ANALYSIS OF REFERENCE IFC MODELS

Robert Amor, PhD, trebor@cs.auckland.ac.nz
Department of Computer Science, University of Auckland, Auckland, New Zealand

ABSTRACT

With the establishment of an open repository for IFC models there is now a growing resource of models available for use by all in the community. This repository should help researchers undertake experiments which are able to be compared, validated and replicated by any other researcher in the community. This is commonly undertaken in other research domains, such as medicine. The repository should also be of benefit to those in the industry who wish to test out software which utilise IFC data models by identifying models which are close to the type of building they wish to work with. In order that researchers, or practitioners, can identify the best model for their particular analysis it is necessary to provide significant meta-data about the models, including analyses of the models by various checkers and IFC analyzers. The current status of the repository is surveyed in this paper with statistics on the wide range of data models available. While this indicates that there is good variability in the just over 100 models currently in the repository, it also points to issues in growing the repository to being a comprehensive resource. This problem will exist unless the community are willing to deposit models into the repository as they are created, and to make them freely available to all to utilise.

Keywords: IFC, Model repository, Model analysis, Model metrics.

1. INTRODUCTION

The Open IFC Model Repository (Amor and Dimyadi 2010) has been inaugurated with over 100 IFC models deposited by the author to date. The 100 models currently deposited are drawn from a personal collection of over 200 IFC models from a range of reference projects, as well as from BuildingSMART's support and certification forums. The approach of collecting IFC files from major research projects and the IFC standard's developer is looking to ensure that those models in the repository are of high quality and likely to be well tested for use in demonstrations and the certification processes. The task of collecting these models highlighted the difficulties in ensuring that standard models are retained past the completion of a particular funded project with several identified repositories of IFC models no longer active. The major sources of IFC models that were accessed for the repository include:

- BLIS[1]: the Building Lifecycle Interoperable Software project which completed in 2002 developed a wide range of certification test files as well as an exchange collection for demonstration scenarios.
- DDS[2]: the Data Design System published a range of IFC models associated with the Munkerud house.
- HITOS[3]: models from Statsbygg (The Norwegian Agency of Public Construction and Property) and their project on the Tromso University College (HITOS).
- BuildingSMART[4]: published a suite of certification models for use prior to the formal certification testing phase, as well as a number of models from various demonstration scenarios.
- KIT[5]: the Karlsruhe Institute of Technology's Institute for Applied Computer Science published a number of semantic data models as part of their programme looking at IFC.

[1] http://www.blis-project.org/
[2] http://www.dds-cad.net/
[3] ftp://ftp.buildingsmart.no/pub/ifcfiles/HITOS
[4] http://buildingsmart-tech.org/certification

- LCie[6]: the 'Life Cycle information exchange' project provided a range of IFC models around a medical clinic example.
- Nemetschek[7]: with their Vectorworks BIM in practice series, publish a number of IFC models within the archive.
- NIST[8]: the National Institute of Standards and Technology's Computer Integrated Building Processes Group has a range of IFC models translated from CIS/2 through their freely available translator.

All IFC models entered into the repository have a range of meta-data describing the parameters of the creation of the model as well as an indicative image of the model (see Figure 1). Each model is also put through the IFC File Analyzer (NIST 2011) and the summary information generated from that software tool is linked to every model.

