Re-engineering the ISO 15926 Data Model: A Multi-level Metamodel Perspective

Andreas Jordan, Matt Selway, Georg Grossmann, Wolfgang Mayer, and Markus Stumptner

University of South Australia
School of Information Technology and Mathematical Sciences
Adelaide, Australia
{andreas.jordan,matt.selway}@mymail.unisa.edu.au
{georg.grossmann,wolfgang.mayer,markus.stumptner}@unisa.edu.au

Abstract. The ISO 15926 standard was developed to facilitate the integration of life-cycle data of process plants. The core of the standard is a highly generic and extensible data model trying to capture a holistic view of the world. We investigated the standard from a software modelling point of view and identified some challenges in terminology, circular definitions and inconsistencies in relationships during the mapping from concepts specified in the standard to an object-oriented model. This makes the standard difficult to understand and more challenging to implement. In this paper we look at mapping the ISO 15926 data model to a *multilevel* metamodel, and aim to formalise critical aspects of the data model which will simplify the model and ease the adoption process.

Keywords: Conceptual modelling, multilevel modelling, metamodel engineering.

1 Introduction

ISO 15926 was developed to capture information that is frequently exchanged by organisations in the process-driven industry. Organisations that share data across the life-cycle of assets spanning design, engineering, operations and maintenance require an infrastructure for interoperabiliy and hand-over of data in an automated fashion. At the heart of the ISO standard is a generic data model (ISO 15926-2) such that any organisation within the process-driven industry can adopt and use. Combined with a reference data library (ISO 15926-4) and a set of initial templates (ISO 15926-7) to facilitate intended use enables information exchange at the *semantic* level.

However, in an effort to make the data model sufficiently generic to be suitable for adoption by a diverse range of organisations across the process-driven industry, the model itself has become modelled in an unorthodox fashion from an ontology engineering perspective because the specification of concepts and their relations do not follow a formal ontology construction methodology [1,2].

The data model contains a number of significant issues from a software modelling point of view. Many of these issues stem from the lack of accountability

A.R. Lomuscio et al. (Eds.): ICSOC 2013 Workshops, LNCS 8377, pp. 248–255, 2014.

in the use of terms such as 'instance', 'entity', 'object' and 'represent' which are used differently by different communities [3]. Others are caused by a tendency to preemptively overspecialise parts of the data model. Moreover, despite the standard purporting to enable information exchange at the semantic level, its documentation illustrating its intended use places constraints on the various classes but these constraints are missing from the data model, which can result in modelling lifecycle information in a way not intended.

In this paper our objective is to map the data model of ISO 15926 into a *multilevel data model* which we refer to as *target model* in the paper. By enabling metalevels of representation we are able to better represent the intended meaning by simpler and consistent naming conventions, and adhering to conventional ontological theories of roles, representation and mereology. Through this re-engineering process we aim to simplify the mapping process for domain experts.

The presence of concept names including terms like *Class_of_Class*, *Class* and *Individual* intuitively suggests a minimum of three metalevels could be constructed. Therefore, the two levels of instantiation as made available by the Object Management Group's (OMG) Meta Object Facility (MOF) will not suffice as we need to support modelling "ontological" classification across more than one type/instance level. For this we employ the notion of multilevel metamodelling such as has been proposed by Atkinson et al [4].

In previous work we investigated the principal application of a multi-level modelling approach which revealed some of the challenges when applying an object-oriented modelling approach on ISO 15926 [5]. This paper goes beyond previous work and focuses on a more detailed discussion from an ontological perspective, in particular on the specifications of *representation*, *part-whole relationships* and *roles*.

Section 2 discusses related work, Section 3 discusses some key aspects of the ISO 15926 data model in more detail, Section 4 briefly discusses some of the key notions of multilevel metamodelling outlining our ideas for re-engineering the flat data model into a multilevel model followed by our conclusion and further work in Section 5.

