Multilevel modeling: What's in a level? A position paper

Mira Balaban[1], Igal Khitron[1], Michael Kifer[2], Azzam Maraee[1]

[1] Computer Science Department, Ben-Gurion University of the Negev, ISRAEL
[2] Computer Science Department, Stony Brook University NY, USA
{mira,khitron,mari}@cs.bgu.ac.il, kifer@cs.stonybrook.edu

Abstract. Multi-Level Modeling (MLM) conceptualizes software models as layered architectures of sub-models whose elements are inter-related by the *instance-of* relation, which is radically different from the traditional two level representations that consist of a *class model* and an *instance model*, where the former represents the schema and the latter the population. Key to the MLM representation is the notion of a *level*, which is the subject of this position paper. We argue that a theory of MLM must have a distinguished notion of *level*, which should have both a syntactic and a semantic status. Moreover, the notion of a level must represent a class model, and the overall MLM theory must be *compositional*, being composed from the theory of class models.

1 Introduction

Multilevel system modeling (MLM) views the enterprise as a layered collection of models that are inter-related by the *instance-of* (or *membership*) relation among objects and classes. The grounds for this approach are both philosophical and pragmatic. On philosophical grounds, researchers have been arguing that faithful modeling of real world domains cannot be restricted by the standard two-layer architecture of the OMG meta-modeling approach. They claim that natural domains have multiple levels of classification, and an artificial restriction to two layers yields models that are too weak [3, 14]. On pragmatic grounds, researchers have argued that a multilevel architecture of models simplifies the management and evolution of complex software [4, 10, 12].

Classes and associations in adjacent layers in the multilevel architecture can be related by the standard *instance-of* relationship, while within each layer, classes are organized via subclass and general association relationships. A multilevel model consists of multiple, possibly overlapping, *ontological dimension*s, each being a multilevel organization of classes and associations, as discussed above. A *potency* mechanism, introduced in [2, 11, 19], is a kind of an inheritance mechanism between levels.

Efforts to formulate multilevel modeling has led to a variety of approaches, including a *category-theory-based* framework [19], a *UML/OCL formulation* [13], multilevel objects and relationships [17, 18], an axiomatic specification of multilevel types [9], a set-theoretic formulation [5], and more. Most approaches

have some form of an *instance-of* relation that stands for set membership, and of *subtyping* relationship that stands for set inclusion.[3]

In [1] the authors claim that a multilevel framework can be either *level blind*, i.e., enable multilevel arbitrary classification structures, as in Telos [16], F-Logic [15] and FOML [6], or *level adjuvant*, i.e., recognize the utility of levels, as in the aforementioned frameworks. In these approaches, a *level* is a numeric attribute defined for all elements in a multilevel model, which enables layering elements based on their level value. Yet, in most approaches the concept of a *level* has no independent syntactic or semantic status.

The main theme of this position paper is the claim that a multilevel theory must have a distinguished notion of *level* with both a syntactic and a semantic status. Moreover, levels should reflect the traditional notion of a *class model*, and the overall theory should be *compositional* with respect to class models, i.e., be constructed—syntactically and semantically—from the theory of class models.

2 Multilevel Formalisms

In this section we analyze several published approaches to MLM with respect to the following three questions: (1) Does the formalism include a clear and well-defined syntax and semantics for multilevel models (the latter meaning well-defined notion of legal instances)? (2) Does the formalism include a distinguished syntactic and semantic notion of a *level*? (3) Is the formalism *compositional* with respect to the class model language?

The category-theory-based formulation of [19]. This approach formalizes the MLM approach of [2, 11] using category theory. It supports a syntactic definition of layered class models, extended with a rich *potency* mechanism that enables cross-level constraints and cross-level instantiation (termed here *single potency*). The paper does not show how subtyping and standard class model constraints, like association classes, GS-constraints, aggregation/composition and inter-association constraints are handled.

With respect to the above three questions: (1) this formalism provides a formal definition of multilevel models, with deep constraints and cross-level instantiation. It is unclear, however, what is the semantics of a multilevel model, i.e., what is the set of legal instances. For example, what is a legal population of a model at some level, say, L3? How it is related, if at all, to the syntactic instances in level L2? (2) This formalism includes a distinguished syntactic notion of *level*: The syntactic elements of a level form a valid instance of the next higher level, that functions as its class model. Yet, the semantic status of a level is left unclear. (3) The formalism is syntactically compositional.

