Automation in Construction 129 (2021) 103756

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Logic representation and reasoning for automated BIM analysis to support automation in offsite construction

Oscar Wong Chong, Jiansong Zhang *

Automation and Intelligent Construction (AutoIC) Lab, School of Construction Management Technology, Purdue University, West Lafayette, IN 47907, United States of America

1. Introduction

In the past decades, a concerning labor shortage is experienced by the labor-driven Architecture, Engineering, and Construction (AEC) industry [1–3]. For instance, in the United States, 80% of the construction companies cannot find trades workers to fill job positions [4]. Just for January of 2020, 267,000 unfilled construction positions were reported in the U.S. [5]. This problem has impacted negatively the productivity in the AEC industry, causing delays and cost overrun in construction projects [6]. In a survey conducted by Construction Labor Market Analyzer, more than 90% of the respondents reported a lower productivity for the years 2014 and 2015 due to the labor shortage [6].

Technologies such as offsite construction and automation allow greater efficiency by automating construction processes, which can be translated into higher productivity in the industry. This increase in productivity can relieve some of the strain imposed by the workforce shortage [7–9]. On the one hand, offsite construction provides many benefits over conventional stick-built construction such as working in a controlled environment, the ability to conduct activities in parallel, and improvement of the built quality [10–12]. Automation in construction, on the other hand, includes the application of technologies such as Computer Numerical Control (CNC), robotics, and other automation machines that could be easier to implement in a factory setting than in onsite construction. For example, these technologies can be adopted to support the fabrication and assembly of building components by automating otherwise manual operations. Moreover, automation can improve safety by saving workers from dangerous and heavy-duty tasks and/or in hazardous conditions.

However, despite these benefits, there are many challenges with the automation of offsite construction in practice. One main challenge is that offsite construction demands more rigorous design, planning [13], and construction requirements than those of onsite construction. For example, the design, manufacturing, and assembly tolerances for offsite construction are tighter than those for stick built because the assembled components need to appropriately fit the prepared foundation rather than the components sequentially erected onsite, in which the latter option allows more flexibility for local adjustments of the connections between the frames and the foundation. In addition, automation requires detailed and precise information such as building information models, material, and building systems to obtain the desired outcomes [14]. These challenges have impeded the wide adoption of automation in offsite construction.

Building Information Modeling (BIM) has the potential to overcome these limitations and enable automation in construction by providing detailed, precise, and complete information as input for automation technologies in the context of offsite construction [14]. However, the support of BIM for offsite construction is still limited in the current digital workflow. For instance, BIM lacks the capability to represent complex buildings or to plan for automated processes in offsite construction [15]. Therefore, to address this gap, the authors propose a method to facilitate the automation of wood construction by automatically analyzing building design information to obtain construction operational level information from the analysis so that it can be further used to feed into construction automation technologies. The proposed method utilizes a logic enabled approach to extract and infer information from IFC-based BIM instance models. The method involves the development of: (1) a set of algorithms (using logic rules) for the automated information extraction and properties inference, (2) representation of the IFC-based BIM instance models into logic facts, and (3) the logic reasoning using the logic rules and facts.

2. Background

2.1. Offsite construction automation

Offsite construction refers to the manufacture and preassembly of building components in a controlled environment, which are then transported and assembled on-site [16]. Offsite construction can be

* Corresponding author.
E-mail addresses: owongcho@purdue.edu (O. Wong Chong), zhan3062@purdue.edu (J. Zhang).

https://doi.org/10.1016/j.autcon.2021.103756
Received 4 September 2020; Received in revised form 9 February 2021; Accepted 11 May 2021
Available online 23 June 2021

Fig. 1. Sample of an IFC instance.

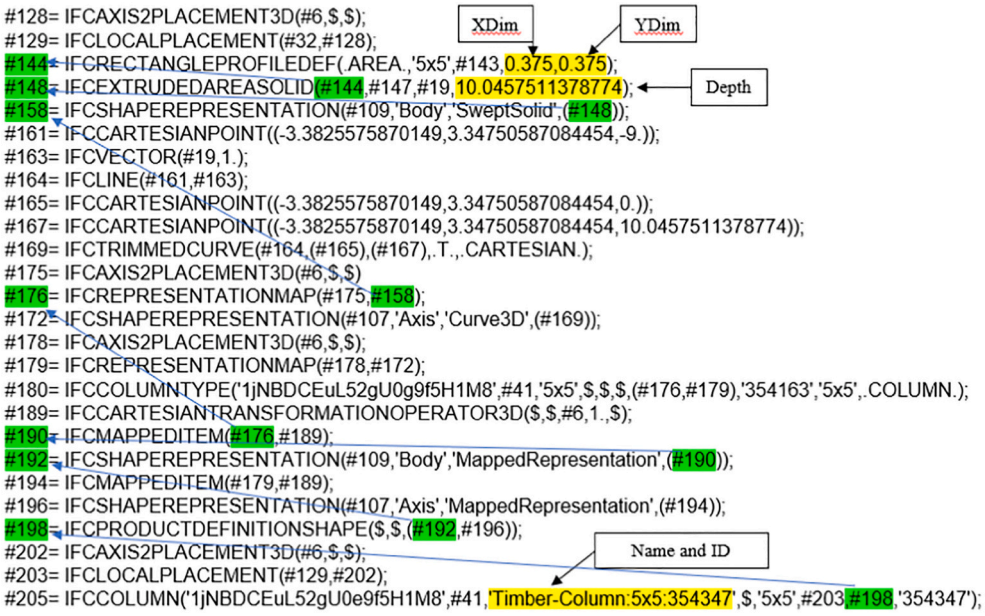

Fig. 2. Tracing pattern of dimensions information for the IfcColumn.

classified as non-volumetric, volumetric, or modular buildings based on the type of element and the level of prefabrication on the building [17]. For instance, the volumetric preassembly consists of structurally enclosed units (modules) [18]. In addition, offsite construction can be realized using concrete, steel, or wood, among other materials. Concrete and steel are predominately used in commercial and industrial facilities such as high-rises, warehouses, and bridges. While wood has been primarily used in residential houses to mid-rise buildings, especially in North America and Europe [19]. For instance, in the United States, nine out of ten houses are built of wood [20]. In addition, wood is a more sustainable and energy efficient material than concrete and steel in terms of the level of carbon dioxide emission [21] and embodied energy [22], respectively. These advantages have made wood one of the most commonly used construction materials.

The adoption of prefabrication and digitalization allows for more automation opportunities [23] in the manufacturing and assembly processes of wood construction. Some commonly implemented automation technologies in offsite wood construction include the use of robots, CNC, and other machines. Industrial arms are the most common type of robots used to automate the assembly and material handling operations in production lines or workstations. For instance, in Willmann et al. [24], a robotic system was used to assemble "The Sequential Roof", a timber roof structure that consists of slat elements. More applications of robotics in the automation of wood construction can be found in [25–27]. Moreover, CNC tools are used to remove layers of material (e.g., drilling, milling, and cutting) to shape the pieces according to designs (i.e., subtractive manufacturing) [28]. The implementation of CNC machines in wood manufacturing lines provides automated prefabrication of building elements (e.g., wood pieces and boards) [27]. Furthermore, other machines such as the semiautomated wood framing machine developed by [29], also facilitate wood framing processes. A common property shared by these automation technologies is that they require reliable and precise digital information as input for its successful operations. This requirement can be fulfilled using BIM as the source of information.

Offsite construction can be partially industrialized [30], meaning that the prefabricated components and assembled units can be treated as manufacturing products instead of a construction product. Consequently, offsite construction opens more opportunities for introducing manufacturing technologies, principles, and methods in the prefabrication of houses. Currently, the adoption of tools and technologies from the manufacturing industry for offsite construction such as production lines, mechanical engineering and manufacturing CAD packages, and other prefabrication-based technologies, creates time and cost savings compared to traditional on-site construction. However, these technologies lack the capacity to analyze building designs, which limit their applicability in the design development stage [14]. This inability of these automation technology to fully integrate with BIM creates inefficiency caused by a bottleneck of information transfer within and between the design and construction phases.

2.2. *Industry class foundation (IFC)*

BIM is a "modeling technology and associated set of processes to produce, communicate, and analyze building models" [14] and it serves as a "shared knowledge resource for information about a facility forming a reliable basis for decisions during its life-cycle from inception onward" [31]. As such, BIM has the potential to improve the collaboration between designers and contractors and to allow a seamless coordination (e.g., data exchange) between design and construction phases in offsite construction projects. However, this potential is not yet fully realized due to proprietary concerns and interoperability difficulties in the AEC industry [32]. As a result, it creates significant re-work, measurable waste, and has impeded the use of BIM to advance automation

[7,33,34]; costing the AEC industry $15.8 billion per year [34]. To improve BIM data exchange, existing approaches focus on data schema standardization and/or term-based semantics of AEC objects [35]. The CIMSteel Integration Standards (CIS/2) for steel construction data and the Industry Foundation Classes (IFC) for building and construction data are two prominent BIM standards. Both standards are defined using the Standard for Exchange of Product (STEP) description methods - ISO 10303 [36].

Industry Foundation Classes (IFC) is a vendor-neutral standard for data exchange developed and maintained by buildingSMART as a solution to the interoperability problem in the AEC industry [37,38]. The IFC schema is written using the EXPRESS data definition language and its data schema architecture contains four conceptual layers, namely resources, core, interoperability, and domain, to describe information such as geometry, material, and relationship of a BIM instance model [39]. IFC contains an essential set of elements such as beam, column, wall, floor, and roof, to describe a building. Furthermore, each element can be represented using different geometric representations, such as swept solid, Boundary representation (B-rep), or body clipping. Moreover, multiple cross-sectional profile definitions exist for each geometric representation. For instance, a column modeled as Swept Solid representation, can have a rectangular profile definition (IfcRectangleProfileDef) or an arbitrary closed profile definition (IfcArbitraryClosedProfileDef) to depict its cross section. These variety of representations come from the 3D modeling approaches adopted by the BIM authoring tools. To illustrate the IFC data model, an example of IFC instances with its corresponding entities, relationship, and tracing pattern for the dimensions of a column element are shown Fig. 1 and Fig. 2, respectively.

The current IFC format presents several limitations. One limitation is the existence of a user defined properties set, which make its use unpredictable. Another limitation is the multiple ways in which a building element can be represented. For example, a wall might be represented as a wall, a thick slab, or an upstand beam [40]. This common misuse introduce subjectivity in the IFC standard. Other limitations of IFC include the needs of improved information description, a more robust way to represent elements, and information and precision loss/inadequacy [41]. These limitations make the standardization of IFC challenging.

Even with the current limitations, the IFC schema is widely accepted as the most promising solution to the interoperability issue faced by the AEC industry [42]. Qualities such as openness, impartiality, and file format simplicity, have made the IFC standard a major focus of BIM research and industry applications. Currently, a growing number of BIM platforms (94) are being certified as IFC compliant; making them compatible with IFC [43].

2.3. BIM to support offsite construction automation

Many research efforts have used BIM-based approaches to improve offsite construction processes at different lifecycle stages. This section presents some of the applications of BIM workflows in offsite construction.

In the design and planning stages of offsite construction, Liu et al. [44] developed a rule-based algorithm and automated BIM-based approach to minimize waste in the design and planning of light-frame boarding (sheathing and drywall), taking into consideration contractors' practical knowledge. Their method helps reduce errors and time consumption in manual modeling of a construction-centric model and it is implemented as a Revit add-on. As an attempt to reduce the design cost, improve the layout accuracy and productivity, Alwisy et al. [13] proposed a framework to automate the design and drafting of wood frame panel modules. The input to the BIM instance model and the corresponding shop drawings for the manufacturing process is generated from 2D CAD layout drawings. Their methodology was implemented using Visual Basic embedded into AutoCAD. In addition, Abushwereb et al. [45] proposed a platform (FrameX) as a Revit add-on to automate the analysis, modeling, and design of light-frame wood structure for offsite construction. Their platform was developed using a rule-based approach that arguably improve time efficiency and accuracy in the early design stage of a project.