2 Related Work

Our work on transforming the flat data model into a multilevel model is based on model-based transformation. We use rules to extract/derive a model that allows more than just the two metamodel levels offered by the MOF framework. This is motivated by the ISO 15926 data model which contains a number of terms such as *Class_of_X* and *Class_of_Class_of_X* which suggests the relationship between these classes represent *ontological instantiation*. In multilevel modelling terms, three ontological model levels would be needed to represent the ISO 15926 data model.

In [4], Atkinson et al introduces a model element termed "clabject" to represent the dual nature of model elements that possess properties of both a 'class' with respect to model elements in the model level below and an 'instance' with respect to model elements in the above model level. The work of De Lara et al in [6] aims to move beyond the limitations of two meta-modelling levels by introducing a framework called MetaDepth that similarly provides an implementable alternative to model-based transformations. Although using similar concepts as [4], the framework extends the *potency* concept to constraints. Additional related work also includes that of Gonzalez-Perez et al [7] whose use of so-called powertypes provide a mechanism to extend the influence of model elements to beyond their immediate model-level. While this approach adopts the concept of *clabject*, the approach differs from [4] and [6] in that the enabling factor for providing a level-agnostic modelling approach employs the powertype pattern as described by Odell in [8].

In addition to the multilevel modelling aspects of this work, we investigate alternative conceptual elements to construct an ontology which incorporates a more detailed ontological theory of roles, representation and mereology than what is currently defined in ISO 15926-2. This is motivated by the need to enable mapping at the semantic level between other standards in the process-driven industry (e.g. MIMOSA's OSA-EAI[1]) and ISO 15926.

In the area of computer science, research into roles began as early as 1977 with Bachman et al's paper (see [9]). The advent of the semantic web has also seen increased interest in developing a robust theory of roles of which a number of contributions have been made, e.g. see [10,11,12]. Mizoguchi's theory of roles introduces a number of additional concepts, namely "Role Holder" (also referred to as a qua-individual in [11]), "Role Concept" and "Role Player" where the Role Holder is a composition of the Role Concept and the Role Player.

3 Discussion on ISO 15926-2 Concepts

Part 2 of the ISO 15926 standard describes the data model comprising some 201 concepts and forms the core of the standard. It provides a generic data model for the representation of life-cycle information[5]. In this section we analyse a number of concepts from the data model, discuss their ontological nature and how an alternate representation can result in a more understandable ontology.

3.1 Modelling in 3D vs 4D

A conceptual model based on a 3D view of the world is fundamentally different to modelling in 4D. One of the most important distinctions is recognising what constitutes *identity* of an *object* [13]. Considered more in-line with a common-sense understanding of the world, the 3D view considers the three spatial dimensions separately from time, and recognises objects as having identity. In contrast, a 4D view treats time as a fourth dimension. The identity of an object is its trajectory through space-time. An example is a person changing as they age. In a

[1] `http://www.mimosa.org/`

3D world-view, we accept that the person changes but their identity does not. In a 4D world-view, the temporal part of a person at time t_1 is not the same as the temporal part of the person at time t_2. In this world-view the identity of the person would need to be determined by summing the temporal parts of the person, summarised by the expression that the person's identity is determined by its "spatio-temporal envelope". While elegant to express in abstract terms, this does not provide an effective way to compute or reference identity.

A challenge in this work is the fact that ISO 15926 is modelled on the 4D world view. While the 4D approach seems ideal for modelling the lifecycle of assets, in terms of implementation and practicality, it becomes challenging to minimize the complexity of queries relating to identity of objects. These types of queries would not be possible using OWL or formulated as a SPARQL query but would necessitate implementation in either a procedural or declarative language. Moreover it makes understanding, applying and modelling in ISO 15926 more complex, particularly when mapping 3D-based standards to it. Therefore, our target model is based on the 3D world view.

