An Approach to Translate Korea Building Act into Computer-readable Form for Automated Design Assessment

Hyunsoo Lee [a], Sangik Lee[a], Seokyung Park[a] and Jin-Kook Lee [a]

[a]Dept of Interior Architecture Design, Hanyang University, Seoul, Republic of Korea
E-mail: hyunsoolee120@gmail.com, kignaseel@gmail.com, seokyung.park529@gmail.com, designit@hanyang.ac.kr

Abstract

This paper aims to describe an approach of implementation for an interpreter that translates the human natural language into a computer-readable form. Selected building permit-related sentences in Korea Building Act are within the scope of this paper. Among the various applications of BIM, the benefit of automated design assessment has been reported by some leading BIM challengers. The related regulations are usually defined in design guidelines, RFPs, Building Code, etc. The regulations are written in human-readable languages, and sometimes their implicit definitions hinder translating into the explicitly defined computer-readable forms. This paper focuses on the development of a translation process of converting Korea Building Act sentences into specific computer-readable forms such as a mid-level pseudo code and BERA (Building Environment Rule and Analysis) language. The scope of the research process and the overview of the approach to the translation from natural language sentences as follows; 1) Classification of the building objects and related properties from Korea Building Act sentences, 2) Classification of methods from predicate in sentences, 3) Parsing natural language sentence within logic rule-based process. In an actual implementation stage, this translation will be demonstrated by an actual GUI-based application.

Keywords -

BIM (Building Information Modeling), Automated Design Assessment, Rule-checking, Logic rule, Computer-readable form

1 Introduction

Rules or regulations in the process of design assessment for acquiring building permits have been written by human and interpreted by domain experts in general. Verifying multiple requirements or design code with given building design model is usually tedious and time-consuming, and sometimes it causes incomplete or contradictory results. With an advent of computer applications, there has been an interest in restructuring design regulation codes into machine- readable forms for improvement of the logical structure of regulatory codes [1].

The application of Building Information Modeling (BIM) in AEC industries has led to supporting computer-interpretable building models and it enables automated design assessment so that it becomes to reduce an error and improve design quality continuously [2,3]. Among various tasks of the automated design assessment process, this paper focuses on an approach and a mechanism to generate computer-readable explicit forms from implicit natural language sentences especially in Building Act of Korea for acquiring building permits. In the perspective of facilitating computer-executable rules from natural language regulations, various applications of automated design review tasks have been explored by several parties in actual projects. Example case studies are as following.

1) Norwegian Project (2009)

The Norwegian project is for development of methods in order to translate and transform building related codes in standard document, national codes and regulations for use in digital rule checker software. This project suggested 6 stages as standardization process; 1) definition of the scope and source for the rule set, 2) computability assessment, 3) committee assessment, 4) logic rule notation, 5) selection of rule format, and 6) implementation of the rule in rule checker software [4].

2) CORENET (2005)

CORENET e-PlanCheck is Singapore's automated code checking system that consists of e-submission system and integrated plan checking. Integrated plan checking is automated checking process for IFC-based files and leading-edge systems that integrate expert knowledge in regulations, AI (Artificial Intelligence), and BIM technologies [5, 6]. For checking the compliance of codes, this project used FORNAX platform to calculate each required condition.

3) SMART Code (2006)

ICC (International Code Council)'s SMARTCode project has developed to automate code compliance checking for the I-Codes and Federal state. The automated code compliance check takes the extracted entity for code checking in STEP file and converted entity information by XML schema. XML extracted from STEP and from documents information could be compared and the legality could be checked [7].

4) GSA (2007)

The GSA and US Courts have supported development of design rules checking of federal courthouses, which is an early example of rule checking applied for automating design guides [8]. Georgia Institute of Technology research team engaged in this project, and they developed a BIM application automating rule checking based on the Courts Design Guide suggesting the prototype of an assessment tool for architects and clients. They interpreted the current rules into a machine readable form for automated rule checking. Through the assessment tool, the appropriate feedback to designers and the adequate data to make the needed assessment were able to be used for getting result of design assessment [9].

