AUTOMATED HAZARD IDENTIFICATION FRAMEWORK FOR THE PROACTIVE CONSIDERATION OF CONSTRUCTION SAFETY

JunHyuk Kwon[1], Byungil Kim[2], SangHyun Lee[3], and Hyoungkwan Kim[4]

[1] MSE student, Department of Civil & Environmental Engineering, University of Michigan, USA
[2] Postdoctoral Fellow, Department of Civil & Environmental Engineering, University of Michigan, USA
[3] Assistant Professor, Department of Civil & Environmental Engineering, University of Michigan, USA
[4] Associate Professor, School of Civil & Environmental Engineering, Yonsei University, Republic of Korea
Send correspondence to junkwon@umich.edu

ABSTRACT: Introducing the concept of construction safety in the design/engineering phase can improve the efficiency and effectiveness of safety management on construction sites. In this sense, further improvements for safety can be made in the design/engineering phase through the development of (1) an automated hazard identification process that is little dependent on user knowledge, (2) an automated construction schedule generation to accommodate varying hazard information over time, and (3) a visual representation of the results that is easy to understand. In this paper, we formulate an automated hazard identification framework for construction safety by extracting hazard information from related regulations to eliminate human interventions, and by utilizing a visualization technique in order to enhance users' understanding on hazard information. First, the hazard information is automatically extracted from textual safety and health regulations (i.e., Occupational Safety Health Administration (OSHA) Standards) by using natural language processing (NLP) techniques without users' interpretations. Next, scheduling and sequencing of the construction activities are automatically generated with regard to the 3D building model. Then, the extracted hazard information is integrated into the geometry data of construction elements in the industry foundation class (IFC) building model using a conformity-checking algorithm within the open source 3D computer graphics software. Preliminary results demonstrate that this approach is advantageous in that it can be used in the design/engineering phases of construction without the manual interpretation of safety experts, facilitating the designers' and engineers' proactive consideration for improving safety management.

Keywords: Construction Safety, Natural Language Processing, Automation

1. INTRODUCTION

The construction industry is well known for the seriousness of its accidents, including fatal injuries. According to the bureau of labor statistics [1], the construction industry occupies 18.5% of all work-related fatalities employing 5.4% of the workforce in the United States. In addition, construction has the highest rate of fatal injuries among all industries in the United States; the fatality incidence rate in construction is peaked up to 9.5% whereas the industry average is 3.5% [1].

Some studies have shown that a fairly large percentage of construction accidents could have been eliminated, reduced, or avoided by making better choices in the design and planning stages of a project [2] [3] [4]. To elaborate, the design of a parapet wall should be 42 inch tall and a fiberglass roof panel should be installed with a guardrail during construction or maintenance in order to reduce fall hazards at work. That is, construction designers can influence construction safety by the proactive considerations in the early phases of a project. This will contribute to the reduction of contractors and workers' mistakes that can lead to accidents.

Despite of such a high reduction potential, however, designers have not focused on the safety of construction workers. Safety is usually not addressed until construction begins and the safety of construction workers is only left up to the contractors [5]. Traditionally, designers have just focused on designing buildings or facilities simply based on building codes and accepted engineering practices. One reason for this lack of designers' attention is that they have a limited experience and knowledge for practical construction. Thus, now the question is how to consider the safety of construction workers in the design phase of a construction project by designers without such experience and knowledge.

The main goal of this study is to provide an automated hazard identification framework for construction designers to enhance construction workers' safety. This framework consists of three major functional elements: (1) automatic extraction of hazard information from textual knowledge such as Occupational Safety and Health Administration (OSHA) regulations; (2) automated activity duration and sequence generation, and (3) visualization for extracted hazard information with work schedule.

This paper is organized as follows. To start, we thoroughly investigate previous research efforts. Then, we describe the automated hazard identification framework, with an emphasis on three main components (information extraction from safety regulations using natural language processing, automatic schedule and sequence generation, and visualization of extracted hazard information). A case study is also conducted to implement the proposed framework. Finally, we conclude this paper by introducing the ongoing and future work.