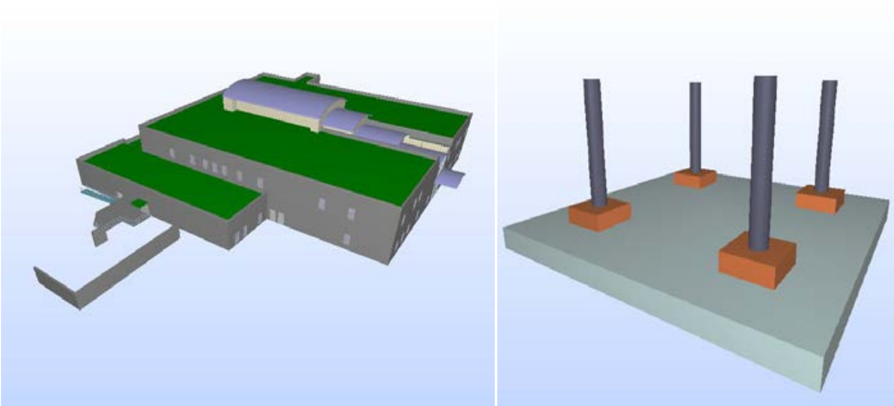

Figure 1: Images of models in the repository.

The repository functionality is under continued development to take into account use cases of those accessing the system, and to establish a usable and intuitive interface. Of particular import in the continued development is the automated generation of analyses whenever a model is deposited, and calculation of a wide range of metrics for every model. Metrics being encoded into the system include those proposed for IFC data models by Amor et al (2007) (which reflect upon standard metrics for the UML model as summarised by Genero et al (2000)) and metrics from Lee et al (2011) which are inspired by a number of earlier projects (e.g., Gielingh 2008, Jeong et al 2009, Kiviniemi et al 2008, and Pazlar and Turk 2008).

2. ANALYSIS OF THE MODEL REPOSITORY

While the IFC models are collected from major projects and the IFC standard's developer who have a significant interest in ensuring that they have robust and well tested models, the fact that a large number of models are now resident in the repository enables analyses to be run to test this assumption. The collection of models also allows a range of analyses to be run to understand the differences that may exist in models of different forms.

The initial analysis run across the repository looks to identify what types of models have been collected and the reported versions of the standard which are represented. Table 1 provides information about the number of models for each IFC version. As can be seen in this table there are a number of models in IFC 2.0 (drawn from the BLIS project) though only a few from the IFC 2.x and 2.x.2 series of releases. As is expected the majority of models are for IFC 2.x.3, the major version used currently by CAD tools and unlikely to be replaced by a new version for a period of years.

[5] http://www.iai.fzk.de/www-extern/index.php?id=1123&L=1
[6] http://www.buildingsmartalliance.org/index.php/projects/activeprojects/140
[7] http://www.nemetschek.eu/
[8] http://cic.nist.gov/vrml/cis2.html

IFC Version	Models
IFC20_LONGFORM	22
IFC2X_FINAL	8
IFC2X2_FINAL	1
IFC2X3	77

Table 1: Number of models by version number.

The models collected reflect the period in which major projects were being run as can be seen in Table 2. The BLIS project produced its models in 2001, many of the 2006 models are the BuildingSMART's published certification files, and the LCie and NIST projects were major contributors to the models which were published in 2010.

Year	Models
2001	22
2002	0
2003	0
2004	0
2005	4
2006	37
2007	5
2008	6
2009	0
2010	34

Table 2: Number of models by year of creation.

While the reported software tool does not seem to be an accurate reflection of a real piece of software, Table 3 shows that a large number of software tools are represented as creators of the models which reside in the repository. Some projects seem to list themselves as the creating tool (e.g., the BLIS project) when it is unlikely that these models were totally hand-crafted by project members.

Software tool	Models
ADT	7
Allplan	8
ArchiCAD	10
BLIS	22
CADstudio	1
EliteCAD	1
ETABS	1
GTSTRUDL	2
IFC	1
IFC Engine	2
NIST	6
Revit Architecture	2
Revit Building	7
Revit MEP	8
Revit Structure	2
SDS/2	8
Tekla Structures	6
TriForma	11
VDI	3

Table 3: Number of models by creating software tool.

One of the aims of the repository is to collect a wide variety, and representative sample, of models for use by researchers and practitioners. An analysis of models in the repository indicates that this is being met at least in terms of the sizes of models collated (see Table 4). From a model with only 54 entities through to a model with close to two million entities there is significant variation in models that can be accessed. The same is true for the size of the files, from 4kb through to over 100Mb.