The rule-based system for design review helps architects to validate building designs effectively. However, current rule-checking software usually requires skilful knowledge of computer programming technique to make new rule-sets to meet architects' demand. That is, current rule-making process is not user-friendly, but rule checking software centred. To resolve this issue, this paper aims to describe a user-friendly approach to the rule-making process based on the sentences in Korea Building Act. This paper also shows a mechanism of such rule-making and partially demonstrates a developed software as an outcome of ongoing research project.

2 Research Scope and Objective

The scope of this study is the approach of translating Korea Building Act sentences composed of natural language into computer-readable forms that can be exported as a rule-set for rule-checking program. We suggest processes of translating into computer-executable form from natural language and its related software interface.

As shown in Fig 1, we suggest the logic rule-based mechanism as follows; 1) Classification of object and its related properties dealt in Korea Building Act rather than a general standard of BIM model such as ifc or Revit schema, 2) Classification of method for checking target building objects and properties, 3) Parsing natural language sentence in Korea Building Act according to logical process. We analyzed Korea Building Act Article 49 ("Egresses from Buildings and Restrictions on their Use"), 53 ("Floor Area Ratio"), 56 ("regulations on basement level), and 64 ("Elevators") and derived 44 regulations with 468 atomic sentences from these articles. The results of each mechanism are managed in web-based database [10]. Also, this web-based database can be applied to "KBimLogic" that is software we have been developing for this research. KBimLogic is GUI-based software that exports executable rule-set file from natural language sentences. Reported rule-set file is going to be used in BIM assessment tool "KBimAssess" for checking building permit in Korea.

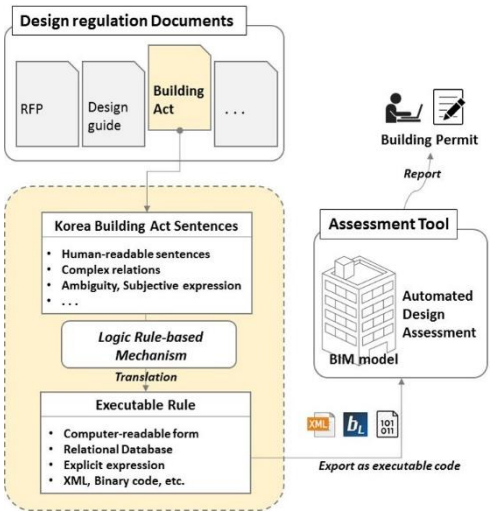

Figure 1. Overview Diagram of Research Scope and Flow

3 Current Rule-making Approaches

Most of the software for performing automated design rule checking is done by software developers, and this approach may lead to lack of reliability of the result [4], because of some limitations. This chapter describes the current state of rule checking software that is executed based on the built-in rule set and its limitation of researches and analysis done before by domain-specific researchers. In addition, this paper suggests building object-oriented approach for code checking compared with previous rule-set oriented approach. Solibri Model Checker (SMC) and Express Data Manager (EDM) are two major systems currently available that provide object-based rule checking system [11].

3.1 Implementation-centred Approach

Developer-based coordination of rule-set in software such as SMC (Solibri Model Checker) [12] is one of the most well-known programs for design assessment. SMC

and its built rules are based on JAVA, and if users require new rules, Solibri developers should develop customized rules for specific purpose. Most of the rule-set in SMC is hard to be coded into the software, and it is not user-friendly language to specify new rules [13]. Users should combine existing rule-set and configure each parameter of rules themselves as shown in Figure 2, otherwise, SMC developers should add new rule-set in the software. Moreover, all rule sets built in SMC cannot be adopted in other rule checking softwares.

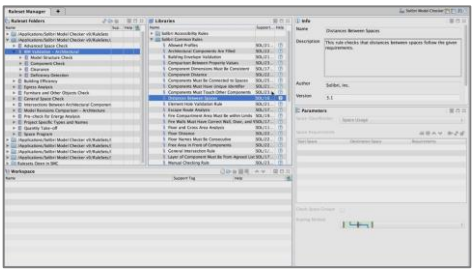

Figure 2. Screen Shot of Rule Manager Workstation in Solibri Model Checker

Therefore, the SMC rule base proved inadequate for checking building regulations and specific building codes since constraints were not able to encode design requirements at the required level in early research of analysis of SMC in terms of automated code checking [11].

