Information and Software Technology 64 (2015) 36–51

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

In defence of deep modelling

Colin Atkinson [a], Thomas Kühne [b,*]

[a] University of Mannheim, B6, C2.11, Mannheim, Germany
[b] Victoria University of Wellington, P.O. Box 600, Wellington 6140, New Zealand

ARTICLE INFO

Article history:
Received 7 August 2014
Received in revised form 30 March 2015
Accepted 31 March 2015
Available online 6 April 2015

Keywords:
Multi-level modelling
Deep modelling
Metamodelling
Ontological classification
Clabjects

ABSTRACT

Context: Since multi-level modelling emerged as a strategy for leveraging classification levels in conceptual models, there have been discussions about what it entails and how best to support it. Recently, some authors have claimed that the deep modelling approach to multi-level modelling entails paradoxes and significant weaknesses. By drawing upon concepts from speech act theory and foundational ontologies these authors argue that hitherto accepted principles for deep modelling should be abandoned and an alternative approach be adopted instead (Eriksson et al., 2013).
Objective: We investigate the validity of these claims and motivate the need to shift the focus of the debate from philosophical arguments to modelling pragmatics.
Method: We present each of the main objections raised against deep modelling in turn, classify them according to the kinds of arguments put forward, and analyse the cogency of the supporting justification. We furthermore analyse the counter proposal regarding its pragmatic value for modellers.
Results: Most of the criticisms against deep modelling are based on mismatches between the premisses used in published definitions of deep modelling and those used by the authors as the basis of their challenges. Hence, most of the criticisms levelled at deep modelling do not actually apply to deep modelling as defined in the literature. We also explain how the proposed alternative introduces new problems of its own, and evaluate its merits from a pragmatic modelling perspective. Finally, we show how deep modelling is indeed compatible with, and can be founded on, classic work in linguistics and logic.
Conclusions: The inappropriate interpretations of the core principles of deep modelling identified in this article indicate that previous descriptions of them have not had sufficient clarity. We therefore provide further clarification and foundational background material to reduce the chance for future misunderstandings and help establish deep modelling as a solid foundation for multi-level modelling.

1. Introduction

The search for improved modelling infrastructures to supersede the traditional four-layer infrastructure underpinning the UML to support domain modelling and model-driven development has been going on since the first UML specification was published [1]. Over that time a range of multi-level modelling approaches based on different modelling architectures have been proposed, from recursively-nested architectures [2], package-based architectures [3] to minimalist architectures [4]. One of the proposed enhancements that has recently gained attention [5] is the "Orthogonal Classification Architecture" (OCA), which separates domain-oriented "ontological" classification relationships from infrastructure-oriented "linguistic" classification relationships and organises them according to the tenets of strict modelling [6]. The number of tools based on the OCA has steadily grown in recent years and the architecture has been used successfully in numerous industry projects and standardizations efforts [7–14].

Although the OCA has become a widely adopted infrastructure for multi-level modelling, it has also been the subject of significant debate. While proponents of the OCA argue that it reduces accidental complexity in multi-level modelling and allows modellers to concisely describe multiple classification levels that often exist in real world domains [15], critics have argued that it is difficult to reconcile with traditional modelling conventions and makes modelling more difficult to understand, especially when combined with the deep instantiation mechanism. To date, this debate has largely focused on syntactical and pragmatic differences between the proposed approaches for multi-level modelling, since it mainly revolves around different strategies for visualising the model elements and their various properties in a multi-level model. However, in a series of recent publications [16,17,4], and in

* Corresponding author.
E-mail addresses: atkinson@informatik.uni-mannheim.de (C. Atkinson), Thomas.Kuehne@ecs.victoria.ac.nz (T. Kühne).

http://dx.doi.org/10.1016/j.infsof.2015.03.010
0950-5849/

particular [18], critics of deep modelling (i.e., the OCA and deep instantiation) have questioned its semantic soundness at a fundamental level and claimed to have uncovered a number of inconsistencies and paradoxes entailed in its use.

If true this would obviously raise serious questions about the usability of the OCA and would significantly impact on a considerable body of work that depends on the OCA as a sound basis [7–14]. However, on closer analysis it turns out that the claimed paradoxes are not in fact paradoxes at all, but rather the result of either –

(a) not applying the OCA's basic premisses and definitions when evaluating it,
(b) not accepting how elements in the OCA are intended to relate to real world entities, and
(c) questioning fundamental tenets of deep modelling based merely on the observation that they are incompatible with other schools of thought.

The main goal of this article is therefore to shed light on the challenges raised in [18] and explain why the claimed paradoxes do not exist (Sections 2 and 3). A second goal of this article is to provide a more cogent justification for the OCA in order to avoid future incorrect interpretations (Section 4). Finally, the third goal is to replace philosophical arguments about the difference between various schools of thought with pragmatic arguments that focus on a modelling framework's impact on modelling practice. We therefore briefly examine Eriksson et al.'s counter proposal [18] and identify potential weakness of the approach for practical modelling scenarios (Section 5).

