

# CONCEPTUAL REQUIREMENTS FOR THE AUTOMATIC RECONSTRUCTION OF BUILDING INFORMATION MODELS FROM UNINTERPRETED 3D MODELS

Claus Nagel, Alexandra Stadler, Thomas H. Kolbe

Institute for Geodesy and Geoinformation Science, Technische Universität Berlin  
Straße des 17. Juni 135, 10623 Berlin, Germany  
{ nagel | stadler | kolbe }@igg.tu-berlin.de

Commission III, WG III/4; Commission IV, WG IV/5 and IV/8

**KEY WORDS:** 3D city models, 3D graphics models, semantic modeling, CityGML, IFC, BIM, 3D building reconstruction, object interpretation, formal grammars.

## ABSTRACT:

A multitude of new applications is quickly emerging in the field of Building Information Models (BIM). BIM models describe buildings with respect to their spatial and especially semantic and thematic characteristics. Since BIM models are manually created during the planning and construction phase, they are only available for newly planned or recently constructed buildings. In order to apply the new applications to already existing buildings, methods for the acquisition of BIM models for built-up sites are required. Primary data source are 3D geometry models obtained from surveying, CAD, or computer graphics. Automation of this process is highly desirable, but faces a range of specific problems setting the bar very high for a reconstruction process. This paper discusses these problems and identifies consequential requirements on reconstruction methods. Above, a two-step strategy for BIM model reconstruction is proposed which incorporates CityGML as an intermediate layer between 3D graphics models and IFC/BIM models.

## 1. INTRODUCTION / MOTIVATION

Building Information Models (BIM) describe buildings with respect to their geometry, topology, and semantic information about all their components. Especially the logical structure and well-defined meaning of the objects are crucial prerequisites for applications which go beyond pure visualization. These application areas include facility management, environmental and energy simulation, urban planning, architecture, civil engineering, and disaster management. The introduction of the US national BIM standard (NBIMS), which itself is based on the ISO standard IFC, lead to a major boost in the development of new applications and software systems. It can be expected that in the future many more applications, owners, and stakeholders will make use and rely on these semantically rich 3D models.

Although BIM models can (and will) be maintained during the entire existence of a building, they are generally only available for newly planned and recently constructed buildings. The reason is that they are manually created by the architects or civil engineers in the planning phase. In order to be able to employ the new BIM applications with existing buildings, BIM models have to be acquired for already built-up sites. This, however, is difficult because the BIM paradigm relies on a component based modelling consisting of walls, slabs, beams, stairs, pipes etc. These components are generally not fully visible or observable in an existing construction. Many elements even will be hidden totally. By using surveying technology like total stations and terrestrial/airborne laser scanners or techniques from photogrammetry, the 3D geometry can be reconstructed to a certain extent. However, only the visible surfaces are registered. This means that neither hidden parts nor the meaning of the surfaces or their belonging to specific object types are acquired. The same situation applies for the multitude of 3D models that are created within CAD and computer graphics systems like Google Sketchup or Autodesk's 3D Studio Max. These models generally consist of geometry and appearance information, but do not represent thematic information and the meaning of the objects. Thus, they also cannot be used for BIM applications.

As a consequence, 3D geometry or graphics models have to be interpreted, modified, and extended to become BIM models. Today, the acquisition of BIM models from observed or modelled 3D geometries is mostly done manually. The automation of this process would reduce efforts and costs substantially. However, automatic reconstruction of semantic building models is known to be a difficult problem that has been investigated by many groups over the last 25 years – with limited success so far. The main reasons result from the high demands on the reconstruction process regarding 1) the definition of a target model which restricts object configurations to sensible building structures and their components, but which is still flexible enough to cover (nearly) all existing buildings in reality; 2) the complexity of input data and reconstructed models; 3) data errors and inaccuracies, uncertainty and ambiguities in interpretation; 4) the reduction of the search space during the interpretation process. It is the purpose of this paper to present and discuss these requirements in more detail and indicate the consequences for any reconstruction process. Furthermore, we propose a two-step reconstruction strategy, which incorporates the OGC standard CityGML as an intermediate layer during the interpretation and final generation of BIM models.

## 2. TWO-STAGE RECONSTRUCTION PROCESS

The starting point of the proposed reconstruction process is pure 3D geometry/graphics models which can be obtained from different sources (cf. fig. 1). On the one hand, these models may be the result of manual design (CAD, computer graphics software). On the other hand, they may be derived from observations and measurements of topographic features in the field of photogrammetry and surveying (cf. fig. 1). Such models mostly result from registration methods for geometry along with methods for data segmentation and 3D geometry reconstruction. Substantial work has already been done in this area and both semi-automatic and automatic systems for 3D geometry reconstruction are available (cf. discussion in Baltsavias, 2003).

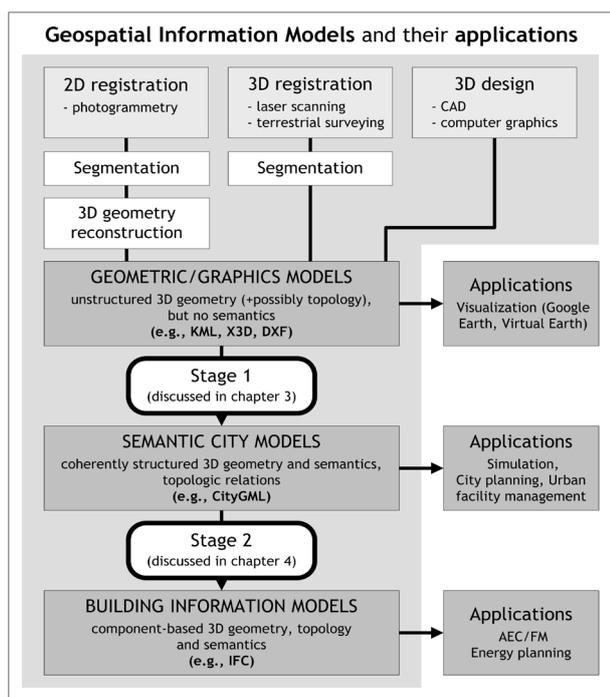


Figure 1. Two-stage reconstruction process.