3.2 Rule Language-based Approach

The EDM Model server [14] is operated with an object-based database and requires EXPRESS and EXPRESS-X that is an object-oriented query language. EXPRESS is a data modelling language defined in ISO10303-11 as an open international standard. Rules are written in this form, so that they can be used in other software that understand EXPRESS-X. Similar to SMC, the EDM Model server consists of a set of built in rules. The EDM Model server supports encoding of wide domain-specific knowledge in flexible way [11], however, a high level of expertise is required users to deal with rules in the EDM model server. To develop new rules, EDM Technology should develop the necessary rules using EXPRESS.

3.3 Building Object-oriented Approach to Generate Computer-readable Forms

Comparing the approaches of current rule-making as mentioned, we suggest building object-oriented and human-friendly approach. Building Act sentence has difficulty for rule checking using combination of rule-set and its parameter in general rule-checking software, because of their complex relations in one sentence and some of subjective expression. We focus on 'building object' defined in Korea Building Act and suggest logic rule-based mechanism for making rule-set file from natural language sentences. As described in Figure 3, human-readable language from Korea Building Act can be raw data for logical process before translated as rule-set file.

Each building object and associated properties in Building Act sentences can be handled intuitively by user who is a non-expert in language programming. This process is executed in GUI program so that users can select each content in natural language form according to the Act sentences' intention and export as computer-readable form.

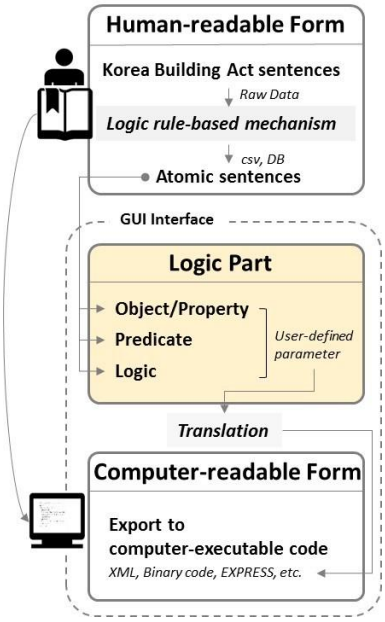

Figure 3. Overall Flow Diagram of Converting Process of Korea Building Act Sentences and its Implementation

4 Logic rule-based Mechanism of Natural Language Sentence of Korea Building Act

Natural language (Human language) includes infinite extensibility of expression and the length of phrases, thus ambiguity and vagueness are inevitable issues [15]. For translation of human-readable language into computer-readable form, logical converting process is necessary. In other words, for the automated design assessment of BIM model, it is important to translate to computer-readable form as restructuring process. This chapter describes

logic rule-based mechanism by three parts in terms of sentence structure: Noun (Object/Property), Predicate (Method), Logic (Parsing process).

4.1 Noun: Building Objects and Properties

In the structure of the Korea Building Act, there are nouns that make up the content in sentences and those definition. In general, the definition of a noun is described in the way of binding to multiple nouns in the name of 'definition' or expressing with the brackets in a sentence that contains the noun and its coverage. Otherwise, the definition of the noun is connected by other clauses. Therefore, definition of the noun is to be also taken into account the connectivity of the relevant regulations, as well as other clause, although exceptional items exist. It requires database shown in Figure 4.

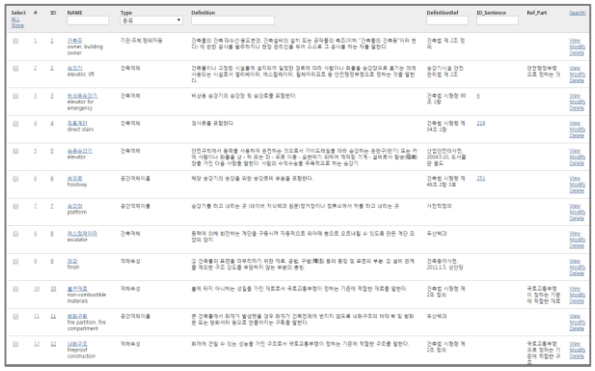

Figure 4. Web-based Database of Object Name and Definition

We extracted the nouns that appear in Building Act sentences and connected them with reference Building Act clause. We set up a relationship between connected nouns and enter the ID (Identity) in web-based database described in Figure 4. The importance of defining a name for each object and setting up a relationship has already been revealed in the previous study [16]. Moreover, name in English is required as name set (Korean-English) for computer-readable form, because there has not been programming language specified by Korean. This database contains ID, name in Korean and English, classified type, definition, and Act clause reference ID.

