Re-engineering the ISO 15926 Data Model: A Multi-Level Metamodel Perspective*,**

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

University of South Australia
School of Information Technology and Mathematical Sciences
Mawson Lakes Blvd, Mawson Lakes, South Australia
joraa001@mymail.unisa.edu.au
{georg.grossmann,wolfgang.mayer,markus.stumptner}@unisa.edu.au
http://kse.cis.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. However the data model is replete with confusing terminology, circular definitions and inconsistencies that make the standard difficult to understand and less likely to be adopted in practice. In this paper we look at re-engineering the ISO 15926 data model as a *multilevel* metamodel, and aim to (re-)formalise critical aspects of the data model which we believe simplifies the model and eases the adoption process.

Keywords: Conceptual modelling, multilevel modelling, metamodel engineering

1 Introduction

ISO 15926 aims to deliver a solution capable of supporting the information representation needs of a broad category of organisations that encompass the process-driven industry. This is in order to support the representation and sharing of life-cycle data across organisations spanning design, engineering, operations and maintenance. At the heart of the 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 between these organisations.

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 overly complex. It is modelled in an unorthodox fashion and from an ontology engineering perspective has fallen victim to a number of classic errors described in literature pertaining to ontology construction methodologies such as found in [7,14].

* The final publication is available at link.springer.com
** DOI: 10.1007/978-3-319-06859-6_22

The data model contains a number of significant issues which further reduce the standard's chances of adoption. Many of these issues stem from the lack of accountability in the use of terms such as 'instance', 'entity', 'object' and 'represent' which are used differently by different communities[17]. Moreover, despite the standard purporting to enable information exchange at the semantic level, its documentation places constraints on the various classes to illustrate its intended use 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 re-engineer the data model of ISO 15926 into a *multilevel data model*. By enabling metalevels of representation we are able to better represent the intended meaning by providing sensible names, 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 (our overarching goal).

The presence of concepts such as "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 [2]. For this we employ the notion of multilevel metamodelling such as has been proposed by Atkinson et al [2].

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 an example of model-based transformation, however we use adhoc techniques to 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 model levels would be needed to represent the ISO 15926 data model.

In [2], Atkinson et al introduces a novel 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 [5] also aims to move beyond the limitations of two meta-modelling levels by introducing a novel framework called MetaDepth that similarly provides an implementable alternative to model-based transformations. Additional related work also includes that of Gonzalez-Perez et al [6] whose use of so-called powertypes that provide a mechanism to extend the influence of model elements to beyond their immediate model-level.

In addition to the multilevel modelling aspects of this work, we investigate alternative conceptual elements to create an ontology which incorporates a more detailed ontological theory of roles, representation and mereology than what is currently defined in ISO 15926-2.

In the area of computer science, research into roles began as early as 1977 with Bachman et al's paper (see [3]). 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 [16], [12], [11]. Mizoguchi's theory of roles introduces a number of additional concepts, namely "Role Holder"(also referred to as a qua-individual in [12]), "Role Concept" and "Role Player" where the Role Holder is a composition of the Role Concept and the Role Player.

3 Ontological Overview of Core Concepts in ISO 15926-2 Data Model

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[10]. 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 [15]. 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 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 t1 is not the same as the temporal part of the person at time t2. In this world-view the identity of the person would need to be determined by summing the temporal parts of the person. From a software-engineering perspective, this makes identifying objects in a 4D view more complex than the 3D view, particularly when using semantic web technologies such as OWL.

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 some procedural/declarative language.

Moreover it makes understanding, applying and modelling in ISO 15926 more complex, particularly when mapping 3D-based standards to it. Therefore, our re-enginneered 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.

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" [9]. An 'Event' and a 'Point_In_Time' are linked via a part-whole relationship where the 'whole' is the 'Point_In_Time' the 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

When it comes to representation, although fundamentally ISO 15926-2's definition is not incorrect, the example illustrating its *intended* use is a cause for concern.

> Definition: A representation_of_thing is a relationship that indicates that a possible_individual is a sign for a thing.

> EXAMPLE 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.

