

DOI: 10.11835/j.issn.2096-6717.2023.069



开放科学(资源服务)标识码 OSID:



# 语义网赋能建筑信息交付及模型数据模式分析

刘吉明<sup>a</sup>, 段立平<sup>a,b</sup>, 林思伟<sup>a</sup>, 缪季<sup>a</sup>, 赵金城<sup>a,b</sup>

(上海交通大学 a. 船舶海洋与建筑工程学院;

b. 上海市公共建筑和基础设施数字化运维重点实验室, 上海 200240)

**摘要:** 针对以建筑信息模型(BIM)进行交付的信息共享模式所依赖的工业基础类(IFC)标准行业适用性不足且难以拓展的问题, 探讨在 IFC 基础上引入语义网实现异源数据集成共享, 并于语义层面实现信息交付。首先, 通过算法解析和模型转化介绍语义化建模方法, 并以二层钢框架厂房结构为例对该方法进行说明; 然后, 通过对转化案例进行数据模式分析, 以验证建筑信息交付的准确性和建筑语义的可传递性。案例实践论证基于 IfcOWL 本体的语义化建模方法的可实施性; 通过分析该语义化模型单元实例的数据模式, 探究制约该语义化建模方法赋能建筑信息交付的关键因素; 针对语义化建模方法所面临的问题, 提出冗余信息规避、领域本体开发和轻量化语义建模的初步解决思路。SPARQL 查询实例表明, 所解析的数据模式对规避冗余信息有效。因此, 该方法在共享和集成建筑多源异构信息方面具有优势, 能有效提升建筑信息管理的智能化水平。

**关键词:** 建筑信息交付; 语义网; 数据模式分析; 工业基础类; 建筑信息模型

**中图分类号:** TU205; TP391 **文献标志码:** A **文章编号:** 2096-6717(2024)01-0244-10

## Building information delivery and model schema analysis empowered by Semantic Web

LIU Jiming<sup>a</sup>, DUAN Liping<sup>a,b</sup>, LIN Siwei<sup>a</sup>, MIAO Ji<sup>a</sup>, ZHAO Jincheng<sup>a,b</sup>

(a. School of Naval Architecture, Ocean and Civil Engineering; b. Shanghai Key Laboratory for Digital Maintenance of Buildings and Infrastructure, Shanghai Jiao Tong University, Shanghai 200240, P. R. China)

**Abstract:** In the construction industry, information sharing by the building information modeling (BIM) generally relies on the industry foundation classes (IFC) schema, but the latter's unsatisfactory adaptability and inextensibility restrain the former. To overcome the limitations, the Semantic Web was introduced based on the IFC schema to realize heterogeneous data integration and sharing and further achieve semantic-level information delivery. Firstly, the semantic modeling method was introduced through algorithm analysis and model transformation. Secondly, this method was used to create a semantic model for a two-story steel frame building. Finally, the schema of this model was verified for accuracy of the building information delivery and the transferability of the building semantics. The practicability of the semantic modeling method with IfcOWL

**收稿日期:** 2023-03-02

**基金项目:** 上海市科技创新项目(21DZ1204600)

**作者简介:** 刘吉明(1998-), 男, 博士生, 主要从事建筑信息智能化研究, E-mail: jimingLiu@sjtu.edu.cn。

赵金城(通信作者), 男, 教授, 博士生导师, E-mail: jczhao@sjtu.edu.cn。

**Received:** 2023-03-02

**Foundation item:** Scientific Research Project of Shanghai Science and Technology Commission (No. 21DZ1204600)

**Author brief:** LIU Jiming (1998-), PhD candidate, main research interest: intelligent building information management, E-mail: jimingLiu@sjtu.edu.cn.

ZHAO Jincheng (corresponding author), professor, doctoral supervisor, E-mail: jczhao@sjtu.edu.cn.

ontology was supported by the modeling case. The key factors that restrain this semantic modeling method from empowering building information delivery was explored by analyzing the schema of the element of the semantic model. And introductory ideas about redundant information avoidance, domain ontology development, and lightweight semantic modeling were proposed to fill the gap. The SPARQL query case shows that the parsed schema is effective for avoiding redundant information. Consequently, this method has advantages in sharing and integrating multi-source heterogeneous building information and can effectively improve the intelligent level of building information management.

**Keywords:** building information delivery; Semantic Web; data schema analysis; industry foundation classes (IFC); building information modeling (BIM)

在推行数字化建造的过程中,项目各方信息流通困难、专业软件数据格式各异、建筑信息多源异构及大数据特征等问题严重阻碍了行业智能化转型。这些涉及信息集成、流通与管理的问题可统称为信息交付问题,而传统碎片化、分散式的交付方式已无法满足数字化建造需求,因此亟需革新建筑信息交付方式以打破信息孤岛壁垒。

建筑信息模型(Building Information Modeling, BIM)技术及其信息交付标准工业基础类(Industry Foundation Class, IFC)的兴起为上述问题带来了转机。IFC是由buildingSMART International(bSI)提出的开放且中立的数据交换格式,其通过描述建筑模型的几何架构、语义架构以及二者的相互关系,实现了对建筑信息的完整描述。也正因如此,IFC被认为适用于建筑全生命周期中的大部分数据交换场景。例如,刘照球等<sup>[1]</sup>基于BIM的模型集成框架、赖华辉等<sup>[2]</sup>基于IFC标准的信息共享方法、马智亮等<sup>[3]</sup>基于IFC标准的建筑能耗监测静态信息模型都有效提升了信息交付效率。然而,由于IFC标准存在编写语言普及程度低和数据架构复杂、难以拓展的问题<sup>[4]</sup>,该技术仍无法解决信息交付难题。

此外,通过建立数学表达式或开发算法来建立数学模型的方法也是一种可行的数据集成方式,比如魏国海等<sup>[5]</sup>的火灾损伤多元信息融合模型、Liu等<sup>[6]</sup>基于数字孪生的信息融合框架,以及张立奎等<sup>[7]</sup>用于桥梁变形重构的基于LSTM神经网络的多元信息融合方法。但现有的数学建模方法无法涵盖所有相关因素,并且所建模型高度抽象,故而也难以在实践中得到推广。

针对以上问题,笔者在IFC标准基础上引入语义网以克服其局限性。首先,介绍语义网核心技术架构,并分析语义网拓展IFC功能的可行性。然后,利用Pauwels等<sup>[8]</sup>提出的转化方法执行建筑信息语义化建模,通过验证所建模型数据传递的准确性来说明该方法的可行性。最后,分析所建模型的数据模式以探讨制约该技术实践应用的关键因素,并基