Besides design and planning, other research works have focused on the optimization of offsite construction. For instance, to optimize offsite construction designs for meeting client expectations, Isaac et al. [46] used a graph-based methodology based on BIM instance models to reduce delays and to avoid incurring additional cost and labors in the prefabrication of modular house modules. Another optimization approach is the integration of manufacturing CAD into the workflow (e.g., SolidWorks and TactonWorks Studio) to automate the selection of module configurations during the design process. This integration method is proposed by [47] to reduce: (1) conflicts between stakeholders, and (2) variability in the downstream process of the prefabrication of building components. Furthermore, Mekawy and Petzold [48] proposed a method to explore and optimize design alternatives of Box Prefabricates using their developed Autodesk Dynamo for Revit package called Box Module Generator. In addition, to reduce waste in the construction phase, Gbadamosi et al. [49] presented a framework for the assessment and optimization of design (BIM instance models) options based on lean principles and design for manufacturing concepts such as ease of assembly, ease of handling, speed of assembly, and assembly waste. Their assembly assessment framework was implemented using Revit, Dynamo Studio, and Microsoft Excel for the design optimization of exterior insulation finish systems.

The use of BIM in other offsite construction processes include quantity take-off and off-site manufacturing. In the first case, Wang et al. [50] proposed a method to reduce manual work in the automated quantity take-off process of wall and floor components using a SQL database and BIM. In the second case, Root et al. [51] proposed a method to better understand the offsite construction process of exterior insulated finishing systems manufacturing using Revit and the Metal Wood Framing plug-in from Strucsoft.

Although many research efforts have contributed to the advancement of BIM and offsite construction, most of the existing efforts only focused on workflows using proprietary BIM platforms and processes in the planning and design stages. The dependency on proprietary BIM applications prevents the realization of a truly seamless BIM interoperability. In addition, the lack of BIM research focused on the construction phase makes the implementation of automation in offsite construction a challenge. To address this limitation, this paper proposes a new methodology for BIM analysis to support wood construction automation, using a logic-based approach and IFC standard.

2.4. Logic representation

Logic has been used for the design of computers and reasoning of computer programs [52]. From a programming language perspective, the direct use of logic is called logic programming [52]. Logic programs consist of rules that establish relations between objects [52]. Formal logic such as predicate logic, allows the representation of knowledge and the derivation of correct conclusions from that knowledge [53]. First-order logic (FOL), a subset of predicate logic, is the most common type of logic representation [54].

2.4.1. First order logic (FOL)

FOL can be used to represent IFC-based BIM information, in the form of logic clauses, which in turn consist of predicates. The relations between the predicates are logically expressed using quantifiers and logic connectives. In FOL, the universal (∀ or for all) and existential (∃ or there exists) quantifiers are used to make assertions about variables in statements [55]. Likewise, the logic connectives: conjunction AND (^), disjunction OR (˅), negation NOT (¬), and implication (→), are used to make logical connections between predicates [55].

Furthermore, logic clauses can be expressed using Horn clauses (HCs). A HC is a conjunction of logic clauses of which at most one literal is positive [56]. HC can be expressed as $H \leftarrow (C_1 \bigwedge C_2 \bigwedge C_3 \cdots \wedge C_n)$, where H is the head of the clause and C_1, C_2, C_3. C_n are goals. This expression implies that the conclusion H holds if all the goals are met. The structure of a HC is simple and sufficiently expressive to represent all types of computations that allows programs to be general and efficient [57]. There are three types of HCs based on the representation structure: facts, rules, and queries [58].

2.4.2. Basic elements of logic programs

A HC can be classified as facts, queries, or rules, based on its structure. In a HC, when only the head H is present, it is a fact. Logic facts are "statement that describe object properties or relations between objects" [54]. It consists of predicates that are composed of a predicate name and one or more arguments. The number of arguments in a predicate is called its arity and the smallest unit of a logic fact is an atom, which consists of only one argument. In the case when a HC only contains a set of goals ($C_1 \bigwedge C_2 \bigwedge C_3 \cdots \wedge C_n$), it is called a query. Queries are used to retrieve information from a logic program and it will reach a conclusion regarding whether relations between objects hold (conjunction of goals). Lastly, a HC is classified as a rule when it contains both the head H and the body (i.e., a set of goals to be evaluated according to the logic facts). If all the goals are met, then the rule evaluates to true. On the contrary, if any of the goals cannot be achieved, the rule evaluates to false.

2.4.3. Second order logic (SOL)

An extension of FOL in terms of expressive power is the second-order logic (SOL). Unlike FOL, whose domain of quantification is the range of individuals, SOL can quantify subsets of individuals with certain properties or relations over the entire domain [59]. Therefore, SOL is particularly useful to find all instances of building components/elements with certain properties from building design logic facts. The application of SOL in logic programming is referred as second-order programming, which is represented by the "find all-solutions" predicates such as findall (Term,Goal,List) [52]. To illustrate this, the clause *column_material:-findall((Column,Materialname),(relassociatesmaterial(Relassociatesmaterial),material(Material),has_relatingmaterial(Relassociatesmaterial,Material),column(Column),has_relatedobjects(Relassociatesmaterial,Column), has_name(Material,Materialname),Materialname == 'lumber'), L)* will function as finding all the column instances from the existing logic facts with the material property 'lumber' and store them in the list *L*.

2.5. Logic reasoning

The use of logic representation and reasoning facilitates the analysis of building design information. Once the logic representation of the IFC data is enabled and the logic rules are defined, the logic reasoning is performed automatically. The essence of logic reasoning relies on the unification function and three deduction rules: 1) identity, 2) generalization, and 3) instantiation [52].

According to Sterling and Shapiro [52], the identity rule consists of the search of logic facts based on queries to determine logical consequences. The second deduction rule is generalization, which relates a logical consequence to an instance of an existential quantified variable for any substitutions. Lastly, the instantiation rule can be used to deduce any instance of a logic fact from a universally quantified fact. The automated deduction in logic reasoning is possible through the unification algorithm [52]. Unification provides efficient pattern matching and variable binding functions [60] to allow the reasoning based on logic rules and facts.

2.5.1. Prolog language

A partial, yet powerful realization of logic programming is through Prolog. Prolog (stands for programming in logic) language is a programming formalism based on the concept of logic programming created in the early 1970s by Alain Colmerauer [52]. The use of Prolog has been proven to be successful in applications such as artificial intelligence (AI), language processing, and expert systems [61,62]. Prolog differentiates from other programming languages through a different programming paradigm: it is declarative as opposed to the more conventional procedural and objected-oriented approaches.

2.5.2. Reasoning using a closed-world assumption

By default, Prolog adopts a closed-world assumption for the logic reasoning. A closed-world assumption considers any unproven assumptions to be false. Therefore, any missing information will be considered as false as well. This implies that the information to be reasoned about needs to be complete and logical.

2.6. Logic-based representation for automated reasoning

Logic-based representation has been widely and extensively used for automated reasoning. Automated reasoning is the ability to make inferences automatically through computing systems and it has been applied to solving many challenging problems in domains of computer science, mathematics, software and hardware verifications, among others [63]. Several disciplines (e.g., sensing, natural language processing, robotics) along with automated reasoning, formed the fundamental building blocks in the conception and rise of AI in the modern era, allowing the creation of intelligent entities that can perceive the environment and perform actions autonomously [64]. In the AEC domain, the application of AI methods (e.g., neural networks, genetic algorithms, and machine learning) have increased exponentially since the early 2000s (e.g., more than 41,827 existing related bibliographic records in the Scopus database) [65]. Although some computing paradigms such as machine learning and other statistical-based computing method are trending, logic-based representation and reasoning through logic programming is a no-less powerful computing paradigm due to its rigor, expressiveness, and ability to draw logical conclusions. Yet, comparatively speaking, it is significantly underexplored. Therefore, the authors are testing the use of logic-based representation and reasoning for automated inferences to unlock the benefits of logic-based automation in the construction domain.

2.7. Logic-based representation and reasoning in the AEC domain

Previous studies have explored the use of FOL in the AEC domain. One of the early applications of FOL is in the area of structural engineering design [66]. More recently, FOL has been used in the representation and reasoning of building design and regulatory information in the area of code compliance checking [55].

2.8. Comparison to ifcOWL ontology

In addition to logic-based approach, semantic modeling can be used to support logical inferences. The most commonly used form of semantic modeling is ontology and it uses sematic web technology [i.e., Web Ontology Language (OWL)] for the representation of things and their relations [67]. In an ontology, knowledge is represented in concept hierarchies and the relationships between the concepts, and axioms [68].

Both, semantic representation and logic representation could be utilized to support the reasoning process and facilitate human interpretation and understanding of the formal representation. However, they differ in their application intent fundamentally. Semantic modeling was originally conceived based on description logic to mainly model knowledge and represent the semantics of a specific domain. As a result, semantic representation (i.e., OWL ontology) relies on rule-based languages such as Semantic Web Rule Language (SWRL) to perform automated reasoning on top of it because it does not permit representation of if-then statements directly [69]. In contrast to semantic modeling, a

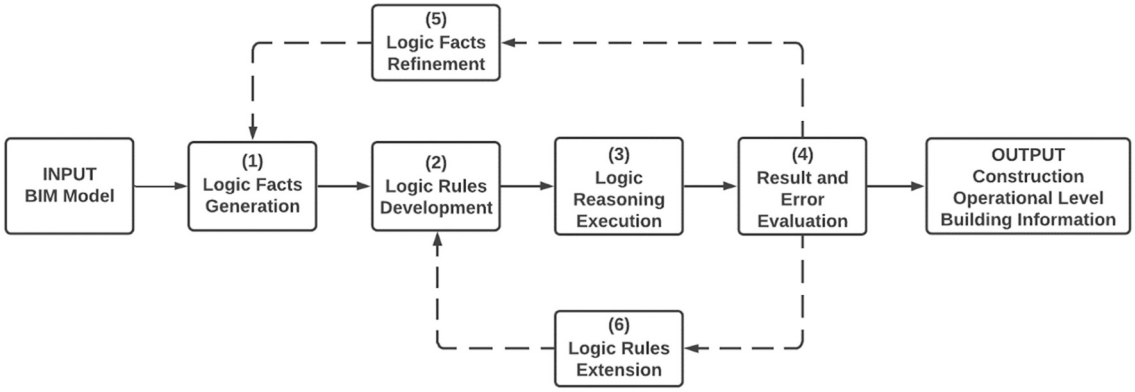

Fig. 3. Proposed method.