2. LITERATURE REVIEW AND RESEARCH OBJECTIVE

2.1 Previous Research Efforts

Efforts have been made to guide construction designers toward identifying the hazards of construction projects. One approach for identifying potential construction hazards is to collect safety knowledge from previous projects and to develop a kind of checklist [6] [7] [8]. Another approach is to utilize a 3D/4D visualizing tool for visually considering construction safety issues in the design phase [9] [10]. Recently, an automated rule-based safety checking system in conjunction with 4D CAD models was developed [11].

The previous studies significantly improved the safety of construction workers when performing design tasks. However, the following issues are still needed to be addressed. Firstly, most approaches depend on manual regulatory interpretation or compliance checking, which is time-consuming and demands the designer's familiarity with safety hazards. Secondly, text-based checklists not only require human interpretation with expert knowledge but also depend on training and educating designers based on empirical studies. Thirdly, the CAD-based approach requires detailed sub-activities and production tasks information that may not be available in the design phase. Finally, although the automated rule-based safety checking coupled with 4D CAD models is the-state-of-the-art approach, this requires human knowledge and experience in safety rules that can be checked in the context of a 4D safety simulation. In addition, an interoperability problem between applications is also identified.

2.2 Need for Automated Hazard Identification and Conformity Checking

Hazard identification and regulation conformity checking in construction are a very complex problem and require a multidisciplinary approach. This difficulty, in one respect, comes from the large amount of non-formalized expert knowledge. In most cases, the conformity-checking process remains manual, and as a result, is not very effective as it is prone to be time-consuming and involve errors from human interpretation. In addition, a lot of results are deduced or interpreted with the help of tacit expert knowledge [12]. In this context, automating the process of conformity checking with regard to safety regulations is a highly desirable in the hazard identification process

To address these limitations, we develop a framework for automated hazard identification that is little dependent on a safety expert's knowledge regarding safety in the regulatory compliance checking process. This framework has the purpose of enhancing the understanding of the safety planning process and safety regulation requirements with visualization, and facilitating the designer's proactive consideration for safety.

3. METHODOLOGY

The proposed hazard identification framework is illustrated in Figure 1. At first, the regulatory rules will be

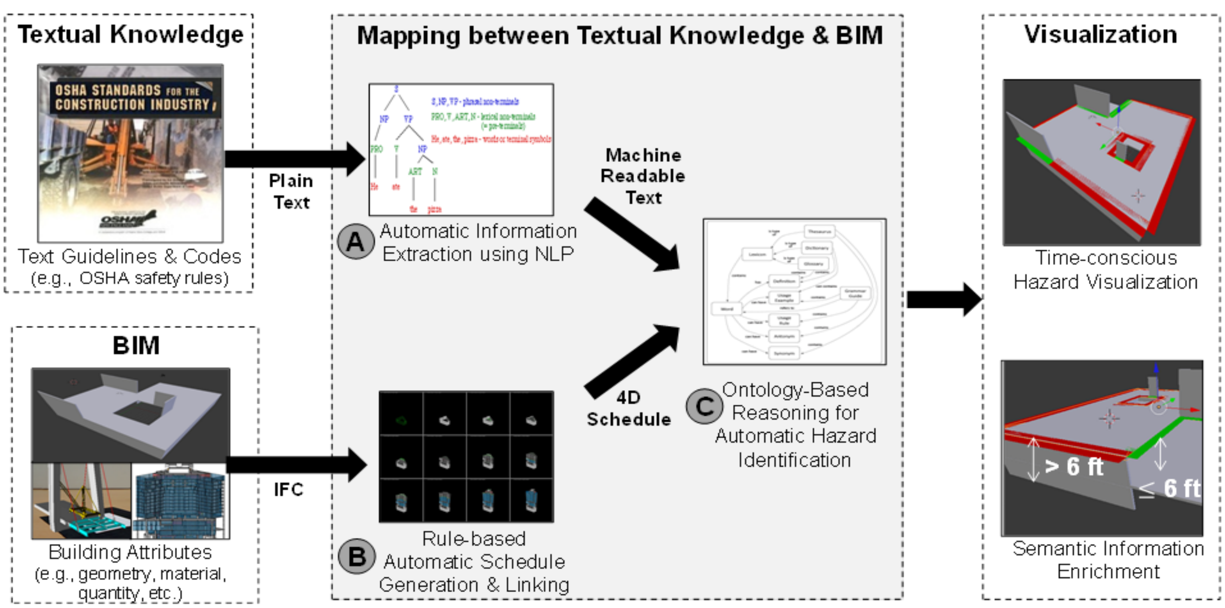

Figure 1. Automated Hazard Identification Framework

extracted and coded from the text-based codebooks using a natural language processing technique and domain ontology. In addition, the attributes of building elements in BIM are extracted to automatically generate schedule. Finally, rules and building attributes are mapped and visualized to identify hazard.