The extracted geometry typically represents the observable surfaces of the topographic features which may range from simple blocks models to complex polyhedral structures depending on the level of detail. The resulting models can comprise additional geometric-topological relations and appearance information. However, they often contain no or just little semantics. The models are suited for visualization purposes in applications such as earth browsers or as input for further data interpretation. They usually base on well-known 3D graphics formats like X3D, DXF, KML, and COLLADA.

The automatic reconstruction of BIM models from 3D geometry models faces a high level of complexity. BIM models represent the built environment as designed rather than observed (Kolbe & Plümer, 2004). They geometrically describe thematic building components as volumetric primitives instead of individual surfaces. Thus, the reconstruction requires the identification and semantic classification of surfaces within the geometry model forming a building element. This is strongly impeded by the fact that geometry models mostly contain unstructured and uninterpreted geometry ("polygon soup").

In order to reduce the complexity of the overall reconstruction process, we propose its decomposition into two individual sub-problems (cf. fig. 1). In a first step, the 3D geometry models are further interpreted in order to structure the geometry and to enrich the models with semantic information which has to go beyond plain semantic tagging of surfaces (Pittarello & de Faveri, 2006). What is needed are spatio-semantically coherent models in such that all semantic components correlate to their geometric counterpart (Stadler & Kolbe, 2007). We choose CityGML as target model for this interpretation stage. CityGML is an international OGC standard for semantic 3D city models which provides a common information model for the representation of topographic features (Gröger et al., 2008). Since it is used to record observable objects it is much 'closer' to surveying and photogrammetric registration methods than BIM. CityGML employs a surface based representation of geometry similar to 3D graphics models which simplifies the interpretation process. Moreover, CityGML already provides an elaborate ontology of the exterior and interior built environment.

In a second step, CityGML models are used as input data for the reconstruction of IFC models (cf. fig. 1). Although CityGML and IFC are targeting different scales and the ontology of each semantic model is tailored to different scopes, both models agree to a great extent in the notion of a building and its semantic decomposition (Benner et al., 2005; Isikdag & Zlatanova, 2009). By using the explicit CityGML semantics as a priori knowledge we can narrow the search space of potential IFC building elements which have to be reconstructed.

A further advantage of the two-stage approach is the creation of a well-defined interface within the reconstruction process. Semantic 3D city models are already important products on their own. By using CityGML as standardized information model the subsequent semantic and structural refinement of 3D geometry models can be made explicit and exchanged between different systems and application areas without information loss. Moreover, semantic 3D city models form the basis for sophisticated analysis tasks in domains like simulation, city planning, and urban data mining. Thus, the reconstruction process might be aborted with the derivation of CityGML models after the first stage. Likewise, the proposed interface allows for starting the reconstruction process from existing CityGML models.

In both stages the interpretation and reconstruction of man-made structures comprises a general object recognition problem which has been topic of intense research for many years (Baltsavias, 2003, Brenner 2003). Generally, there are two opposed strategies for object recognition which can be classified as bottom-up and top-down approaches. Bottom-up approaches are data-driven, i.e., geometric primitives are directly extracted from the input data which are aggregated to form more complex structures. This is followed by the rule-based identification and combination of semantic objects. Often these rules are expressed by means of a constrained decision tree. Problems arise if objects cannot be observed in the input data due to errors or incompleteness. At this point prototypical 3D models of man-made structures have to be introduced. In contrast, top-down respectively model-driven approaches start from generating hypotheses for 3D models which are based on a predefined set of prototypes. The hypotheses can comprise arbitrary aggregations and combinations of prototypes. By verifying the hypotheses against the input data the prototypical 3D models can be subsequently refined in order to best match the input data. Thus, the verification process has to be controlled by a strong inference strategy. Finally, hybrid approaches combine both strategies. On the one hand, the instantiation of prototypes is oriented at the input data. On the other hand, model-driven hypotheses are generated taking into account identified prototypes in order to explain the input data.

In chapters 3 and 4 the conceptual requirements on the reconstruction will be explained for the two stages followed by a discussion on the inherent complexity and demands on a reconstruction strategy in chapter 5. Finally, in chapter 6 we draw a conclusion and give a brief outlook.

### 3. STAGE 1: GRAPHICS MODEL → CITYGML

In the first stage of our proposed reconstruction process a purely geometric graphics model (e.g., KML) is converted to a semantically enriched boundary model (e.g., CityGML) (see fig. 2). Work done in this field mostly concentrates on specific aspects of the reconstruction process. Thiemann & Sester (2004) propose an interpretation of geometric building models, which separates single semantic components. Schmittwilken et al. (2007) reconstruct stairs from uninterpreted laser scan point clouds. Dörschlag et al. (2007) reconstruct semantic building

models according to building component hypotheses. Anyway, this approach currently adheres to specific forms of geometry and LOD handling (in the following referred to as “Replace geometry” and “Automatic LOD recognition”). In the following we defocus from specific approaches but will give a broader overview of possible interpretation variants and emerging requirements for the reconstruction process. We will start the discussion with an investigation of input and target model characteristics in terms of structure and data accuracy.

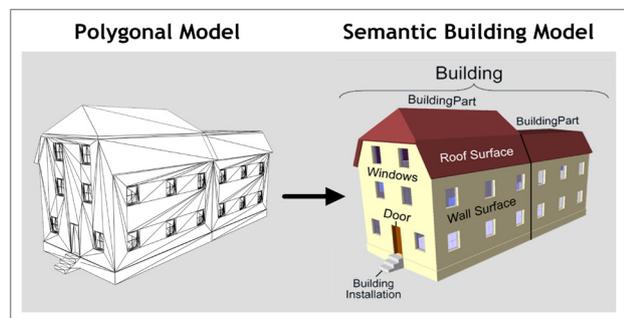


Figure 2. Deriving a semantically structured boundary model from a polygonal model.