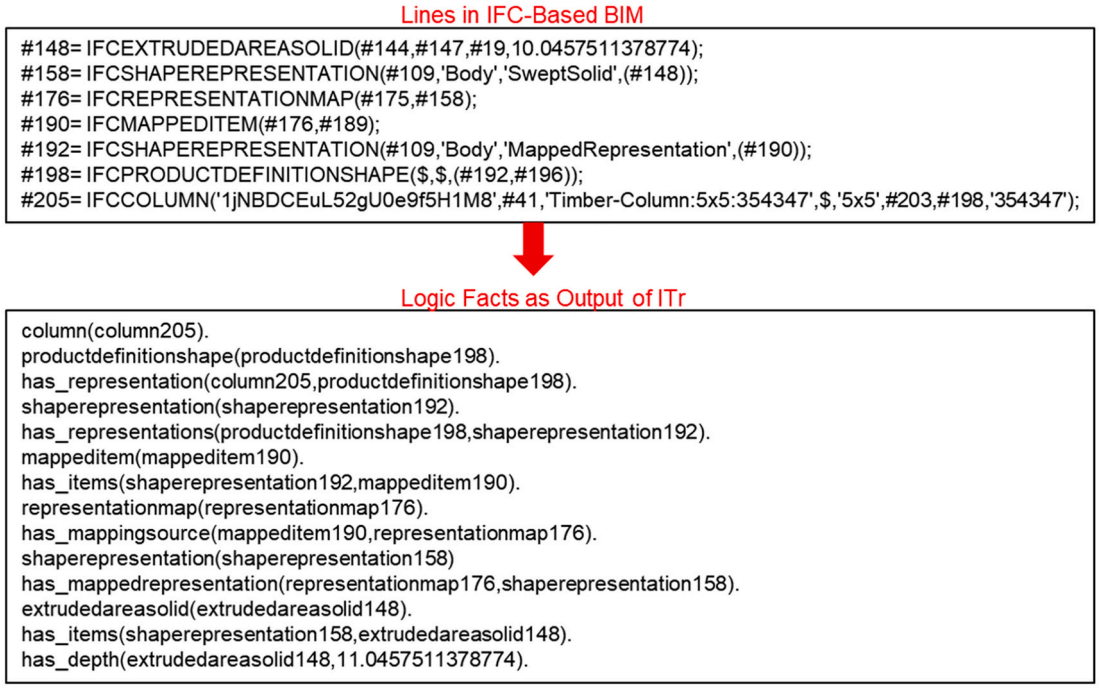

Fig. 4. Sample of IFC transformation into logic facts.

logic-based representation has been used extensively in automated reasoning. The use of logic-based representation in automated theorem provers has solved many famous scientific problems such as the problem of Robbins conjecture in 1966 [70]. In addition, logic-based representation has been successfully used to verify the correctness of computer programs and to assist in the construction of mathematical proofs [71]. Despite these achievements, it is still underexplored in the automation of architecture, engineering, and construction projects, compared with other popular approaches such as statistical machine learning and semantic modeling.

In an effort to link IFC standard and semantic web technologies to support flexibility, interoperability, and reusability of data and data exchange [72], and also to support logic-based reasoning, the ifcOWL ontology was created. The ifcOWL ontology and IFC schema are more similar in structure [73] compared to that between logic facts and IFC. In addition, the conversion from the IFC EXPRESS schema to ifcOWL ontology seems to be more direct and easier than the transformation of the IFC schema to logic facts from the logic-based approach. However, despite these advantages, there are challenges in its implementation for logic inference; for example: 1) the size of an ontology is considerably large and complex to load and use due to its numerous classes, objects and data properties, and logical axioms, and 2) the time efficiency is comparatively low due to the large number of assertions in the loading (to reasoning engine) and reasoning processes [73,74]. Moreover, an ifcOWL ontology needs separate and additional rule representations (e.g., SWRL) in the reasoning process, which are also limited in terms of expressivity, simplicity, and computational logic reasoning performance compared to the use of first-order logic and second-order logic approaches. For example, SWRL is not as effective as FOL in decidability (i.e., true or false decision problems). In addition, it is easy to find all instances of building components/elements with certain properties from the building design using SOL, which in contrast is not feasible with SWRL.

3. Proposed method

In order to extract and analyze building design information (e.g., geometric and physical properties) from BIM to support automation in offsite wood construction, the authors propose a novel data-driven method to address this challenge. The proposed method is based on

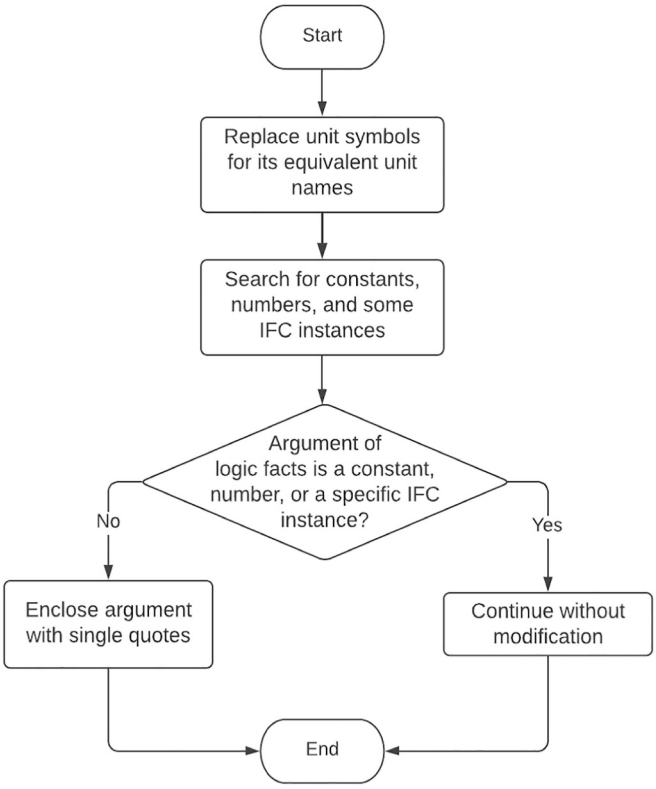

Fig. 5. Refinement algorithm flow chart.

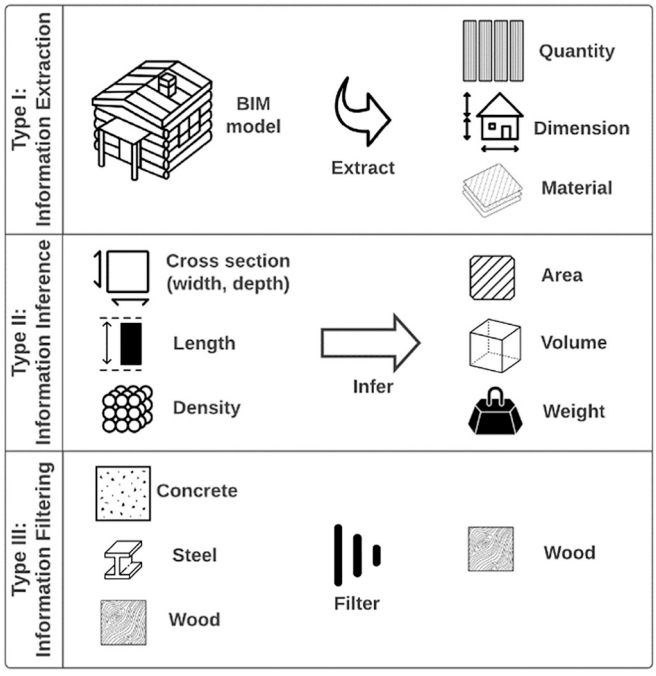

Fig. 6. Types and application examples of logic rules.

the use of logic representation and reasoning of BIM information in an automated fashion. Moreover, the proposed method allows performance improvement by iteratively refining and extending the refinement algorithm and logic rules, respectively. The proposed method consists of six main steps (Fig. 3): (1) Logic Facts Generation, (2) Logic Rules Development, (3) Logic Reasoning Execution, (4) Result and Error Evaluation, (5) Logic Facts Refinement, (6) Logic Rules Extension. Following, detailed explanations of the six steps along with some implementation examples to illustrate the specifics of the proposed method are provided.

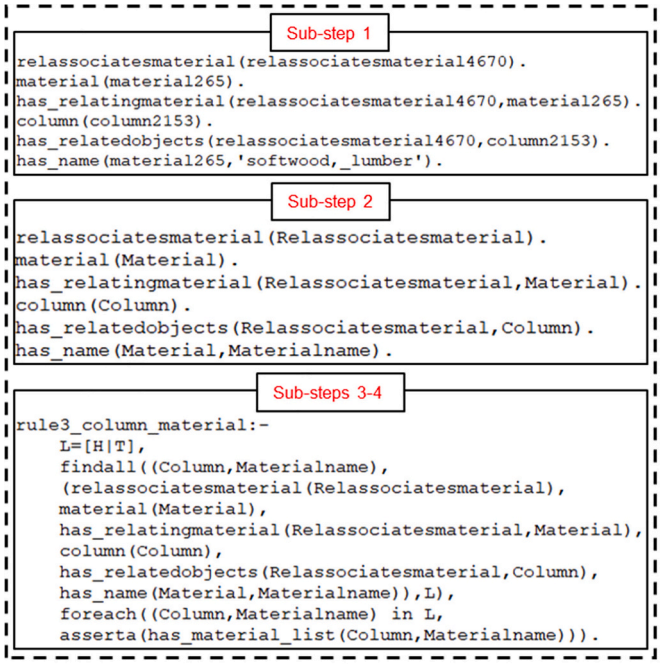

Fig. 7. Logic rule development for extracting the material information of columns.

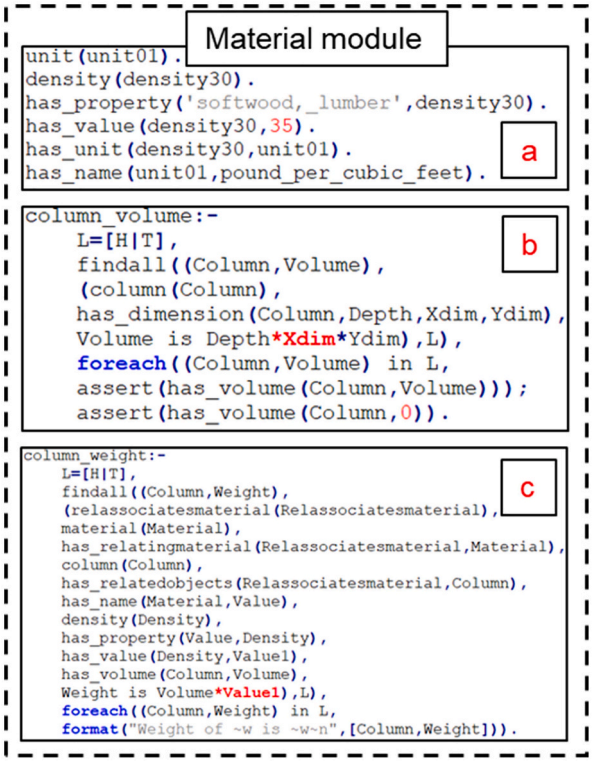

Fig. 8. Type II logic rule sample: a) material module, and rule to derive b) volume and c) weight for columns.

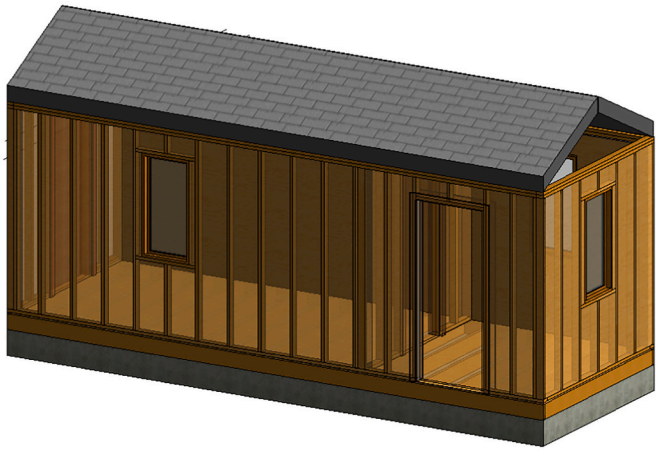

Fig. 9. BIM Dev model for algorithm development.

3.1. Logic facts generation

The generation of logic facts consists of three sub-steps: sub-step (1) BIM to IFC export, sub-step (2) IFC to logic facts conversion, and sub-step (3) logic facts refinement.