于分析的数据模式提出冗余信息规避的可行方法。

## 1 语义网核心技术与IFC功能拓展

语义网是万维网之父Berners-Lee等<sup>[9]</sup>于1998年提出的一种网络技术,旨在解决“传统网络仅作为给用户文档的媒介而无法直接处理数据信息”的问题。该技术的主旨是通过机器可理解的统一信息描述来共享暂存信息,而这一功能特征恰与建筑业亟需解决的信息交付难题高度契合。正因如此,语义网赋能建筑信息交付得到学者广泛探究。其中,专研建筑信息智能化的国际组织bSI不仅成立链接数据工作组(Linked Data Working Group, LDWG)来实现IfcOWL本体的标准化,同时还将语义网纳入其技术路线图<sup>[10]</sup>。

### 1.1 语义网核心技术

Berners-Lee最早提出了语义网基本架构,其逐步发展完善为现行的语义网堆栈<sup>[11]</sup>。该技术架构下的资源描述框架(Resource Description Framework, RDF)、本体与网络本体语言(Web Ontology Language, OWL)以及SPARQL(SPARQL Protocol and RDF Query Language)三项核心技术为其数据表示、逻辑推理和知识查询功能提供支持。

RDF是万维网联盟(World Wide Web Consortium, W3C)提出的以相互关联的有向标记图描述资源信息的标准框架。通过以主体-谓词-客体的三元组表达式描述资源,实现对多源异构数据的统一表达,进而为语义化建模和知识查询推理奠定数据基础。

本体和OWL是语义网体系中赋予模型高级知识逻辑的关键模块。本体原为哲学领域概念,知识工程引入该术语以创建自动推理信息模型。OWL特指W3C推行的OWL2网络本体语言,是一种旨在表述事物、事物组及事物间关系所蕴含知识的陈述性语言。在语义网架构中,把基于OWL概念创建的RDF图称为OWL本体,用以弥补RDFS

(Resource Description Framework Schema) 语义丰富度不足的缺陷。图 1 所示是 W3C 倡议的 OWL2 本体的分类架构和主要配置文件<sup>[12]</sup>, 其中 OWL2 DL 本体在知识推理工具中应用最广。其原因在于 OWL2 DL 可产生与 SROIQ 描述逻辑兼容的语义, 且根据相似理论<sup>[13]</sup>, 其推论足够可信, 故推理工具可参照描述逻辑的实践经验进行开发。

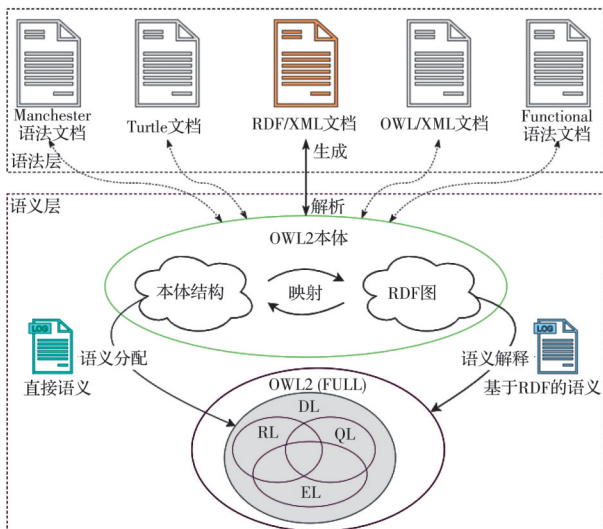


图 1 OWL2 本体分类及语义层次关系

Fig. 1 Classification and semantics hierarchy relationships of OWL2 ontology

SPARQL 是基于图模式匹配的 RDF 查询语言, 可跨数据源作联合查询并返回 RDF 图。SPARQL 可对匹配图模式作增删查改, 进而实现 RDF 数据集的动态更新。除查询外, 知识推理也是语义网的重要功能。语义网不仅有源自 RDFS 和 OWL 的推理逻辑, 还可借由 SWRL、SPIN 和 SHACL 规则语言实现高级知识推理。此外, 辅助本体开发的本体推理也是一种知识推理形式<sup>[14]</sup>。

### 1.2 语义网拓展 IFC 信息交付功能

在实际信息交付场景中, IFC 标准所面临的主要问题是数据模式难以拓展和信息提取修改困难。通过分析各问题成因以及语义网处理该问题的优势, 论证了语义网拓展 IFC 信息交付功能的可行性。

拓展性问题是指难以扩充既定的 IFC 数据模式来满足个性化的信息表述需求, 该问题主要源于定义数据模式的 EXPRESS 语言。首先, 这是一种不常用的声明性语言, 难以被建筑业用户所掌握; 其次, 该语言可定义的信息类型繁杂(可同时定义实体、关系、规则及复杂数据结构), 故由其声明的 IFC 数据模式复杂且难以扩展。bSI 针对该问题提供了自定义属性集和数据字典的解决方案, 但前者难以实现术语对齐, 后者缺乏灵活性而无法提供有效帮助。语义网能够集成本体和描述逻辑对各领域信

息进行个性化知识表示, 然后依托本体映射实现语义层面的一致性匹配, 进而将 IFC 兼容信息与其他信息进行统一表达。因此, 语义网技术可有效解决 IFC 标准面临的拓展性难题。

IFC 标准的另一个问题是缺乏高效的信息查改方法。IFC-SPF 文件是基于 IFC 数据模式编写的纯文本文件, 该文件不仅可读性差, 而且难以对 IFC 语句进行修改。Mazairac 等<sup>[15]</sup>曾为此提出了一种检索 IFC 信息的查询语言 BIMQL, 但该方法不仅只能用于信息查询, 还难以被用户掌握, 故未得到广泛应用。在语义网被引入建筑领域后, Krijnen 等<sup>[16]</sup>随即提出利用 SPARQL 来弥补 BIMQL 存在的不足。其原因在于: 首先, SPARQL 用法与 SQL 语言类似, 容易被用户所掌握; 其次, SPARQL 可直接对 RDF 图进行增删查改; 并且用户还能通过 SPARQL 对 SPARQL 端点中的 RDF 数据进行远程访问, 有助于开发项目信息管理的云服务。