Depending on data origin, there are various ways for the generation of graphics models. See fig. 1 for an illustration of different generation processes. In the following four categories of graphics models are examined, which are relevant in practice:

**Photogrammetric models** result from interpreting aerial or satellite images. Main goal is the reconstruction of roof structures, which is done either manually (photogrammetric stereo processing) or automatically (Fischer et al., 1998). Since facades are often occluded, their location is mainly approximated by extruding the eaves outline to the ground. Therefore, facades are typically displaced and all structural information is missing. Furthermore, positioning errors of images, limited image resolution and interpretation errors lead to inaccuracies in extracted roof structures. Models may even be incomplete due to occlusions, which are mainly caused by vegetation or shadows.

**Airborne laser scan models** are based on point clouds from laser scan flights. To reduce data volume and to smooth the resulting models, the initial point clouds are approximated by adjusting planes. These planes are the basis for the reconstruction of roof structures, which is mostly done automatically (Milde et al. 2008). Here too, facades are generally occluded by roof overhangs. Therefore, laser scan data is often combined with existing building footprints (e.g., cadastre data) resulting in facades, which are free from any structures (like balconies or window offsets). Random errors in the raw data are reduced by the use of adjusting planes. Like for photogrammetric models, incompleteness may be caused by occlusions.

**CAD and planning models** are based on building plans or surveys. They are highly detailed and often even include interior structures. Depending on the modelling software, resulting models are either component based or boundary representations. Here, also invisible objects may be part of the model (e.g. power supply, if contained in the underlying building plans). Typical modelling errors like overshoots, undershoots, self-intersections and permeations are significant, since they affect the topology of the model.

**Visualization models** are graphics models that are produced with the primary goal of (fast) 3D visualisation. The models contain few geometric details but often rely on structural information from façade images, which are projected onto the models for visualization purposes. Since models are explicitly built

for visualization, only visible parts are trustworthy. Underneath the surface one has to reckon with coarse errors (e.g., overlapping objects like visualised in fig. 3). For the sake of modelling simplicity, the geometric composition of visualization models may conflict with semantic structures (e.g., representing all facades of multiple aligned buildings as one big polygon).

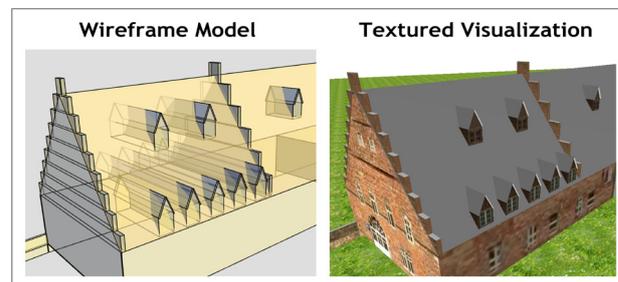


Figure 3. Dormers are extruded through the whole building (left). Visualization of the building does not reveal overlapping building and dormer bodies (right).

Regarding their structure, all presented models already comprise or may be transformed to polygons. In order to be able to deal with data inaccuracy and incompleteness, hypotheses must be also accepted even if they are not fully verified. Refer to chapter 5 for detailed discussions and strategies. Some of the models only comprise information, which is visible from the outside. This has influence on the target model’s LOD.

By interpreting the input model, we want to generate a topologically sound and semantically structured boundary model. Geometry and semantics have to be structured coherently with link in between to ensure a consistent data model, which forms a convenient basis for data analysis. CityGML as target data format fulfills all these requirements (Kolbe, 2009).

#### Characteristics of CityGML

CityGML is a standardized information model which puts focus not only on the objects’ geometry but also on their semantics, topology, and appearance. Key features of CityGML are:

- Objects may comprise coexisting geometric representations for different levels of detail (LOD concept).
- Topological relations between objects are realized by links between identical geometries (XLink concept).
- Variable complexity in the structuring of geometry and semantics – preferably coherent structures (spatio-semantic coherence, see Stadler & Kolbe, 2007).
- Aggregation hierarchies on the part of both geometry and semantics support complex object structures (hierarchical structuring).

A possible drawback when it comes to model interpretation is the versatile data model of CityGML: the same objects can be expressed in different ways allowing for ambiguity in modeling.

Fig. 4 shows the structural differences of input and target model. Whereas input models consist of unstructured geometry without further semantic information, the target models should consist of spatio-semantically coherent structures.

The structure and mechanisms of CityGML describe a generic way to define general characteristics of urban objects. E.g., a building is composed of wall, roof, and ground surfaces. In order to be able to ensure a correct interpretation of all these surfaces, profound knowledge is essential. Therefore, it is required to complement CityGML by additional constraints representing typical configurations, e.g., wall surfaces have to be upright and perpendicular to ground surfaces. Additional constraints may arise from user requirements for the target model.

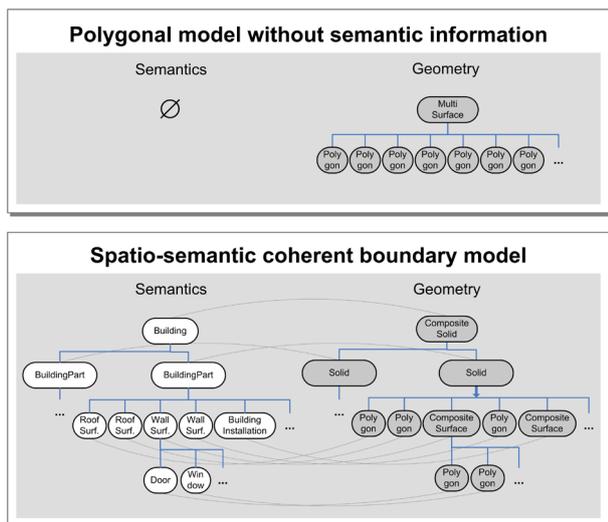


Figure 4. Semantics and geometry of input and target models.