Based on the name database, we classified objects and properties by the logic rule criteria that is specified by Building Act's characteristic. In Korea Building Act, most of clauses deal with in-memory level objects for rule-checking. Figure 5 shows the classification of object (a) and property (b). The classification of objects can be separated into two parts; 1) target object for assessment, and 2) non-target object. In target object criteria, there are two categories; 1) BIM-enabled object: space object and building object, 2) Non-BIM object: circulation, geometric information, and so on, that exists in memory. Also we extract the properties required for assessment. The classification of property consists of two parts; 1) Instance level of property: The contained properties in this level can be derived from ifc instance criteria. Through the relation between multiple objects, this property such as inclusion, distance, connectivity, and direction, and so on, can be verified. 2) Class level of property: The contained properties in this level can be derived from ifc schema or class. This property can be generated in a BIM model basically, such as name, height, length, and area, and so on.

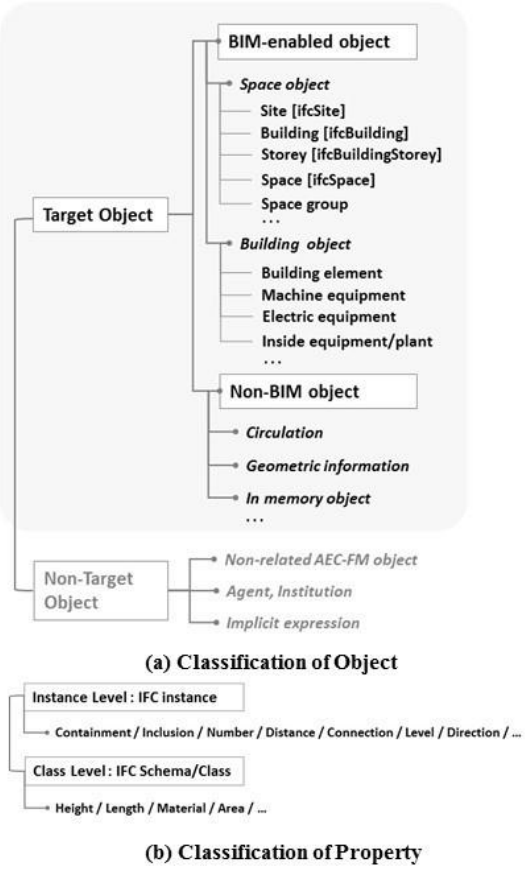

Figure 5. Classification of Object and Property Defined in Korea Building Act

4.2 Predicate: Classification of Rules

This chapter describes the classification of methods related to predicate part of sentence. The main issue of method for rule-checking is querying object and then verifying its property. As shown in Figure 6, the method can be applied as an intermediate role for mapping between component of sentences and BIM model's

objects and properties.