除以上优势外, 语义网的知识推理功能有助于高效利用建筑信息。在项目管理过程中, 所集成数据增多, 隐含知识也随即陡增。而挖掘隐含知识对于开发项目管理服务是有益的。如, Wu 等<sup>[17]</sup>将卷积神经网络识别的人的施工行为信息与人和构件空间关系信息集成, 再引入从施工技术规范中转化的 SWRL 规则, 识别了危险施工行为。因此, 语义网有助于让 IFC 表达的信息切实服务于实际工程。

## 2 IFC 标准赋能 BIM 信息集成共享

鉴于在 IFC 模型基础上进行语义化建模, 故对 IFC 模型的生成方式和信息转化效率进行分析。

### 2.1 模型视图定义

模型视图定义 (Model View Definition, MVD) 是对 BIM 模型作数据筛选输出 IFC 模型的过滤器。其兴起的原因在于随着 BIM 信息交付所涉及领域的不断扩张, IFC 标准面向不同专业的数据冗余问题愈发突出, 因此, bSI 推出 MVD 作为 IFC 标准的子集以适应不同应用需求。

bSI 在其 MVD 数据库收录了 MVD 的所有可用版本, 而在建筑业 BIM 软件中得到应用的分别是 IFC 2×3 协调视图 2.0 (Coordination View 2.0, CV 2.0)、IFC4 参照视图及 IFC4 设计传递视图。其中 CV 2.0 在各类软件中兼容性最好, 应用也最为广泛。但 bSI 正在积极推进 IFC4 标准相关 MVD 的研发与推广, 从视图所涉范围的广度和更新速度来看, IFC4 全面取代 IFC 2×3 标准是必然趋势。

### 2.2 不同 IFC 模型数据转化效率对比

为对比经不同 MVD 输出 IFC 模型的数据转化

效率,选取如图 2 所示的双层钢框架厂房结构作案例分析。利用集成式 IFC 解析器(<https://github.com/Autodesk/revit-IFC>)从 Revit 中以 CV 2.0、参照视图和设计传递视图分别输出 IFC 模型,将其分别读取并汇总信息,如表 1 所示。在所有输出模型中,参照视图(结构)丢失信息最为严重,该视图因忽略楼板的信息输出,直接导致实体总数异常;

CV2.0 因对梁实体的区分只考虑构件截面类型而忽视建模方法与梁长差异,故输出的实体关系信息偏少;至于设计传递视图,由于其材料关系定义采取实例与材性映射方式,因而此关系数目多于其他输出结果,从语义化建模角度来看,这种表达形式能够涵盖更完备的构件信息,有助于提高语义化建模精度。

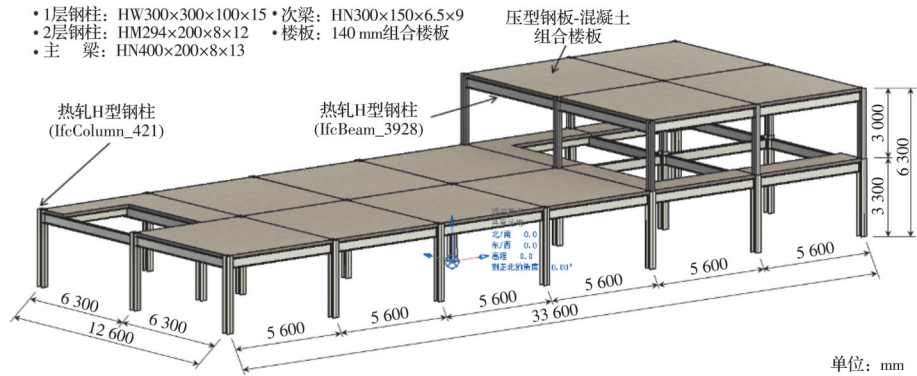


图 2 双层钢框架厂房 BIM 模型结构示意图

Fig. 2 BIM model of two-story steel frame building

表 1 钢框架厂房各 MVD 输出模型组成信息汇总

Tab. 1 Composition information of each MVD output model of the steel frame structure

MVD 名称	实体总数	关系总数	实体类型	材料关系	属性关系	类型关系	
CV 2.0	149	476	3类	2	429	21	
参照视图	建筑	149	620	3类	2	574	36
	结构	132	537	2类	1	509	20
	建筑服务	149	620	3类	2	574	36
设计传递视图	149	764	3类	146	574	36	

针对该案例的楼板信息转化结果,进一步对比各输出模型转化效率。楼板采用压型钢板-混凝土组合楼板,Revit 建模时采用双层组合形式。在楼板的信息转化过程中,参照视图与 CV 2.0 只能将其简化为单层混凝土构件;并且前者还缩减了楼板的实体关系,尤其是材料关系的输出,导致建筑信息严重丢失。而设计传递视图对于组合楼板的信息转化效果明显优于前述视图,该视图不仅有效保留了构件应有的实体关系,还能将组合楼板以双层形式正常输出。因此,设计传递视图输出模型在保真度和信息丰富度方面有明显优势。

### 3 建筑信息语义化建模及其转化机制

语义化建模旨在以 RDF 对 IFC 实体信息作资源虚拟化。针对此转化需求,Karan 等<sup>[18]</sup>通过集成自定义 BIM 本体与地理信息系统(Geographic Information Systems, GIS)本体来处理 IFC 属性,进而生成 BIM 与 GIS 集成的 RDF 模型;W3C 则致力

于研发以 BOT (Building Topology Ontology) 为代表的轻量化本体<sup>[19]</sup>,并且由 Bonduel 等<sup>[20]</sup>开发出的 IFCToLBD 转换器实现了基于 BOT 本体的 RDF 建模;此外,Niknam 等<sup>[21]</sup>也提出了一种可用于建筑信息共享的通用本体 BIMSO。与上述 3 种方法相比,Pauwels 等<sup>[8]</sup>倡议的基于 IfcOWL 本体的转化方法能够对 IFC 兼容的所有建筑信息执行语义化建模,在适用性与技术成熟性方面更具优势,因此应用其开发的 IFCToRDF-0.4 转换器(<https://github.com/pipauwel/IFCToRDF>)作语义化建模。