Observing the main characteristics of input and target models, we can assess the following critical issues in the interpretation process of stage 1:

**Spatio-semantic coherence**

We intend to map geometry and implicitly contained semantic information of graphics models to the class hierarchy of CityGML. Therefore, it is necessary to explain implicit contents by appropriate hypotheses. The resulting target model shall be spatio-semantically coherent. Consequently, when interpreting input models, our hypothesis has to consist of both a semantic structure and the corresponding geometric representation. The respective reconstruction rules describe the interaction of semantic components as well as appropriate geometric structures. E.g., we interpret vertical surfaces, which have direct connections to two other vertical surfaces and one horizontal surface, as walls. In case of a verification of the hypothesis, they will be stored as thematic surfaces of type “WallSurface” and attached with a MultiSurface geometry (lodXMultiSurface).

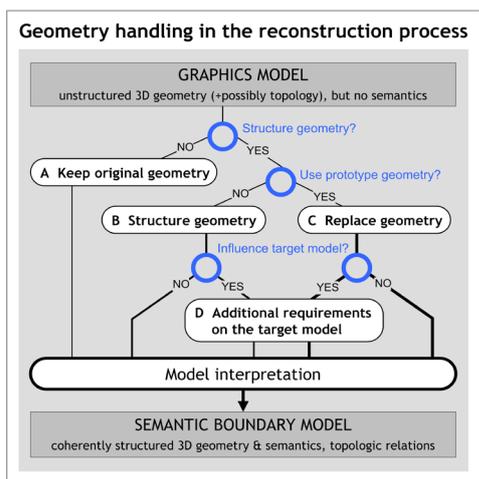


Figure 5. Alternative methods of geometry handling in the reconstruction process. Decision points are marked with circles. Increasing line weights stand for growing degrees of geometric structuring.

**Geometry handling**

There are several ways to deal with geometry when transforming a fine grained (but mostly error-prone) polygonal model into a sound CityGML model. Depending on the degree of geometry adaptation, we distinguish four different approaches:

- A Keep original geometry** – geometry remains unchanged. We merely attach semantic information to polygons. An example is given by Pitarello & de Favari (2006).
- B Structure geometry** – we transform the unstructured collection of polygons into a well-formed composition of geometries (e.g., Solids, CompositeSurfaces). The structuring of geometry still has no effect on coordinate values.
- C Replace geometry** – depending on the target model’s requirements, it might be necessary to adapt the input model’s geometry. Such geometry substitutions may also result in topological changes. E.g., when deriving a LOD1 building from a detailed photogrammetric model, the roof has to be flattened and window apertures in walls will be filled (c.f. Thiemann & Sester, 2006).
- D Additional requirements on the target model** – Extra knowledge about buildings can be expressed as additional requirements for the interpretation process. The resulting constraints will exceed those stated by CityGML. Examples are minimal or maximal dimensions, parallelism, rectangularity, maximal amount of related objects, etc.

Approaches A and B are pure model interpretations in such that the input model’s geometry is inherited. Since coordinates remain the same, these approaches are relevant for interpretation of legal data, where geometric changes may be prohibited. A possible drawback of retaining original geometry is the disability to reduce data load by joining coplanar faces, to remove data noise, and to resolve geometric and topological errors. Approaches C and D incorporate geometry in building hypothesis, i.e., replacing input geometry by geometry of prototypes. Depending on model constraints, this may cause considerable geometry modifications. Main benefit of both approaches is the big influence on the resulting model (see fig. 5).

**LOD concept**

The simultaneous representation of multiple LODs is a fundamental concept of CityGML. Since graphics models do not follow the same LOD definition as CityGML, the question arises how to decide on the appropriate target model LOD. Following possibilities are feasible:

- **Automatic LOD recognition** – the input model’s granularity allows for drawing conclusions about sensible target LODs. E.g., if the graphics model does not contain window setoffs or molded roof structures, it will make no sense to choose a high LOD for the target model.
- **User input** – alternatively, we can ask the user to specify the target model’s LOD. Problems may arise, if the input model does not fulfill the requirements of the chosen LOD.
- **Build a LOD series** – having specified one appropriate target LOD, we can think about covering also all lower LODs. The result is a LOD series with explicit linkage between multiple LOD representations.

For different LODs, the underlying geometric and semantic structure varies considerably. Therefore, the chosen target LOD has big influence on the hypotheses chosen for the model interpretation process. Consequently, it might be beneficial to use LOD adapted interpretation methods, which build on each other: prototypes of one LOD aggregate to prototypes of the next lower LOD. It has to be investigated whether an a priori generalization of the input model’s geometry is sensible, if the target LOD is lower than the input model would allow for.

**Topologic relations**

CityGML represents topology by explicit links between geometries that are part of several objects. E.g., two buildings might share a common side wall or a specific geometry might be used in more than one LOD representation. By referring to

the XLink concept of CityGML, there are two benefits for the interpretation process:

- Aggregated hypotheses for possible configurations of neighboring prototypes can share some geometries, making topological adjacency relations explicit.
- When using the same geometric entity for several LOD representations, correspondences can be recorded explicitly.

#### Aggregation of objects – Hierarchical structuring

CityGML employs aggregation hierarchies regarding both geometric and semantic objects allowing for different degrees of object aggregation. E.g., openings (like windows and doors) are part of thematic surfaces (walls, roof, ground), which are aggregated to building parts; building parts are again aggregated to form whole buildings. This advises a hierarchical strategy for model interpretation relying on hypotheses with increasing refinement (c.f. Dörschlag et al. 2007). An essential aspect is the direction of the interpretation process. When interpreting a building model, one can initially introduce a hypothesis for the whole building, go on with searching for building parts, afterwards distinguish between different thematic surfaces and finally end up with interpreting single polygons or vice-versa (top-down or bottom-up approach).