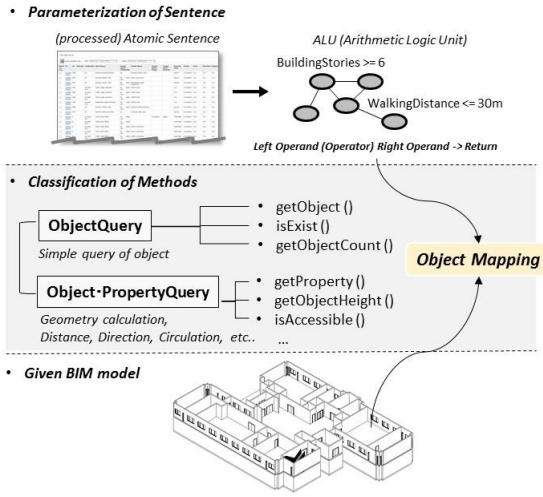

Figure 6. Association the Method with ALU and BIM model

Table 1. Classification of Basic Methods for Korea Building Act Assessment

Instance	Property Type	Property Character	Specified Method
Obj query		Object query	getObject()
		Existence	isExist()
		Number	getObjectCount()
Obj, Prop query	Basic property	Property	getProperty()
		Material	getMaterial()
		Usage	getObjectUsage()
	Geometry	Height	getObjectHeight()
		Length	getObjectLength()
		Area	getObjectArea()
		Gradient	getObjectGradient()
	Complex compute	Material type	getMaterialType()
		Illuminance	getSpaceIlluminance()
		Structure	getObjectStructure()
		Fire resistant	isFireReistant()
		Fire proof	isFireProof()
		Fire compartment	isFireCompartment()
	Relation	Inclusion	hasObject()
		Distance	getObjectDistance()
		Connection	isConnectedTo()
		Circulation	isAccessible()
		Direction	getDoorDirection()

The second stage of classification is the types of property, and third stage is about the attribute of property. In this stage, representative rule methods are defined according to the property's attribute. The methods can be expanded by various parameters.

4.3 Logic: From Sentences to Executables

There are condition clause and content clause in one Building Act sentence usually. In this case, target building model should satisfy conditional clause before checking content clause. One clause is generally composed of noun (subject, object) and verb (predicate). In logic rule-based criteria, an atomic sentence is a type of declarative sentence which is either true or false and which cannot be broken down into other simpler sentences [17]. In other word, atomic sentence can be expressed an each single S (subject) + O (object) + V (verb) structure. We aim to convert natural language sentence to the arithmetic logic unit (as independent syntactical units) for restructuring sentences. Figure 7 describes the logic rule-based parsing process of natural language sentence.

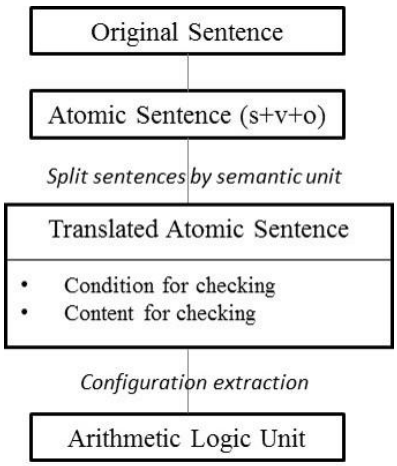

Figure 7. Parsing Process of Sentence (from original sentence to ALU)

As an example of parsing process of sentence, we choose Building Act, Article 64, (1) clause that is about condition of elevator installation. The following 1 to 6 describes each parsing process.

1. Original Sentence

A project owner of a building (excluding buildings prescribed by Presidential Decree) with six or more floors and a total floor area of 2,000 square meters or more shall have an elevator installed therein. In such cases, the size and structure of elevators shall be prescribed by Ordinance of the Ministry of Land,

Infrastructure and Transport. (Building Act, Article 64, (1), Elevators) [18].

2. Atomic Sentence

A building with six or more floors and a total floor area of 2,000 square meters or more shall have an elevator installed therein.

3. Translated Atomic Sentence (TAS)

TAS 1 (Condition)
-A building has six or more floors and a total floor area of 2,000 square meters or more
TAS 2 (Content)
-A building shall have an elevator installed therein.

4. Configuration Extraction from TAS

TAS 1 (Condition)
-Number of Floors (more than 6)
-Total Floor Area (more than 2000 ㎡)
TAS 2 (Content)
-Install (Elevator)

5. Arithmetic Logic Unit

- getBuildingFloor()>=6,
- getTotalFloorArea()>=2000 ㎡
- isExist(Elevator)

6. Expression of Method and Relation

IF
getBuildingFloor()>=6
AND getFloorArea()>=2000 ㎡
THEN {isExist(Elevator)}

Parsing the natural language sentence is necessary process of translation for computer-readable form. This process is now in progress by manual, however, the eventual goal is automated sentence parsing. As shown in Figure 8, 448 atomic sentences derived from our scope are managed in web-based database. Each atomic sentence and its reference Act clause have their own ID so that we can track other clause sentences in relation.

Figure 8. Web-based Database of Building Act Sentences and Those Derived Atomic Sentences

5 Implementation

This chapter describes an approach to implementation as GUI program for generating computer-executable file. Figure 9 shows overview diagram of KBimLogic program and its functional modules that customize logic rule-based process we mentioned in chapter 4.