3.1 Automatic Hazard Information Extraction Using Natural Language Processing (NLP) Techniques

In this process, as denoted by A in Figure 1, hazard information is extracted from textual safety and health regulations by using NLP techniques—a syntactic parsing and a semantic analysis with domain-specific ontologies [13]. A syntactic parsing is a method to perform sentence analysis according to the rules constructing sentences, whereas a semantic analysis using ontology is capable of extracting information recognizing the meaning of the domain-specific terminology and contexts.

Information systems increasingly depend on ontology to structure data in a machine-readable format and to ensure satisfactory performance. There have been significant research efforts to automate the code compliance checking process [13] [14]. However, these approaches/tools require the manual extraction of rules from regulatory documents, as well as the manual encoding of these rules. Information Extraction (IE), a subfield of NLP, aims at extracting structured data/information from unstructured text automatically. This task has been carried out in this research using the General Architecture of Text Engineering (GATE) system, which has been widely and successfully used in IE [13].

3.2 Automatic Schedule and Sequence Generation

A 4D CAD model integrated with construction schedule is needed to identify hazard information and safety requirements, which vary over time (e.g., different construction processes space utilizations that can have different hazards). In this framework, spatial, geometry, quantity, relationship, and material layer set information that is stored in BIM is automatically extracted (denoted by B in Figure 1) based on sequencing rules (e.g., 2nd floor installation after 1st floor). Then, using the extracted information, task durations can be inferred from productivity and estimating database (e.g., RS Means). Finally, task schedule data integrated with building element can be generated.

3.3 Ontology-based Reasoning and Visual Representation of Hazard Information

The domain textual or tacit knowledge plays an important role in this information extraction & conformity-checking process. To take the domain knowledge into account, ontologies specific to a construction project is pursued as denoted by C in Figure 1. Ontology is represented as a set of concepts within a domain and the description of the relationships between the concepts [15]. As the ontology provides a formal and machine manipulatable model of the domain knowledge as well as a means for encoding/decoding meanings of terms, an effective knowledge representation and sharing system can be built [16].

In this framework, the extracted hazard information is visually integrated into the geometry of building elements and related construction activities data in industry foundation class (IFC) using an internal conformity-checking algorithm within the open source 3D computer graphics software, Blender. Blender features a variety of format conversions, including IFC format and is operable at diverse platforms (e.g., Windows, OS X, Linux, etc.) without any compatibility issue.

It also enables not only a 3D modeling and simulation but also task automation and customization using built-in scripts. In the proposed framework, the user imports a BIM model into Blender whose scripts then, check whether the properties in building elements meet the safety requirements extracted from regulations (e.g., Occupational Safety Health Administration (OSHA) Standards). After running the checking function, results will be visualized (e.g., a building element will turn green in case of conformity while it will be red otherwise).

4. CASE STUDY

As a proof-of-concept, we conducted a case study using the proposed framework. This study put a focus on fall hazard identification. Falls are the most common causes of serious injuries and deaths. Of all occupational fatalities in construction workers, 34.1% were due to falls [1]. Providing designers with information on fall hazards can give a great opportunity to reduce and prevent falls from heights. In this regard, OSHA Standard 1926.501(b)(2) (Figure 2) was selected as a sample text.

Fall Prevention Regulations

– "Leading edges."
  - 1926.501(b)(2)(i)
  - Each employee who is constructing a leading edge 6 feet (1.8 m) or more above lower levels shall be protected from falling by guardrail systems, safety net systems, or personal fall arrest systems.
  - 1926.501(b)(2)(ii)
  - Each employee on a walking/working surface 6 feet (1.8 m) or more above a lower level where leading edges are under construction, but who is not engaged in the leading edge work, shall be protected from falling by a guardrail system, safety net system, or personal fall arrest system.
  - 1926.502(b)(1)
  - Top edge height of top rails, or equivalent guardrail system members, shall be 42 inches (1.1 m) plus or minus 3 inches (8 cm) above the walking/working level
  - 1926.502(b)(2)
  - Midrails, screens, mesh, intermediate vertical members, or equivalent intermediate structural members shall be installed between the top edge of the guardrail system and the walking/working surface when there is no wall or parapet wall at least 21 inches (53 cm) high.