#### Ambiguities in data modeling

The generic character of CityGML allows for various modeling variants for the same city object. E.g., dormers can be modeled as building installations or as part of the building using roof and wall surfaces. They may comprise different complexities in terms of geometry and semantics. Consequently, there will be multiple valid hypotheses for the interpretation of objects. Only with the existence of an appropriate weighting function, we can determine the most likely object representation. In chapter 5 we will explicate requirements on the weighting function.

### 4. STAGE 2: CITYGML → IFC

The second stage of our proposed two-stage strategy aims at automatically reconstructing IFC models from CityGML input models. The CityGML models can result from the previous interpretation stage. Alternatively, already existing models can be directly fed into this second stage.

CityGML and IFC vary substantially in many aspects. A fundamental difference arises from their distinct modelling paradigms which are due to the way 3D models are acquired in the GIS domain respectively in the field of BIM and Computer Aided Architectural Design (CAAD). In GIS, 3D objects are derived from surface observations of topographic features based on sensor-specific extraction procedures. Features are hence described by their observable surfaces applying an accumulative modelling principle. In contrast, BIM models reflect how a 3D object is constructed. They follow a generative modelling approach and focus on the built environment rather than on topography. Therefore, BIM models are typically composed of volumetric and parametric primitives representing the structural components of buildings (Kolbe & Plümer, 2004). Fig. 6 exemplifies the implications of both modelling approaches.

The process of reconstructing a component-based volume model from a surface model requires the instantiation and rule-based combination of volumetric building objects such as wall, slab, and roof elements which are most likely to explain the given input model. A key aspect to the identification of the proper IFC primitives to be instantiated from the input surfaces is semantic information. This comprises the thematic classification of surfaces as well as the meaning and function of objects and their interrelationships. Both CityGML and IFC provide elaborate semantic models of the exterior and interior

built environment. This a priori knowledge allows for reducing the search space of potential IFC elements. For example, a CityGML *WallSurface* object is most likely to be mapped to an *IfcWall* respectively *IfcWallStandardCase* element.

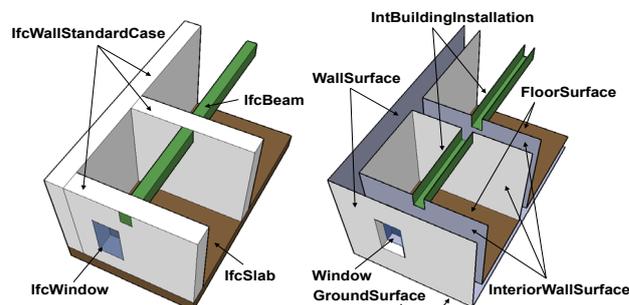


Figure 6. Snapshot of a building storey modeled in IFC (left side) and CityGML (right side).

The generation of hypotheses for IFC elements draws its complexity from the fact that building components generally can only be observed in parts and often are not observable at all. Since CityGML is used to model observed topographic features, only the visible parts are represented in the input data and can be used as a starting point for reconstructing IFC elements. Furthermore, for each visible part of a building component even two or more surfaces might be observable which are represented as individual semantic objects in CityGML. This leads to a high combinatorial complexity for the matching of IFC elements based on CityGML entities. Only in rare cases IFC elements can be directly reconstructed from a single CityGML feature. In fig. 6, a corresponding  $1:1$  matching relation can be found for the *IfcWindow* element and its CityGML *Window* counterpart.

More often we have to deal with  $n:1$  matching relations between CityGML and IFC entities in such that two or more input surfaces have to be identified to form a single IFC element. First, this is typically the case for wall components as shown in fig. 7. In this example, the interior wall surfaces  $I_{11}$  to  $I_{13}$  and  $I_{21}$  to  $I_{22}$  represent two separate wall objects  $W_1$  and  $W_2$  and have to be mapped to corresponding *IfcWall* elements in the target model. Furthermore,  $n:1$  relations occur for components which penetrate other components and hence are partly concealed and non-observable. An example for this is a single ceiling beam which continues over two or more rooms as depicted in fig. 6. Since this beam is observable from both rooms it may be represented in CityGML as two thematic *IntBuildingInstallation* objects with individual surface geometries. Thus, besides semantic information a process for identifying input surfaces to be aggregated to a single IFC element has to additionally analyze geometric-topological relations between the object geometries such as parallelism, perpendicularity, distance, and adjacency.

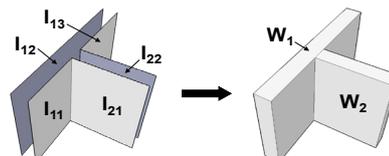


Figure 7.  $n:1$  match between CityGML wall surfaces and reconstructed IFC entities.

The reverse  $1:m$  relation results from splitting a CityGML object into two or more IFC elements. A split is performed, for example, for a CityGML *WallSurface* object spanning a complete building façade. In contrast to CityGML, IFC buildings are structured in storeys. This requires the partitioning

of the façade surface into one or more *IfcWall* elements per storey. The number  $m$  of resulting IFC elements cannot be determined a priori due to allowed modelling ambiguities in IFC. The same is true for  $n:m$  matching relations. An  $n:m$  matching for separate input wall surfaces with more than one possible IFC element hypotheses is illustrated in fig. 8. As this example shows, the reconstruction of IFC elements will most often lead to more than one valid hypothesis for the same configuration in the input data. The reconstruction strategy therefore has to deal with competing hypotheses.

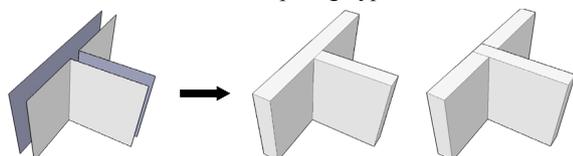


Figure 8.  $n:m$  match between CityGML wall surfaces and reconstructed IFC entities which are valid and hence competing hypothesis explaining the input model.