From an ontological perspective, "a representation is not embodied unless it becomes a representated thing"[15]. According to Mizoguchi, "a *representation* is composed of two parts: form and content. The content is hidden and it is a proposition which the author of the representation would like to convey through the representation"[15]. The example is misleading in that it suggests that the nameplate itself is the representation used to identify a pressure vessel which is not the case. In actual fact, the nameplate is not the representation at all, it is simply a representational medium. 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 ambigious in at least two senses, the first is whether the symbols comprising the serial number that is the representation of the pressure vessel or the representation of the *identification* of the pressure vessel? A second 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 mean anything 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 semantic interpretations of mereological relations, it is necessary to first distinguish between the different types that exist. Winston et al in [18] 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 [18] 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 arrangment between its members are implied and therefore does not satisfy the functional criteria. Since both part and whole attributes are of type 'possible_individual' means that dissimilar objects can be involved in this type of relation and by definition the possible_individual's involved in the part/whole relation are separable.

arrangement_of_individual is a specialisation of the concept 'Composition_Of_Individual' by restricting the *range* of the attribute "whole"to types of *arranged_individual*. Here we must provide the definition of an arranged_individual as given by ISO 15926[9]:

> "An arranged_individual is a possible_individual that has parts that play distinct roles with respect to the whole. The qualities of an arranged_individual are distinct from the qualities of its parts."

Our re-engineering involves representation in a 3D model with time and so do 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[15].

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 esential/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

In our ontology, we adopt the theory of roles from Mizoguchi[15] and redefine the role-related concepts in ISO 15926 based on this theory. It is our intention that we 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 different types of roles and 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 define any specific semantics over their intended use. Further to the issue of comprehensibility the term 1 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

Table 1. UML 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

to Mizoguchi's *role concept* in [16], 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 [16].

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 [1,2,5,6]. The major 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[2]). 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[2]. Proposed frameworks for multilevel modelling supporting more than two instantiation levels have been around for more than a decade (e.g. see [1,5]). 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. 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
relationship	A relationship is something that one thing has to do with another.
class_of_relationship	A class of relationship is a class_of_abstract_object whose members are members of relationship
class_of_class_of_relationship	A class_of_class_of_relationship is 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 idiosynchrasies 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. Atkinson, C.: Meta-Modeling for Distributed Object Environments. In: Proceedings First International Enterprise Distributed Object Computing Workshop (1997)
2. Atkinson, C., Kühne, T.: The essence of multilevel metamodeling. In: UML 2001The Unified Modeling Language. Modeling Languages, Concepts, and Tools, pp. 19-33. Springer Berlin Heidelberg (2001)
3. Bachman, C., Manilal, D.: The role concept in data models. In: Proceedings of the third international conference on Very large data bases. Vol. 3 pp464-476 (1977)
4. Boella, G., van der Torre, L., Verhagen, H.: Roles, an Interdisciplinary Perspective. Applied Ontology. Vol. 2 pp 81-88 (2007)
5. De Lara, J., Guerra, E.: Deep meta-modelling with METADEPTH. J. Objects, Models, Components, Patterns. In: 48th International Conference, TOOLS 2010 1–20 (2010)
6. Gonzalez-Perez, C., Henderson-Sellers, B.:A powertype-based metamodelling framework. Software & Systems Modeling. vol. 5 pp72–90 (2006)
7. Guarino, N, Welty, C.: Evaluating ontological decisions with OntoClean. Comms ACM vol. 45, 61–65 (2002)
8. Int. Stds. Org.: Industrial automation systems and integration – Product data representation and exchange(1984)
9. Int. Stds. Org.:Industrial automation systems and integration Integration of lifecycle data for process plants including oil and gas production facilities Part 2: Data model (2003)
10. Jordan, A., Grossmann, G., Mayer, W., Selway, M., Stumptner, M.: On the application of software modelling principles on ISO 15926. In: Proceedings of the Modelling of the Physical World Workshop, pp. 3:1–3:6. ACM, New York (2012)
11. Loebe, F.:Abstract vs. social roles Towards a general theoretical account of roles. Applied Ontology. Vol. 2 pp127-158 (2007)
12. Masolo, C., Guizzardi, G., Vieu, L., Botazzi, E., Ferrario, R.:Relational roles and qua-individuals. In AAAI Fall Symposium on Roles, an Interdisciplinary Perspective pp 103–112 (2005)
13. Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., Guarino, N.:Social roles and their descriptions. In: 9th International Conference on the Principles of Knowledge Representation and Reasoning (KR2004), pp. 267–277. Canada (2004)
14. Mizoguchi, R.: Tutorial on ontological engineering Part 2: Ontology development, tools and languages. New Generation Computing vol. 22, 61–96 (2004)
15. Mizoguchi, R.: Tutorial on ontological engineering Part 3: Advanced course of ontological engineering. New Generation Computing Vol. 22, 198–220 (2004)
16. Mizoguchi, R., Kozaki, K., Kitamura, Y.:Ontological Analyses of Roles. In: Computer Science and Information Systems. Vol (2012)
17. Smith, B.: Against Idiosyncrasy in Ontology Development. In Proceedings of the fourth International Conference, FOIS 2006, pp. 15–26
18. Winston, M.E., Chaffin, R., Herrmann, D.:A taxonomy of Part-Whole Relations. Cognitive Science. Vol. 11 pp417–4444 (1987)