对于已开发的 IfcOWL 本体而言,IFC4 对应本体相较于 IFC 2x3 对应本体在类层次关系及公理数目上更具优越性,再结合前文各 IFC 模型的转化效果,最终选定设计传递视图输出模型进行语义化建模。建模步骤分为 IFC-SPF 文件向 Turtle 文件转化和引入 IfcOWL 本体实现高阶语义赋能。在 JDK17 的编译环境中启动转换器并执行文件转化,然后将初始 RDF 模型导入 Protégé 本体编辑器,最后用 IfcOWL 本体为该模型赋予建筑语义内涵。

#### 3.1 IFC to RDF 转化算法解析

为明晰转换器的工作原理,通过 IntelliJ IDEA 将其运行于 maven 项目,并对其转化机制作源码分析。如图 3 所示,主程序涵盖了 IFC 模型转化从参数输入到初始 RDF 模型输出的全过程;程序根据命令行参数自动识别转化模式,依次调用 showFile()、setup()与 convert()方法执行文件读取、规范设置和格式转化,并最终输出初始 RDF 模型。其 setup()

方法用于确定 IFC 标准的版本并导入为文件转化提供规范参照的序列化文件;convert()方法则实现了文件解析、本体注册、资源创建以及三元组编写。鉴于此二者在文件转化中的重要作用,图 3 也对其运行流程予以展示。为提升所绘流程图的可读性,

对某些不可能发生的条件分支进行甄别。例如,在转化模块的红框标识步骤,参照合法 IFC 语句,不会将 Type 类型实体作为其实体声明,把程序针对该情况对应的条件分支予以省略。

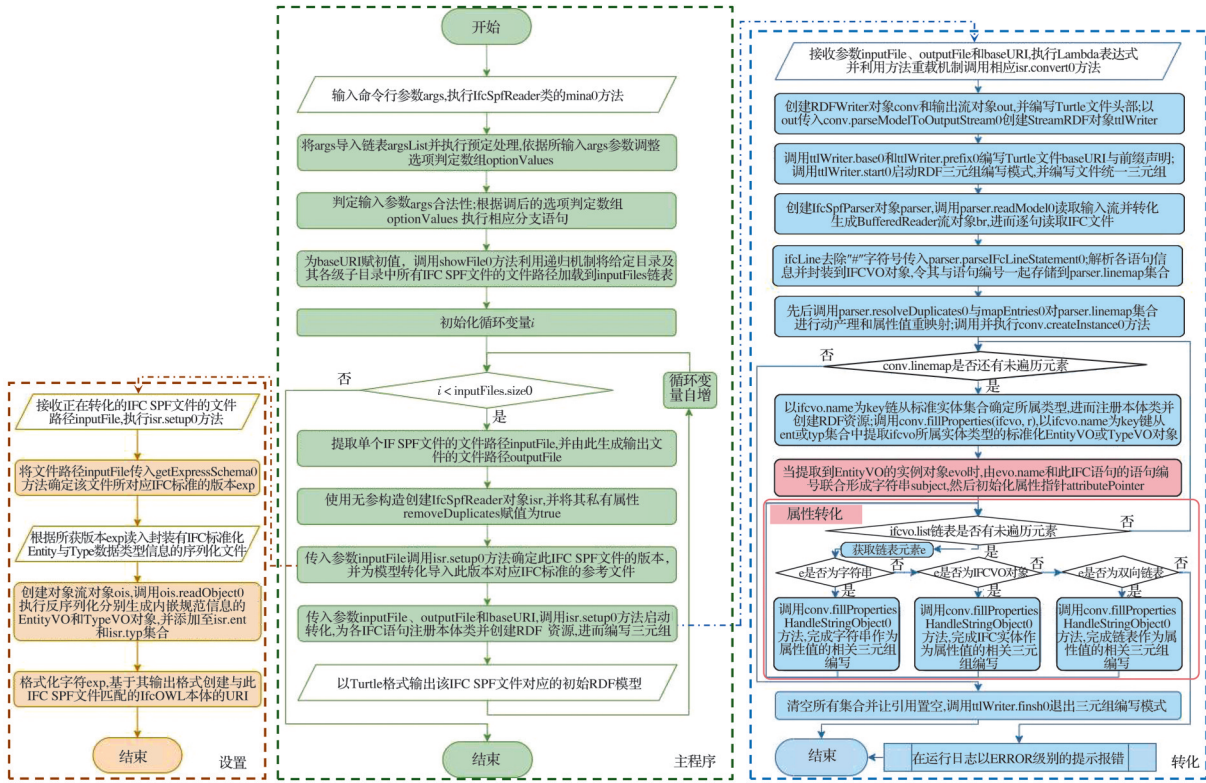


图 3 IFCtoRDF-0.4 转换器运行流程  
Fig. 3 Flow chart of IFCtoRDF-0.4 convert

转换器执行文件转化的核心步骤包括参考标准导入、IFC 语句解析和三元组编写。在设置模块中,通过反序列化 IFC 标准的序列化文件导入参考标准;在转化模块加粗蓝框标识步骤中,调用 parseIfcLineStatement() 解析 IFC 语句;三元组则在转化模块不同步骤中根据其类别调用特定方法予以编写。为分析语句解析的基本逻辑,绘制如图 4(a) 所示的算法伪代码。图示 switch 分支语句的前两项分别完成语句编号和实体名称的解析,后两项则联合完成括号内各属性值的解析。在该算法中,程序主要利用 IFC 语句的特定字符对控制条件 state 作调控,进而执行恰当的分支语句以完成语句解析。另外,对于属性值为数组甚至是多维数组的复杂情形,程序则利用栈数据结构和双向链表数据结构进行分层解析。

IFC 语句对应三元组主要源于属性信息的转化,如图 3 中属性转化子模块所示,程序将依据属性值所属类型来执行相应分支语句,进而编写三元组。分析了字符串属性值转化分支的运行逻辑,并

将该分支调用的 fillPropertiesHandleStringObject() 方法的算法伪代码绘图如图 4(b) 所示。此算法的核心在于识别属性类别以及创建并添加属性资源集合键值对,主要遵循从字面值赋值到顶层属性赋值这一自下而上的逻辑。对于转化过程中遇到的非常规情况,将以不同级别示警在运行日志中予以提示。其余转化分支虽然具体流程与前述算法略有区别,但转化逻辑类似,可参照理解。

利用此转换器输出双层钢框架厂房的初始 RDF 模型并导入 Protégé 应用程序,由于转化过程已基于 Apache Jena 框架实现本体类的注册,故可引入 IfcOWL 完成建筑语义赋能。