The UML/OCL constraint-based approach of [13]. This paper suggests formulation of MLM using UML/OCL. The idea is to join metamodels of several levels

[3] Some languages that support these relationships are Telos [16] and the executable logic language FOML [6], which is derived from Flogic [15].

into a single model, using a concept of *superstructure*, which is modeled using generalizations, associations and UML/OCL constraints.

With respect to the above three questions: (1) Syntax and semantics are given by UML/OCL; (2) elements are assigned a level, but there is no syntactic unit that combines all elements with a common level value, and assigns it a semantic status; (3) the formalism is not compositional.

The multilevel objects and relationships approach of [17,18]. This approach suggests a representation that is based on constructs that integrate views at different abstraction levels. An overall model consists of interrelated *clabject hierarchies*, in which clabjects are classified into different abstraction levels. The representation is formulated and implemented in F-logic [15].

With respect to the above three questions, the approach is similar to the previous one: (1) syntax and semantics are given by the definition of abstraction hierarchies and the F-Logic translation; (2) elements are assigned to an abstraction level, but there is no syntactic/semantic unit of level; (3) the formalism is not compositional.

The FOL axiomatization of multilevel domains in [9]. This paper presents an axiomatic FOL specification of a domain of types. The domain is structured in linearly ordered levels, starting from an initial level of individuals. A level is a set of all types whose elements belong to a lower level.

Concerning our three questions: (1) the paper deals only with axiomatization of sets of instances at each level and their types; (2) there is a distinguished semantic notion of level; (3) the compositionality issue does not apply, since this approach lacks a concrete language for MLM.

The set-theoretic MLM theory in [5]. The formal executable theory for MLM is based on set-theoretic semantics of class models [7,8]. Class models are combined to form *ontological dimensions*, which are the backbone of a multilevel model. The formal theory has a *direct semantics* for multilevel models, which accounts for complex model interactions, including deep constraints and inference. Legal instances are defined using *Herbrand instances* in which the syntactic symbols comprise the semantic domain. The MLM theory is seamlessly embedded in a provably correct way into the executable logic language FOML [6].

Regarding our three questions, the answers are all positive, owed in part to the distinguished syntactic and semantic status of the concept of a *level*. A level is represented by a *schema*, which is a pair of an *instance model* and a *class model*. The instance model is syntactically included in the class model, using a mapping, called a *Herbrand mediator*, of objects to classes, links to associations, and property renaming. This enables renaming of properties and classes between levels, which is absolute necessity for practical applications of MLM. In addition, the instance model of a level is required to be a *partial instance* of the class model in the next higher level. This partiality is important for proper semantics of potency, but we have no room to expand on this issue here.

The semantics of a multilevel ontological dimension defines a legal instance of a syntactically correct dimension as a sequence of legal instances of the class models in the schemas, where (1) an instance of the class model in level i must include the instance model in level $i-1$ (which is a partial instance of this class model); and (2) satisfy all cross-level constraints in higher level schemas. Therefore, the formalism has a well-defined syntax and semantics, a distinguished notion of *level*, and is compositional with respect to class models.

3 On the Necessity of Level-based Compositional Theory for Multilevel Modeling

Much of the literature on MLM is plagued by a confusion between true multilevel modeling in which levels correspond to class models with each level being a partial instance of the next, and modeling that is simply based on formalisms with higher-order features where classes are uniformly treated as objects that can be members of other (meta)classes (like in F-logic [15], RDF, Telos [16], etc.). With respect to the approaches discussed in Section 2, an example of the former kind of true MLM is [5], while [13] is an example of the latter kind of approaches (which, in our view, are not truly MLM). Approaches in the middle of the spectrum, which attempt to include the concept of a level (with partial or no semantics), are [9, 19]. [17, 18] has a level notion that prevents compositionality with respect to class models.