3.1.1. BIM to IFC export

The first sub-step consists of exporting the BIM instance model from the BIM authoring tool to an IFC format. Many BIM authoring tools such as Autodesk Revit, Graphisoft ArchiCAD, and Trimble SketchUp are IFC certified and therefore have built-in functions for exporting their respective BIM instance models to the IFC format [43]. In addition, different parameters can be adjusted such as the specific IFC schema version and model view definitions when exporting to the IFC format. For the implementation of the proposed method, the IFC 2 × 3 and MVD CV2.0 were selected.

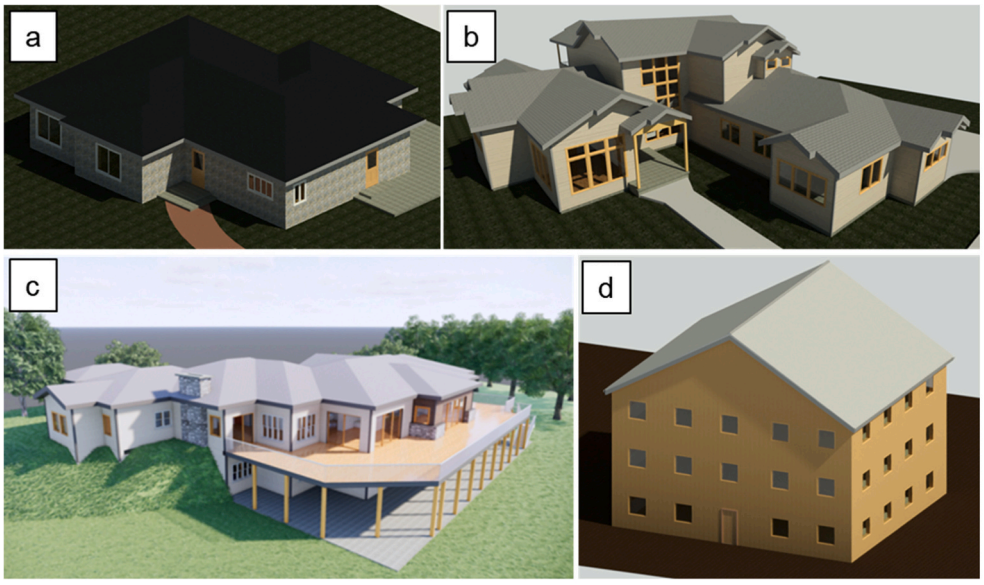

Fig. 10. BIM testing models: a) T1, b) T2, c) T3, and d) T4.

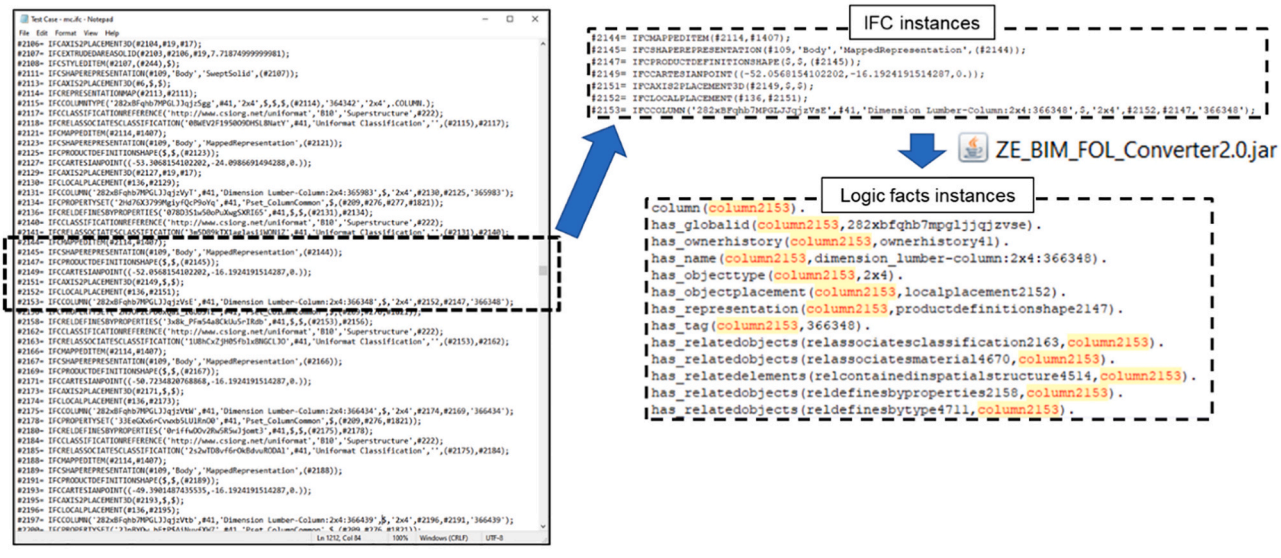

Fig. 11. Sample conversion from IFC to logic facts.

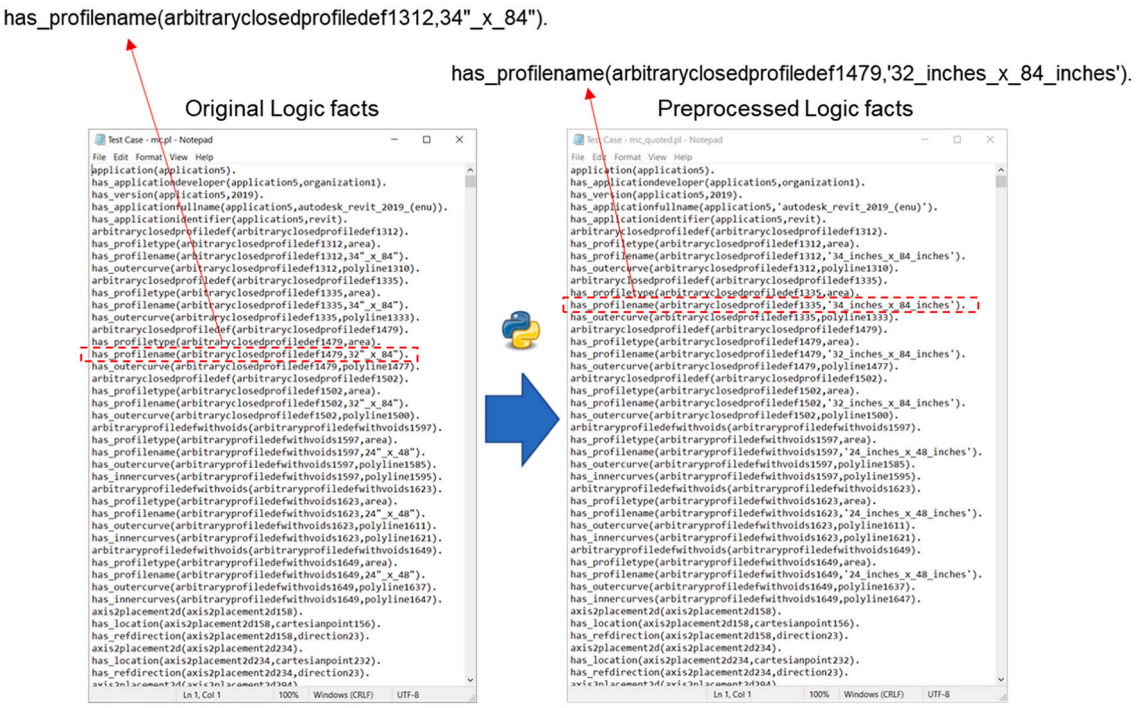

Fig. 12. Sample of preprocessed logic facts.

3.1.2. IFC to logic facts conversion

In this sub-step, the IFC data is represented using logic facts. The conversion of the IFC model into logic facts is performed using the open source application ZE_BIM_FOL_Converter2.0, developed by [75].

The prototype system ZE_BIM_FOL_Converter2.0 implements an information extraction (IE) and information transformation (ITr) algorithms of the building design information (IFC) [75]. In Fig. 4, an example of the IFC data conversion into logic facts by applying the IE and ITr algorithms is provided.

3.1.3. Logic facts refinement

The last sub-step refines the raw logic facts converted from the previous sub-step (IFC to logic facts conversion), to make them suitable for the logic reasoning and compliant with the Prolog syntax requirements. Two types of logic facts refinement include: 1) conversion of unit symbols to their corresponding unit names, and 2) addition of single quote symbol (*'string'*) to enclose strings in logic fact arguments such as constant, numbers, and some IFC instances.

The unit of measurements in the logic facts are expressed in symbolic forms, which are inherited from the original BIM authoring software. For instance, the unit for a window's dimensions, are commonly expressed as 24" × 48" in the original BIM software. Accordingly, the predicate *has_objecttype(window1724,24"_x_48")* also contains the symbol for inch (") in its second argument. By refining it, the unit symbols are changed to its equivalent unit name. For example, the window dimensions *24"_x_48"* is changed to *24_inches_x_48_inches*.

Arguments that start with a number followed by strings are not compliant with the Prolog syntax requirements. In this case, single quote is used to enclose the string arguments (*'string'*) of the logic facts to make the logic facts compliant with the Prolog syntax. For instance, without enclosing the second string argument of the predicate *has_globalid(window1724,2pbnvhtcp1kavcyplsrmwh)*, which is an instance of the global ID for the *window1724*, it will result in a syntax error because the argument starts with a numeric value which is not allowed in Prolog syntax. By enclosing the arguments with single quotes, Prolog treats the argument as a string.

To this end, a refinement algorithm was developed using the regular expression (re) module in Python (version 3.7) to refine the logic facts information (Fig. 5).

3.2. Logic rules development

For the development of the logic rules, a knowledge engineering approach is applied. Knowledge engineering is concerned with "data and information representation and encoding methodologies" to meet the needs of the user. Particularly, two heuristics from knowledge engineering were applied: 1) heuristic of specific situations and 2) heuristic of situation comparison [76]. The first heuristic uses specific cases to acquire the target knowledge. Relating to the logic rules development, specific predicates corresponding to building components (e.g., wall and floor) and building elements (e.g., column and beam) were used to develop the extraction and inference rules. For example, the predicate *has_representation(column2118,productdefinitionshape2112)*, which translates to column2118 has representation productdefinitonshape2112, is transformed to the general form *has_representation(Column,Productdefinitionshape)* by replacing the specific arguments with variables, that will serve as a template to capture the predicate *has_representation* for any columns and *productdefinitionshapes* instances of the logic facts. The second heuristic applies when comparing predicates that are shared by different building components. For instance, IfcSlab can be used to model roofs and floors. Therefore, to clarify if the predicate *slab(Slab)*, is a roof or slab, additional predicates such as *has_predefinedtype(Slab,floor)* and *has_predefinedtype(Slab,roof)* allow its correct classification, respectively.

Logic rules are developed to allow information extraction and filtering, as well as inference of new information from the building design logic facts. Logic rules in this paper can be divided into three types, based on the purpose (Fig. 6): information extraction (IE) type (type I), information inference type (type II), and information filtering type (type III). The type I logic rules are used to extract information such as quantities, dimensions, and materials of the building components/elements (i.e., wall, roof, column) from the building design. Therefore, to capture the various IFC entities representation in a BIM instance model [i.e., building components/elements (e.g., wall, column),

Table 1
List of rules developed using the Dev model.