	Entities	File size (kb)	Entities/kb
Minimum	54	4	13.2
Maximum	1,904,600	109,996	22.8
Average	92,351	5,132	18.0
Std. Dev.	267,259	14,937	2.2

Table 4: Statistics on models in the repository.

The calculation in Table 4 of the number of entities per kilobyte of file size is perhaps not surprising in its low variability. What is perhaps of more interest is the consistency of this ratio no matter how small or large a model is. Figure 2 indicates that there is basically a linear relationship for very small files through to the very largest files for entities per kilobyte of file size.

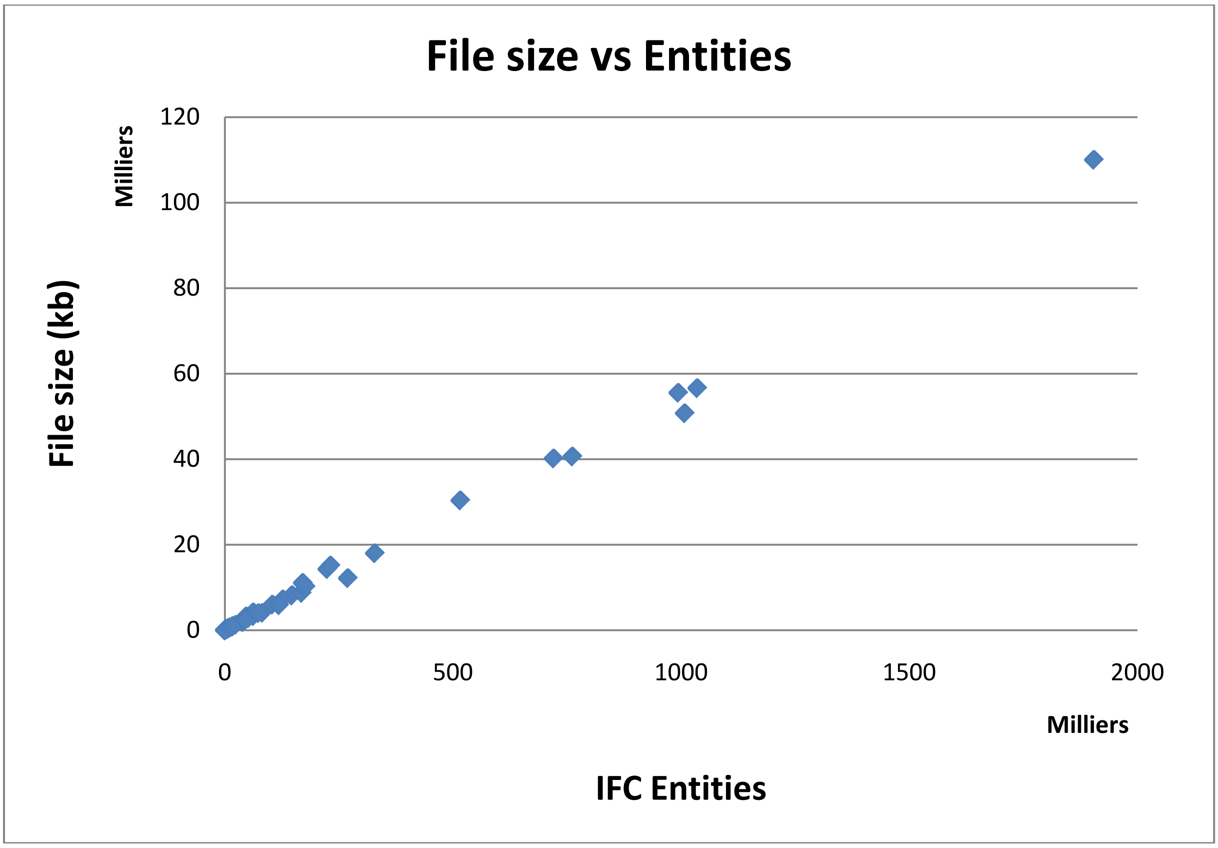

Figure 2: Analysis of trend in file size as the number of entities increases.

