Information Infrastructure for Regulation Management and Compliance Checking

Shawn Kerrigan[1], Gloria Lau[1], Liang Zhou[2], Gio Wiederhold[2] and Kincho H. Law[1]
[1]Department of Civil and Environmental Engineering
[2]Department of Computer Science
Stanford University
Stanford, CA 94305-4020
Email Contact: law@ce.stanford.edu

1. Introduction

Developers and owners of facilities are facing more and more complicated regulatory and code requirements. These regulatory requirements do provide social benefits as they implement more stringent environmental laws (e.g., for hazardous waste management), and improve safety and accessibility (e.g., OSHA and ADA guidelines.) Regulations are typically specified by Federal as well as State governmental bodies and are often amended by local counties or cities. Regulations emanating from diverse agencies often overlap; because settings and objectives differ they may be inconsistent. As new issues of public safety or fairness arise new regulations are promulgated; old regulations are rarely removed. The distributed responsibilities increase the complexity of dealing with regulations. The scope of concern and the terminology used to express those concerns differs among agencies. It is nearly impossible for affected individuals and organizations to find and understand all applicable regulations. We believe that information technology (IT) has the potential to mitigate and help solve many of these problems.

In this research, we are developing a formal information infrastructure that will enhance the access and retrieval of government regulations. In the information service framework, government regulations will be made available on-line, and tools will be provided to locate, merge, compare, and analyze the information. Furthermore, we plan to illustrate the usage of the code representation by developing tools for compliance checking applications.

Building an integrated and distributed information management system for government regulation compliance checking is a complex task. In order to deal with fundamental research issues, it is necessary to focus on specific applications In this research, we focus on two sets of regulatory information: (1) disabled accessibility requirements and guidelines for building design, and (2) environmental regulations for hazardous waste management. For the disabled accessibility regulations, we focus on the *Americans with Disabilities Act Accessibility Guide* (ADAAG), the *Uniform Federal Accessibility Standards* (UFAS) and the *California Building Code* (or Title 24). For the environmental regulations, we currently deal with the *Code of Federal Regulation, Title 40* (Part 260-265 on hazardous waste management).

2. Research Methodology

We partition the research on regulatory management and compliance checking into six phases:

1. **Repositories:** In the first phase of the project, our goal is to investigate appropriate infrastructure and tools that will help build an online repository and populate them with available regulations for our pilot applications. The regulations to be collected emanate at the federal as well as at the state and local levels.

2. **Access Tools:** In the second phase, tools will be developed that will allow access of the regulation text and the other information gathered for the regulation repositories via the Internet. An Internet-enabled interface for the regulation repository will be developed. Our plan is, as much as possible, to build linkages to various public and proprietary services.

3. **Ontology Development:** This stage involves the development of a vocabulary and structural thesauri to represent the terms in the regulatory domain. The ontology allows the development of logical rules about the relationships and requirements of different regulations.

4. **Integrated Access:** The fourth task is to enable integrated access to the regulation repositories. Integrated access allows the retrieval of regulations based on the content or relationships between the regulations, as opposed to simple text matching. These access tools would assist with sophisticated retrieval of information, relying on the ontology that could help understand the content and relationships of regulations.

5. **Analysis Tools:** The idea behind developing analysis tools is that they make it possible to determine complementary or conflicting regulations based on the interpretation for these regulations we developed using our ontology. Analysis tools serve two purposes. First, they can help validate and improve the quality of the ontology by comparing the results of analysis tools with the results of expert or legal analysis. Second, they are useful for checking the content of regulations within a domain or across different domains of federal, state and local regulations.

6. **Compliance Checking:** An important objective of this research is to provide the means to interface the regulations with usage such that the regulations are not passive but active documents that can be dynamically linked to application programs for users to search and access regulations, to perform compliance check and for supporting human functions such as design activities.

Our current work focuses on the first step of building the repository. In addition, we have begun partial developments towards the second task in putting together simple access tools and the third task in extracting structural thesauri and ontology development.

3. Building the repository

The goal of the repository is to provide regulations in a standard, easy-to-use format, and to develop meta-data that describe the content of the regulations. First, starting with regulations in a variety of formats (pdf, text or html) the regulations are converted into an XML format. The XML regulation is tagged to show the basic document structure. Once the XML format and the structure of the regulation can be discerned, meta-data are added to the regulation. A regulation section can be tagged with related information, such as previous court ruling and expert interpretation which can be made accessible via the regulation section itself. The meta-data make regulations more convenient to search and easier for users to understand their content and relationships. Because of the volume of regulations involved in each domain, it is important to make the process in extracting and adding meta-data to the XML-coded regulations as automated as possible. Two specific approaches are being investigated:

1. One type of meta-data that can be added automatically is to extract links between the regulation provisions. Regulations tend to be very heavily cross-referenced, so identifying the links between them is important. Using a pattern-matching approach allows the system to extract references to related provisions and add these links as meta-data to the regulation provision.

2. A second type of meta-data that can be generated automatically is to identify what concepts a particular provision covers. Identifying the concepts in regulation provisions is important for matching the related provisions that do not explicitly reference each other. In this work, a commercial software package by Semio Corporation is used to facilitate the extraction of concepts from the voluminous amount of regulations.