Rule	Rule Type*	Component/Element	IFC Object Source	Geometric Representation	Profile Definition	Target Information
1	I	Interior wall	IfcWallStandardCase	Swept solid	Rectangle	Dimensions
2	I					Quantity
3	I					Material
4	I	Opening (Interior wall)	IfcOpeningElement			Quantity
5	I			Swept solid	Rectangle	Dimensions
6	II					Area
7	II					Total area
8	II	Interior wall	IfcWallStandardCase			Area
9	II					Net area
10	I	Exterior wall		Swept solid	Rectangle	Dimensions
11	I					Quantity
12	I					Material
13	I	Opening (Exterior wall)	IfcOpeningElement			Quantity
14	I			Swept solid	Rectangle	Dimensions
15	II					Area
16	II					Total area
17	II	Exterior wall	IfcWallStandardCase			Area
18	II					Net area
19	I	Floor	IfcSlab	Swept solid	Rectangle	Dimensions
20	I					Quantity
21	I					Material
22	I	Opening (Floor)	IfcOpeningElement			Quantity
23	I			Swept solid	Rectangle	Dimensions
24	II					Area
25	II					Total area
26	II	Floor	IfcSlab			Area
27	II					Net area
28	I	Roof		Swept solid	Rectangle	Dimensions
29	I					Quantity
30	I					Material
31	II					Area
32	I	Column	IfcColumn	Swept solid	Rectangle	Dimensions
33	I					Quantity
34	I					Material
35	II					Volume
36	II					Weight
37	I	Beam	IfcBeam	Swept solid	Rectangle	Dimensions
38	I					Quantity
39	I					Material
40	II					Volume
41	II					Weight
42	I	Opening	IfcOpeningElement			Quantity
43	I			Swept solid	Rectangle	Dimensions
44	III	Floor	IfcSlab			Material

*Type I: extraction; Type II: inference; Type III: filtering

geometric representations (e.g., Brep, SweptSolid), and cross-sectional profile definitions (e.g., rectangular, circular)], the corresponding logic rules need to be created to account for such combinations. The type II logic rules are used to infer physical properties (e.g., area, volume, and weight) from the building components/elements. These properties are derived from information extracted using type I logic rules and a supporting material module. The supporting material module contains material and density information for wood materials, which are encoded as logic facts. Lastly, the type III logic rules are used to filter out building components/elements of other materials (e.g., concrete and steel) from those of wood materials, and filter out unrelated object types such as recess and corner board.

The development process for logic rules consists of four sub-steps: 1) relevant logic facts identification, 2) terms replacement, 3) logic facts connection, and 4) SOL application. In the first sub-step, the relevant logic facts are identified and serve as (a) constraints in the unification process with only the relevant logic facts. For example, the set of relevant logic clauses [i.e., *relassociatesmaterial(relassociatesmaterial4670* to *has_name(material265,'softwood,_lumber')*] as shown in Fig. 7, are used to extract the material information for column instances; and (b) sub goals that need to be met for the logic rules to succeed. In the terms replacement sub-step, constants and numbers are replaced by variables (first letter upper-cased). The third sub-step consists of joining the logic facts using the logical conjunction (,). Finally, in the last sub-step, the all-solution predicates (i.e., findall, bagof, and setof) are applied to extend the rule for all building component/element instances. To illustrate the process, the development of a logic rule to extract the material, and to infer the volume and weight for columns with swept solid geometric representation and rectangular profile definition are summarized in Fig. 7 and Fig. 8, respectively. Similarly, different logic rules are developed for columns with other types of geometric representations (e.g., B-rep) and profile definitions.

3.3. Logic reasoning execution

In addition to the logic facts and logic rules, supporting modules were developed to provide reasoning support: a) unit conversion, which are logic clauses that convert any length unit to feet unit length; and b) material properties, which contains the material type and density information. Once, all the components and information are ready, the logic reasoning is performed in an automated way. First, the logic facts are loaded into the logic reasoner. Then, the logic rules are executed internally along with the supporting modules by the logic reasoner. The reasoner applies the unification and substitution functions to prove the goals defined in the logic rules. Finally, the results are returned in terms of logic rules' success or failure.

Table 2
Results of the Dev model (gold standard).

Component/ Element	Item	No. of Relevant LC	No. of Correctly Extracted LC	No. of Extracted LC
Exterior wall	Quantity	4	4	4
Exterior wall	Dimensions (Length, height, width)	12	12	12
Exterior wall	Material	12	12	12
Exterior wall	Opening quantity	5	5	5
Exterior wall	Opening dimensions (Length, height, width)	15	15	15
Interior wall	Quantity	1	1	1
Interior wall	Dimensions (Length, height, width)	3	3	3
Interior wall	Material	3	3	3
Interior wall	Opening quantity	1	1	1
Interior wall	Opening dimensions (Length, height, width)	3	3	3
Floor	Quantity	1	1	1
Floor	Dimensions (Length, height, width)	3	3	3
Floor	Material	1	1	1
Floor	Opening quantity	–	–	–
Floor	Opening dimensions (Length, height, width)	–	–	–
Roof	Quantity	2	2	2
Roof	Dimensions (Length, height, width)	6	6	6
Roof	Material	6	6	6
Roof	Opening quantity	–	–	–
Roof	Opening dimensions (Length, height, width)	–	–	–
Column	Quantity	60	60	60
Column	Dimensions (Length, height, width)	180	180	180
Column	Material	60	60	60
Beam	Quantity	50	50	50
Beam	Dimensions (Length, height, width)	150	150	150
Beam	Material	50	50	50
Total/Average		628	628	628

3.4. Result and error evaluation

The performances of the extraction algorithms are evaluated using the precision, recall, and F1-measure. The precision performance metric measures how good the extraction result is and it is defined as the ratio of the number of correctly extracted logic fact instances in the total number of logic fact instances extracted (Eq. (1)). Correspondently, recall is a measurement of the degree of coverage in the target information and the definition is the ratio of the number of correctly extracted logic fact instances in the total number of existing logic fact instances in the source file (Eq. (2)). The use of precision or recall alone does not provide a full-picture evaluation, therefore a third performance metric called F1-measure is used to provide an overall performance based on the precision and recall. The F1-measure is defined as the harmonic mean between the precision and recall (Eq. (3)).

$$Precision\ (P) = \frac{Correctly\ extracted\ LF}{Total\ LF\ extracted} \tag{1}$$

$$Recall\ (R) = \frac{Correctly\ extracted\ LF}{Total\ exiting\ LF} \tag{2}$$

$$F1\ measure = \frac{2PR}{P+R} \tag{3}$$

3.5. Logic facts refinement

This step applies when additional logic predicates that are not suitable for the logic reasoning are detected and were not captured in the logic facts refinement sub-step of Step 1. Those predicates make the related logic rules yield unexpected results. Therefore, the refinement algorithm needs to be extended to modify those predicates, making them suitable for the logic reasoning. For example, the numbering of the predicate *has_segments#(Term1, Term2)*, makes the logic reasoning ineffective because the numbered predicates (i.e., *has_segments1,..,has_segments#*) are treated as independent predicates and therefore the reasoner is unable to unify all the instances related to that predicate.

3.6. Logic rules extension

At the beginning, the proposed method contains only the set of logic rules created based on the development BIM instance model. For example, the proposed method can be used to analyze columns with swept solid geometric representation (used in the development model) and not for columns with geometric representations that were not present in the development model (e.g., B-rep). To improve the performance, supplementary logic rules are gradually added to the existing rule set for those cases in which the current logic rules fail to capture. Once the newly added logic rules are incorporated to the set of rules, the method will be able to analyze effectively the previously incorrectly identified geometric representations, as well as materials information. This process improves the performance and overall effectiveness of the proposed method by iteratively and accumulatively increasing the number of logic rules in the set.

4. Experimental implementation and validation

In this section, the experimental implementation of the proposed method along with the results are presented. Five BIM instance models were used for this purpose, one for the development and four for the validation. Additionally, the time efficiency of the automated analysis was also measured.

4.1. Implementation software

The Prolog programming language implementation used for the proposed method is B-Prolog. B-Prolog uses HC representation and it was selected because of its inherent reasoning capabilities, compatibility with C and Java programming languages, and is based on classic Prolog [55].

4.2. Test cases

The five BIM instance models used for the experiments are described here. The proposed method was initially implemented using the development model (Dev), a wood frame structure modeled in Autodesk Revit software (version 2019) as shown in Fig. 9. This Dev model serves as a gold standard.

For the testing and validation, four additional BIM instance models, namely T1, T2, T3, and T4, were used (Fig. 10) to: (1) evaluate the robustness of the proposed method, and (2) to test the effectiveness of the developed logic rules on unseen models. The first three models (T1-

Table 3
Result for the model T1.

Component/Element	Item	No. of Relevant LC	No. of Correctly Extracted LC	No. of Extracted LC	P (%)	R (%)	F (%)
Exterior wall	Quantity	12	12	30	40	100	57.1
Exterior wall	Dimensions (Length, height, width)	36	36	90	40	100	57.1
Exterior wall	Material	60	60	78	76.9	100	87
Exterior wall	Opening quantity	14	14	14	100	100	100
Exterior wall	Opening dimensions (Length, height, width)	42	42	42	100	100	100
Interior wall	Quantity	30	30	30	100	100	100
Interior wall	Dimensions (Length, height, width)	90	81	81	100	90	94.7
Interior wall	Material	90	90	90	100	100	100
Interior wall	Opening quantity	17	17	17	100	100	100
Interior wall	Opening dimensions (Length, height, width)	51	45	45	100	88.2	93.8
Floor	Quantity	1	1	1	100	100	100
Floor	Dimensions (Length, height, width)	3	0	0	0	0	0
Floor	Material	3	3	4	75	100	85.7
Floor	Opening quantity	–	–	–	–	–	–
Floor	Opening dimensions (Length, height, width)	–	–	–	–	–	–
Roof	Quantity	11	11	11	100	100	100
Roof	Dimensions (Length, height, width)	33	0	0	0	0	0
Roof	Material	33	33	33	100	100	100
Roof	Opening quantity	–	–	–	–	–	–
Roof	Opening dimensions (Length, height, width)	–	–	–	–	–	–
Column	Quantity	–	–	–	–	–	–
Column	Dimensions (Length, height, width)	–	–	–	–	–	–
Column	Material	–	–	–	–	–	–
Beam	Quantity	–	–	–	–	–	–
Beam	Dimensions (Length, height, width)	–	–	–	–	–	–
Beam	Material	–	–	–	–	–	–
Total/Average		526	475	566	83.9	90.3	87.0

Table 4
Result for the model T2.

Component/Element	Item	No. of Relevant LC	No. of Correctly Extracted LC	No. of Extracted LC	P (%)	R (%)	F (%)
Exterior wall	Quantity	69	69	69	100	100	100
Exterior wall	Dimensions (Length, height, width)	207	99	99	100	47.8	64.7
Exterior wall	Material	276	276	276	100	100	100
Exterior wall	Opening quantity	90	90	90	100	100	100
Exterior wall	Opening dimensions (Length, height, width)	270	264	264	100	97.8	98.9
Interior wall	Quantity	63	62	62	100	98.4	99.2
Interior wall	Dimensions (Length, height, width)	189	183	183	100	96.8	98.4
Interior wall	Material	189	186	186	100	98.4	99.2
Interior wall	Opening quantity	30	30	30	100	100	100
Interior wall	Opening dimensions (Length, height, width)	90	90	90	100	100	100
Floor	Quantity	4	4	4	100	100	100
Floor	Dimensions (Length, height, width)	12	12	12	100	100	100
Floor	Material	12	12	12	100	100	100
Floor	Opening quantity	–	–	–	–	–	–
Floor	Opening dimensions (Length, height, width)	–	–	–	–	–	–
Roof	Quantity	16	9	9	100	56.3	72
Roof	Dimensions (Length, height, width)	48	27	27	100	56.3	72
Roof	Material	34	27	27	100	79.4	88.5
Roof	Opening quantity	–	–	–	–	–	–
Roof	Opening dimensions (Length, height, width)	–	–	–	–	–	–
Column	Quantity	8	8	8	100	100	100
Column	Dimensions (Length, height, width)	24	24	24	100	100	100
Column	Material	8	8	8	100	100	100
Beam	Quantity	9	9	9	100	100	100
Beam	Dimensions (Length, height, width)	27	27	27	100	100	100
Beam	Material	9	9	9	100	100	100
Total/Average		1684	1525	1525	100	90.6	95.0

T3) consisted of wood residential projects and T4 is a three stories commercial wood building project. The four BIM instance models were created in Autodesk Revit to an level of development/detail (LOD) 300 (T1-T3) and LOD 400 BIM (T4), respectively.