We claim that MLM is much more than just a class model with higher order features and constraints. It is a collection of *levels* bound together by the *instance-of* relation, where each level is a class model, not just a numeric attribute. Besides the instance-of relations, *inter-level mediation* includes terminology changing mappings that permit renaming of concepts. Thus, levels should be expressed both syntactically and semantically as an integral part of the structure. A compositional[4] MLM theory on top of class models turns MLM into a natural extension of the class model in two ways: (1) a class model is a degenerate case of a simple MLM; (2) a composite multilevel model consists—syntactically and semantically—of class models and the theory of MLM should reuse the theory of class models. Thus, in [5] we define consistency, finite satisfiability and optimization of MLM as natural extensions of their class model variants, and show that intra-layer correctness and optimization can be checked using class model algorithms.

References

1. C. Atkinson, R. Gerbig, and T. Kuehne. Comparing Multi-Level Modeling Approaches. In *MULTI 2014*, 2014.

[4] By *compositionality* we mean that the semantics of MLM consists of the semantics of the class models of the layers, properly stitched together by inter-level relations.

2. C. Atkinson and T. Kühne. The essence of multilevel metamodeling. In *UML*, pages 19–33. Springer, 2001.
3. C. Atkinson and T. Kühne. Rearchitecting the uml infrastructure. *ACM Trans. Model. Comput. Simul.*, 12(4):290–321, Oct. 2002.
4. C. Atkinson and T. Kühne. Reducing accidental complexity in domain models. *Software & Systems Modeling*, 7(3):345–359, 2008.
5. M. Balaban, I. Khitron, M. Kifer, and A. Maraee. Formal executable theory of multilevel modeling. In *CAISE*, 2018.
6. M. Balaban and M. Kifer. Logic-Based Model-Level Software Development with F-OML. In *MoDELS 2011*, 2011.
7. M. Balaban and A. Maraee. Finite Satisfiability of UML Class Diagrams with Constrained Class Hierarchy. *ACM TOSEM*, 22(3):24:1–24:42, 2013.
8. M. Balaban and A. Maraee. UML Class Diagram: Abstract syntax and Semantics. `https://goo.gl/UJzsjb`, 2017.
9. V. A. Carvalho and J. P. A. Almeida. Toward a well-founded theory for multi-level conceptual modeling. *Software & Systems Modeling*, 17(1):205–231, 2018.
10. J. de Lara, E. Guerra, R. Cobos, and J. Moreno-Llorena. Extending deep meta-modelling for practical mde. *The Computer Journal*, 57(1):36–58, 2014.
11. J. de Lara, E. Guerra, and J. S. Cuadrado. Model-driven engineering with domain-specific meta-modelling languages. *SoSyM*, 14(1):429–459, 2013.
12. J. de Lara, E. Guerra, and J. S. Cuadrado. When and how to use multilevel modelling. *ACM Trans. Softw. Eng. Methodol.*, 24(2):12:1–12:46, Dec. 2014.
13. M. Gogolla, M. Sedlmeier, L. Hamann, and F. Hilken. On metamodel superstructures employing UML generalization features. In *MULTI'2014*, 2014.
14. B. Henderson-Sellers. *On the mathematics of modelling, metamodelling, ontologies and modelling languages*. Springer Science & Business Media, 2012.
15. M. Kifer, L. G., and J. Wu. Logical foundations of object-oriented and frame-based languages. *Journal of the ACM*, 42(4):741–843, 1995.
16. J. Mylopoulos, A. Borgida, M. Jarke, and M. Koubarakis. Telos: Representing knowledge about information systems. *ACM TOIS*, 8(4):325–362, 1990.
17. B. Neumayr, K. Grün, and M. Schrefl. Multi-level domain modeling with m-objects and m-relationships. In *APCCM 2009*, 2009.
18. B. Neumayr, C. G. Schuetz, M. A. Jeusfeld, and M. Schrefl. Dual deep modeling: MLM with dual potencies and its formalization in F-Logic. *SoSyM*, pages 233–268, 2018.
19. A. Rossini, J. de Lara, E. Guerra, A. Rutle, and Y. Lamo. A graph transformation-based semantics for deep metamodelling. In *Applications of Graph Transformations with Industrial Relevance*. Springer, 2012.