BIM IN 2007 – ARE WE THERE YET?

Robert Amor, Ying Jiang, Xiaofan Chen
Department of Computer Science, University of Auckland, New Zealand

ABSTRACT: As the prevalence of BIM increases in A/E/C-FM disciplines it is timely to review the standards that are being utilised and how well they are serving the discipline. The analysis presented analyses the most common standard, the IAI's IFC, from a meta-level and asks questions about the evolving model from the viewpoint of metrics for data models as well as a low level analysis of the accuracy and correctness of implementations of the data model interpreters. Metrics applied to the evolving versions of the IFC schema can indicate the trajectory of the schema and profile areas which may be of concern in the maintenance of the schema and applications that have to utilise the schema. Analysis of the approaches to importing and exporting data for design tools, based on the schema, help indicate how market ready the technology really is. Where commercial projects are starting to rely on the standards as a mechanism to reliably transfer semantically correct information there must be guarantees of the accuracy of the data as it is manipulated by these design tools.

KEYWORDS: data model, metrics, translator, data management.

1 INTRODUCTION

There has undoubtedly been progress in the development of BIM for A/E/C-FM industries, especially in the last decade. The IFC standard (IAI 2007) has been adopted by the majority of the large CAD vendors, and in several countries there are strong government-level policies in place to ensure the adoption of this standard. There is also strong support for the use of BIM by many in the industry as reflected in supportive articles appearing in the various trade magazines and Internet commentary sites (e.g., http://www.laiserin.com/ and http://www.aecbytes.com/).

However, there has not been as much research-based critical analysis of the adoption of these standards and their impact on the industry. To help understand where current BIM is taking the industry there needs to be analysis of the evolving schemas and the approaches to the management of the data against these schemas.

One form of analysis is through the measures of the structure of the data model which can be achieved through utilising standard data model metrics. These metrics provide commentary on the strength of a data model and its maintainability over the long term. A second form of analysis is the accuracy and correctness of the design tools interpreters when they interact with the data models. This analysis provides an insight to the problems that will be encountered when actually using the current BIM approaches. In this paper we apply metrics designed for UML class diagrams to the evolving IFC schema to understand its trajectory. We also examine the results of data model manipulation through various commercial CAD systems in order to understand how they maintain the integrity of BIM data.

2 IFC DATA MODEL METRICS

One approach to understanding the strength of a data model and its evolution is to apply metrics to that model. One of the main motivations in applying these metrics is to ensure the maintainability of the data model and software systems which interact with the data model. Some of the main aspects of maintainability are: the ease of understanding of the data model; the amount that the data model changes over time; how easy it is to test the correctness of the data model; and how to ensure compliance with its specification. The metrics therefore provide viewpoints on the evolving complexity of the model, and in a standard software development project would help drive the refactoring of the data model.

2.1 *Metrics for data models*

There are a wide range of metrics which have been mooted for data models (Piattini et al 2002), but for this work we look at metrics applicable to UML class diagrams as developed by the research community and summarised by Genero et al (2000). Metrics for UML class diagrams will be applicable in this domain as UML class diagrams are analogous to the EXPRESS-G diagrams developed for the IFC efforts.

The metrics calculated and the motivation for applying them (Genero et al 2000) are as follows:

- Overall number of classes: Provides a simple measure of the global complexity of the data model. An indicator of the coverage provided by a model.
- Depth of the inheritance tree: Provides an indication of how much impact changes in the model are likely to have. As the average depth increases the impact of

changes near the root of the tree have a greater impact on the model. If this number increases over versions of a model it indicates a model and software which will require greater work to maintain under change.
- Number of children: Provides a direct measure of the reuse being made in the model but as the average number grows also indicates greater difficulties in maintaining the model and code as changes in the parent impact a greater number of child classes.
- Number of associations in a class: Provides a measure of the reusability of a class, where in general the greater the number of associations the less reusability exists for that class. The greater the number of associations the greater the complexity of a class and the greater the difficulty to understand the class definition correctly.
- Number of dependencies in: Provides a measure of how many classes rely on a particular class. The greater the number of dependencies in the greater the complexity of the model.

In addition to these measures we introduce measures specifically related to the property set construct found in recent versions of the IFC data model. Of interest are the following:
- Overall number of property sets defined: Provides a simple measure of the complexity of the data model.
- Ratio between property set data and class data: Provides an indicator of the balance between the formally defined data model and the informally defined model.

2.2 *Application of metrics to the IFC schema*

In order to calculate the metrics described in section 2.1 the EvaSys system (Ma et al 2006) was extended to report on the structure of the schema loaded alongside the data files it analyses. This analysis was run on the long-form version a set of final versions of the IFC schema which have been published by the IAI (2007) over the last decade. The results of which are presented in the following tables.