The analysis in Figure 3 shows that the density of entities in a file is pretty uniformly distributed around 18 entities per kilobyte.

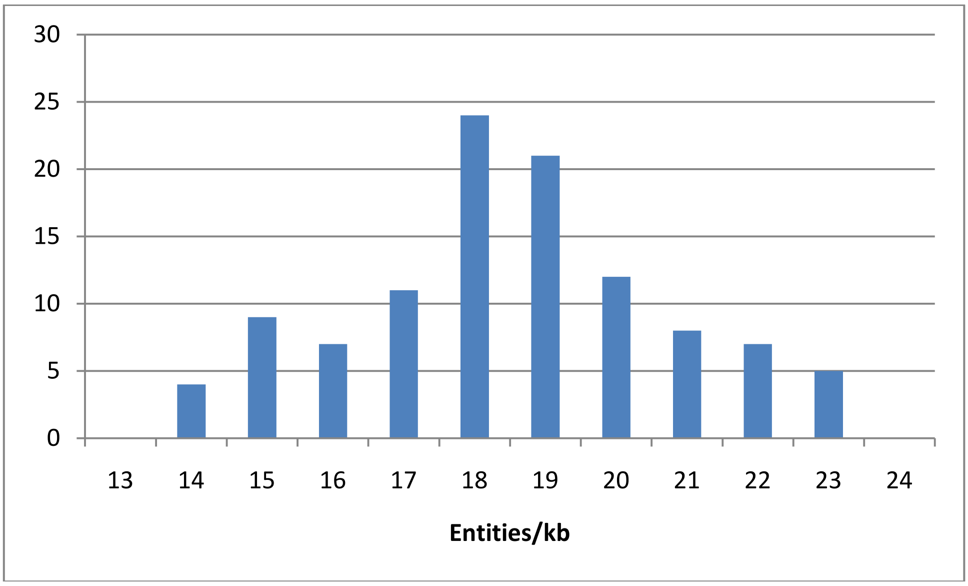

Figure 3: Distribution in density of entities in a file.

Very few of the models in the repository had optimised versions alongside the original files. Automated optimisation of a model is a function that is being built into the repository, to ensure that all models entered have a corresponding optimised version. When tracked over time this will provide an indication of whether the software tools are becoming more efficient at generating their IFC models. For the models which do have an optimised version (all created by the Solibri IFC Optimizer) we can see that there are typically fairly major gains to be found from an optimisation process. The gains found in these optimisations, while not consistent, are significantly more than reported in Pazlar and Turk (2007). Across all the entities in these eight models, even though they represent the output of five different CAD tools, there is approximately a 50% benefit from optimisation techniques on the number of entities which need to be generated and manipulated by a software tool handling an IFC model.

Model name	**Original Entities**	**Optimised Entities**	**Reduction**
Ettenheim-GIS-05-11-2006.ifc	515,973	328,567	36%
miniExample20080731_CoordView_SweptSolid.ifc	163	117	28%
2010-03-01 Project.ifc	1,007,858	269,460	73%
Med_Dent_Clinic_Arch.ifc	231,525	170,640	26%
Med_Dent_Clinic_Combined.ifc	1,904,600	1,035,240	46%
Med_Dent_Clinic_MEP_Elec.ifc	720,473	118,840	84%
Med_Dent_Clinic_MEP_Mech.ifc	993,676	761,667	23%
Med_Dent_Clinic_MEP_Plumb.ifc	17,294	16,988	2%

Table 5: Impact of optimisation on models.

