PhD Thesis Review

Reviewer: doc. RNDr. Petr Šaloun, Ph.D.
Title: Towards OntoUML for Software Engineering:
Subtitle: Transformation of OntoUML into Relational Databases
Author: Zdeněk Rybola
Institution: Czech Technical University in Prague, Faculty of Information Technology

Up-to-datedness of the dissertation

All software engineering areas are very popular in recent years, particularly the Model-Driven Development (MDD). The research area of the thesis is actual, and the goals are solved.

Formal structure and organization of the dissertation

Thesis is 244 pages long, there is the running example supplement, 111 pages long, both written in English. The thesis consists of eight chapters, three appendices, bibliography, and few more chapters dealing with publications of the author. They are written well, the chapter's sequence is easy to follow and particular problems are placed and discussed well.

The introduction introduces topic, motivation, overview of the related work and author's previous results and contributions.

The state of the art gives an overview of the MDD approach, and the process of forward engineering based on transformations of conceptual models into their realization. The UML introduction, particularly the class diagrams, then OCL, its syntax, semantics, and an overview of some tools supporting the definition, validation and transformation of OCL constraints. Then an overview of some available tools for modeling in UML and OCL are discussed. Chapter three deals with the complex overview of UFO and OntoUML, as the next chapter introduces author's approach to the transformation of an OntoUML PIM into its realization in a relational database. Chapters 5, 6 and 7 are focused on transformations of OntoUML PIM into UML PIM, UML PIM into RDB PSM, and RDB PSM into ISM respectively. Chapter eight summarizes the results and gives few topics for further research.

The running example of a Library OntoUML PIM is clear and it contains details of the transformation of OntoUML PIM into its proper realization in a relational database, moreover it contains links to the thesis, where is it necessary. Say in short, the original OntoUML PIM model, and the other consecutive models are created in the running example by applying the transformation proposed in the thesis.

Completion of the dissertation objectives

They are three research questions in the thesis:

1. Is it possible to use OntoUML for conceptual data modelling?
2. Is it possible to transform an OntoUML model into a relational database model and generate SQL scripts from it?
3. Is it possible to realize all the implicit constraints defined by the types of universals and relations used in the OntoUML model in the relational database?

Assessment of the methods used in the dissertation

Methods used in the thesis are common in software engineering area. For example, the method for the transformation of OntoUML PIM into its realization in RDB, allows reusing the existing
know-how for the transformation of UML into RDB, and its division into three steps includes an optional refactoring and optimizations of the models between these individual steps.

**Evaluation of the results and contributions of the dissertation**

The contributions of the thesis can be summarized as follows:

1. Complex overview of OntoUML and UFO-A concepts is given – including the distinction of universals and individuals, together with the classification of the object types, moments and part-whole relations.

2. Example of a complex OntoUML conceptual model from the library domain is provided. This model contains most of the OntoUML and UFO-A concepts to demonstrate their gradual transformation.

3. Proposition of the method of the transformation of OntoUML PIM into the UML PIM is proposed, preserving the semantics defined by the OntoUML universal and relation types.

4. Proposition of the method of the transformation of UML PIM into the RDB PSM is proposed, preserving all the semantics defined in the derived UML PIM model – focusing on the special multiplicity constraints, and on the constraints derived from the initial OntoUML model.

5. Proposition of the method of the transformation of RDB PSM into ISM consisting of the SQL creation scripts is proposed, preserving the semantics derived from the OntoUML universal and relation types used in the initial OntoUML PIM. An overview of the available realizations of these constraints is provided and the advantages and disadvantages of each approach are discussed.

**Remarks, objections, notes, and questions for the defense**

The core of thesis is contained in chapters 4 to 7, and it was published in various international research conferences. There is presented a large running example the transformation of OntoUML PIM into its proper realization in a relational database.

Questions:

1. Could you estimate the benefits of the solutions and usability of the proposed implementations in practice? Choose a proper metrics for it.

2. Since you published your first results as early as 2011, did some of your partial results have not been part of standard procedures today?

**The overall evaluation of the dissertation**

The research results were published in journals and conference proceedings on the international level. The author of theses proved the ability to conduct research and achieve scientific results. In accordance with par. 47, letter (4) of the Law Nr. 111/1998 (The Higher Education Act) I do recommend the thesis for the presentation and defense with the aim of receiving the Ph.D. degree.

Ostrava, November 30, 2017

[Signature]

doc. RNDr. Petr Šaloun, Ph.D.
PhD Dissertation Review

PhD Candidate: Ing. Zdeněk Rybola
Title: “Towards OntoUML for Software Engineering: Transformation of OntoUML into Relational Databases”
Submitted to: Faculty of Information Technology, Czech Technical University in Prague, Czech Republic.

Overall Evaluation:

The work of Mr. Rybola aims at exploring the use of the Ontologically Well-Founded Conceptual Modeling language OntoUML in a Model-Driven Software Engineering Process. In particular, the thesis proposes a transformation of OntoUML models into UML Platform Independent Models (PIM), then into Relational Databases Platform Specific Models (PSM) and, finally, from the latter into SQL Implementation Specific Models (ISQ).

The thesis develops a general introduction to OntoUML in chapter 3 and also complements it with a sufficiently interesting running example. These two pieces of the work are methodologically useful and indeed compile in one place some of the most recent material on OntoUML that was somewhat scattered in different publications. I do not agree, however, in claiming these as contributions of the thesis. In my opinion, these are part of the research strategy/methodology employed here. Accordingly, this is something I would change in the text of the thesis. In any case, although there are some small corrections to be made in the description of the categories underlying UFO and OntoUML, the candidate demonstrated in these parts of the thesis a solid understanding of the subject. In summary, despite disagreeing with these accounts as contributions, I think this was a useful methodological strategy adopted by the candidate.

The actual bulk of thesis' contribution is developed in chapters 5-7. Here, once more, I do not agree with the entire set of mapping choices made by the candidate. For example, during the defense, it would be nice to hear from him on the specific choices made by the author when mapping roles, relators and material relations. In that subject, it seem clear to me that the most obvious proxy for modeling roles would be the explicit representation of qua individuals in the model (a UFO concept intimately tied to roles and relators). Moreover, the choice of eliminating relators when these do not have explicit attributes seems to be a poor one, given that, at minimum, relators help to eliminate ambiguity in cardinality constraints in material relators, which is something that can influence mapping choices henceforth. Nonetheless, I think what is elaborated in these chapters suffices to show that the candidate has acquired the ability to conduct independent research.

I am convinced that this thesis pursues a worthwhile goal and that it accomplishes these objectives to an adequate level. Moreover, I am convinced that the results presented here will stimulate discussions and will potentially be useful to the conceptual modeling, ontology and information systems engineering communities. Furthermore, the thesis uses a reasonable research methodology to pursue its goals. Generally speaking, the text of the thesis is well
written and well structured. Finally, the results presented here are also sufficiently original.

In terms of the limitations of this dissertation in its current presentation, I must say that perhaps its biggest weakness is the lack of a more profound discussion and in-depth analysis of related work. Although I am unaware of existing proposals in this direction that are as explicit and comprehensive as the one proposed here, there are, in the literature, partial attempts to map OntoUML to other PSM/ISM solutions (including Relational Databases). These should have been discussed in more details in the text and should have been the subject of a more comprehensive comparative analysis here.

In terms of results, an important limitation of the approach is the lack of some automated support to complement the proposals put forth here, as this limits the applicability of these results. Given that it has not been possible to develop this for the thesis, I would expect a more elaborated discussion on minimal requirements for such a support to be elaborated in the final chapter of the dissertation. In fact, a second noticeable limitation of the text presented there is the lack of a more detailed and farseeing discussion on the section of final considerations. Despite comprehensive, the work presented here is still partial as an OntoUML mapping as it does not cover essential aspects of practical OntoUML modeling such as the mapping of domain constraints, higher-order types, events, or even primary meta-properties of part-whole relations (e.g., transitivity, anti-symmetry, weak supplementation). It would have been important to acknowledge that and to deal explicit with these limitations in the final reflections on what has been achieved here and what needs to be pursued in future works that will build on these results.

Finally, I must say that I enjoyed reading Mr. Rybola's work. In reading his text, one can clearly perceive his acquired passion for the subjects addressed here and I think this is an important quality attribute for a PhD thesis. Moreover, I am personally very happy to see a work that advances the state of art in ontology-driven conceptual modeling, in general, and in OntoUML, in particular.

With all these points considered, I believe the candidate demonstrated the ability to conduct research and achieve scientific results and, hence, I do recommend the thesis for the presentation and defense with the aim of receiving the Ph.D. degree.

Prof. Dr. Giancarlo Guizzardi
Free University of Bozen-Bolzano, Italy
Prof. Ivan Luković, Ph.D. in Computer Science
University of Novi Sad
Faculty of Technical Sciences
Department of Computing and Control
Trg Dositeja Obradovića 6
21000 Novi Sad, Serbia

Novi Sad, 3-DEC-2017

Review of Dissertation Thesis of Zdeněk Rybola
„Towards OntoUML for Software Engineering: Transformation of OntoUML into Relational Databases“

Motives, Research Topics, and Formal Structure

The author finds roots of his Ph.D. research in the fact that software engineering is a demanding discipline that deals with complex systems, with a goal to ensure high-quality software implementations as complex systems. Nowadays, various software development approaches have been developed for complex software development. The author based his research on the extensive usage of Model Driven Software Development (MDSD) paradigm, as well as in a modest sense the Domain Specific Modeling (DSM) paradigm, in order to support conceptual modeling of software systems and than generating software implementations by means of a chain of model-to-model and model-to-code transformations.

The main concern of the author's approach is how to provide the appropriate meta models for conceptual modeling of Information Systems (IS) and their database structures and then, how to generate database implementations in the SQL language for a target Relational Database Management System (RDBMS). To the best of my knowledge, current software engineering approaches still offer not too much at the level of abstract design of IS models and their transformations to the program code implementations. Therefore, providing the most appropriate method for the development of an IS and its database structure is always a hot and interdisciplinary oriented research topic in the computer science and software engineering.

As a meta model for conceptual IS modeling, the author selected OntoUML, a highly expressive abstract modeling syntax based on the Unified Modeling Language (UML) and Unified Foundational Ontology (UFO). Author embedded OntoUML into his MDSD approach for generating SQL program code for target RDBMSs, by means of chain of transformations starting from OntoUML designers' specifications. The first transformation is a transformation from OntoUML specification as a Platform Independent Model (PIM) to a UML Class Model specification as a PIM model. Then, the second transformation is from the UML Class Model specification to a Relational Data Model specification as a Platform Specific Model (PSM). Finally, the third transformation in the chain is from Relational Data Model specification to a RDBMS SQL program code as an Implementation Specific Model (ISM). Through his work, the author paid a special care to various database constraint types, and ways how to handle constraint specifications in his approach. As OntoUML cannot express all the constraints as a structure, which is a typical limitation of such approaches, the author introduce the OCL language into his approach to cover specifications of more complex types of constraints that cannot be expressed by structural concepts, only. The author analyzed related works and found that there were no similar approaches, already formulated in the available literature.

The text report of this Ph.D. thesis has been well structured, in general. The text is written clearly, in a good English, and I followed it easily. The report has been equipped with all necessary components.
Up-to-datedness of the dissertation – Thesis Relevancy

According to all my previous considerations, I believe that this Ph.D. thesis present a research in a quite relevant topic of computer science and software engineering, with very good potentials for practical applicability of the obtained research results.

Fulfilling the main objectives

I believe that the author has successfully completed a research plan presented in his thesis and, by this, justified designed thesis goals in a satisfactory extent. The author identified and presented the main problems in MDSD approaches and their applications in the IS and database development, analyzed (however, modestly) how other approaches deal with such problems and proposed his own approach to address them in a relatively broad extent.

Method Appropriateness

To complete the designed goals, the author has selected quite appropriate and up-to-date computer science paradigms, methods, and information technologies. Potentials for applications and obtaining effective results of MDSD and Domain Specific Modeling (DSM) paradigms are much wider, than the current academy and industry communities are giving. I believe that the author has recognized such issue in his thesis, in a very good manner.

Evaluation of Main Results and Contributions

As already notified in the introductory part of this report, the main results are that the author has proposed a novel approach to the conceptual modeling of database structures of ISs and transformations of conceptual IS specifications to the ISM models, based on the SQL language. From my point of view, the main goal addressed in this thesis was to justify that OntoUML was a suitable meta modeling paradigm for conceptual modeling of the IS database structures that could be combined effectively together with MDSD approaches for generating SQL program code for target DBMSs.