The hypothesis generation further comprises the instantiation of the potential IFC entity in such that it best fits the spatial properties of the related CityGML objects. This spatial fitting requires parameter estimation for the rotation, scale, and translation transformations as well as the element shape with respect to the identified input surfaces. According to the generative modelling paradigm, CAAD models usually employ the *Constructive Solid Geometry* (CSG) for shape representation. The primitives are defined through shape parameters which depend on the type of the IFC entity. For example, an *IfcWallStandardCase* is given by the wall height, the wall thickness, and the wall offset from axis which describe a vertically extruded solid. As a consequence of this parametric description there are strong implicit geometric constraints for the resulting CSG primitive such as parallelism and perpendicularity of wall surfaces. The estimation of shape parameters is supported by additional appearance information of observable surfaces provided by the input CityGML model. This comprises texture images or arbitrary sensor data which can be analyzed in order to deduce information about the interior build-up of components.

In contrast, CityGML employs the *Boundary Representation* (B-Rep) for the modelling of object geometry which is defined as the accumulation of all surfaces enclosing the volume of an object. Problems arise from the fact that man-made objects may have deviations from the idealized CSG shape used in building construction, e.g., opposite surfaces of a real-world wall often do not adhere to parallelism as the wall thickness changes over height. Since deviations are observable and hence incorporated into the B-Rep model, there is no set of parameters for an ideal CSG primitive which strictly explains the input data. Thus, the hypothesis generation must employ a non-strict matching of strong CSG primitives. Instead, primitives could be weakened as proposed by (Brenner, 2004). However, the concept of weak CSG primitives has not been considered for IFC so far.

The conversion of B-Rep geometries to corresponding CSG representations is ambiguous in general. Consequently, deriving CSG primitives from measured surface data is also ambiguous. In the context of reconstructing IFC from CityGML this uncertainty is even increased. The building components are only observable in parts and, thus, the CSG primitives cannot be derived from closed volumes. One has to keep in mind that the resulting ambiguity cannot be resolved without extra knowledge. The following fig. sketches a simple scene to illustrate this problem which is based on fig. 8. Even if the shown  $n:m$  matching problem is resolved there are still remaining geometric ambiguities for modelling the wall connection.

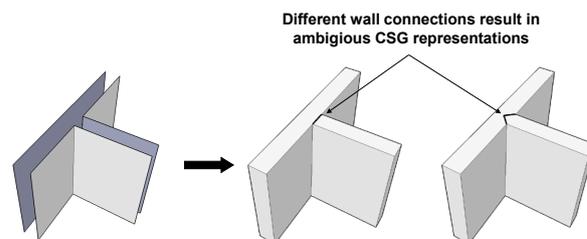


Figure 9. The conversion of B-Rep geometries to corresponding CSG representations is ambiguous and cannot be resolved without a priori knowledge or assumptions.

The parameter estimation for both transformation and shape has to obey additional contextual constraints. Unary constraints affect the interpretation process of a single IFC entity. For example, wall elements must usually meet a reasonable minimum and maximum height and length. Often the parameters defining an IFC element cannot be estimated isolated from other building components. For example, walls on the same storey usually share the same parameters for wall height and offset from ground and adjoining walls are likely to have the same wall thickness. Such many-to-many constraints do not only influence the parameter estimation of a single IFC entity but also mutually affect and dynamically change the parameters of previously generated hypotheses for other building elements.

Whereas both the implicit geometric constraints of primitives and the unary contextual constraints usually impair the best fit to the input data, mutual constraints often aim at aligning primitives. A best fit does not necessarily imply a correct alignment of the reconstructed IFC element. If the alignment is enforced after the fitting operation, the best fit property will most likely be lost (Brenner, 2004). Thus, the parameter estimation of potential IFC primitives has to ensure a best fit and the correct alignment at the same time. This results in complex element hypotheses which have to establish interdependencies. On the other side, mutual constraints also facilitate the unification of parameters over several primitives which helps in reducing the number of overall model parameters and, thus, in simplifying the final hypothesis of the target model (Fischer et al., 1998).

Purely geometric-topological constraints on primitives cannot prevent unreasonable instantiations and combinations of IFC elements. For example, they do not express that roof elements may not be instantiated at the bottom of a building. In fact, the IFC data model itself does not formally specify rules on how to combine building components in order to form a valid building. However, the reconstruction of valid building models is to be considered the main target of the interpretation process. What is needed is a framework providing enhanced model expressiveness to describe the structure of buildings and to incorporate semantic constraints on IFC entities and their aggregations in addition to geometric-topological constraints. By this means, not only the geometric and syntactic correctness of the generated hypotheses with respect to the IFC data model can be evaluated but also their semantic and structural validity. Consequently, the creation of an enhanced model for buildings reflecting common structural, functional, and physical agreements in BIM related fields such as architecture and structural engineering is a key requirement for coming from CityGML to IFC.

By analyzing typical structures and configuration patterns of building components in existing IFC datasets, we can derive a priori likelihoods for the instantiation of single IFC elements as well as for their valid combination. These a priori likelihoods can be fed back into the enhanced model of the built environment and, thus, can be introduced as a priori knowledge into the hypothesis generation process. This knowledge can be

utilized, e.g., in interpreting non-observable components. Although, generally, non-observable parts cannot be reconstructed because they are not represented in the input CityGML model, an enhanced model of the built structure helps in detecting them. For example, if a ceiling surface and the corresponding floor surface of two rooms on top of each other have a distance considerably larger than the usual thickness of a slab element, the instantiation of a single *IfcSlab* entity is syntactically correct but most likely a false interpretation. Using structural patterns we can rather assume that the ceiling is suspended. By applying stochastic information, we can even reconstruct probable configurations of IFC entities explaining the observed surfaces, for example two *IfcSlab* entities and the *IfcSpace* in between.

The verification of generated hypotheses requires the backward projection from IFC to CityGML which comprises the conversion of both geometry and semantics (Benner et al., 2005; Isikdag & Zlatanova, 2009). Although the B-Rep can be obtained automatically and unambiguously from CSG, each IFC element is transformed to a set of surfaces describing a closed volume. One has to remind that this transformation does not reflect the input data which only contains observable surfaces. It still has to be examined whether the non-observable surfaces have to be removed in order to get stable verification results.