### 4 语义化建筑信息模型数据模式分析

为探讨基于 IfcOWL 语义化建模方法未得到有效使用的可能原因,对语义化模型单元实例的数据模式进行分析。鉴于 RDF 模型数据模式的通用性,此内容也将为建筑信息的数据存取与逻辑推理研究提供技术参照。

```

算法 1: IFC 语句解析算法
输入: 已去除 IFC 语句中“#”字符的字符串 ifcLine
输出: IFC 语句编号为 key, 对象 ifcvo 为 value 的键值对
1 begin
2   创建对象 ifcvo, 创建 StringBuffer 对象 sb
3   strLine = ifcvo.fullLineAfterNum ← getIFC(ifcLine)
4   初始化变量 i, state, clCount; 创建链表 list 和栈对象 stack
5   for i < ifcLine.length() do
6     ch ← getLineCh(strLine, i)
7     switch state do
8       // 省略各分支跳出本循环且执行变量 i 自增的叙述
9       case 0 do 判断 ch 类型, 执行对应分支语句完成
                ifcvo.lineNum 的赋值、state 的更新以及有效字符 ch
                向 StringBuffer 对象 sb 的添加
10      case 1 do 判断 ch 类型, 执行对应分支语句完成
                ifcvo.Name 的赋值、state 的更新以及有效字符 ch 向
                StringBuffer 对象 sb 的添加
11      case 2 do 判断 ch 类型, 执行对应分支语句; 通过栈数
                据结构和双向链表的结合, 提取 IFC 语句中多层嵌套的
                IFC 实体, 并按顺序将该信息存储到链表 list 中; 嵌套
                IFC 实体的信息提取完成后更新参数 state
12      case 5 do 判断 ch 类型, 执行对应分支语句完成嵌套实
                体的有效字符 ch 向 StringBuffer 对象 sb 的添加和
                state 的更新
13      end
14    end
15  end

```

(a) IFC 语句解析算法的伪代码

```

算法 2: 字符串分支属性转化算法
输入: RDF 资源 r, EntityVO 对象 evo, 联合字符串 subject, 属性
        指针 attributePointer 和链表元素 e
输出: 更新后的属性指针 attributePointer
1 begin
2   if e 为字符串 || ‘@’ || ‘#’ then
3     attributePointer++
4   else
5     if e 为 IFC 规范中 Type 实体名 then
6       记录 Type 实体类型, conv.typeRemembrance ← e
7     else
8       由属性指针 attributePointer 从 evo 获取该位置属性实
            体名称, 进而注册本体属性 prop 并获取其值域 range,
            将元素 e 移除多余字符并转化为字面值字符串 lStr
9       if range 为 ENUMERATION 子类 then
10        调用 conv.addEnumProperty(r, prop, range, lStr)
11        依次编写三元组 r prop r1
12      else if range 为 Select 子类 || 不为 OwlList 子类 then
13        调用 conv.createLiteralProperty(r, prop, range, lStr)
14        依次编写 r1 rdf:type range, r1 的赋值三元组
            以及 r prop r1
15      else
16        解析出错, 在运行日志以 WARN 或 ERROR 示警
17      end
18    attributePointer++
19  end
20 end
21 end

```

(b) 字符串分支属性转化算法的伪代码

图 4 转化模块关键算法的运行逻辑

Fig. 4 Execution logic of key algorithms in convert module

### 4.1 建筑单元实例属性分析

建筑单元实例是指具体构件的实例对象。图 5 所示为分析案例涉及到的梁、板、柱构件的继承关

系;图示层次关系采取了与 IFC 模式一致的类继承方式,故而两种信息表达方式在类层次关系上具有一致性。

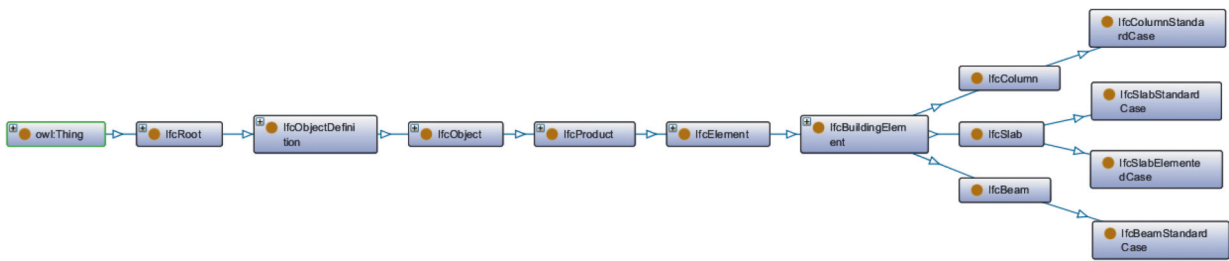


图 5 RDF 模型中建筑单元的类继承关系

Fig. 5 Class hierarchy relationships of building elements in RDF model

在 RDF 模型中,建筑单元实例的语义化描述均由图 6 所示语义层次关系表示。如图 6 所示,单元实例所具有的属性皆为对象属性,各属性所在三元组主体(定义域)皆为实例对象本身,三元组客体(值域)为字面值或 RDF 资源。其中赋值三元组直

接阐述了建筑单元的名称、标签和类型信息,3 个 RDF 子图涵盖了业主历史、对象布置和几何表示的相关信息。此外,图 6 所示图例说明了在执行数据模式分析时使用的各节点符号的具体含义。