The results presented in the thesis may be of a significant value both for academy (scientific) community, and for various software industry applications. The author presented here a real life case study, in which he justified the applicability of his approach in complex software projects. In computer science and software engineering, it is always important and hard issue.

Influence on Further Science Development

The author has opened new research questions and issues in his thesis. Some of them he has successfully resolved and thus created a valuable contribution to the academy society.

Some questions and issues, normally, still remain for further research. From my point of view, the author selected here a “never-ended” interdisciplinary research topic that comprises various research areas of software modeling and programming techniques. The author created by his research very nice potentials for formulating new research topics. By this, the Ph.D. thesis of Zdeněk Rybola could be a good basis for further research activities and creating new Ph.D. theses.

Thesis Creativity

This research thesis presents a quite creative research (scientific) work. The author proved his good knowledge of theoretical concepts, showed his ability to combine existing knowledge, create a new approach and justify its potentials for practical application, and by this create a new value that can be recognized in a scientific community.
Conclusion

I am really honored to declare that

Zdeněk Rybola, as the author of this Ph.D. thesis, proved the ability to conduct research and achieve scientific results.

In accordance with par. 47, letter (4) of the Law Nr. 111/1998 (The Higher Education Act) I do recommend the thesis for presentation and defense with the aim of receiving the Ph.D. Degree.

Ivan Luković
Some Drawbacks

As in any scientific work, there are some remarks that I would like to report here about drawbacks I have noticed in the thesis. The first, let me notify that those are really not strong findings. They do not influence seriously a significant value reported by this research work. I would appreciate if the author could just receive this message as an idea how to improve his future scientific work. It is important, as any Ph.D. candidate declare by this him or herself as a future research and academic worker, and above all a mentor of future Ph.D. candidates.

I find a composition of the Introduction chapter as a little bit strange to me. As the research motives are perfectly introduced and explained, the research goals, expected results, and values coming from those results are discussed with no enough details. The author even do not declare explicitly what are his research goals and, even more, what are the values coming from these goals, if they will be addressed in the proper way.

Section 1.2 is aimed at the discussion of a research problem. However, the author does not formulate an exact problem he wants to face. Everything written in Section 1.2 is fine. However, I could not understand from this text, what is a problem, and why in the previous works of other authors, such things have not be already done. Author should discuss here how previous works have contributed to this problem, and why the author believes that his approach may benefit.

Section 1.3, Related Work, is really poor. It should be definitely structured as a complete, new chapter, with much more considerations of already done works in regard to the author's thesis. Comparisons of related works with the author's approach are important here.

Section 1.4 also needs much more details to be presented. It is a crucial section of the author's work. The author says here: "This thesis provides only an overview of the OntoUML language and a theoretical proposition of the transformation methods. No exact methodology is defined and no tooling support is implemented for the proposed transformations. The implementation of the tool supporting the proposed transformation, as well as an extensive experimental evaluation of the proposed realizations are subjects for the future research."

I am sorry to say that I really cannot agree with this statement. This thesis is created at Czech Technical University and Faculty of Information Technology. Sorry to say, but it is not a thesis from Faculty of Philosophy, or from humanities, in any way. Therefore, the author was supposed to create some supporting tool, and give some basic experimental evaluation. From my point of view, this is the first urgency of the author's future work, and my advice for the author is to overcome this drawback as soon as possible in his upcoming activities, and to demonstrate the results of such activities at some internationally recognized conferences and in journals.

Future research work should be better elaborated. A value of some Ph.D. thesis is not just in what an author proposes, but also in what are the conceptual (not only technical) perspectives for future research that he or she opens. The author opens in his work really nice research perspectives. However he is supposed to elaborate them properly, as he is supposed as a Ph.D. to be a supervisor of future Ph.D. candidates. Therefore, future work should give a brief elaborate of at least two future Ph.D. works.

I believe that the author should have given more attention, particularly in the concluding section, about practical applicability of his approach. What are the pros and contras for applications in some real software projects of his approach? I believe that a wider consideration of various types of "untypical" constraints is really necessary for a practical usability of such approaches. Therefore, the author has a good chance to open this question as a new research perspective.