One of the ways this impacts on the 3D model, is the handling of the concept *Possible_Individual* and its subtypes. This concept is defined as "A "thing" that exists in space and time. This includes "things" which are *imaginary* or *possibly exist* in the past, present or future." [14]

The subtypes of *Possible_Individual* include *Physical_Object*, *Event*, *Period_in_Time* and *Point_in_Time*. These subtypes are treated the same in ISO 15926 due to the 4D world view. However, in the 3D world view, *Physical Objects* and *Events* must be treated differently. Furthermore, in ISO 15926 a *Point_in_Time* is an *Event* and those "events that are not points in time are spatial parts of a *Point_in_Time*, defining the time of the event" [14].

An *Event* and a *Point_in_Time* are linked via a part-whole relationship where the *whole* is the *Point_in_Time* an *Event* occurs. By adopting a 3D world view in our multilevel model, we separate events and temporal concepts such that we can use the more intuitive notion of an event occurring at a certain point in time.

3.2 Representation of Concepts in the Real World

The specification of *representation* is particularly challenging because specifying the user's intention is not straight forward. Its definition in [14] is given as "A *representation_of_thing* is a relationship that indicates that a *possible_individual* is a sign for a *thing*.". To illustrate its meaning, it's accompanying example is as follows "The relationship between a *nameplate* with its *serial number* and other data, and a particular pressure vessel (*materialized_physical_object*) is an example of *representation_of_thing* that is an *identification*." [14]

According to Mizoguchi et al in [13] he states a representation is only embodied when it becomes a represented thing and consists of two parts, form and content. The previous example can be misleading in that it can be interpreted in different ways: For example, the nameplate itself could be the representation and then be used to identify a pressure vessel. However, the intention of the user might be to use the nameplate only as the medium which holds an identification number rather than representation of the pressure vessel. It is the symbols comprising the serial number that is the representation of the pressure vessel. In ISO 15926-2, *Identification* is defined as a subclass of *Representation_of_Thing*. The example is ambiguous in at least two senses, the first is whether the symbols comprising the serial number are the representation of the pressure vessel or the representation of the *identification* of the pressure vessel. Another ambiguity relates to the inclusion of the terms "other data" mentioned in the example. Does the "other data" form part of the representation of the *Identification* of the pressure vessel or part of the *representation* of the pressure vessel itself? "Other data" could refer to any property of the pressure vessel, e.g., max pressure rating, in which case, the "other data" does not form part of the representation of either the identification of the pressure vessel nor the pressure vessel itself. To disambiguate these types of issues ISO 15926-2 needs to be supplemented with a more formal notion of representation.

3.3 Mereology - Part/Whole Relations

In order to adequately represent the different interpretations of mereological relations, it is necessary to first distinguish between the different types that exist. Winston et al in [15] identifies six distinct kinds of part-whole relations:

- Component/Integral Object E.g./ handle-cup
- Member/Collection E.g. tree-forest
- Portion/Mass E.g. slice-pie
- Stuff/Object E.g. steel-bike
- Feature/Activity E.g. paying-shopping
- Place/Area E.g. oasis-desert

Three key characteristics are used to distinguish each type of part-whole relation. They are *functional* roles such as 'an impeller is part-of a pump', the *similarity* of the parts with respect to the whole such as 'a molecule of water is a part of water' and lastly whether the parts are *separable* from the whole. ISO 15926-2 also contains mereological relations and we apply the criteria outlined in [15] to determine which category ISO 15926's part-whole relations belong to.

Composition_of_Individual is the most abstract part-whole relation. We argue that it fits the "Member/Collection" relation type as no arrangement between its members is implied and therefore it does not satisfy the functional criteria. Since both part and whole attributes are of type *Possible_Individual*, dissimilar objects can be involved in this type of relation and by definition the *Possible_Individuals* involved in the part/whole relation are separable.