Figure 2. OSHA Standards 1926.501(b)(2) (Subpart M. Fall protection: Leading edges)

It should be also noted that automatic schedule and sequence generation processes (B in Figure 1) was excluded from the scope of the case study. A BIM model of a 1-story office building that features an irregular floor configuration was created in Autodesk® Revit® 2011

4.1 Automatic Hazard Information Extraction

Safety regulatory rules were extracted and coded from the construction standards information identified in

OSHA codebooks using natural language processing techniques and domain ontology.

We adapted information extraction rules provided by El-Gohary and Zhang [13], and then the modified rules were executed in GATE, a rule-based Java Annotation Patterns Engine (JAPE) transducer. Each rule has a left-hand side and a right-hand side. The left-hand side contains information about certain facts' and objects' pattern used to match the text having target information (information to be extracted). The right-hand side encodes the actions to be taken when the target information is matched. When any rules whose left-hand sides are matched in this manner at a given time, their right-hand sides are executed. This continues through the entire set of rules constructed in the module by us.

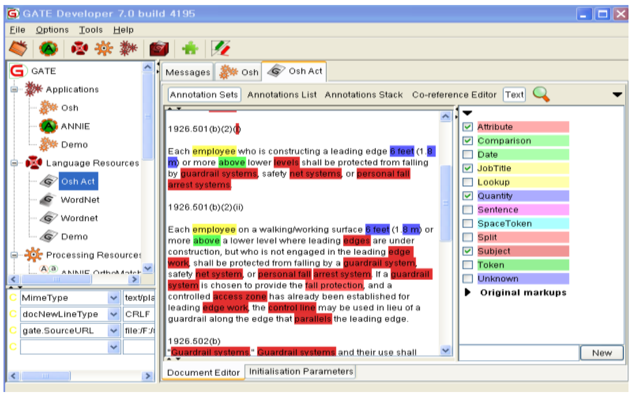

Figure 3. Result of information extraction

A generic ontology, WordNet, was used to represent the extracted information because it sufficiently provides computer manipulation for this study. In addition, a tuple format (e.g., subject, compared, comparison, quantity, protection) was used as an intermediate processing step because it conveniently provides ordering along with optimal performance.

The extracted results are summarized in Table 1. As shown in the table, the information extracted from "a leading edge 6 feet (1.8 m) or more above lower levels shall be protected from falling by guardrail system, safety net system, or personal fall arrest systems" is organized in a tuple format as follows: <(Subject: a leading edge), (Compared: Lower level), (Comparison: above), (Quantity: 6 feet), (Protection: Guardrail system, Safety net system, Personal fall arrest system)>.

Four sentences (18 tuples of hazard information) were randomly selected for the evaluation of extraction correctness. The recall and precision were evaluated by counting all of informational components; subject, compared, comparison, quantity, and protection. Recall is the number of correct results divided by the number of results that should have been extracted while precision is the number of correct results divided by the number of all returned results. Out of 18 tuples in this study, all were correctly extracted except one (i.e., Recall is 94.4% as shown in Table 2).

Table2. Preliminary Results of information extraction

Recall	Precision	F-measure
94.4%	100%	97.1%

Specifically, "42 inches" are captured instead of "42 inches plus or minus 3 inches." Precision is 100%. Therefore, the results of the evaluation indicate that the proposed IE approach is effective in extracting fall hazard information from OSHA Standards.

4.2 Visual Representation of Fall Hazard Information

Extracted fall hazard information was visualized in an IFC-based viewer, Blender for interoperability issue.

Leading edges and openings to be potentially occurred of falls was identified as following steps:

1) Initially assume that all slab edges are hazardous.
2) Consider safe areas to be places where slab edges have a shallow height (less than 6ft) from adjacent lower slabs.
3) Identify the walls/rails in contact with the slabs using a clash detection algorithm.
4) Consider safe areas to be places where slab edges are protected by walls/rails identified in the previous step.