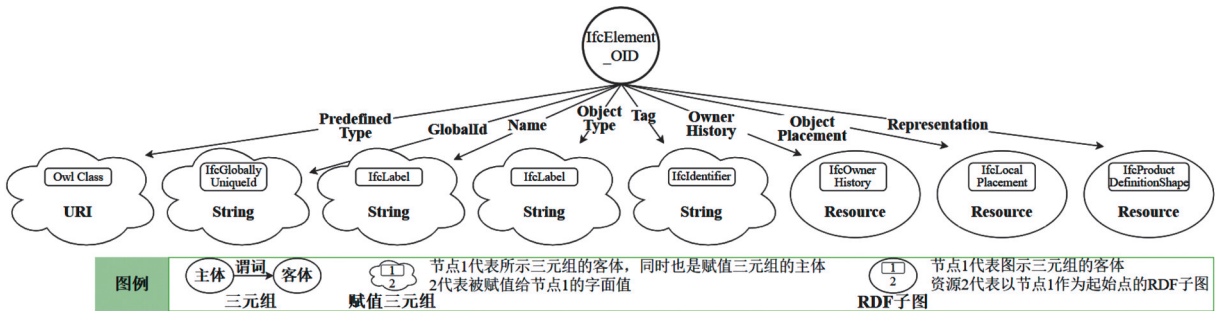


图 6 构件实例对象的属性架构

Fig. 6 Object property schema of building elements

### 4.2 梁柱单元数据模式分析

梁柱单元是钢框架结构主要的构件形式,因此,选取图 2 中标识的单元实例 IfcBeam\_4072 和

IfcColumn\_421 进行数据模式分析,通过分析所选单元数据模式的建筑内涵来验证信息传递的准确性和语义可传递性。由于业主历史属性的层次关

系简单,因此数据模式分析主要聚焦于对象布置和几何表示属性。

4.2.1 梁单元数据模式

图7展示了 IfcBeam\_4072 对象布置属性的数据模式。该属性主要用于声明对象的局部坐标系,进而确定构件单元的空间方位。单元实例通过 objectPlacement 属性与自身局部坐标系关联。然后,该局部坐标系通过 placementRelTo 属性依次建立起与其他局部坐标系的相对关系,直至关联到模型的全局坐标系。图示 IfcAxis2Placement3D 实体源自规范 ISO 10303-42,用于以单个笛卡尔点和两个正交轴来声明空间方位或放置坐标系。图示该实体均用于放置三维坐标系,其中笛卡尔点为坐标

系原点,参考方向为局部坐标系 X 方向,轴方向为局部坐标系 Z 方向。此外,ISO 10303-42 规定 IfcAxis2Placement3D 代表的局部坐标系与整体坐标系一致时,参考方向和轴方向可以省略。正因如此,图示 inst:IfcAxis2Placement3D\_171 实体和 inst:IfcAxis2Placement3D\_32 实体均只有 location 属性。对笛卡尔坐标和方向向量的数据模式进行简化,具体形式可参见文献[22]中图4所示的链表结构。依据图7所示的对象布置属性数据模式,IfcBeam\_4072 的局部坐标系原点的绝对坐标为(5 600, 3 150, 5 960),局部坐标 X 轴和 Z 轴的方向向量分别为(0, -1, 0)与(0, 0, 1)。

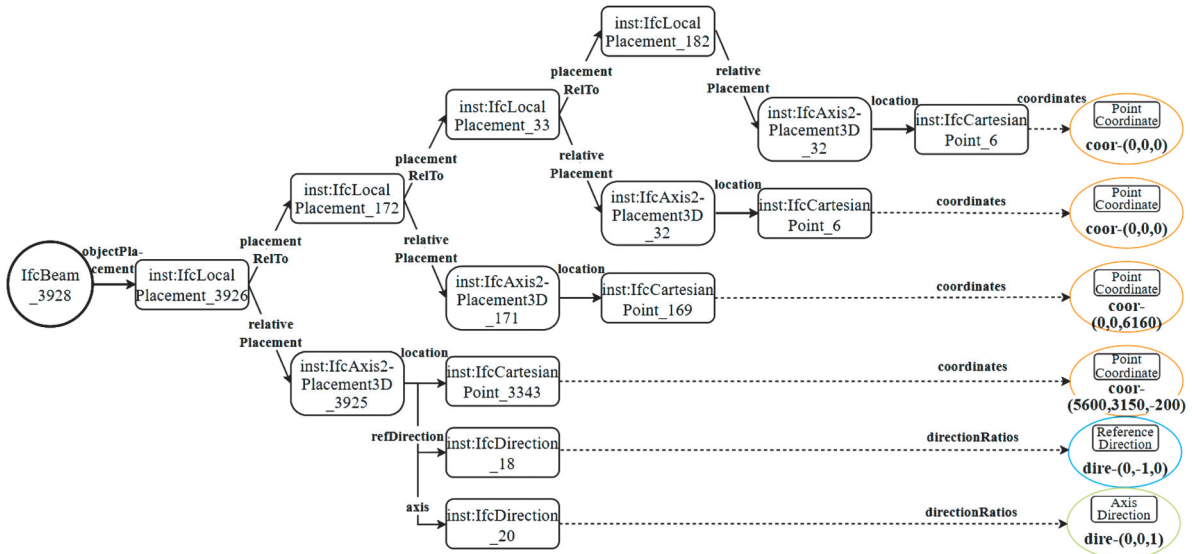


图7 IfcBeam\_4072 对象布置属性数据模式

Fig. 7 Object placement property schema of IfcBeam\_4072

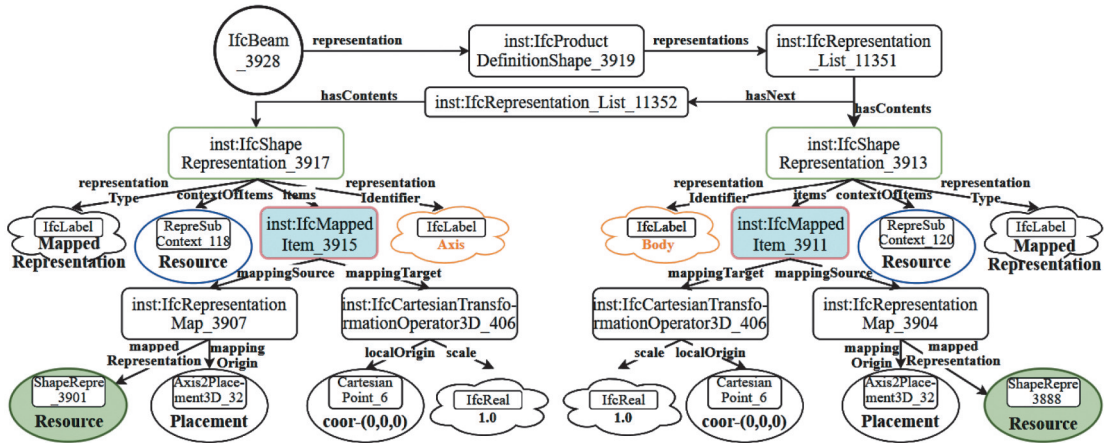
图8为 IfcBeam\_4072 的几何表示属性的数据模式。IfcShapeRepresentation 实体是表达单元实例几何形状的关键,其 representationType 属性声明了单元形状表示的几何模型类型,representationIdentifier 属性声明了单元的表示形式。如图8(a)所示,该构件的模型类型为映射表示,即此单元实例是通过实体映射生成的。在映射表示中,MappingTarget 属性用于表述生成实体对映射实体执行的空间变换。依据所选单元的缩放比例和转换坐标系可知,该单元的创建形式采取了不包含笛卡尔变换的直接映射,这与该构件在 Revit 中的建模方式契合。图示 Axis 模块描述了构件在局部坐标系下的方位信息。如图8(b)所示,梁单元起点为(-3 150, 0, 200),终点为(3 150, 0, 200)。根据图2所示构件参数及构件所属局部坐标系方位可知,该信息表述是准确的。图示 Body 模块描述了单元实体建模的信息。如图8(c)所示,梁单元的拉伸长度为 5 980.6