4.3. From BIM instance model to logic facts (step 1)

First, each BIM instance model was exported from Revit to IFC format. Next, the IFC models were converted to logic facts using the ZE_BIM_FOL_Converter2.0 prototype as shown in Fig. 11. The top right of Fig. 11 shows some IFC instances from an IFC file; while the bottom right of Fig. 11 shows corresponding logic facts successfully converted. Then the logic facts were preprocessed using the Python algorithm (Fig. 5). Some sample logic facts before and after the preprocessing are shown in Fig. 12.

4.4. Logic rules development and logic reasoning execution (step 2–3)

The logic rules algorithms to extract and infer information from BIM instance models were developed using a randomly selected subset from

Table 5
Result for the model T3.

Component/Element	Item	No. of Relevant LC	No. of Correctly Extracted LC	No. of Extracted LC	P (%)	R (%)	F (%)
Exterior wall	Quantity	97	97	97	100	100	100
Exterior wall	Dimensions (Length, height, width)	291	291	291	100	100	100
Exterior wall	Material	427	427	427	100	100	100
Exterior wall	Opening quantity	58	58	58	100	100	100
Exterior wall	Opening dimensions (Length, height, width)	174	174	174	100	100	100
Interior wall	Quantity	115	115	149	77.2	100	87.1
Interior wall	Dimensions (Length, height, width)	345	336	336	100	97.4	98.7
Interior wall	Material	345	345	345	100	100	100
Interior wall	Opening quantity	33	33	33	100	100	100
Interior wall	Opening dimensions (Length, height, width)	99	99	99	100	100	100
Floor	Quantity	12	12	12	100	100	100
Floor	Dimensions (Length, height, width)	36	36	36	100	100	100
Floor	Material	41	41	41	100	100	100
Floor	Opening quantity	–	–	–	–	–	–
Floor	Opening dimensions (Length, height, width)	–	–	–	–	–	–
Roof	Quantity	4	4	4	100	100	100
Roof	Dimensions (Length, height, width)	12	12	12	100	100	100
Roof	Material	4	4	4	100	100	100
Roof	Opening quantity	–	–	–	–	–	–
Roof	Opening dimensions (Length, height, width)	–	–	–	–	–	–
Column	Quantity	17	17	26	65.4	100	79.1
Column	Dimensions (Length, height, width)	51	51	57	89.5	100	94.4
Column	Material	17	17	26	65.4	100	79.1
Beam	Quantity	–	–	–	–	–	–
Beam	Dimensions (Length, height, width)	–	–	–	–	–	–
Beam	Material	–	–	–	–	–	–
Total/Average		2178	2169	2227	97.4	99.6	98.5

Table 6
Result for the model T4.

Component/Element	Item	No. of Relevant LC	No. of Correctly Extracted LC	No. of Extracted LC	P (%)	R (%)	F (%)
Exterior wall	Quantity	12	12	18	66.7	100	80
Extterior wall	Dimensions (Length, height, width)	36	36	54	66.7	100	80
Extterior wall	Material	36	36	42	85.7	100	92.3
Extterior wall	Opening quantity	39	39	609	6.4	100	12.0
Extterior wall	Opening dimensions (Length, height, width)	117	117	2295	5.1	100	9.7
Intterior wall	Quantity	12	12	12	100	100	100
Intterior wall	Dimensions (Length, height, width)	36	36	36	100	100	100
Intterior wall	Material	60	60	60	100	100	100
Intterior wall	Opening quantity	41	41	41	100	100	100
Intterior wall	Opening dimensions (Length, height, width)	123	123	123	100	100	100
Floor	Quantity	3	3	8	37.5	100	54.5
Floor	Dimensions (Length, height, width)	9	9	24	37.5	100	54.5
Floor	Material	3	3	8	37.5	100	54.5
Floor	Opening quantity	–	–	–	–	–	–
Floor	Opening dimensions (Length, height, width)	–	–	–	–	–	–
Roof	Quantity	2	2	2	100	100	100
Roof	Dimensions (Length, height, width)	6	6	6	100	100	100
Roof	Material	4	4	4	100	100	100
Roof	Opening quantity	–	–	–	–	–	–
Roof	Opening dimensions (Length, height, width)	–	–	–	–	–	–
Column	Quantity	538	538	538	100	100	100
Column	Dimensions (Length, height, width)	1614	1614	1614	100	100	100
Column	Material	538	538	538	100	100	100
Beam	Quantity	490	490	607	80.7	100	89.3
Beam	Dimensions (Length, height, width)	1470	1005	1356	74.1	68.4	71.1
Beam	Material	490	490	607	80.7	100	89.3
Total/Average		5679	5214	8602	60.6	91.8	73.0

Table 7
Summary of the testing results.

#	Model	P (%)	R (%)	F1 (%)
1	Dev*	100	100	100
2	T1	83.9	90.3	87.0
3	T2	100	90.6	95.0
4	T3	97.4	99.6	98.5
5	T4	60.6	91.8	73.0

* Development model.

the Dev model logic facts. A total of 44 logic rules were developed initially, in which 26 correspond to extraction rules, 17 correspond to inference rules, and 1 filtering rule (Table 1). After the logic rules development, the logic rules were tested using the Dev model in the logic reasoning step. These steps were further applied to the four test models (T1-T4) in a similar way.

4.5. Result and error evaluation (step 4)

In this section, the experimental results for the testing of the five

Table 8
List of new rules developed from the testing models (T1-T4).

Rules	Rule Type*	Component/Element	IFC Object Source	Geometric Representation	Profile Definition	Target Information
1	I	Interior wall	IfcWallStandardCase	Sweptsolid	Arbitrary closed	Dimensions
2	I			Clipping	Rectangle	
3	I			Clipping	Arbitrary closed	
4	I		IfcWall	Clipping	Rectangle	
5	I					Material
6	I	Exterior wall	IfcWallStandardCase	Clipping	Rectangle	Dimensions
7	I			Clipping	Arbitrary closed	
8	III					Material
9	III					
10	III	Exterior wall	IfcWall			Object type
11	III	Opening	IfcOpeningElement			
12	I	Floor	IfcSlab	Sweptsolid	Arbitrary closed	Dimensions
13	III					Material
14	III					
15	III					
16	I	Roof	IfcRoof	Facetedbrep		Dimensions
17	I		IfcSlab	Sweptsolid	Arbitrary closed	
18	I		IfcSlab			Material
19	I	Opening	IfcOpeningElement	Sweptsolid	Arbitrary closed	Dimensions
20	I			Facetedbrep		
21	III	Column	IfcColumn			Material
22	III					
23	I	Beam	IfcBeam	Sweptsolid	Rectangle (Assembly)	Dimensions
24	I			Sweptsolid	Arbitrary closed	
25	I			Clipping	Arbitrary closed	
26	I			Clipping	Rectangle	
27	I			Facetedbrep		
28	III					Material

*Type I: extraction; Type II: inference; Type III: filtering

Table 9
Time performance results.

Model	Total Number of LC	Number of Relevant LC	Time (s)
Dev	11,647	628	0.11
T1	62,877	526	0.82
T2	212,787	1684	6.49
T3	548,053	2178	39.88
T4	524,561	5679	78.24

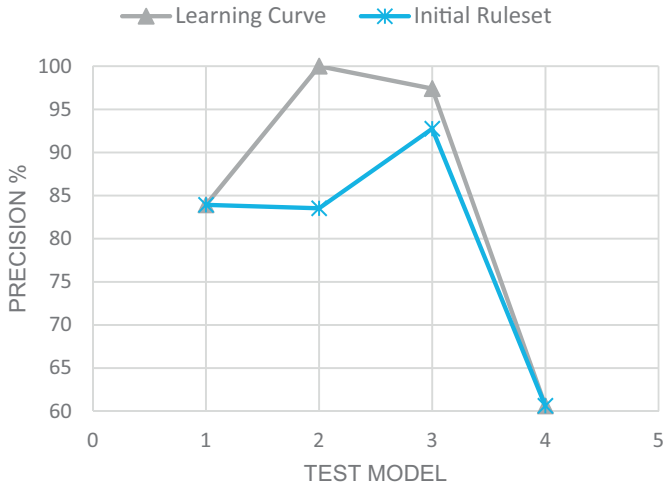

Fig. 13. Learning curve for precision.

models based on the performance metrics (precision, recall, and F1 measure) were presented. First, the detailed results for the development model Dev are shown in Table 2. Following, the detailed result for the validation tests of the proposed method using the four BIM instance models (T1-T4) are presented in Table 3, Table 4, Table 5, and Table 6,

Fig. 14. Learning curve for recall.

respectively. Each table contains the building component/element (column 1) and the target information (column 2) defined. Column 3 to 5 show the total number of logic clauses present in each converted model, the number of correctly extracted logic clauses, and the number of extracted logic clauses, respectively. The last three columns present the precision P, recall R, and F1 measure, respectively.

A summary of the performance result in terms of precision, recall, and F1-measure for the development model and the four test cases is shown in Table 7. The output is a report in a comma-separated values file (.csv) that contains the result of the analysis.

4.6. Performance improvement: Algorithm refinement and logic rules extension (step 5–6)

After the evaluation step of each test model, the refinement

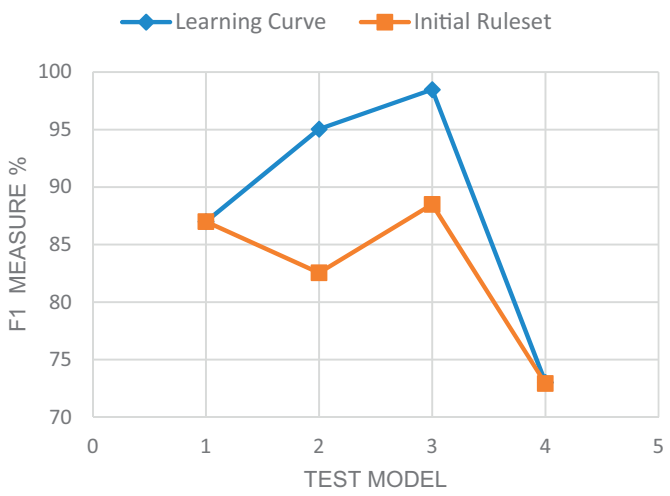

Fig. 15. Learning curve for F1 measure.

algorithm and logic rules were extended to improve the reasoning performance. For the refinement algorithm, two specific predicates *has_cfsfaces#* and *has_segments#* used in the B-rep and swept solid geometric representations, respectively, were modified to facilitate the implementation of the inherent unification and SOL functions of Prolog. This is done by extending the refinement algorithm in Step 5, which eliminates the numbering sequence of the predicates (*has_cfsfaces* and *has_segments*). Then, in Step 6, 28 new logic rules (18 extraction rules and 10 filtering rules) were added to the initial set of logic rules, to improve the reasoning capability of the ruleset (Table 8). Examples of geometric representation and profile definition variations encountered for walls, roofs, and floors information, include B-rep and arbitrary closed profile, respectively. After application of Step 5 and 6, the proposed method was able to extract and infer all the information from the BIM instance models.