When identifying the unprotected leading edges and openings in the slabs, we used the concept of an open source collision detection library, such as RAPID [17], because the wall-slab relationships are not explicit in IFC data. Once the function of hazard identification is implemented, the following requirements need to be met:

1) Walking/working surfaces (horizontal and vertical surfaces) with an unprotected side or edge that is 6

Table 1. Automatically Extracted Fall Hazard Information

Subject Extracted	Compared Extracted	Comparison Extracted	Quantity Extracted	Protection Extracted
A leading edge	Lower level	above	6 feet	Guardrail system Safety net system Personal fall arrest system
Walking/working surface	Lower level	above	6 feet	Guardrail system Safety net system Personal fall arrest system
Top edge height	Walking/working level	above	42 inches	—
Midrail, screens, mesh, vertical member	Walking/working level	at least	21 inches	—

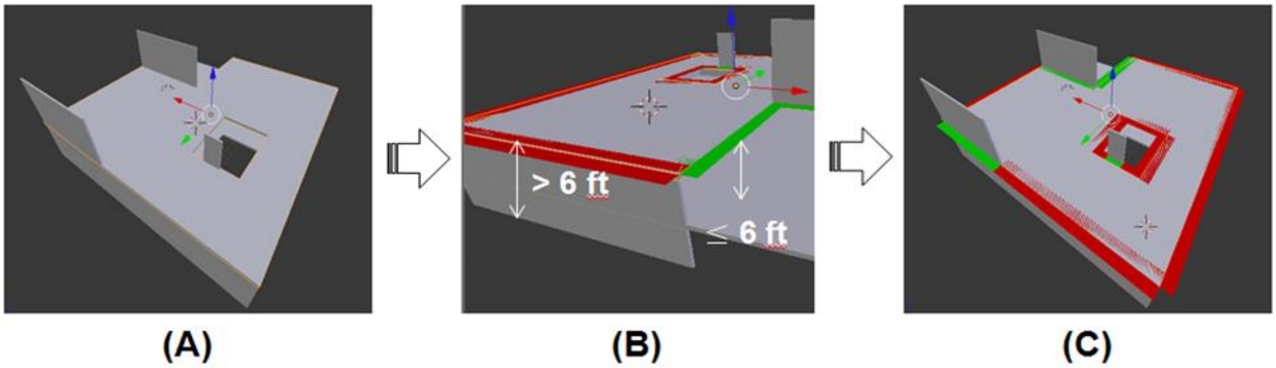

Figure 4. Visualization of the unprotected leading edges and openings

feet (1.8 m) or more above a lower level shall be identified.

2) Working surfaces on, at, above, or near wall openings (including those with chutes attached) where the outside bottom edge of the wall opening is 6 feet (1.8 m) or more above lower levels, and the inside bottom edge of the wall opening is less than 42 inch (1.07 m) above the walking/working surface, shall be identified

The Blender visualized the BIM model with fall hazard information as shown in Figure 4. Red areas show all of the components that potentially violate the applied rule sets. Also, why such violations happen is presented using the extracted rules. For example, red areas in Figure 4 (B) and (C) are caused by violating a rule, "a leading edge 6 feet (1.8 m) or more high from lower levels shall be protected from falling by guardrail system". On the other hands, green areas in Figure 4 (B) and (C) show that an edge is less than 6 feet high from lower levels, which conforms OSHA standards. With such visualization, designers can immediately see which areas need to be changed or not. In this edge example, the slab edges that are not appropriately protected to meet the safety requirement need to be revised. Or designers can ask for installing fall protection equipment on the job site.

5. DISCUSSION AND CONCLUSION

In this paper, we presented an automated hazard identification framework for supporting design for construction safety. We focused on three research aspects: (1) fall hazard information extraction using natural language processing from textual regulation documents, (2) automatic activity schedule generation, and (3) a visual representation of extracted fall hazard information on a building design model. As a preliminary study, we implemented (1) hazard information extraction from OSHA standard using a syntactic parsing technique and (2) visual representation of fall hazards information integrated in a 3D building model. Initial experiment results indicate that automated hazard identification by information extraction and regulation conformity checking is capable of representing fall hazard information graphically without the tacit knowledge of a safety expert. As such, our proposed framework has the potential to enhance earlier decision making for construction safety, as visually represented hazard information makes a user little dependent on a safety expert's knowledge and provides a graphical representation of safety requirements with respect to building elements or activities.