Table 1. Simple measures of the IFC schema.

IFC version	**1.5.0**	**1.5.1**	**2.0**	**2.x**	**2.x.2**	**2.x.2 Add1**	**2.x.3**
Classes	184	186	290	370	623	629	653
- Root classes	33	29	46	70	104	105	101
- Inherited classes	151	157	244	300	519	524	552
Attributes	413	432	818	965	1268	1279	1320
- Optional	104	120	321	507	562	568	599
- Required	309	312	497	458	706	711	721
Average attributes per class	2.24	2.32	2.82	2.61	2.04	2.03	2.02

Table 1 indicates that the complexity of the IFC data model has increased significantly over the major versions of the schema. There exists over three times as many classes as in the original IFC model. This is a strong indicator for the growth in the domain coverage of the IFC data model. An issue for the community now becomes the effort involved in understanding the complete IFC model in order to implement it correctly. The likelihood of schema errors increases with model complexity as does the likelihood of implementation errors. It is interesting to note the decline in number of base attributes (defined as attributes with a simple type) per class which is an indicator of the amount of information defined per class. An indication that information is being inherited through the class hierarchy effectively but also that there is less new information being added in the new specialised classes.

Table 2. Associations in the IFC schema.

IFC version	**1.5.0**	**1.5.1**	**2.0**	**2.x**	**2.x.2**	**2.x.2 Add1**	**2.x.3**
Average depth of inheritance tree	2.18	2.14	2.68	2.69	3.26	3.27	3.49
Maximum depth of inheritance tree	5	5	8	7	8	8	8
Average number of children	0.82	0.84	0.84	0.81	0.83	0.83	0.85
Maximum number of children	16	19	27	16	26	26	28
Average number of associations	3.88	4.02	4.80	4.33	4.14	4.15	4.43
Average number of dependencies in	3.44	3.51	4.58	4.01	4.14	4.15	4.43

Table 2 also indicates the increasing complexity of the IFC schema. The average depth of the inheritance tree has increased by over one and is still increasing slowly. This means that changes to the model will, on average, affect a greater number of classes in an implementation of the model, and hence from a maintenance point of view indicates greater effort to maintain IFC compliant software under change. The average number of direct children per class is not increasing, which is a good sign from a maintenance and development point of view. The average number of associations and in-dependencies is increasing slowly, also reflecting a more complex model and the likelihood of difficulties in understanding the full schema.

Table 3. Property sets in IFC.

IFC version	**2.x.2**	**2.x.2 Add1**	**2.x.3**
Unique properties	1527	1797	1791
Maximum properties for a class	136	151	151
Average properties per class	5.15	6.87	6.65
Attributes	1268	1279	1320
Average attributes per class	2.04	2.03	2.02
Ratio of attributes to properties per class	0.40	0.30	0.30

The data on properties is a little sparse for trends to be emerging as yet. However, the ration between attributes and properties per class gives a very clear indication as to where information about an object is likely to be found. The class with 151 properties (IfcPerformanceHistory) is defined in the IFC schema with a very small number of attributes. So here, and for many similar classes, the question might be asked as to why the published properties are not incorporated into the specification of the class (as optional attributes if need be).

3 DESIGN TOOLS' IFC TRANSLATORS

The ability of design tools (CAD, simulation, etc) to correctly handle IFC data is another major aspect of the maturity of the BIM marketplace and the confidence of industry to work with these data models. The major CAD tools and a growing number of simulation and analysis tools are providing IFC import and export. There is also a growing library of case studies of the use of IFC for real-life projects (IAI 2007). However, the analysis in section 2 would indicate that this will be a major development task for those prepared to handle IFC data in regards to having to manipulate a schema of great complexity.

In the last year a range of researchers have reported on the manipulation of data related to IFC. Lipman (2006) has investigated the possibilities and difficulties of mapping between CIMsteel and IFC representations. Pazlar and Turk (2006) have looked in particular at geometric data exchange utilising the IFC data model. Ma et al (2006) and Amor and Ma (2006) have focused on issues in the preservation of the semantics of IFC data when mapped through a design tool's internal representation. All of these studies indicate that the translation of IFC data through a design tool is lossy. This recent work, focused on the IFC data model, reflects research undertaken over the last two decades which identified a range of data mappings which are impossible to accurately implement and which raised issues as to what could be achieved with object-based data transformations (Banerjee et al 1987, Lerner and Habermann 1990, Eastman 1992, Zicari 1992, Amor 1997, Atkinson et al 2000, Amor and Faraj 2001, and Grundy et al 2004).