Besides these automated procedures, manual processes are required, for examples, to examine the validity of the concepts extracted, to add expert and legal interpretation, etc.. Once the repository becomes fully functional on-line, we plan to examine and compare the regulations stored in the semi-structured (XML) database. A scalable approach based on knowledge composition will be employed to provide a systematic framework for management and comparison of regulations [SKC 1999, Wiederhold 1994]. Domain-specific terminologies and vocabularies will be extracted from the repository, and a structured thesaurus will be developed. Sections in the regulations will be linked to concepts in the thesaurus and also to related sections. Upon user queries on a particular concept in the provisions, the system will respond with related sections and present the user with conflicts among multiple sources of regulations, if there is any.

4. Compliance Checking

One of research objectives is to develop practical means to link the regulatory information with application programs for users to search and access regulations and to perform compliance check. Figure 1 illustrates a partially automated on-line compliance checking process using the Internet and web-based technologies [Han 2000]. The code-checking program examines the design data submitted on-line and summarizes the results in a generated web page. The web page contains a graphical representation of the building model along with "redline" information with hyperlinks to specific comments. For the example shown, the comments have hyperlinks to the actual building code document provisions (in this case, the Americans with Disabilities Act Accessibility Guide (ADAAG)).

In our current research, we have also been actively developing a simulation-based approach for performance evaluation of the provisions as specified by the intent of the accessibility guidelines [Han 2000]. Specifically, our work includes formalizing the tools required to simulate the accessibility routes and motion of wheelchairs. Such simulation tools may, in the future, play a role in performance based code evaluation and complement the text-based prescriptive code. It should be noted that it is unlikely that a simulation tool could cover all possible scenarios, as given in regulation codes and provisions. However, the proposed infrastructure will attempt to take advantage of simulation tools when they become available to provide appropriate evaluation (and possibly critics) of the sufficiency and performance of the regulations.

5. Summary and Discussion

This research aims to develop an information infrastructure for regulation management and compliance checking. In this paper, we briefly summarize our research methodology and planned activities. We also report some of the results related to building a regulation repository and

3

CLIENT SIDE   SERVER SIDE
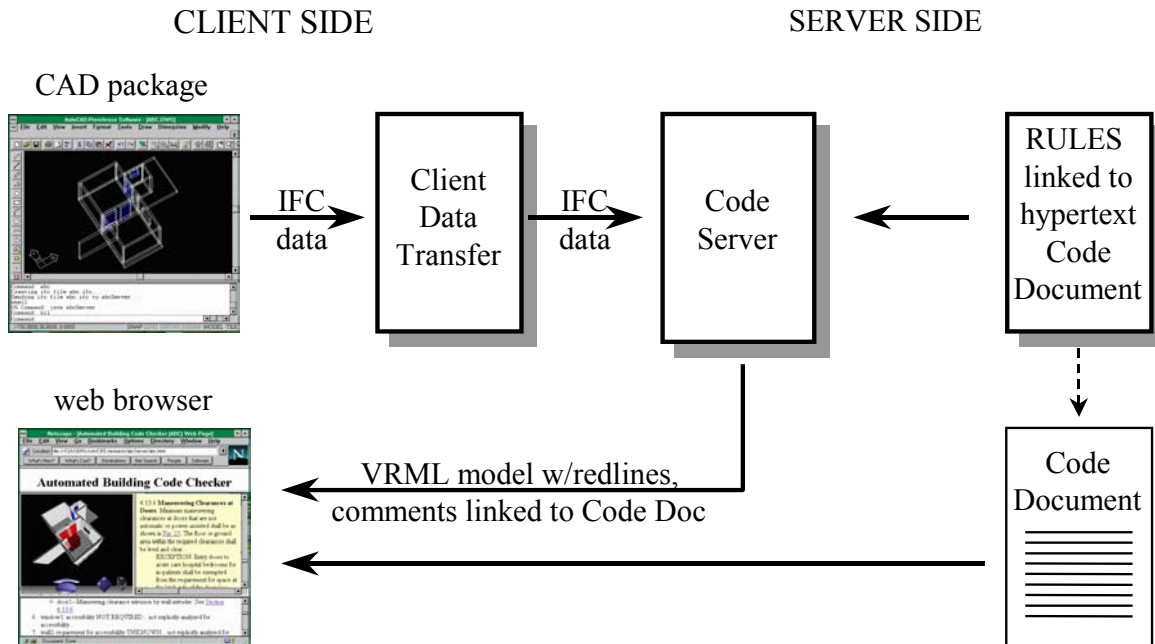
CAD package

Figure 1: An On-Line Compliance Checking Process

describe a possible application scenario for compliance checking of a design. While current research focuses on handicapped accessibility codes and hazardous waste regulations, the issues addressed and the tools developed should also found applications to other governmental regulations such as occupational safety and health rules, energy compliance codes, etc..

Acknowledgements

References:

Han, C.S., *Computer Models and Methods for a Disabled Access Analysis Design Environment*, Ph.D. Thesis, Department of Civil and Environmental Engineering, Stanford University, June, 2000.

SKC Project: Description of Scalable Knowledge Composition Project, Computer Science Department, Stanford University, http://www-db.stanford.edu/SKC/, 1999.

Wiederhold, G., "An Algebra for Ontology Composition," *Proceedings of the 1994 Monterey Workshop on Formal Methods*, US Naval Postgraduate School, Monterey, CA pp. 56-61, 1994.