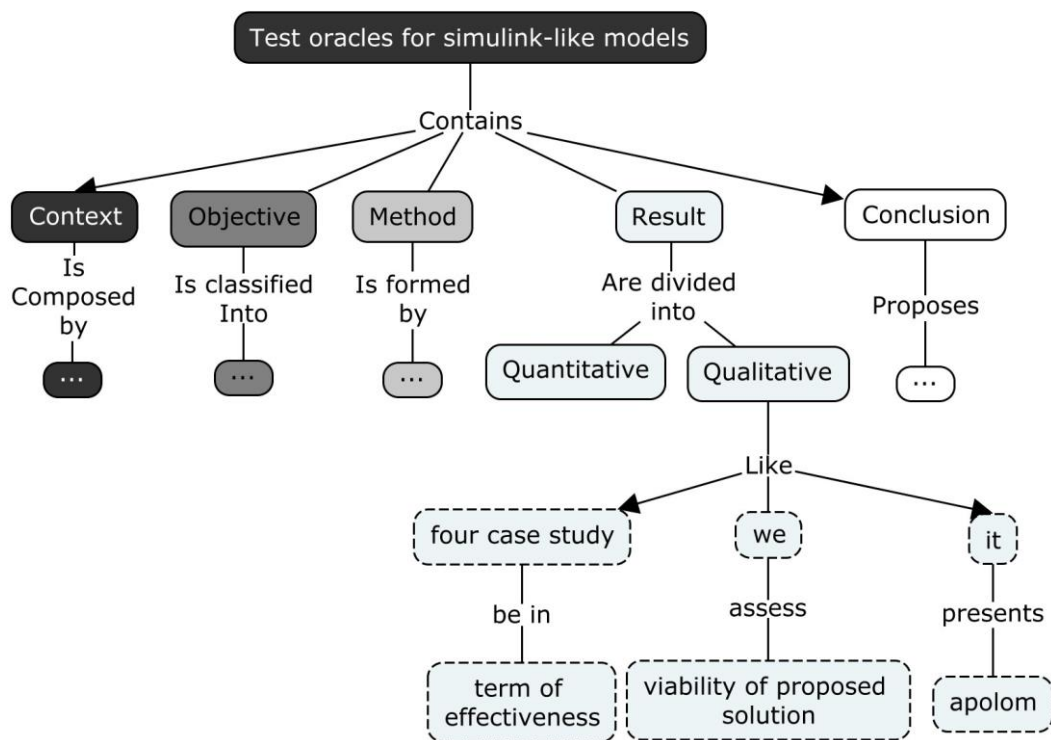
Classification:

- **RQ1 –** The concepts presented in the Concept Map represent ideas and information related to the element "Result" of the correspondent abstract.

Strong Disagree	Disagree	Undecided	Agree	Strongly Agree

- **RQ2 –** The relationships (links) adequately connect the concepts related to the element "Result" of the correspondent abstract.

Strong Disagree	Disagree	Undecided	Agree	Strongly Agree

- **RQ3 –** The Concept Map cover all information related to the element "Result" of the correspondent abstract.

| Strong Disagree	Disagree	Undecided	Agree	Strongly Agree
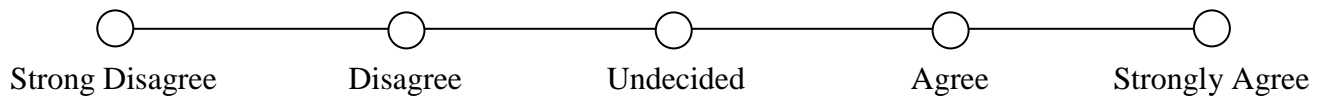
Paper 5 – Knowledge management initiatives in software testing: A mapping study