4.7. Time performance testing

A time performance testing of the proposed method implementation was empirically performed. The results show that the processing time increases with the total number of logic clauses and the number of relevant logic clauses (Table 9). The total number of logic clauses correspond to the loading time of the logic facts into the Prolog system and the number of relevant logic clauses are related to the searching and matching of the logic clauses according to the relationship from the logic rules. The longest analysis time corresponds to the T4 model and it took around a minute to complete, despite having to load 524,561 logic clauses and to analyze 5679 relevant clauses. The experiments were conducted using a laptop with a random access memory (RAM) of 15.7 gigabytes and an Intel(R) Core(TM) i7-9750H processor with 2.60 GHz of central processing unit (CPU) speed.

4.8. Discussion

A closer comparison on the performances between the implementation of the method using a learning curve and the implementation with only the initially developed logic rules are shown in Figs. 13, 14, and 15. The results showed an overall increase in precision, recall, and F1 measure for the test models (T1-T4), except for the model T4 because it contained 2295 recesses that are mapped as openings in the IFC model, causing the overall performance to drop.

Besides the information extraction and inference, the algorithm was also able to classify all components' properties and relations correctly. Properties and relations include the function of walls (i.e., internal or external) and slabs (i.e., roof or floor), material properties set, relationship between walls and openings, and others. Lastly, although the derived information is presented in a general file format (.csv), which can be used for quantity takeoff and fabrication purposes (e.g., cutting), it can be easily extended to other file formats (e.g., .nc for CNC code and .py for robotic controllers) based on the requirements of the downstream applications.

5. Contributions

The proposed method contributes to the body of knowledge in four main ways. First, it provides a promising way, in terms of recall, precision, and time efficiency, to analyze BIM instance models and infer construction level information to facilitate wood construction automation, by using logic representation and reasoning based on the powerful but underexplored FOL and SOL. Furthermore, the generated information (e.g., length of the wood elements) can be further transferred as input to downstream applications (e.g., CNC or robots) to perform operations (e.g., cutting) in a continuous and digital workflow. Second, this research confirmed the robustness of logic representation and reasoning to represent and reason about IFC schema information in the AEC domain, such a new application area is demonstrated in addition to the existing automated code compliance checking research [69]. Third, this research leveraged SOL for efficient information extraction from IFC-based BIM instance models by extending the binary nature of the FOL (i.e., satisfy or fail) into set quantification (i.e., all instances of a query). Fourth, unlike existing BIM Application Programming Interface (APIs) that are integrated in proprietary software to achieve data extractions from the software's native formats, the proposed approach utilizes IFC as the BIM input data representation, which contributes to the standardization and interoperability of information exchange in the AEC domain.

6. Conclusions, limitations, and future direction

The objective of this research was to analyze building design information using logic representation and reasoning and to infer construction level information for supporting wood construction automation during preconstruction. To accomplish this, a novel data-driven method was proposed that utilizes FOL and SOL approaches to reason about logic enabled BIM instance models. More specifically, the proposed method can be used for automated information extraction (e.g., quantities and dimensions) and information inference (e.g., area and weight) from the logic enabled IFC representation of building objects. The proposed method was successfully implemented using B-Prolog and a development BIM instance model. In addition, the proposed method was tested and validated in four unseen BIM instance models. The experimental results have shown that the proposed method can achieve high performance in precision, recall, and F1 measure. These experiments confirmed that the iterative nature of the proposed method serves to continuously improve the results until a satisfactory performance is achieved. By adding more supplementary logic rules to the current set of rules, the analysis capabilities of the proposed method become more robust every time rules are created and added based on previously unseen cases of IFC object instances. The research has also empirically shown that the use of logic-enabled algorithm is time efficient.

The impact of this work in the AEC domain could be far-reaching. First, the proposed method opens the door to greater automation in decision making and computational tasks in the AEC domain by providing a ready-for-reasoning representation of building information. Second, the application of this work could be extended to support automation in the construction domain in general such as in robotic automation, lean construction, or lifecycle analysis, among others, by customizing reasoning rules. Third, the proposed method builds on the open standard IFC and logic representation, which favor the interoperability and human readability (as long as the predicate names are meaningful), respectively.

One limitation of this study is the use of bounding box to determine the dimensions of irregular shaped BIM objects. In many cases, wood elements in construction such as studs are parallelepiped and the dimensions of the bounding box matched with the real ones, however, in some cases, the use of the bounding box dimensions can be inaccurate, especially for irregularly shaped elements. This limitation becomes important when computing the volume and weight of irregularly shaped elements because the magnitude of the error increases with the degree of irregularity. Another limitation is that the quality and quantity of the extracted information depend on the level of development/detail of the BIM instance model. For example, in an LOD 300 BIM instance model, the columns (studs) information would not be modeled, therefore the method would not be able to capture that information. Lastly, despite that the proposed method is applicable to any level of prefabrications of offsite construction, the test cases were focused on wood construction mainly. Despite these limitations, the proposed method has shown to be promising and reliable for automated BIM information analysis.

Further work is needed to improve the comprehensiveness of the proposed implementation for irregularly shaped objects. Additionally, the logic-enabled BIM information is lacking in terms of the semantic information for offsite construction and the level of details of BIM instance models. Therefore, to leverage these two aspects, in future work, the authors will propose a knowledge model that would further help with the analysis of wood construction. In terms of the level of details, one direction is to analyze information from BIM instance models with LOD 400, which provide a richer content information for fabrication and/or construction phase. Another alternative is to complement the current method with shop drawings and/or specifications information. This way will provide additional information that are missing from LOD 300–350 BIM instance models.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the National Science Foundation (NSF). This material is based on work supported by the NSF under Grant No. 1827733. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. Also, the authors would like to thank Professor Clark Cory and Dr. Temitope Akanbi for sharing the BIM instance models (T1-T3) and T4, respectively, that contributed in the experimental section of this research. The first author was supported by an Education Support Leave from the Technological University of Panama.

References

[1] D. Castro-Lacouture, Construction automation, in: Springer Handb. Autom, Springer, 2009, pp. 1063–1078. ISBN: 978–3–540-78831-7.
[2] C.Z. Li, R.Y. Zhong, F. Xue, G. Xu, K. Chen, G.G. Huang, G.Q. Shen, Integrating RFID and BIM technologies for mitigating risks and improving schedule performance of prefabricated house construction, J. Clean. Prod. 165 (2017) 1048–1062, https://doi.org/10.1016/j.jclepro.2017.07.156.
[3] F. Craveiro, J.P. Duarte, H. Bartolo, P.J. Bartolo, Additive manufacturing as an enabling technology for digital construction: a perspective on Construction 4.0, Autom. Constr. 103 (2019) 251–267, https://doi.org/10.1016/j.autcon.2019.03.011.
[4] Associated General Contractors, Eighty Percent of Contractors Report Difficulty Finding Qualified Craft Workers to Hire as Firms give Low Marks to Quality of New Worker Pipeline. https://www.agc.org/news/2019/08/27/eighty-percent-contractors-report-difficulty-finding-qualified-craft-workers-hire-0, 2019 (accessed February 3, 2021).
[5] U.S. Bureau of Labor Statistics, Job Openings and Labor Turnover, Economic News Release. https://www.bls.gov/news.release/jolts.a.htm, 2020 (accessed August 8, 2020).
[6] Construction Industry Resources LLC, Construction Productivity in an Imbalanced Labor Market. www.myCLMA.com, 2016.
[7] McGraw-Hill Construction, The Business Value of BIM for Construction in Major Global Markets: How Contractors around the World Are Driving Innovation with Building Information Modeling, Bedford, MA, 2014. http://static.autodesk.net/dc/content/dam/autodesk/www/solutions/building-information-modeling/construction/business-value-of-bim-for-construction-in-global-markets.pdf.
[8] S.H. Ghaffar, J. Corker, M. Fan, Additive manufacturing technology and its implementation in construction as an eco-innovative solution, Autom. Constr. 93 (2018) 1–11, https://doi.org/10.1016/j.autcon.2018.05.005.
[9] M. Pan, T. Linner, W. Pan, H. Cheng, T. Bock, A framework of indicators for assessing construction automation and robotics in the sustainability context, J. Clean. Prod. 182 (2018) 82–95, https://doi.org/10.1016/j.jclepro.2018.02.053.
[10] R.M. Lawson, R.G. Ogden, Sustainability and Process Benefits of Modular Construction, in: 8th CIB World Build, Congr, Salford, United Kingdom, 2010, pp. 38–51. http://www.irbnet.de/daten/iconda/CIB18783.pdf.
[11] McGraw-Hill Construction, Prefabrication and Modularization, McGraw Hill Construction, Bedford, 2011.
[12] KPMG, Smart Construction: How Offsite Manufacturing can Transform our Industry. kpmg.com/uk, 2016.
[13] A. Alwisy, S. Bu Hamdan, B. Barkokebas, A. Bouferguene, M. Al-Hussein, A BIM-based automation of design and drafting for manufacturing of wood panels for modular residential buildings, Int. J. Constr. Manag. 19 (2019) 187–205, https://doi.org/10.1080/15623599.2017.1411458.
[14] R. Sacks, C. Eastman, G. Lee, P. Teicholz, BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors, 3rd ed., Hoboken, New Jersey Wiley, 2018 https://doi.org/10.1016/S0926-5805(02)00090-0.
[15] S. Meschini, K. Iturralde, T. Linner, T. Bock, Novel Applications Offered by Integration of Robotic Tools in BIM-Based Design Workflow for Automation in Construction Processes, CIB* IAARC W119 CIC 2016 Workshop, 2016.
[16] C. Goodier, A. Gibb, Future opportunities for offsite in the UK, Constr. Manag. Econ. 25 (2007) 585–595, https://doi.org/10.1080/01446190601071821.
[17] A. Gibb, Standardization and pre-assembly - distinguishing myth from reality using case study research, Constr. Manag. Econ. 19 (2001) 307–315, https://doi.org/10.1080/01446190010020435.
[18] N. Lu, T. Korman, Implementation of building information modeling (BIM) in modular construction: benefits and challenges, in: Constr. Res. Congr 2010, 2010, pp. 1136–1145.
[19] R.M. Lawson, R.G. Ogden, R. Bergin, Application of modular construction in high-rise buildings, J. Archit. Eng. 18 (2012) 148–154, https://doi.org/10.1061/(ASCE)AE.1943-5568.0000057.
[20] S.W. Nunnally, Construction Methods and Management, Seventh, 2007. ISBN: 0135000793.
[21] J. Nässén, F. Hedenus, S. Karlsson, J. Holmberg, Concrete vs. wood in buildings - an energy system approach, Build. Environ. 51 (2012) 361–369, https://doi.org/10.1016/j.buildenv.2011.11.011.
[22] J. Glover, D.O. White, T.A.G. Langrish, Wood versus concrete and steel in house construction, J. For. 100 (8) (2002) 34–41.
[23] M. Chui, J. Mischke, The impact and opportunities of automation in construction, Voices Infraestruct. (2019) 7, https://www.mckinsey.com/~/media/McKinsey/Industries/Capital Projects and Infrastructure/Our Insights/The impact and opportunities of automation in construction/The-impact-and-opportunities-of-automation-in-construction.pdf?shouldIndex=false.
[24] J. Willmann, M. Knauss, T. Bonwetsch, A.A. Apolinarska, F. Gramazio, M. Kohler, Robotic timber construction - expanding additive fabrication to new dimensions, Autom. Constr. 61 (2016) 16–23, https://doi.org/10.1016/j.autcon.2015.09.011.
[25] H. Hasan, A. Reddy, A. Tsayjacobs, Robotic Fabrication of Nail Laminated Timber, Proceedings of the 36th International Symposium on Automation and Robotics in Construction (ISARC), 2019, https://doi.org/10.22260/isarc2019/0162.
[26] S. Asbjørn, A. Oded, E. Phillip, P. Luka, S. Florin, G. Fabio, K. Matthias, Topology optimization and robotic fabrication of advanced timber space-frame structures, in: Robot. Fabr. Archit. Art Des 2016, 2016, pp. 190–203, https://doi.org/10.1007/978-3-319-04663-1.
[27] P. Eversmann, F. Gramazio, M. Kohler, Robotic prefabrication of timber structures: towards automated large-scale spatial assembly, Const. Robot. 1 (2017) 49–60, https://doi.org/10.1007/s41693-017-0006-2.
[28] Thomas Staff Writer, Understanding CNC Machining, Articles. https://www.thomasnet.com/articles/custom-manufacturing-fabricating/understanding-cnc-machining/, 2021 (accessed February 2, 2021).
[29] P. Martinez, M. Livojevic, P. Jajal, D.R. Aldrich, M. Al-Hussein, R. Ahmad, Simulation-driven design of wood framing support systems for off-site construction machinery, J. Constr. Eng. Manag. 146 (2020), 04020075, https://doi.org/10.1061/(asce)co.1943-7862.0001853.
[30] I.J. Ramaji, A.M. Memari, Product architecture model for multistory modular buildings, J. Constr. Eng. Manag. 142 (2016) 1–14, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001159.
[31] NIMBS Committe, National Building Information Modeling Standard, Nbim, 2007, p. 180, https://doi.org/10.1017/CBO9781107415324.004.
[32] B. Becerik-Gerber, K. Kensek, Building information modeling in architecture, engineering, and construction: emerging research directions and trends, J. Prof. Issues Eng. Educ. Pract. 136 (2010) 139–147, https://doi.org/10.1061/(asce)ei.1943-5541.0000023.
[33] C. Eastman, P. Teicholz, R. Sacks, K. Liston, BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors, 2nd ed, Wiley, Hoboken, NJ, 2011.