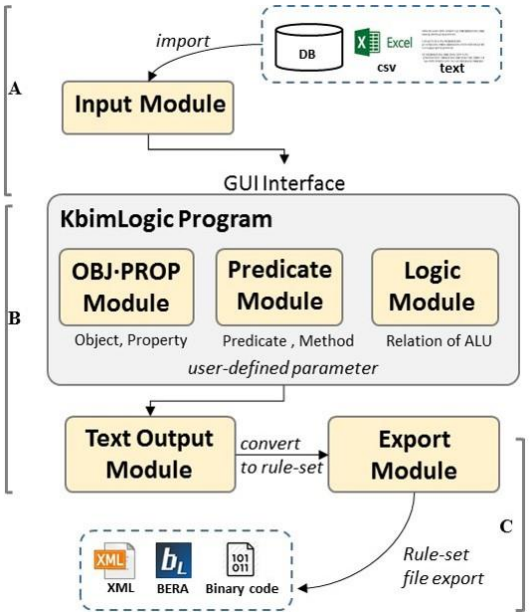

Figure 9. Overview of KBIMLogic Program with associated functional modules. A: Human-readable language part (original sentences of Korea Building Act), B: Intermediate part between Human-readable and Computer-readable code, C: Computer-readable part (XML, Binary code, BERA Language etc.)

1. Input Module: Human-readable part

Input module is the first part (Human-readable) that imports natural language sentences in form of database, csv, or text file. In this implementation, we linked web-based database and data source of the input module with SQL database table, so that parsed atomic sentences for condition and content in Building Act that a user wants to translate can be imported

2. KBimLogic: Intermediate part

1) OBJ·PROP Module

Object and Property module basically imports the object and property of clause that user selects in input module. Also, among the imported objects and properties, user can choose for specific purpose or enter manually.

2) Predicate Module

In predicate module, user can select appropriate type of method according to condition and content of clause sentences.

3) Logic Module

By selected type of method in the predicate module, user can define the type of relation between condition and content of clause.

3. Text Output Module: Intermediate part

Text output module functions as translating human-readable form to computer-readable one. Selected objects, properties, their method, and relation are converted to script language such as pseudo code form.

4. Export Module: Computer-readable part

Export module exports the script code converted in the text output module to computer executable rule-set file. Form of rule-set can be as XML, binary code, or BERA (Building Environment Rule and Analysis) Language.

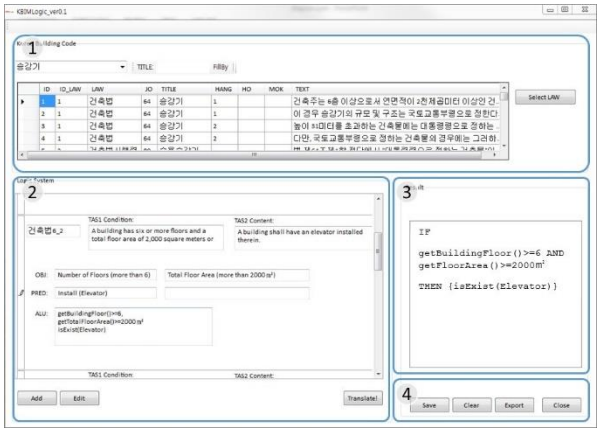

Figure 10. An Example Screen Shot of KBimLogic Software v0.1 (under the development). 1) Input Module, 2) OBJ • PROP Module, Predicate Module, Logic Module, 3) Text Output Module, and 4) Export Module.

Figure 10 is a screen shot of implemented KBimLogic software that is composed of logic rule-base functional modules described in Figure 9. KBimLogic is implemented in Visual Studio C# programming environment using SQL data server. Users can choose Building Act clause which they want to translate shown in Figure 10, number 1 area of KBimLogic interface. In number 2 area, user can select object, property, and method intuitively. In other word, user can restructure a sentence with a combination of each component. In number 3 area, selected each option is translated as pseudo code, then exported as rule-set file in number 4 area.