The ongoing research focuses on the following aspects. We are working on extracting hazard information and checking safety regulation conformity using construction domain ontology because ontology can play an important role in conceptual modeling and requirement analysis. Then, the framework will be tested through diverse case studies and by domain experts. We also have a plan to expand the scope of the module to include additional types of hazard information. This will extend the applicability and usability of the framework. In addition, the automatic implementation will be our future work.

REFERENCES

[1] U.S. Bureau of Labor Statistics, *Industry Injury and Illness Data - 2010*, viewed August 16 2012, http://www.bls.gov/iif/oshsum.htm

[2] Hecker, S., Gambatese, J., and Weinstein, M. "Designing for Worker Safety", *Professional Safety*, pp.32-44, Sep. 2005.

[3] Behm, M. "Linking Construction Fatalities to the Design for Construction Safety Concept", *Safety Science* Vol.43, pp.589-611, 2005.

[4] Gibb, A, Haslam, R., Hide, S., Gyi D., "The Role of Design in Accident Causality", *Proceedings of the Designing for Safety and Health in Construction Research and Practice Symposium*, Eugene, OR: University of Oregon Press, pp. 11-21, 2004.

[5] Prevention through Design – Design for Construction Safety, viewed August 16 2012, http://www.designforconstructionsafety.org/

[6] Gambatese, J., Behm, M., and Rajendran, S., "Designer's role in construction accident causality and prevention: perspectives from an expert panel", *Safety Science*, Vol.46(4), pp.675-691, 2008.

[7] WORKCOVER New South Wales, *Chair: Safety in Design Tool*, 2001.

[8] Cooke, T., Lingard, H., Blismas, N., Stranieri, A., "ToolSHeD: The development and evaluation of a decision support tool for health and safety in construction design, Engineering", *Engineering, Construction and Architectural Management,* Vol. 15(4), pp.336-351, 2008.
[9] Sacks, R., RozenFeld, O., "Spatial and Temporal exposure to Safety Hazards in Construction", *J of Constr Engrg & Mngmt,* Vol.135(8), pp.726-736, Aug. 2009.
[10] Benjaoran, V. and Bhokha, S., "An integrated safety management with construction management using 4D CAD model", *Safety Science*, Vol.48, pp.395-403, 2010.
[11] Zhang, S., Teizer, J., Lee, J.K., Eastman, C.M., Venugopal, M., "Build Information Modeling (BIM) and Safety: Automatic Safety Checking of Construction Models and Schedules", *Automation in Construction*, Elsevier, (in press), 2012.
[12] Yurchyshyna, A., Faron-Zucker, C., Thanh, N., and Zarli, A., "Knowledge capitalisation and organisation for conformance checking model in construction", *Int. J. Knowledge Engineering and Soft Data Paradigms*, Vol. 2(1), pp.15-32, 2010.
[13] El-Gohary, N. and Zhang, J., "Automated Information Extraction from Construction-related Regulatory Documents for Automated Compliance Checking", *Proceedings of the 2011 CIB World Congress*, 2011.
[14] Eastman, C., Lee, J., Jeong, Y., and Lee, J., "Automatic rule-based checking of building designs." *Automation in Construction,* Vol.18(8), pp.1011-1033, 2009.
[15] Akinci, B., Karimi, H., Pradhan, A., Wu, C., and Fichtl, G., "CAD and GIS Interoperability through Semantic Web Services", ITcon, Vol. 13, pp.39-55, 2008.
[16] Chandrasekaran, B., Josephson, J.R. and Benjamins, V.R, "What are Ontologies?, and Why do we need them?", *IEEE Intelligent Systems*, January/February, pp.20-26, 1999.
[17] Nepal, M.P, Staub-French, S., Zhang, J., Lawrence, M., and Pottinger, R., "Deriving Construction Features from an IFC Model", *CSCE 2008 Annual Conference*, Jun. 2008.