[34] M.P. Gallaher, A.C. O'Connor, J.L. Dettbarn, L.T. Gilday, Cost Analysis of Inadequate Interoperability in the U.S. Capital Facilities Industry, National Institute of Standards and Technology (NIST), 2004. http://www.bentleyuser.dk/sites/default/files/nist_report.pdf.
[35] E.A. Poirier, D. Forgues, S. Staub-French, Dimensions of interoperability in the AEC industry, in: Constr. Res. Congr 2014, 2014, pp. 1987–1996, https://doi.org/10.1061/9780784413517.176.
[36] U. Isikdag, G. Aouad, J. Underwood, S. Wu, Building information models: a review on storage and exchange mechanisms, in: Proceedings of the CIB W78's 24th International Conference on IT in Construction, 2007, pp. 135–143.
[37] buildingSMART, Industry Foundation Classes (IFC). https://www.buildingsmart.org/standards/bsi-standards/industry-foundation-classes/, 2019 (accessed July 28, 2019).
[38] K. Afsari, C. Eastman, D. Shelden, Cloud-based BIM data transmission: current status and challenges, in: Proc. 33rd Int. Symp. Autom. Robot. Constr, 2016, https://doi.org/10.22260/isarc2016/0129.
[39] Y. Arayici, T. Fernando, V. Munoz, M. Bassanino, Interoperability specification development for integrated BIM use in performance based design, Autom. Constr. 85 (2018) 167–181, https://doi.org/10.1016/j.autcon.2017.10.018.
[40] J. Steel, R. Drogemuller, B. Toth, Model interoperability in building information modelling, Softw. Syst. Model. 11 (2012) 99–109, https://doi.org/10.1007/s10270-010-0178-4.
[41] R. Ren, J. Zhang, H.N. Dib, BIM interoperability for structure analysis, in: Proc., Constr. Res. Congr, ASCE, Reston, VA, 2018, pp. 470–479.
[42] J. Wu, J. Zhang, New automated BIM object classification method to support BIM interoperability, J. Comput. Civ. Eng. 33 (2019), https://doi.org/10.1061/(ASCE)CP.1943-5487.0000858, pp. 04019033 1–17.
[43] BuiildingSMART, Certified Software, Software Certification. https://www.buildingsmart.org/compliance/software-certification/certified-software/, 2020 (accessed June 8, 2020).
[44] H. Liu, G. Singh, M. Lu, A. Bouferguene, M. Al-Hussein, BIM-based automated design and planning for boarding of light-frame residential buildings, Autom. Constr. 89 (2018) 235–249, https://doi.org/10.1016/j.autcon.2018.02.001.
[45] M. Abushwereb, H. Liu, M. Al-Hussein, A knowledge-based approach towards automated manufacturing-centric BIM: wood frame design and modelling for light-frame buildings, in: Modul. Offsite Constr. Summit Proc., Banff, AB, 2019, pp. 100–107, https://doi.org/10.29173/mocs82.
[46] S. Isaac, T. Bock, Y. Stoliar, A methodology for the optimal modularization of building design, Autom. Constr. 65 (2016) 116–124, https://doi.org/10.1016/j.autcon.2015.12.017.
[47] P. Jensen, T. Olofsson, H. Johnsson, Configuration through the parameterization of building components, Autom. Constr. 23 (2012) 1–8, https://doi.org/10.1016/j.autcon.2011.11.016.
[48] M. Mekawy, F. Petzold, Exhaustive exploration of modular design options to inform decision making, in: A. Fioravanti, S. Cursi, S. Elahmar, S. Gargaro, G. Loffreda, G. Novembri, A. Trento (Eds.), Proc. 35th Int. Confer- Ence Educ. Res. Comput. Aided Archit. Des. Eur, Rome, Italy, 2017, pp. 107–114.
[49] A.Q. Gbadamosi, A.M. Mahamadu, L.O. Oyedele, O.O. Akinade, P. Manu, L. Mahdjoubi, C. Aigbavboa, Offsite construction: developing a BIM-based optimizer for assembly, J. Clean. Prod. 215 (2019) 1180–1190, https://doi.org/10.1016/j.jclepro.2019.01.113.
[50] M. Wang, S. Ahn, Y. Zhang, M. Sadiq Altaf, M. Al-Hussein, Y. Ma, Automatic material estimation by translating bim data into erp readable data for panelized residential construction, in: Modul. Offsite Constr. Summit Proc, 2019, pp. 9–16, https://doi.org/10.29173/mocs71.
[51] S.P. Root, J. Cribbs, A.D. Chasey, Case Study: Off-site manufacturing of EIFS Panelized Wall Assemblies to Gain Efficiency in Construction Sequencing, Modular and Offsite Construction (MOC) Summit Proceedings, 2019, pp. 357–364, https://doi.org/10.29173/mocs114.
[52] L. Sterling, E. Shapiro, The Art of Prolog: Advanced Programming Techniques, Second Edi, The MIT Press, Cambridge, Massachusetts, 1994. ISBN: 9780262193382.
[53] A.-L. Johansson, A. Eriksson-Granskog, A. Edman, Prolog Versus You An Introduction to Logic Programming, Uppsala, Sweden, 1989, https://doi.org/10.1017/CBO9781107415324.004.
[54] M.P.H. Uszkoreit, M.V.W. Wahlster, M.J. Wooldridge, B.G. Buchanan, P.J. Hayes, J.A. Hendler, N. Jennings, H. Kamp, R. Kowalski, H. Levesque, S. Oviatt, An Introduction to Language Processing with Perl and Prolog, 2006, https://doi.org/10.1007/3-540-34336-9.
[55] J. Zhang, N.M. El-Gohary, Semantic-based logic representation and reasoning for automated regulatory compliance checking, J. Comput. Civ. Eng. 31 (2016) 1–12, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000583.
[56] J.A. Makowsky, Why horn formulas matter in computer science: initial structures and generic examples, J. Comput. Syst. Sci. 34 (1987) 266–292, https://doi.org/10.1007/3-540-15198-2_24.
[57] M. Triska, The Power of Prolog, Self-published, 2005.
[58] N.-F. Zhou, B-Prolog User's Manual (Version 8.1): Prolog, Agent, and Constraint Programming, 2014.
[59] J. Väänänen, Second-Order and Higher-Order Logic, The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/fall2020/entries/logic-higher-order/, 2019.
[60] G. Luger, W. Stubblefield, AI Algorithms, Data Structures, and Idioms in Prolog, Lisp, and Java, 2009, https://doi.org/10.1017/CBO9781107415324.004.
[61] G. Ritchie, Prolog Step-by-Step, 2002.
[62] M.A. Covington, D. Nute, A. Vellino, Prolog Programming in Depth, Knowledge Creation Diffusion Utilization, 1995, p. 515.
[63] F. Portaro, Automated Reasoning, The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/sum2011/entries/reasoning-automated/, 2011.
[64] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Second, Pearson Education, Inc, Upper Saddle River, New Jersey, 2003. ISBN: 0-13-790395-2.
[65] A. Darko, A.P.C. Chan, M.A. Adabre, D.J. Edwards, M.R. Hosseini, E.E. Ameyaw, Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities, Autom. Constr. 112 (2020) 103081, https://doi.org/10.1016/j.autcon.2020.103081.
[66] K. Jain, K.H. Law, Knowledge Representation with Logic, Stanford Univ, Stanford, CA, 1989.
[67] OWL Working Group, Web Ontology Language (OWL), OWL, 2012. https://www.w3.org/OWL/ (accessed February 1, 2021).
[68] N.M. El-Gohary, T.E. El-Diraby, Domain ontology for processes in infrastructure and Construction, J. Constr. Eng. Manag. 136 (2010) 730–744, https://doi.org/10.1061/(asce)co.1943-7862.0000178.
[69] J. Zhang, N.M. El-Gohary, Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking, Autom. Constr. 73 (2017) 45–57, https://doi.org/10.1016/j.autcon.2016.08.027.
[70] B.I. Dahn, Robbins algebras are Boolean: a revision of McCune's computer-generated solution of Robbins problem, J. Algebra. 208 (1998) 526–532, https://doi.org/10.1006/jabr.1998.7467.
[71] E. Kotelnikov, Automated Theorem Proving with Extensions of First-Order Logic, Chalmers University of Technology, 2018. https://research.chalmers.se/publication/504640/file/504640_Fulltext.pdf.
[72] P. Pauwels, W. Terkaj, EXPRESS to OWL for construction industry: towards a recommendable and usable ifcOWL ontology, Autom. Constr. 63 (2016) 100–133, https://doi.org/10.1016/j.autcon.2015.12.003.
[73] E. González, J.D. Piñeiro, J. Toledo, R. Arnay, L. Acosta, An approach based on the ifcOWL ontology to support indoor navigation, Egypt. Inform. J. (2020), https://doi.org/10.1016/j.eij.2020.02.008.
[74] P. Pauwels, T. Krijnen, W. Terkaj, J. Beetz, Enhancing the ifcOWL ontology with an alternative representation for geometric data, Autom. Constr. 80 (2017) 77–94, https://doi.org/10.1016/j.autcon.2017.03.001.
[75] J. Zhang, N.M. El-Gohary, Automated extraction of information from building information models into a semantic logic-based representation, in: 2015 ASCE Intl. Work. Comput. Civ. Eng, ASCE, Reston, VA, 2015, pp. 173–180, https://doi.org/10.1061/9780784479247.022.
[76] P.H. Winston, Artificial Intelligence, Wesley Publishing Company, United Kingdom, 1992. ISBN: 0201533774.