2. Validity challenges to the OCA

The orthogonal classification architecture (OCA) is a modelling framework intended to enhance domain modelling with support for multiple classification levels. Designed to improve on the traditional four-layer architecture by the OMG [19], it separates domain-oriented "ontological" classification relationships from infrastructure-oriented "linguistic" classification relationships and organises them according to the tenets of strict modelling [6]. The combined use of the OCA with deep instantiation [20] is often referred to as deep (meta-) modelling [9].

The central objection to the OCA laid out by Eriksson et al. [18] is a set of claimed problems which they characterise as manifestations of a fundamental "*paradox*" inherent to the approach. This "*paradox*", which they refer to as the "*linguistic/ontological metamodelling paradox*", is expressed in terms of their understanding of the OCA shown in Fig. 1. This is an exact reproduction of Fig. 2. from Eriksson et al. [18] which is claimed to be a "*slightly modified*" version of a figure that first appeared in one of the first papers introducing the OCA [6, Fig. 3].

We argue that changing the associations between "Object", "Class", and "Metaclass" in the original diagram into what appear to be "instance-of" arrows is more than a *slight modification* and may in fact be the source of some interpretations of the OCA that are incompatible with the latter's tenets (c.f. Section 2.2). The original use of associations implied that, for example, instances of "Object" are ontological instances of "Class" instances. Eriksson et al.'s modification changes this meaning to implying that, for example, "Object" itself is an ontological instance of "Class". This is a fundamental change in semantics, not just a slight modification. We nevertheless use the overall structure of Fig. 1 in order to refer to the main principles of the OCA. In particular, Fig. 1 shows how the linguistic levels are organised horizontally, with the elements in the left-hand column (i.e., linguistic level L_0, c.f.

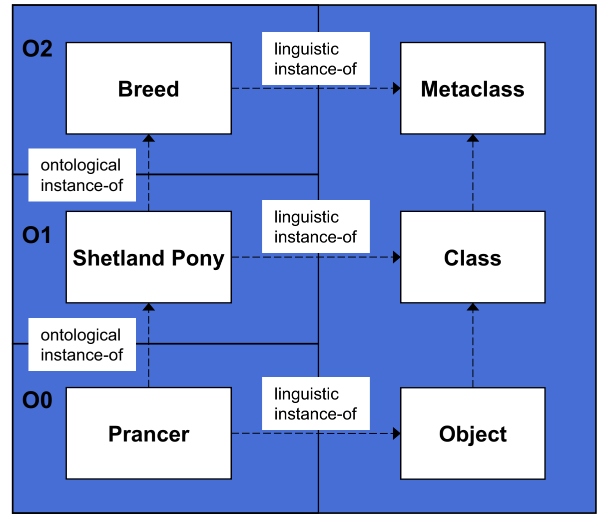

Fig. 1. Reproduction of Fig. 2 from [18].

Fig. 4) being instances of linguistic types in the right-hand column (i.e., linguistic level L_1, c.f. Fig. 4), and the ontological levels are organised vertically (within L_0) with elements in an ontological level (e.g. O_0) being instances of elements in the ontological level immediately above them (e.g. O_1).

In their article [18], Eriksson et al. basically advance four fundamental challenges to the validity of the OCA, relating to the –

C1 location of domain metatypes,
C2 location of the infrastructure element "Metaclass",
C3 correspondence of model elements to real world entities, and
C4 omission of supertypes.

In order to investigate the validity of these challenges, we present them in the following subsections as they were raised and then evaluate their merits.

2.1. Challenge 1: Location of domain metatypes

2.1.1. Claimed problem

Eriksson et al. claim that the OCA gives rise to a paradox regarding the location of domain metatypes. In the context of Fig. 1 they identify the following problem with the way the OCA handles the domain metatype "Breed" –

> *...a class in a (say UML) model (here Shetland Pony) is argued to be an instance of another class (here Breed) that is therefore "meta" to the first class (Shetland Pony) but, paradoxically, cannot be considered to be part of the underpinning metamodel.* [18, p. 2101]

By "*underpinning metamodel*" Eriksson et al. are referring to an M_2 level as it occurs in the OMG's four-layer architecture that would contain at least "Object" and "Class" of Fig. 1 (c.f. Fig. 2). They furthermore explain that –

> *...the Breed class is a metatype with respect to the Shetland Pony class. But clearly such a Breed metatype could not be expected to be part of the M2 layer defining the UML [c.f. Fig. 2], i.e. one would not expect nor wish to find a Breed class alongside a Class class in an M2-level metamodel of a general purpose modelling language like the UML.* [18, p. 2101]

– assuming that as a type for "Shetland Pony", the type "Breed" must be located in the same level that contains the language definition (here the UML's M_2 level).