One of the shortcomings of the work undertaken in the last year has been the high level at which the analysis was undertaken, providing a summary of the issues that exist (e.g., lost objects), but not enough detail to understand exactly where issues were prevalent. The analysis tool reported by Ma et al (2006) has been further extended to provide very detailed reports on the differences between two IFC data files. The results of this detailed analysis are presented below.

3.1 *Testing process for IFC data files*

The testing process has been structured to ensure repeatability and to utilise standard data models and translators where possible. This is a little problematic in that there are no repositories of standard data files to be used consistently by researchers working in this area and the translators for the various design tools are constantly changing.

To ensure that data files being tested are quality files and correctly specified we have attempted to source files directly from the IAI where possible. For IFC 2x3 this has been eased by the open publication of the initial conformance test files for use by all tool developers (IAI 2007). For earlier versions of the IFC we have sourced test files from the various IAI road shows as exemplars public promoted by the IAI. We have also sourced a set of IFC 2x2 test files crafted for a Masters in Engineering project in Fire Engineering (Dimyadi 2006).

To ensure that the IFC import and export process is well crafted we have chosen to utilise the commercial CAD tools as our design tools. These tools have the longest history of IFC translator development and are all certified by the IAI as conformant to the IFC specification.

To ensure that we are focusing on a process that can be related to a single design tool we have structured the test as follows. For every IFC data file in our repository we have imported it directly into the CAD tool and then immediately exported the model as an IFC file again. No manipulations were made to the model after import into the CAD tool. When exporting the model we chose the default settings for the IFC export offered by each of the CAD tools.

3.2 *Results of round trip translation of IFC data files*

As reported in previous research papers (Pazlar and Turk 2006, Ma et al 2006, and Amor and Ma 2006) there are significant differences between the original IFC data file and the file after being exported from the CAD package. Table 4 gives an indication of the level of differences which occur in these data files for one typical CAD system. In almost all cases there are less objects with preserved GUIDs than with changed GUIDs. The number of objects in the two files can be vastly different, and in many cases property set information is not preserved across an import and export.

Table 4. IFC 2x3 test file results for one CAD system.

IFC 2x3 Test Files	**Objects with GUIDs**		**Objects with same GUIDs**	**Object with no GUIDs**		**Objects with property sets**	
	In	**Out**		**In**	**Out**	**In**	**Out**
beam_profile_basic_rev_1	209	219	110	1130	2566	12	12
beam_profile_para_ac_1	496	252	126	2634	3145	122	0
brep_beams_opening_ben_1	62	62	7	1212	5435	0	0
col_brep_opening_ben_1	24	24	8	4639	1558	0	0
col_profile_clip_ben_1	14	14	7	123	267	0	0
columns_basic_all_1	22	14	7	124	595	4	0
curtain_wall_basic_rev_1	137	119	60	1009	1288	1	0
DoorOperationsPlacementInsideWall_rev_1	167	183	68	639	571	14	18
doors_explicit_geom_all_1	26	32	7	5694	4239	0	3
extruded_beam_open_tek_1	26	24	6	247	2537	0	0
extruded_slab_openings_all_1	14	14	4	128	110	0	0
mem_profile_basic_tek_1	16	22	11	233	1309	0	0
OpeningsInExtrudedColumns_rev_1	61	53	24	230	609	9	4
railing_brep_ac_1	18	10	5	618	525	2	0
railing_extrusion_tek_1	17	22	11	554	1958	0	0
ramp_geometry_ben_2	10	10	5	6879	4030	0	0
RampAsContainer_rev_1	28	29	16	185	179	3	3
roof_with_openings_ben_1	18	18	5	129	327	0	0
RoofWithGeometry_rev_1	119	111	60	3447	3436	8	8
slab_profile_basic_ac_1	24	12	6	210	226	3	0
slab_recess_tek_1	14	14	5	71	300	0	0
stair_geom_ac_1	12	8	4	4853	4653	1	0
stair_geometry_ben_1	11	11	6	1842	1174	0	0
wall_layers_number_1	56	30	15	646	437	13	0
wall_L-shape_all_1	26	18	7	164	140	4	0
wall_opening_straight_ac_1	97	77	15	2210	2071	17	6
wall_recess_ben_1	42	42	9	402	510	0	0
window_brep_ac_1	63	50	10	2326	2334	11	5
windows_placement_inside_wall_all_1	61	61	13	337	694	5	5

Drilling down to categorise the differences for a particular instance of a translated IFC data file reveals a plethora of relatively minor through to major differences. While the majority of differences are minor as in Table 5, a major concern must be the number of objects whose GUID are not preserved during a translation. While most of the objects with changed GUIDs are not representing physical objects in a building (e.g., IfcRelAssociatesMaterial, IfcRelDefinesByProperties, IfcPropertySet) there are still a small number of changed or dropped objects that do represent physical constructs within the building model (e.g., IfcBeam, IfcCurtainWall).