The quality of the reconstructed IFC model depends to a great extent upon the quality of the input model. As semantic information is a premise for the identification of IFC elements, the proposed interpretation process requires the input model to provide a coherent representation of semantics and geometry (cf. section 3.3). An additional structural model of the built environment allows for detecting and possibly correcting errors within the input model such as falsely classified or even missing building components. In general, the reconstruction of the interior built environment implies an input CityGML model conformant to the quality requirements defined by LOD4. However, lower LODs may also serve as input data resulting in IFC building models which are only represented by their exterior wall and roof structure. Such building models can already be sufficient in BIM applications for which the building interior is negligible or can be used as templates for architectural interior design. Furthermore, they support current efforts to continue IFC from the building scale to the city scale and, thus, to use IFC for virtual 3D city modelling.

## 5. STRATEGY REQUIREMENTS

From the specifications of the graphics models, CityGML, and IFC we can define our input and target models (see chapters 3 and 4). Both CityGML and IFC specify the thematic structuring of objects in a formal way. The model semantics are defined by the international standards in the respective domains, in particular the ISO 19100 standards family and STEP. Syntactically, the models are described using formal concepts such as UML, XSD, and EXPRESS which provide a generic description of objects and their relations. However, they are not meant to qualify objects and inter-object relations in order to restrict the modeling to only sensible object configurations. In order to carry out the interpretation process, we have to increase the expressiveness of the CityGML and IFC modeling frameworks in terms of physical, functional, and semantic / logical properties. A very promising way to formulate respective constraints is by the use of formal grammars.

*Formal grammars* originate from linguistics. They define the symbols of a language together with the rules to compose and verify valid sentences (Chomsky 1959). The symbols of the language represent the features or components and the

production rules define the valid combinations of complex configurations, i.e. hypotheses for the objects to be reconstructed. Different types of formal grammars are being used for object recognition for a long time now (Tsai & Fu, 1980). *Attributed grammars* allow to parameterize components and to define functions and constraints on the combination of parameters from different components. *Stochastic grammars* assign a priori probabilities to the occurrence of components and the applicability of rules and compute the overall probability for each reconstructed object (given by all symbols of a sentence). Starting from 1971, *shape grammars* have been used to describe valid combinations of geometric primitives (Stiny & Gips, 1971). *Split grammars* as defined by (Wonka et al., 2003) are inspired by shape grammars and describe patterns for spatial decomposition of geometric objects, here applied to building façade reconstruction.

Recently, combinations of the different types of grammars and Monte Carlo strategies for the generation of hypotheses have been proposed for façade reconstruction (Ripperda, 2008; Reznik & Mayer, 2008; Hohmann et al., 2009), the reconstruction of roofs (Milde et al., 2008), stairs and entrance areas of buildings (Schmittwillken et al. 2007). In (Dörschlag et al., 2007) it is proposed that building components are represented as prototype constraint graphs which are composed according to the rules of an attributed grammar in order to form complex building hypotheses called reconstructed constraint graphs. All these concepts allow for a high flexibility with respect to reconstructable objects. However, means for the handling of errors and unobservabilities of expected components are mostly missing yet. This will have to be solved in order to become applicable in productive environments.

When working with real-world data, one has to deal with uncertainties regarding geometry and semantics. Geometric errors are often caused by measuring or modeling inaccuracies. Both measuring and modeling imply generalization processes due to the mapping of infinitely detailed structures to models consisting of a finite number of parameters. E.g., walls in a building model may be described by a single thickness parameter, although their thickness varies in reality. Semantic errors mostly result from ontological inconsistencies. A building might be called building, house, or man-made structure and might be decomposed into building parts which are horizontal segments (floors) or vertical ones according to the underlying ontology.

Further problems for object interpretation may arise from model incompleteness. This implies both incomplete object information (missing geometry or semantics) and unavailability of object parts, e.g., due to occlusions. IFC beams might be partially observable, but for complete representation, their geometry has to be extrapolated. In order to deal with emerging uncertainties, grammars must become robust with respect to errors.

In chapters 3 and 4 we pointed out the problem of modeling ambiguities. As a consequence of their generic character, both CityGML and IFC allow for alternative modeling variants of the same real-world building which differ in terms of geometric and semantic complexity. This implies the existence of various possible interpretation results. Thus, a grammar is required that allows for multiple disjunctive production rules. Consequently, the grammar can produce alternative hypotheses and interpretations. In order to be able to choose the “best” of all competing hypotheses, a weighting function is required that takes into account following two aspects: 1) Goodness of fit, 2) Complexity of the hypothesis. The weighting function has to balance between both aspects in order to avoid overfittings.

Furthermore, it should have a defined semantics in the sense that it specifies a meaning for the “best” hypothesis. By choosing probability theory, the best matching means the most probable interpretation of the examined situation. Thus, the interpretation process amounts to a maximum a posteriori (MAP) estimation, finding the most probable model

$$\hat{M} = \arg \max_{M_i} P(M_i | D) \quad (1)$$

where  $M_i$  are the different hypotheses, and  $D$  the given input data. Possible frameworks are:

- Minimum description length principle (MDL), based on information/probability theory (Grünwald et al., 2005)
- Akaike Information Criterion (AIC) (Akaike, 1974)

Both MDL and AIC are grounded in the concept of entropy, which specifies the amount of information contained in a message reflecting the model complexity. Frankly speaking, the common idea is to choose the *simplest suitable* model (see also Fischer et al. 1998, Dörschlag et al. 2007).

Concerning the interpretation sequence, our target models impede the pure application of both top-down and bottom-up strategies. The model complexity leads to an infinite number of possible building hypotheses, which argues against a top-down approach. The inaccuracy and incompleteness of input models prevents a pure bottom-up approach. Therefore, we have to go for a mixed approach, which compensates weaknesses of both single approaches. Since the grammar will include disjunctive rules, a combinatorial complexity is induced, rendering the interpretation process NP-complete. Thus, strong heuristics are required, which have to cut down search space substantially.