mm,在梁局部坐标系下,拉伸起点为(-2 990.3, 0, 0),拉伸方向为(1, 0, 0)。与初始 Revit 模型对比可知,该信息表述准确无误。图示 IfcProfileDef 实体声明了构件的截面形状,鉴于建筑构件截面形式多样,不对截面的语义层次关系进行展开。

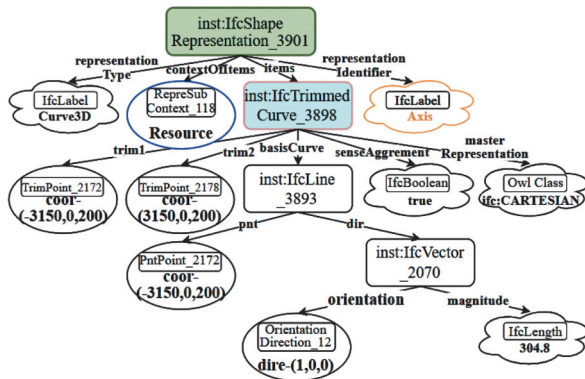
4.2.2 柱单元数据模式

在建立的钢框架语义化模型中,柱单元的数据模式与梁单元具有相似性。柱单元对象布置属性数据模式的分析结果表明,IfcColumn\_421 的局部坐标系与结构整体坐标系一致,原点的绝对坐标为(-16 800, 6 300, 0)。

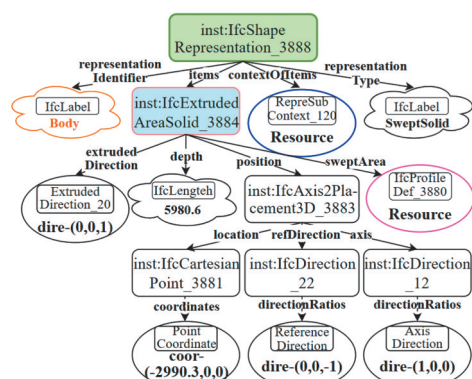
图9展示了 IfcColumn\_421 几何表示属性的数据模式,所示层次关系仅保留与 IfcBeam\_4072 几何表示属性数据模式不同的部分。相比于梁单元实例,IfcColumn\_421 的数据模式仅含有 Body 模块。依据 IfcExtrudedAreaSolid 实体的信息可知,该构件的拉伸长度为 3 300 mm。依据 inst:IfcAxis2



(a) IfcBeam\_4072 几何表示属性数据模式总体架构



(b) RDF 资源 ShapeRepre\_3901 的数据模式



(c) RDF 资源 ShapeRepre\_3888 的数据模式

图 8 IfcBeam\_4072 几何表示属性数据模式

Fig. 8 Representation property schema of IfcBeam\_4072

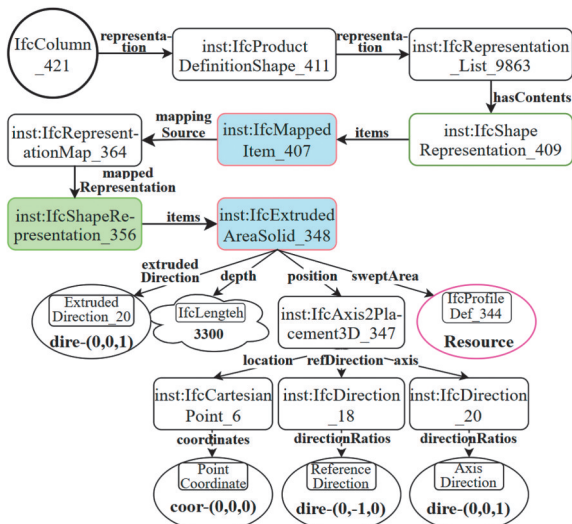


图 9 IfcColumn\_421 几何表示属性数据模式

Fig. 9 Representation property schema of IfcColumn\_421

Placement3D\_347 的转换坐标系可得,在柱局部坐标系下,构件拉伸方向为(0, 0, 1),拉伸起点为(0, 0, 0)。再根据图 2 所示构件参数和此构件所属局部坐标系方位可知,该信息表述是准确的。

#### 4.2.3 数据模式冗余信息分析与规避

对于项目信息管理而言,几何表示属性中由

IfcGeometricRepresentationContext 及其子类实例描述的背景信息是冗余的。以图 10 所示的 inst:IfcGeometricRepresentationSubContext\_118 为例,它不但与每一个 IfcShapeRepresentation 实体重复关联,同时所表示的信息也缺乏实际用途,故而导致所创建模型的信息冗余。此外,语义化模型中存在表达相同含义的冗余信息。如图 8 所示,Axis 模块的 OrientationDirection 和 Body 模块的 ExtrudedDirection 均描述了梁单元的拉伸方向。

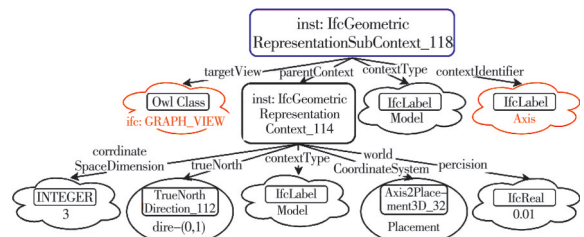


图 10 几何表示实体中背景信息的数据模式

Fig. 10 Schema of geometric representation context

为 SPARQL 语句设立精确图模式可有效规避冗余信息。比如,可为梁柱单元及其拉伸长度查询设立如下 SPARQL 语句:



```
SELECT ?element ?depth WHERE{
?element (((representation/ representation)/
hasContents)/ items)/ mappingSource)/ items)/
depth ?depth
}
```

