Formalization of ORM Revisited

Terry Halpin

INTI International University, Malaysia and LogicBlox, Australia
terry.halpin@logicblox.com

Abstract. Fact-oriented modeling approaches such as Object-Role Modeling (ORM) and Natural Language Information Analysis Method (NIAM) enable conceptual information models to be expressed using graphical diagrams that may be assigned formal semantics by mapping them onto sets of logical formulae. Various formalizations for such mappings exist. This paper extends such previous work by providing a new approach to formalizing second generation ORM (ORM 2). We show that the metalevel association between semantic value type and data type must be a mapping relationship rather than a subtyping relationship, and we axiomatize a special representation relationship to support this mapping at the instance level. Our new formalization includes coverage of preferred reference schemes and additional constraints introduced in ORM 2. Other issues examined briefly include the use of finite model theory, sorted logic, and practical choices for implementing certain kinds of logical formulae as constraints or derivation rules.

1 Introduction

For conceptual data modeling, fact-oriented approaches such as Object-Role Modeling (ORM) [11,15] and Natural Language Information Analysis Method (NIAM) [23] differ from Entity Relationship Modeling (ER) [3] and class modeling in the Unified Modeling Language (UML) [22] by uniformly modeling facts as unary or longer relationships that are instances of fact types (e.g. Person smokes, Person was born on Date) instead of modeling only some facts as relationships and others as instances of attributes (e.g. Person.isSmoker, Person.birthdate). This attribute-free approach enables all facts, constraints, and derivation rules to be verbalized naturally in sentences easily understood and validated by nontechnical business users using concrete examples, and promotes semantic stability, since one never needs to remodel existing structures in order to add facts about attributes [9]. ORM's graphical notation for data modeling is also far more expressive than that of industrial ER diagrams or UML class diagrams [15].

To ensure that fact-oriented models and queries are unambiguous and executable, it is essential to formalize them in terms of underlying logics. The first formalization of a fact-oriented approach appeared in 1989 in our doctoral thesis [5], which provided an algorithm to map an extended version of NIAM to predicate logic in which deduction trees (semantic tableaux augmented by natural deduction) were used to test whether a conceptual schema is strongly satisfiable (consistent when populated) and

P. Herrero et al. (Eds.): OTM 2012 Workshops, LNCS 7567, pp. 348–357, 2012.