- **Abstract**

Context: Software testing is a knowledge intensive process, and, thus, Knowledge Management (KM) principles and techniques should be applied to manage software testing knowledge. Objective: This study conducts a survey on existing research on KM initiatives in software testing, in order to identify the state of the art in the area as well as the future research. Aspects such as purposes, types of knowledge, technologies and research type are investigated. Method: The mapping study was performed by searching seven electronic databases. We considered studies published until December 2013. The initial resulting set was comprised of 562 studies. From this set, a total of 13 studies were selected. For these 13, we performed snowballing and direct search to publications of researchers and research groups that accomplished these studies. **Results**: **From the mapping study, we identified 15 studies addressing KM initiatives in software testing that have been reviewed in order to extract relevant information on a set of research questions. Conclusions**: Although only a few studies were found that addressed KM initiatives in software testing, the mapping shows an increasing interest in the topic in the recent years. Reuse of test cases is the perspective that has received more attention. From the KM point of view, most of the studies discuss aspects related to providing automated support for managing testing knowledge by means of a KM system. Moreover, as a main conclusion, the results show that KM is pointed out as an important strategy for increasing test effectiveness, as well as for improving the selection and application of suited techniques, methods and test cases. On the other hand, inadequacy of existing KM systems appears as the most cited problem related to applying KM in software testing.
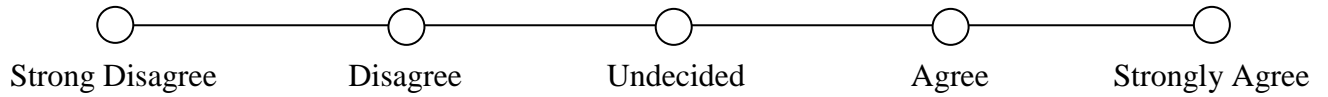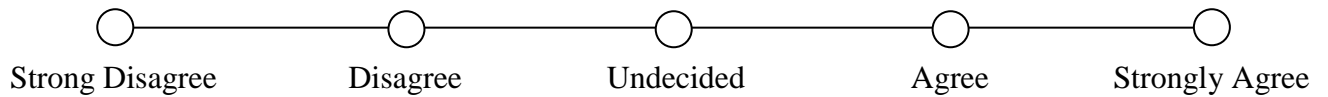
- **Concept Map**
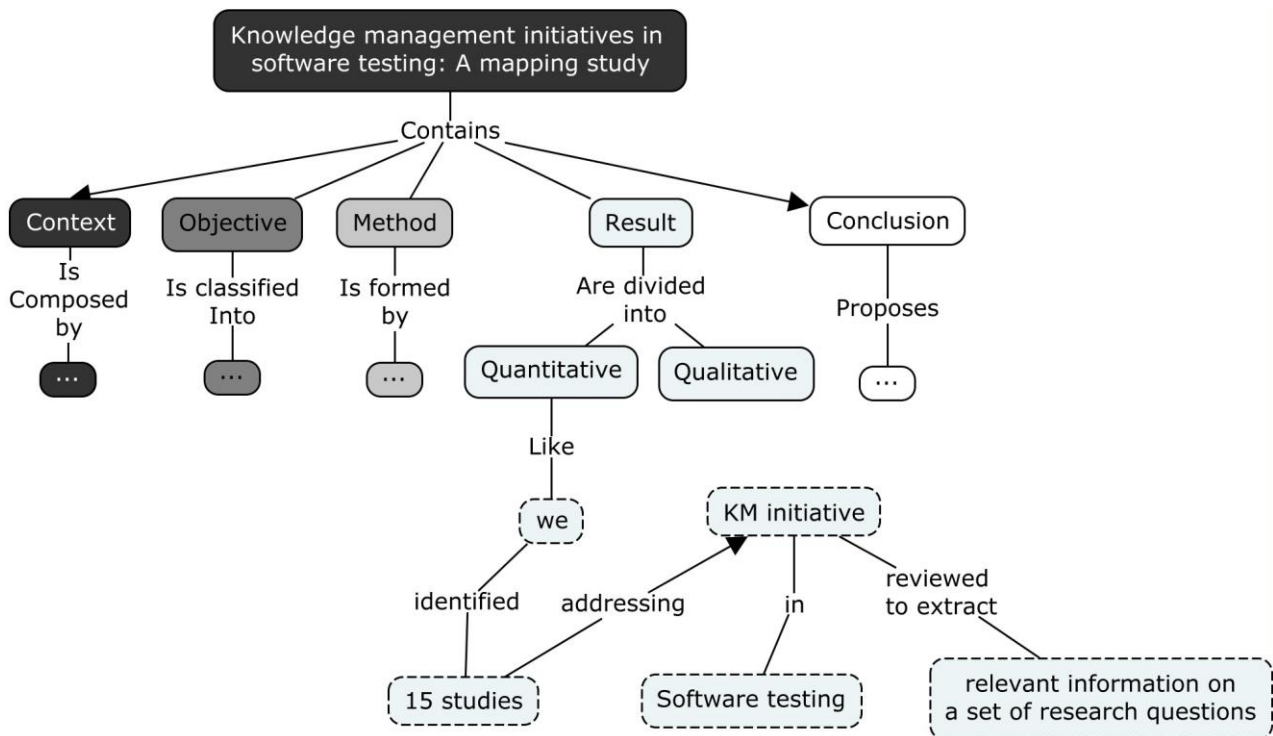
Classification:

- **RQ1 –** The concepts presented in the Concept Map represent ideas and information related to the element "Result" of the correspondent abstract.

| ○ ——— ○ ——— ○ ——— ○ ——— ○ |
| Strong Disagree   Disagree   Undecided   Agree   Strongly Agree |

- **RQ2 –** The relationships (links) adequately connect the concepts related to the element "Result" of the correspondent abstract.

| ○ ——— ○ ——— ○ ——— ○ ——— ○ |
| Strong Disagree   Disagree   Undecided   Agree   Strongly Agree |

- **RQ3 –** The Concept Map cover all information related to the element "Result" of the correspondent abstract.

| ○ ——— ○ ——— ○ ——— ○ ——— ○ |
| Strong Disagree   Disagree   Undecided   Agree   Strongly Agree |
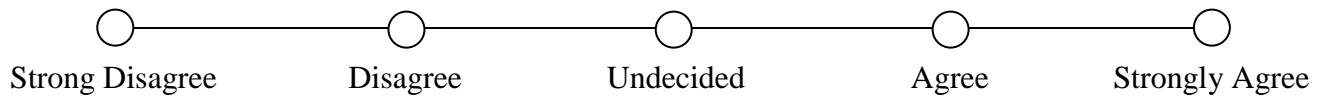
Paper 6 - The Use of Visual Text Mining to Support the Study Selection Activity in Systematic Literature Reviews: A Replication Study
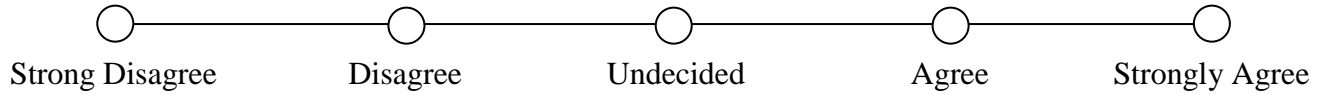
- **Abstract**