上述语句对谓词属性进行了简写,即省略了谓词的前缀与后缀。依据此精确图模式匹配,便能有效规避冗余信息。此外,基于此精确图模式编写 DELECT 语句可移除用户不需要的冗余信息。笔者将此技术应用到所开发的 SemBIMCURD-SJTU 软件中。

通过以上对单元实例的对象布置和几何表示属性数据模式的分析,语义化模型对于建筑信息交付的准确性和建筑语义的可传递性得到验证。但不容忽视的是,采用的基于 IfcOWL 的语义化建模方法仍然存在弊端。具体来说,IfcOWL 完全复现了 IFC 标准,这使得所建立的语义化模型数据模式复杂,并且存在大量冗余信息。此外,引入 IfcOWL 本体作语义赋能后的语义化模型体量较大,难以发挥该模型在知识查询与推理方面的优势。

## 5 结论

通过案例分析验证了语义化模型进行建筑信息交付的数据准确性和语义可传递性,进而论证了该建模方法的可行性。此外,通过对所转化模型数据模式进行建筑内涵分析发现,于 IfcOWL 的语义化建模方法尚未在建筑业全面推广有如下原因:

1) 建筑业通用本体不够完善,IfcOWL 本体数据模式完整继承了 IFC 标准,这种复杂的数据层次关系制约了模型知识查询与推理功能的实践。

2) IFC 模型向 RDF 模型转化数据冗余现象突出,完全转译 IFC 模型信息的语义化建模手段信息传递效率较低。

3) 建筑项目的信息管理具有大数据特征,对所有建筑信息执行该语义化建模的经济性和可行性不足。

针对上述问题,除本文提及的冗余信息规避方法外,还可以结合语义化模型各模块的建筑信息内涵,通过领域本体开发和轻量化语义建模来实现高效的信息表达和交付,进而将该技术推广实践于安全风险分析、建筑性能管理和数字孪生建模等涉及异源数据集成的应用领域。

## 参考文献

- [1] 刘照球, 李云贵, 吕西林, 等. 基于 BIM 建筑结构设计模型集成框架应用开发[J]. 同济大学学报(自然科学版), 2010, 38(7): 948-953.
- [2] 赖华辉, 邓雪原, 刘西拉. 基于 IFC 标准的 BIM 数据共享与交换[J]. 土木工程学报, 2018, 51(4): 121-128.
- [3] 马智亮, 滕明焜, 任远. 从 BIM 模型提取建筑能耗监测静态数据的方法[J]. 哈尔滨工业大学学报, 2019, 51(12): 187-193.
- [4] PAUWELS P, ZHANG S J, LEE Y C. Semantic Web technologies in AEC industry: A literature overview [J]. Automation in Construction, 2017, 73: 145-165.
- [5] 魏国海, 刘才玮, 曹永升, 等. 钢筋混凝土梁火灾损伤的多元信息融合模型[J]. 土木与环境工程学报(中英文), 2022, 44(6): 153-161.
- [6] LIU Z S, SHI G L, ZHANG A S, et al. Intelligent tensioning method for prestressed cables based on digital twins and artificial intelligence [J]. Sensors, 2020, 20(24): 7006.
- [7] 张立奎, 段大猷, 王佐才. 基于 LSTM 神经网络的多源数据融合桥梁变形重构方法[J]. 土木与环境工程学报(中英文), 2022, 44(3): 37-43.
- [8] PAUWELS P, TERKAJ W. EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology [J]. Automation in Construction, 2016, 63: 100-133.
- [9] BERNERS-LEE T, HENDLER J, LASSILA O. The Semantic Web [J]. Scientific American, 2001, 284(5): 34-43.
- [10] International BuildingSMART. BuildingSMART - Technical Roadmaps [EB/OL]. [2023-03-02]. <https://www.buildingsmart.org/about/technical-roadmap/>.
- [11] GILCHRIST A, ZENG M L, CLARKE S D, et al. Logic and the organization of information - an

- appreciation of the book of this title by Martin Frické. A set of short essays [J]. *Journal of Information Science*, 2013, 39(5): 708-716.
- [12] W3Working GroupC OWL. OWL2 web ontology language document overview (second edition) - W3C Recommendation, 11 December, 2012 [EB/OL]. [2023-03-02]. <https://www.w3.org/TR/owl2-overview/>.
- [13] Schneider M. OWL 2 web ontology language RDF-based semantics (second edition)-W3C recommendation, 11 December, 2012 [EB/OL]. [2023-03-02]. <https://www.w3.org/TR/2012/REC-owl2-rdf-based-semantics-20121211/>.
- [14] HU Z Z, LENG S, LIN J R, et al. Knowledge extraction and discovery based on BIM: A critical review and future directions [J]. *Archives of Computational Methods in Engineering*, 2022, 29(1): 335-356.
- [15] MAZAIRAC W, BEETZ J. BIMQL - An open query language for building information models [J]. *Advanced Engineering Informatics*, 2013, 27(4): 444-456.
- [16] KRIJNEN T, BEETZ J. A SPARQL query engine for binary-formatted IFC building models [J]. *Automation in Construction*, 2018, 95: 46-63.
- [17] WU H T, ZHONG B T, LI H, et al. Combining computer vision with semantic reasoning for on-site safety management in construction [J]. *Journal of Building Engineering*, 2021, 42: 103036.
- [18] KARAN E P, IRIZARRY J, HAYMAKER J. BIM and GIS integration and interoperability based on semantic web technology [J]. *Journal of Computing in Civil Engineering*, 2016, 30(3): 04015043.
- [19] RASMUSSEN M H, LEFRANÇOIS M, SCHNEIDER G F, et al. BOT: The building topology ontology of the W3C linked building data group [J]. *Semantic Web*, 2021, 12(1): 143-161.
- [20] BONDUEL M, ORASKARI J, PAUWELS P, et al. The IFC to linked building data converter: Current status [C]//6th linked data in architecture and construction workshop. 2018: 34-43.
- [21] NIKNAM M, KARSHENAS S. A shared ontology approach to semantic representation of BIM data [J]. *Automation in Construction*, 2017, 80: 22-36.
- [22] PAUWELS P, KRIJNEN T, TERKAJ W, et al. Enhancing the ifcOWL ontology with an alternative representation for geometric data [J]. *Automation in Construction*, 2017, 80: 77-94.

(编辑 胡英奎)