However, *Composition_of_Individual* is also a catchall for other types of part-whole relations. An example of *Composition_of_Individual* is that a grain of sand is part of a pile of sand, which is a portion/mass relationship. Therefore, when mapping a specific instance of *Composition_of_Individual* into the target model, the entities that constitute the part and whole must be reasoned over in order to determine the relationships correct classification.

Arrangement_of_Individual is a specialisation of the concept 'Composition_of_Individual' that restricts the range of the 'whole' to an *Arranged_Individual*, which is defined in [14] as "A possible_individual that has parts that play distinct roles with respect to the whole."

Therefore, we argue that *Arrangement_of_Individual* be classified as a component/integral object relationship. By classifying the way in which part/whole relationships are utilised in ISO 15926 we can better support the mapping of other standards to ISO 15926 through our target model.

Our target model involves representation in a 3D model with time and so does not consider temporal events to contribute to the identity of an object. Instead we employ Mizoguchi's approach by treating a continuant as a role in the context of a process[13].

3.4 Roles

A generally acceptable informal definition of a *role* is an entity that is played by another entity in some context. From a pragmatic perspective we believe the role theory of Mizoguchi is suitable to implement our domain ontology intended to be used to map to ISO 15926-2. When trying to define the characteristics of roles, the differing theories generally agree on a number of fundamental characteristics.

- Rigidity i.e. whether a role is essential/non-essential to all its instances.
- Externally founded i.e. roles require external concepts to define them
- Dynamicity i.e. entities can stop and start playing one or more roles

We suggest to adopt the theory of roles from Mizoguchi[13] and redefine the role-related concepts in ISO 15926 based on this theory. The goal of this change is to provide a more formal, robust and intuitive framework that appeals to a commonsense understanding of roles and remove confusing terminology such as *Class_Of_Possible_Role_And_Domain*. ISO 15926-2 contains a number of types that represent roles. These are given first-class status in the model where five are specialisations of the entity *Class* and the remainder are reified relationships. However ISO 15926 does not give detailed semantics for their intended use. Further to the issue of comprehensibility the term *role* is used interchangeably to refer to different *kinds* of roles whose semantics are quite different (see Table 1). ISO 15926-2's definition of role is loosely analogous to Mizoguchi's *role concept* in [10], however this is where the similarity ends. Although ISO 15926 provides documentation on intended use, the data model does not adequately contain the necessary semantics/constraints to properly enforce the use of roles. Therefore, we believe the data model would benefit by introducing a more robust theory of roles such as that by Mizoguchi in [10].

4 Multilevel Modelling

Research addressing a number of limitations to the UML began as early as 1997 and has continued through to the present day, e.g. see [16,4,6,7]. The major

Table 1. Role *Kinds* used in ISO 15926

Role Kind	Semantics
UML	Appear at either end of an association between two class objects
Description Logic	Binary relationships which are interpreted as sets of pairs of individuals and permit the establishment of role hierarchies
Mizoguchi	Defines roles as a composition of a role concept and a potential player of the role within a context.
Activity	Describes an ISO 15926 'Role_And_Domain' that occurs in the context of an activity

issues surround UML's instantiation mechanism when needing to model more than two model levels which restrict the ability of classes to influence the semantics of objects past a single model level[4]). Since UML's adoption by the Object Management Group (OMG) in 1997, it has become the standard modelling language. Although the UML has shown significant value in many areas particularly in the field of software engineering, despite its ubiquity a number of limitations have been identified along the way. A key limitation relates to the *instantiation* mechanism which can only carry information concerning attributes and associations across a single level[4]. Proposed frameworks for multilevel modelling supporting more than two instantiation levels have been around for more than a decade (e.g. see [16,6]).

While OWL-DL supports *punning*, it's semantics restrict its ability to enforce two key properties of multilevel modelling, i.e. 'potency' and 'level'. Moreover, its accompanying rule language, the Semantic Web Rule Language (SWRL) does not support rules between classes. For these reasons we are required to use a more expressive language with which to implement our multilevel metamodel.