Background: Systematic literature reviews (SLRs)are an important component to identify and aggregate research evidence from different empirical studies. One of the activities associated with the SLR process is the selection of primary studies. The process used to select primary studies can be arduous, particularly when the researcher faces large volumes of primary studies. Aim: An experiment was conducted as a pilot test to compare the performance and effectiveness of graduate students in selecting primary studies manually and using visual text mining (VTM) techniques. This paper describes a replication study. Method: The same experimental design and materials of the previous experiment were used in the current experiment. **Result: The previous experiment revealed that VTM techniques can speed up the selection of primary studies and increase the number of studies correctly included/excluded (effectiveness). The results of the replication confirmed that studies are more rapidly selected using VTM. We observed that the level of experience in researching has a direct relationship with the effectiveness.** Conclusion: VTM techniques have proven valuable in the selection of primary studies.
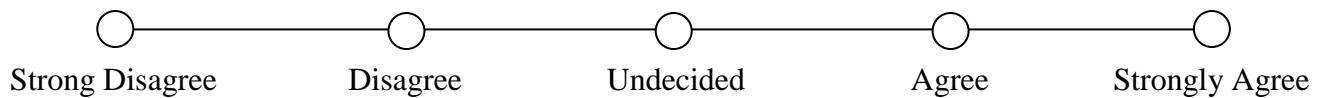
- **Concept Map**
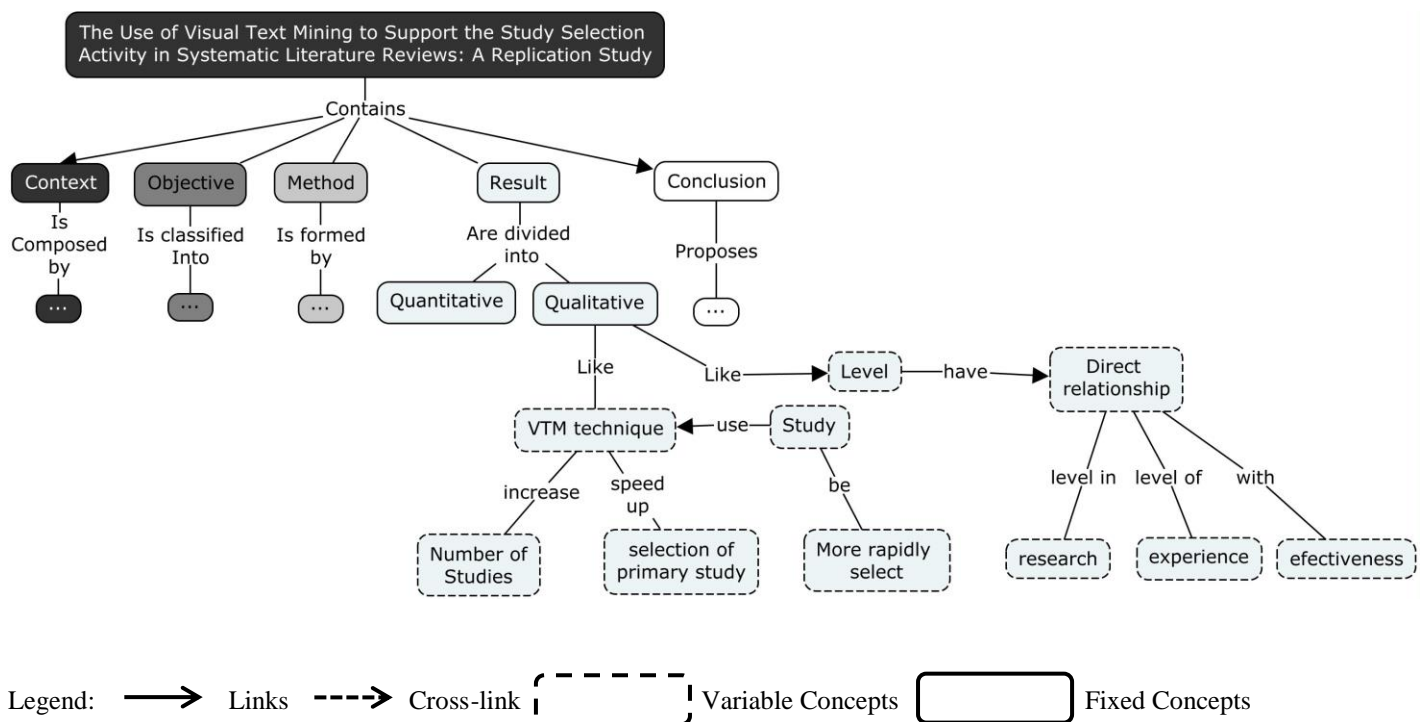
Legend: → Links   --→ Cross-link   [Variable Concepts]   [Fixed Concepts]

Classification:

- **RQ1 –** The concepts presented in the Concept Map represent ideas and information related to the element "Result" of the correspondent abstract.

| Strong Disagree | Disagree | Undecided | Agree | Strongly Agree |

- **RQ2 –** The relationships (links) adequately connect the concepts related to the element "Result" of the correspondent abstract.

| Strong Disagree | Disagree | Undecided | Agree | Strongly Agree |

- **RQ3 –** The Concept Map cover all information related to the element "Result" of the correspondent abstract.

| Strong Disagree | Disagree | Undecided | Agree | Strongly Agree |
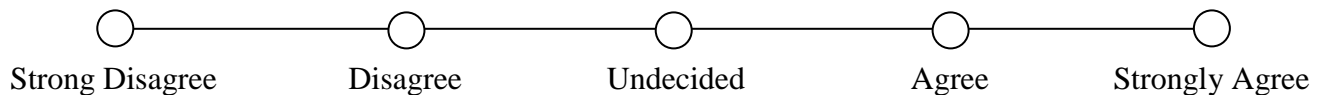
End of Task.

After you have performed the above task, please answer the question on Appendix C.

Appendix C

Instructions:
Please answer the following questions.

Do you have some consideration about this research?

_____

_____

_____

_____

_____

_____

Thank you for your time!