## 6. CONCLUSIONS

The reconstruction of BIM models from uninterpreted 3D geometry/graphics models sets the bar very high for an automated interpretation process. In order to reduce the overall complexity and to increase the flexibility of the reconstruction process, we have proposed a subdivision into two major stages 1) from graphics models to CityGML building models, and 2) from CityGML to IFC building models. IFC and CityGML are appropriate target models for reconstruction, as these models are well-defined and their instances are usable in a broad range of applications. However, in order to be able to reconstruct either CityGML or IFC from 3D graphics models, stronger concepts than the pure data models from the specifications are required which restrict reconstructed objects to sensible building structures. Formal grammars seem to be a promising approach to express valid aggregations of components adhering to functional and logical constraints, but will have to combine geometric shapes, attributes, constraints between attributes of different objects, and stochastic aspects like uncertainty and a priori probabilities. Although formal grammars have been used for building reconstruction, there are no formalisations for CityGML or IFC available so far. Another challenge is the definition of the weighting function being key to the search for the optimal interpretation. Finally, strong heuristics are required in order to cope with the huge search space. In the future, we will investigate these issues in both reconstruction stages.

## REFERENCES

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6).

Baltsavias, E. 2003. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. In: *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 58, 2004.

Benner, J., Geiger, A., Leinemann, K. 2005. Flexible Generation of Semantic 3D Building Models. In: *Gröger, G. et al. (Eds.): Proceedings of the 1st International Workshop on Next Generation 3D City Models, Bonn 2005*.

Brenner, C. 2003. Building Reconstruction from Laser Scanning and Images. In: *Proc. of the ITC Earth Observation Science Department Workshop on Data Quality in Earth Observation Techniques, Enschede, The Netherlands, 2003*.

Brenner, C. 2004. Modelling 3D objects using weak CSG primitives. In: *Stilla, U. et al. (Eds.). PIA07. Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 35*.

Chomsky, N. 1959. On Certain Formal Properties of Grammars. *Information and Control*, 2, 137-167

Dörschlag, D., Gröger, G., Plümer, L. 2007. Semantically Enhanced Prototypes for Building Reconstruction. In: *Stilla, U. et al. (Eds.). Proc. of PIA'07, Intern. Archives of ISPRS, 36*.

Fischer A., Kolbe T. H., Lang F., Cremers A.B., Förstner W., Plümer L., Steinhage V. 1998. Extracting buildings from aerial images using hierarchical aggregation in 2d and 3d. In: *Computer Vision and Image Understanding*, 72(2).

Grünwald, P. D., Myung, I. J., Pitt, M. A. (Eds.) 2005. *Advances in Minimum Description Length*. MIT Press.

Gröger, G., Kolbe, T.H., Czerwinski, A., Nagel, C. 2008. *OpenGIS City Geography Markup Language (CityGML) Encoding Standard*. OGC Doc. No. 08-007r1.

Hohmann, B., Krispel, U., Havemann, S., Fellner, D. 2009. CITYFIT – High-Quality Urban Reconstruction by Fitting Shape Grammars to Image and derived Textured Point Clouds. In: *Proceedings of the International Workshop 3D-ARCH 2009, Trento 2009*.

Isikdag, U., Zlatanova, S. 2009. Towards Defining a Framework for Automatic Generation of Buildings in CityGML Using Building Information Models. In: Lee, Zlatanova (eds.), *3D Geo-Information Sciences*, Springer

Kolbe, T. H., Plümer, L. 2004. Bridging the Gap between GIS and CAAD. *GIM International 2004, Vol. 18, No. 7*.

Kolbe, T. H. 2009. Representing and Exchanging 3D City Models with CityGML. In: Lee, Zlatanova (eds.), *3D Geo-Information Sciences*, Springer

Milde, J., Zhang, Y., Brenner, C., Plümer, L., Sester, M. 2008. Building Reconstruction using a structural description based on a formal grammar. In: *Int. Archives of ISPRS, Vol. XXXVII. Part B3b. Beijing 2008*

Reznik, S., Mayer, H. 2008. Implicit Shape Models, Self-Diagnosis, and Model Selection for 3D Facade Interpretation. *Photogrammetrie, Fernerkundung, Geoinformation* 3/2008.

Ripperda, N. 2008. Grammar Based Facade Reconstruction using rjMCMC. *Photogrammetrie, Fernerkundung, Geoinformation* 2/2008.

Pittarello, F., de Faveri, A. 2006. Semantic description of 3D environments: a proposal based on web standards. In: *Proc. of the 11th int. conference on 3D web technology, Columbia, Maryland 2006*.

Schmittwilken, J., Saatkamp, J., Förstner, W., Kolbe, T. H., Plümer, L. 2007. A Semantic Model of Stairs in Building Collars. *Photogrammetrie, Fernerkundung, Geoinformation* 6/2007

Stadler, A., Kolbe, T. H. 2007. Spatio-semantic Coherence in the Integration of 3D City Models. In: *Proceedings of the 5th International Symposium on Spatial Data Quality, Enschede 2007*.

Stiny, G., Gips, J. 1971. Shape Grammars and the Generative Specification of Painting and Sculpture. *IFIP Congress, North Holland*

Thiemann, F., Sester M. 2004. Segmentation of Buildings for 3D-Generalisation. In: *Proceedings of the Workshop on generalisation and multiple representation, Leicester 2004*.

Thiemann, F., Sester M. 2006. 3D-Symbolization Using Adaptive Templates. In: *Proceedings of ISPRS Technical Commission II Symposium, Vienna 2006*.

Wonka, P., Wimmer, M., Sillion, F., Ribarsky, W. 2003. Instant Architecture. *ACM Transaction on Graphics*, 22(4).