Table 5. Examples of differences found between translated IFC data files.

Exemplar observed differences	**In**	**Out**
Representational accuracy	4.154093800000022	4.1540938
Instantiated data	Date: 0	Date: 1172711486
Type changes	IFCLENGTHMEASURE(0.)	IFCREAL(0.)
Updated information	Version: 9.1	Version: 9.0
Changed representations	SweptSolid	MappedRepresentation

4 CONCLUSIONS

Applying meta-level analysis to the evolving IFC schema and data files of IFC data provides us with insight into the evolution and status of IFC as a BIM standard. This analysis shows, unsurprisingly that the IFC schema has become more complex over the years as it has been extended to cover a larger segment of the A/E/C-FM domains. While some aspects of this complexity are understandable in a mature model there are measures of the schema indicating complexity which is not necessary. The number of associations and dependencies between classes (including inheritance depth) are manageable by standard modelling and refactoring techniques which can be applied to schema. Working to reduce these numbers will ensure that the complexity and maintainability of the evolving IFC schema will not impact as severely on the community which has to implement the final specifications.

Analysis of the translators used in commercial CAD systems indicates that there are a range of serious issues which need to be addressed in the certification process for IFC as well as the accuracy of existing translators. Semantic integrity of the data represented in the IFC data model has to be maintained as the model moves between design tools. Analysis to date indicates that this does not happen in many circumstances.

REFERENCES

Amor, R. (1997). A Generalised Framework for the Design and Construction of Integrated Design Systems, PhD thesis, Department of Computer Science, University of Auckland, Auckland, New Zealand, 350 pp.

Amor, R. and Faraj, I. (2001). Misconceptions about Integrated Project Databases, ITcon journal, ISSN 1400-6529, 6, http://www.itcon.org/2001/5/, 57-68.

Amor, R. and Ma, H. (2006). Preservation of Meaning in Mapped IFCs, Proceedings of EC-PPM 2006, Valencia, Spain, 13-15 September, 233-236.

Atkinson, M.P., Dmitriev, M., Hamilton, C. and Printezis, T. (2000). Scalable and Recoverable Implementation of Object Evolution for the PJama1 Platform, Persistent Object Systems, 9th International Workshop, POS-9, Lillehammer, Norway, 6-8 September, 292-314.

Banerjee, J., Kim, W., Kim, H. and Korth, H. (1987). Semantics and Implementation of Schema Evolution in Object-Oriented Databases, Proceedings of the 1987 ACM SIGMOD international conference on Management of data. San Francisco, USA, 311-322.

Dimyadi, J.A.W. (2006). Generating FDS Fire Simulation Input using IFC-based Building Information Model, Research project report, Master of Engineering in Fire Engineering, University of Canterbury, New Zealand.

Eastman, C.M. (1992). A data model analysis of modularity and extensibility in building databases, Building and Environment, 27(2), 135-148.

Genero, M., Oiattini, M. and Calero, C. (2000). Early Measures for UML class diagrams, L'Objet, Vol 6, No 4.

Grundy, J.C., Hosking, J.G., Amor, R.W., Mugridge, W.B. & Li, Y. (2004). Domain-Specific Visual Languages for Specifying and Generating Data Mapping Systems, Journal of Visual Languages and Computing. 15(3-4), 243-263.

IAI (2007). International Alliance for Interoperability, web site last accessed 15/3/2007, http://www.iai-international.org/.

Lerner, B.S. and Habermann, A.N. (1990). Beyond schema evolution to database reorganization, Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA), Ottawa, Canada, October, 67-76.

Lipman, R.R. (2006). Mapping Between the CIMsteel Integration Standards (CIS/2) and Industry Foundation Classes (IFC) Product Model for Structural Steel, Conference on Computing in Civil and Building Engineering, Montreal, Canada, 14-16 June.

Ma, H., Ha, K.M.E., Chung, C.K.J. and Amor, R. (2006). Testing Semantic Interoperability, Proceedings of the Joint International Conference on Computing and Decision Making in Civil and Building Engineering. Montreal, Canada, 14-16 June.

Pazlar T. and Turk Z. (2006). Analysis of the geometric data exchange using the IFC, Proceedings of EC-PPM 2006, Valencia, Spain, 13-15 September, 165-171.

Piattini, M., Genero, M. and Calero, C. (2002). Data Model Metrics, Handbook of Software Engineering and Knowledge Engineering, Vol 2, ISBN: 981-02-4974-8.

Zicari, R. (1992). A Framework for schema updates in an object-oriented database systems, Morgan Kaufmann Series In Data Management Systems, 146-182.