Of the 201 concepts comprising the ISO 15926 data model, 81 of the concepts are prefixed with either class_of_X or class_of_class_of_X. This seems to imply that there exists a minimum of three *logical* levels of instantiation. The definitions of (most of) these 81 classes seem to also support this view, e.g. consider the three classes listed in Table 2.

Table 2. Logical/Ontological instantiation

Concept	Definition in ISO 15926-2 [14]
relationship	"something that one thing has to do with another"
class_of_relationship	"a class_of_abstract_object whose members are members of relationship"
class_of_class_of_relationship	"a class_of_class whose members are instances of class_of_relationship"

5 Conclusion and Future Work

In this paper we discussed a number of issues concerning the complexities and modelling idiosyncrasies of ISO 15926. We proposed the use of a number of alternative theories covering roles, parthood and representation based on a 3D world view intended to make the data model easier to understand and to implement through re-engineering into a multilevel metamodel. Future work consists of implementing additional rules to discriminate between concepts that represent linguistic classification through the use of structural information of the class hierarchy.

References

1. Guarino, N., Welty, C.A.: Evaluating ontological decisions with OntoClean. Commun. ACM 45(2), 61–65 (2002)
2. Mizoguchi, R.: Tutorial on Ontological Engineering: Part 2: Ontology Development, Tools and Languages. New Generation Comput. 22(1), 61–96 (2003)
3. Smith, B.: Against Idiosyncrasy in Ontology Development. In: Proc. of FOIS 2006, pp. 15–26. IOS Press (2006)
4. Atkinson, C., Kühne, T.: The Essence of Multilevel Metamodeling. In: Gogolla, M., Kobryn, C. (eds.) UML 2001. LNCS, vol. 2185, pp. 19–33. Springer, Heidelberg (2001)
5. Jordan, A., Grossmann, G., Mayer, W., Selway, M., Stumptner, M.: On the Application of Software Modelling Principles on ISO 15926. In: Proc. of MODELS Workshop on Modelling of the Physical World (MOTPW 2012). ACM (2012)
6. de Lara, J., Guerra, E.: Deep meta-modelling with metaDepth. In: Vitek, J. (ed.) TOOLS 2010. LNCS, vol. 6141, pp. 1–20. Springer, Heidelberg (2010)
7. Gonzalez-Perez, C., Henderson-Sellers, B.: A powertype-based metamodelling framework. Software and System Modeling 5(1), 72–90 (2006)
8. Odell, J.J.: Power types. Journal of Object-Oriented Programming 7(2), 8 (1994)
9. Bachman, C.W., Daya, M.: The Role Concept in Data Models. In: Proc. of VLDB 1977, pp. 464–476. IEEE (1977)
10. Mizoguchi, R., Kozaki, K., Kitamura, Y.: Ontological Analyses of Roles. In: Proc. of FedCSIS 2012, pp. 489–496 (2012)
11. Masolo, C., Guizzardi, G., Vieu, L., Bottazzi, E., Ferrario, R.: Relational roles and qua-individuals. In: Proc. of AAAI Fall Symposium on Roles, an Interdisciplinary Perspective, pp. 103–112. AAAI Press (2005)
12. Loebe, F.: Abstract vs. social roles - towards a general theoretical account of roles. Applied Ontology 2(2), 127–158 (2007)
13. Mizoguchi, R.: Tutorial on Ontological Engineering: Part 3: Advanced Course of Ontological Engineering. New Generation Comput. 22(2), 193–220 (2004)
14. I.S.O.: ISO 15926– Part 2: Data Model (2003)
15. Winston, M.E., Chaffin, R., Herrmann, D.: A Taxonomy of Part-Whole Relations. Cognitive Science 11(4), 417–444 (1987)
16. Atkinson, C.: Meta-Modeling for Distributed Object Environments. In: Proc. of EDOC 1997. IEEE (1997)