



Review report of a final thesis

Reviewer: Ing. Marek Suchánek
Student: Bc. Richard Husár
Thesis title: Searching Inside (Onto)UML Structural Conceptual Models
Branch / specialization: Knowledge Engineering
Created on: 11 May 2021

Evaluation criteria

1. Fulfillment of the assignment

- [1] assignment fulfilled
- ▶ [2] **assignment fulfilled with minor objections**
- [3] assignment fulfilled with major objections
- [4] assignment not fulfilled

I evaluate the assignment of the diploma thesis as fulfilled; I have only minor objections. The student got acquainted with UML and XMI, according to the assignment. Still, a description of OntoUML is missing in work (although he probably also worked with it during the implementation, respectively with UML stereotypes). Similarly just Neo4j is described and the reason of why it is chosen is not clear. The design and implementation of both parts are done, but the description of testing could be more elaborate. The use case scenario shows how the developed tools can be used with Enterprise Architect and OpenPonk UML models. The solution is then briefly discussed in conclusion.

2. Main written part

60/100 (D)

The text of the thesis is structured according to the assignment and procedure of the thesis; the parts follow each other logically. Apart from the OntoUML section, nothing significant is missing and contains relevant information. The structure of the work could be improved (paragraphs with only literal quotations, single subsection, etc.).

The work is written in English without major grammatical issues. On the other hand, the work lacks care in terms of typography: dash vs hyphen, no-break-space before cite/reference, code fragment as a block without a caption (p. 14), Figure 3.11 with unnecessary overflow, no syntax highlighting in Listing 3.2, big blank space on page 38, unresolved reference (??) on page 39, terms from XML in the text should be monospaced everywhere (inconsistent, somewhere in italics, somewhere in quotes, somewhere normal text), etc. Another issue is with names and capitalization, sometimes it is inconsistent but sometimes wrong everywhere, including headings (e.g. Openponk vs OpenPonk, Python vs python - or is it a snake, Neo4j vs neo4j (somewhere is also Noe4j or

just Neo - is that the same thing?), Github vs GitHub, Enterprise architect vs Enterprise Architect, Repocribo vs repocribo, Stackoverflow vs Stack Overflow, etc. Some abbreviations are not listed (EA, MOF, GPL, ACID, SQL, etc.) There are strange text wraps in the list on page 30. The last problem is pictures - for diagrams, it is suitable to use vector format, Enterprise Architect can also export SVG using a plugin or at least background-less PNG. Some of the figures are too small; however, they could be re-arranged and enlarged (using the same space in the thesis).

The work uses 26 citation sources, which are relevant to the topic. Almost all online sources do not list the author. Images taken or created according to the template/description should use citations (for example, 1.1, 1.2, and 1.10 - but for some, it is correct, for example, 1.6). Blocks of literal quotations in the State-of-the-art chapter are, in my opinion, overused.

3. Non-written part, attachments

85 /100 (B)

The non-written parts include the implementation of the parser and integration with Repocribo. Both are successful; I also appreciate the work with Docker. The code is quite well structured due to the nature of the implementation (parsers are very specific). However, there are small possibilities for improvement (loading the configuration from a file or environment variables, using f-string universally, code style according to PEP8, etc.). The `__pycache__`, `dist`, `.egg-info`, and other generated folders should not be copied and included in the attachment.

4. Evaluation of results, publication outputs and awards

95 /100 (A)

The results of the work are ready for basic use and further expansion and improvement. The use of the tool brings an exciting possibility to conceptual modelling, which is also taught at FIT CTU. The integration of concepts from different models is an important topic that is likely to be even more exposed in the future.

The overall evaluation

85 /100 (B)

Overall, I evaluate the work with grade B. Although I have many complaints about the text of the work, the evaluation was influenced by the implementation and its usability in conceptual modeling, including teaching at FIT CTU.

Questions for the defense

- Why did you choose Neo4j compared to other graph databases? Would it be possible to work with the specified SPARQL endpoint in general? What would be the advantages and disadvantages?
- What will be the next steps for further development and application of the work results in practice?

Instructions

Fulfillment of the assignment

Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.

Main written part

Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies?

Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 26/2017, Art. 3.

Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.

Non-written part, attachments

Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.

Evaluation of results, publication outputs and awards

Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

The overall evaluation

Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.