6 Summary

This paper introduced a logic rule-based mechanism for translating natural language sentences in Korea Building Act into the executable and computer-readable form, and demonstrated its implemented software as a part of still ongoing project. We suggested classification of building objects, derived properties, and methods from sentences in Korea Building Act, as well as logical paring process of the natural language sentences. The automated process of natural language by computer per se has not been completed yet, however, it is important to secure the integrity of flawless logical process of natural language sentences. Further development of structured logic rule-based mechanism is strongly expected by continuous research in the future. This could have a significant positive impact on the development of automated design assessment tools to meet a wide range of design assessment requirements, not only national legislation.

Acknowledgment

This research was supported by a grant (14AUDP-C067809-02) from Architecture & Urban Development Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

References

[1] C. Eastman, Jae-min Lee, Yeon-suk Jeong, Jin-kook Lee, Automatic rule-based checking of building designs, *Automation in Construction*, 18 (2009):1011–1033, 2009.

[2] Chuck Eastman, Paul Teicholz, Rafael Sacks, Kathleen Liston, *BIM HANDBOOK-A guide to Building Information Modeling for Owners, Managers, Designers, Engineers, and Constructors*, 1.6, 5.2-5.3, John Wiley & Sons, Inc, New Jersy, 2007.

[3] Jin-Kook Lee, Charles M. Eastman, Yong Cheol Lee, Implementation of a BIM Domain-specific Language for the Building Environment Rule and Analysis, *Journal of Intelligent & Robotic Systems*, DOI 10.1007/s10846-014-0117-7, 2014.

[4] Eilif Hjelseth, Foundation for development of computable rules, Presented at *CIB-W78*, pages 1-10, Turkey, Istanbul, 2009

[5] Jungsik Choi, Inhan Kim, An Approach to Share Architectural Drawing Information and Document Information for Automated Code Checking System, *TSINGHUA SCIENCE AND TECHNOLOGY*, Volume 13, Number S1, pages 173-178, 2008.

[6] CORENET e-PlanCheck: *Singapore's Automated Code Checking System*, AECbytes "Building the Future" Article, pages 1-8, 2005.

[7] ICC, Smart Code Project Definition, On-line: http://web.stanford.edu/group/narratives/classes/0809/CEE215/ReferenceLibrary/International%20Code%20Council%20%28ICC%29/SMARTcodes%20fact%20sheet%2011-01-07.pdf, Accessed: 10/01/2015

[8] The National 3D-4D-BIM Program Office of the Chief Architect Public Buildings Service U.S. General Services Administration, "BIM Guide Series 02 v096", General Services Administration, On-line: http://www.gsa.gov/bim, Accessed: 28/12/2014

[9] GSA Project- BIM Enabled Design Guide Automation, On-Line: http://dcom.arch.gatech.edu/gsa/, Accessed: 20/01/2015

[10] Open BIM based Technological Environment for Building Design Quality Enhancement, On-line: www.designitlab.kr/bim, Accessed: 22/01/2015

[11] Lan Ding, Robin Drogemuller, Julie Jupp, Mike A Rosenman, John S Gero, Automated Code Checking, *Clients Driving Innovation International Conference*, QUT Digital Repository, On-line: http://eprints.qut.edu.au/27228, Accessed: 01/01/2015

[12] Solibri Model Checker® (SMC), an IFC-based rule checking BIM tool Available, On-line: http://www.solibri.com, Accessed: 28/12/2014

[13] Håvard Bell, Lars Bjørkhaug, Eilif Hjelseth, Standardized Computable Rules, Standards Norway, Strandveien 18 P.O.Box 242, NO-1326 Lysaker, Norway, 2009.

[14] EDMmodelServer(ifc)™, One-line: http://www.epmtech.jotne.com/edmmodelserver-ifc, Accessed: 29/11/2014

[15] Nawari O. Nawari, Automated Code Checking in BIM Environment, *14th International Conference on Computing in Civil and Buildng Engineering*, Moscow, Russia, 2012.

[16] Jin-Kook Lee, Jaemin Lee, Yeon-suk Jeong, Hugo Sheward, Paola Sanguinetti, Sherif Abdelmohsen, Charles M. Eastman, Development of space database for automated building design review systems, *Automation in Construction* 24 (2012) 203–212, 2012.

[17] Definition of Atomic Sentence, On-line: http://en.wikipedia.org/wiki/Atomic_sentence, Accessed: 14/01/2015.

[18] Statutes of the Republic of Korea, On-line: http://elaw.klri.re.kr/, 20/01/2015