Review of Daniel Varro's MTA Doctoral Thesis:
Design and Analysis of Precise Model Transformations for Model-Driven Development

November 30, 2012

Daniel Varro's doctoral thesis focuses on a core issue of Model-Driven Development (MDD), the design and analysis of model transformations. Modern system and software development has been transformed significantly during the past two decades. The main driver for this transformation has been the rapid penetration of information technology in all categories of engineered systems. The effects of this change are increased heterogeneity and the need for turning system design into a cross-disciplinary activity. Domain-Specific Modeling Languages, model integration across design domains, formal verification of model properties, and model synthesis have become an essential tool for modern system and software design.

This trend is the general technical context for the thesis. Model transformations are the workhorse for MDD. Soundness of the technology infrastructure for their design, implementation and verification has major impact on progress. In summary, selection of the topic is excellent and well aligned with mainstream technology trends.

The thesis addresses four challenges:

1. Precise specification of model transformations
2. Design of model transformations
3. Scaling and integrating model transformations
4. Verification and validation of model transformations

As Varro correctly observed in Section 1.2 of his Thesis Summary, model transformations are arguably more complex and general than compilers, so the list of challenges he addressed covers a tremendous space. In spite of this broad agenda and ambitious list of research interest, he was able to make meaningful contributions, partially because he carefully narrowed down the research goals and partially because the field is relatively new and lacks maturity.

In *Challenge 1*, Varro correctly observes that much progress is needed in finding abstraction levels for specifying model transformations. His primary goal of achieving reusability is highly desirable and represents a step toward achieving compositionality in the specification of model transformations.

In *Challenge 2,* he observes that designing model transformations is hard therefore design technology needs to be elevated to be a major research topic. While this is certainly true, the specific problem he mentions – difference of the transformation language from the source and

target languages - is arguable. It is quite common in MDD that the source and target languages have deep semantic differences, or, the transformation together with the semantics of the target language defines the semantics for the source language. In these cases design of the transformations is inherently hard and requires deep understanding of the different semantic domains. The most effective help for the designers is a unified formal framework and mathematical domain for specifying semantics for the target and source DSMLs and for the transformation itself. This approach would not eliminate the hardness of designing model transformations, but would make the reasoning and tool support for validation and verification much easier. It must be mentioned that this argument does not challenge in any way Varro's contributions in his Thesis, but simply points to other interesting problems and approaches.

Challenge 3 calls for addressing scalability of model transformations and targets a technical solution that enables embedding transformations as building blocks in complex information systems. This goal is certainly practical and progress in the area would advance acceptance of MDD in industry.

Challenge 4 exposes one of the hardest problems in model transformations -validation and verification. Making correctness of model transformations part of the research agenda is certainly an excellent choice and increases value of the overall research contribution.

In the following I summarize my assessment on the novelty and significance of the scientific results.

1. **Specification Techniques for Model Transformations**
   The thesis addresses Challenge 1 by developing a hybrid model transformation language framework based on graph pattern matching, ASM-based control flow specification and introduction of higher-order transformations. Specification of VCTL is completed by defining formal semantics for graph patterns.
   a. While the method of graph pattern matching for model transformations is not new, the expressiveness of the graph pattern matching sublanguage of VCTL exceeds current state-of-the-art, therefore 1/1 point of the thesis contains new scientific results.
   b. Similarly, separating and making the control flow explicit in transformations is a know technique for achieving determinism. However, specification of the control flow using ASMs and the elegant way of integrating graph matching rules into ASM formalism is new, therefore 1/2 point of the thesis is a new scientific result.
   c. Increasing the level of abstraction by introducing generic transformations is a new idea and establishes opportunity for writing reusable transformations. Thesis 1/3 is a new scientific result.

Results summarized in Thesis 1 are addressing the fundamental goal of Challenge 1 by improving the precise specification of reusable model transformations.

2. Design Techniques for Model Transformations
   The thesis addresses Challenge 2 by developing a technique for deriving model transformation by example. The proposed approach is based on the application of inductive logic programming.
   a. The developed model transformation by example method is new and represent significant contribution to the state-of-the-art. The thesis provides convincing analysis for its applicability. Varro correctly points out that the example based approach is limited to certain problem categories and completeness of the transformation cannot be addressed in the framework. A bit more discussion on the characterization of the candidate problem categories would be helpful. This comment does not influence my assessment that the result is significant and receives well deserved attention from the research community. Consequently, thesis 2/1 is a new scientific result.
   b. Theses 2/2 and 2/3 focus on the logic and mechanism of the proposed partial automation process. In my assessment the two points should be combined into a single contribution point since 2/2 is a required setup for the automation process in 2/3. The two points together represent an innovative application of inductive logic programming to a new area, therefore they can be considered as new scientific result.

3. Efficient Execution Strategies for Model Transformations
   The thesis addresses Challenge 3 by the development of three methods to improve scalability. The contributions are practical by nature, but this does not decrease their value, since applicability of MDD methods would not become reality without significant advancement in engineering techniques used in their implementation.
   a. The introduced model-sensitive approach for generating search plans is an innovative idea. The discussion in the thesis makes the case for improvement intuitively clear, however evidence for actual improvements obtained through experiments would be quite useful. Since the proposed method clearly makes sense and improvement can be expected, 3/1 can be considered as novel contribution.
   b. Introduction of incremental pattern matching and casting the solution in the RETE algorithm framework is an excellent idea. Effectiveness of the developed method was demonstrated experimentally. Accordingly, thesis point 3/2 can be accepted as novel scientific contribution.
   c. The concept of model transformation plugins is a nice practical approach and without any doubt it has engineering significance. Generating EJB3 plugins from

transformation specification is a complex problem and the solution described in the thesis is elegant and effective. Accordingly, thesis 3/3 is accepted as novel contribution.

4. Termination Analysis of Model Transformations
   The thesis addresses Challenge 4 by developing sufficient condition for the termination of graph transformations. Given the general undecidability of the termination problem, the sufficient condition derived by using cardinality Petri net abstraction is a valuable achievement.
   a. Abstracting the graph transformation systems with cardinality Petri net represents an innovative insight and can be considered as new result. Closely related to this abstraction is the proof that the derived cardinality Petri net simulates the graph transformation system (Theorem 6.9). I recommend combining 4/1 and 4/2 into a single contribution that is indeed novel and advances the theoretical foundations of graph transformation based model transformations.
   b. The development of a sufficient condition for termination (point 4/3) extends the state-of-the-art and significantly improves Varro's earlier results.

In summary, the Thesis represents a significant body of work that advances the state-of-the-art and state-of-the-practice of model transformations. Publications provide evidence of the acceptance of the results by the scientific community. Through numerous EU program participation, the results are in the process of transitioning to industry and making impact in the future.

Based on the observations above, I found the results sufficient for the MTA Doctor degree and recommend the thesis for public defense.

Janos Sztipanovits, External Member of MTA
E. Bronson Ingram Distinguished Professor of Engineering
Joe B. Wyatt Distinguished University Professor of Engineering
Vanderbilt University
Nashville, TN USA