Also of interest as an analysis over the models in the repository is the correctness of the models in relationship to the IFC standard specification. While this functionality is not currently automated, having the repository allowed an analysis of the models with the Solibri Model Checker (Solibri 2011). This identified that, when checked, many models have issues of critical severity (e.g., intersecting spaces or gaps between components) and all had issues of moderate severity (e.g., missing expected components). However, for many of the models, which were developed for testing aspects of IFC handling in preparation for certification, this is not surprising. These models do not represent a complete building, but a range of components of a certain type to be tested by an application's import handler (e.g., a range of doors in a single wall).

3. DISCUSSION

The slowly growing open repository of IFC models is providing the community with a wide variety of models as was the goal when first envisaged. It is including models from several versions of the IFC standard and created by a wide range of software tools and the major CAD packages. It also has an enormous variation in the size of the models available from millions of entities down to just a few entities in the most basic of possible models.

While there are currently only just over 100 models in the repository the author will be able to increase this to over 200 with currently identified IFC models. However, it is likely that a one-man-band will be insufficient to move the repository to a scale which will be sustainable over time. To become a truly useful resource there needs to be some input from others in the community to ensure that the repository reflects models from all major IFC-based projects, and hence the wide variety of analyses that are represented across these various projects.

Even with just over 100 models there are a range of analyses which have been undertaken across the repository. This showed remarkable consistency in the ratio between number of entities in a model and the resulting file size across all scales of models. The analysis of optimisation on a small set of models shows an enormous potential to improve the content of models, with around 50% of entities being able to be removed in an optimisation process. This finding warrants further investigation across the complete repository. The analysis of issues within the models identified that there are no perfect models by this measure, though in many cases this is explained by the emphasis on test cases for certification which do not sit within the context of a complete building.

The open repository is still being developed with further functionality added and with improvements to its usability. The most useful additions in progress are the automation of a range of analyses on a model when it is deposited into the repository. Once these are in place then a more comprehensive analysis of IFC models past and present will be a possibility and will provide a more significant picture of the progress of IFC models and the handling of them by CAD and other software tools.

REFERENCES

Amor, R. and Dimyadi, J. (2010) "An Open Repository of IFC Data Models and Analyses to Support Interoperability Deployment", *Proceedings of CIB W78*, Cairo, Egypt, 16-18 November, pp. 1-11.

Amor, R., Jiang, Y. and Chen, X. (2007) "BIM in 2007 - are we there yet?", *Proceedings of CIB W78 conference on Bringing ITC knowledge to work*, Maribor, Slovenia, 26-29 June, pp. 159-162.

Genero, M., Oiattini, M. and Calero, C. (2000) "Early Measures for UML class diagrams", *L'Objet*, 6(4).

Gielingh, W. (2008) "An assessment of the current state of product data technologies", *Computer-Aided Design*, 40(7), 750–759.

Jeong, Y.S., Eastman, C.M., Sacks, R. and Kaner, I. (2009) "Benchmark tests for BIM data exchanges of precast concrete", *Automation in Construction*, 18(4), 469–484.

Kiviniemi, A., Tarandi, V., Karlshøj, J., Bell, H. and Karud, O.J. (2008) "Review of the development and implementation of IFC compatible BIM", Erabuild Funding Organizations, Europe.

Lee, G., Won, J., Ham, S. and Shin Y. (2011) "Metrics for Quantifying the Similarities and Differences between IFC Files", *Journal of Computing in Civil Engineering*, 25, 172-181.

NIST (2011) "IFC File Analyzer", available at http://cic.nist.gov/vrml/cis2.html

Pazlar, T. and Turk, Ž. (2007) "Evaluation of IFC Optimization", *Proceedings of CIB W78 conference on Bringing ITC knowledge to work*, Maribor, Slovenia, 26-29 June, pp. 61-65.

Pazlar, T. and Turk, Ž. (2008) "Interoperability in practice: Geometric data exchange using the IFC standard", *Journal of Information Technology in Construction*, 13, 362–380.

Solibri (2011) "Solibri Model Checker v6.2", available at http://www.solibri.com/