

Supervisor's statement of a final thesis

Czech Technical University in Prague

Faculty of Information Technology

Student: Bc. Filip Masri
Supervisor: Ing. Jan Šedivý, CSc.
Thesis title: Automatic ontology learning from semi-structured data
Branch of the study: Knowledge Engineering

Date: 30. 5. 2017

<p><i>Evaluation criterion:</i></p> <p>1. Difficulty and other comments on the assignment</p> <p><i>Criteria description:</i> 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.)</p> <p><i>Comments:</i> The ontology generation is a current research topic. Many companies and institutions try solving it. The ontology creation from semistructured data is a little simpler but still a challenging problem. Finding a solution requires knowledge in many different disciplines. The main difficulty is the significant variability of the web content to process, and an enormous problem is also an objective comparison of ontologies or performance estimation. These topics are still a long lasting research problem. The thesis is a contribution showing some possible solutions.</p>	<p><i>The evaluation scale: 1 to 5.</i></p> <p>1 = extremely challenging assignment, 2 = rather difficult assignment, 3 = assignment of average difficulty, 4 = easier, but still sufficient assignment, 5 = insufficient assignment</p>
<p><i>Evaluation criterion:</i></p> <p>2. Fulfilment of the assignment</p> <p><i>Criteria description:</i> 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.</p> <p><i>Comments:</i> The thesis meets the assignment statements completely.</p>	<p><i>The evaluation scale: 1 to 4.</i></p> <p>1 = assignment fulfilled, 2 = assignment fulfilled with minor objections, 3 = assignment fulfilled with major objections, 4 = assignment not fulfilled</p>
<p><i>Evaluation criterion:</i></p> <p>3. Size of the main written part</p> <p><i>Criteria description:</i> 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.</p> <p><i>Comments:</i> The extent of the final thesis is adequate. The thesis is logically broken into sections with an appropriate size. The text is adequately and clearly describing the thesis problem.</p>	<p><i>The evaluation scale: 1 to 4.</i></p> <p>1 = meets the criteria, 2 = meets the criteria with minor objections, 3 = meets the criteria with major objections, 4 = does not meet the criteria</p>
<p><i>Evaluation criterion:</i></p> <p>4. Factual and logical level of the thesis</p> <p><i>Criteria description:</i> 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.</p> <p><i>Comments:</i> The thesis does not contain any factual errors. Indeed, some of the described facts may be interpreted differently by different experts. Subjectivity is a problem of all projects working with the human knowledge where we always can rely only on statistically supported judgements.</p> <p>The whole thesis is written very logically. The chapters explain systematically the solution.</p>	<p><i>The evaluation scale: 0 to 100 points (grade A to F).</i></p> <p>95 (A)</p>
<p><i>Evaluation criterion:</i></p> <p>5. Formal level of the thesis</p> <p><i>Criteria description:</i> Assess the correctness of formalisms used in the thesis, the typographical and linguistic aspects, see Dean's Directive No. 14/2015, Article 3.</p>	<p><i>The evaluation scale: 0 to 100 points (grade A to F).</i></p> <p>100 (A)</p>

Comments:

The thesis is written in English, which is superb for this type of work. Since the author is not a native speaker we may have some objections or find better formulations, but the thesis is understandable, which is the main point.

Evaluation criterion:

The evaluation scale: 0 to 100 points (grade A to F).

6. Bibliography

100 (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 whole text is referring to the list of literature. Especially at the beginning of the project the author has gathered a lot of literature and selected the best-known solutions. The initial review helped in choosing the first solution. The list of literature contains 32 references including the relevant ones. I have not noticed a violation of the citation ethics. The format of citation complies with the traditional way of referencing.

Evaluation criterion:

The evaluation scale: 0 to 100 points (grade A to F).

7. Evaluation of results, publication outputs and awards

100 (A)

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 author has picked efficient solutions and has put together a program creating ontologies from semistructured data found on the web. He has demonstrated the quality of the program on different data sets. He has suggested criteria for comparing ontologies. For better comprehension, he has designed a visualization depicting the ontology structure. While solving the thesis problem he has used several programming languages, and he has proved himself to be an experienced programmer as well as researcher.

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 thesis is one of the several steps for collecting the general knowledge. This work will continue. We expect to use the outcomes for practical solutions too.

Evaluation criterion:

The evaluation scale: 1 to 5.

9. Activity and self-reliance of the student

9a:

1 = excellent activity,
2 = very good activity,
3 = average activity,
4 = weaker, but still sufficient activity,
5 = insufficient activity

9b:

1 = excellent self-reliance,
2 = very good self-reliance,
3 = average self-reliance,
4 = weaker, but still sufficient self-reliance,
5 = insufficient self-reliance.

Criteria description:

Review student's activity while working on this final thesis, student's punctuality when meeting the deadlines and consulting continuously and also, student's preparedness for these consultations. Furthermore, review student's independency.

Comments:

The student was very focused while working on the thesis. He has been solving the problems very systematically. He was dealing with all details. We had regular weekly meetings The student has reported the results to a shared document. He was working very independently.

Evaluation criterion:

The evaluation scale: 0 to 100 points (grade A to F).

10. The overall evaluation

100 (A)

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:

Filip was one of the best students with high focus and systematic approach. He has written an excellent thesis. I hope, we find a similarly gifted follower to continue on our path of converting the general knowledge to computer processable form.

Signature of the supervisor: